Proceedings of the 1st European conference on disability, virtual reality and associated technologies (ECDVRAT 1996)

Book

Published Version

Conference Proceedings

Sharkey, P., ed. (1996) Proceedings of the 1st European conference on disability, virtual reality and associated technologies (ECDVRAT 1996). ICDVRAT. The University of Reading, Reading, UK, pp258. ISBN 070491140X Available at https://centaur.reading.ac.uk/27456/

Publisher: The University of Reading

The 1st European Conference on Disability, Virtual Reality and Associated Technologies

Proceedings

Paul M. Sharkey
Conference Chair/Editor

8, 9, 10 of July, 1996
Maidenhead, UK

ECDVRAT '96

The papers appearing in this book comprise the proceedings of the First European Conference on Disability, Virtual Reality and Associated Technologies, held on the 8th, 9th and 10th of July, 1996 at the Holiday Inn, Maidenhead, UK. The papers presented reflect the authors' opinions and are published as presented and without change (formatting excepted). Their inclusion in this publication does not necessarily constitute endorsement by the editor, by ECDVRAT, or by the University of Reading.

Please use the following format to cite material from this Proceedings:

> Author(s), "Title of paper", in Proceedings of the 1st European Conference on Disability, Virtual Reality and Associated Technologies, Paul M Sharkey, Editor, page numbers, Maidenhead, UK, July 1996.

ISBN 0 7049 1140 X

Published by

The University of Reading

For information, contact:
ECDVRAT, Department of Cybernetics, University of Reading, Whiteknights, Reading, RG6 6AY, UK
Phone: +44 (0)1734 318 219 Fax: +44 (0)1734 318 220

Printed in the UK.

Contents

Keynote Addresses

Session I Communication and Language Session Chair: Paul Sharkey

Session II Virtual and Enhanced Environments Session Chair: William Harwin

Session III Rehabilitation I Session Chair: John Wann

ISBN 0 7049 1140 X

Session IV Technology — Session Chair: Penny Probert

Session V Visual Impairment, Ambisonics and Mobility — Session Chair: Paul Wilson

Session VI Rehabilitation II — Session Chair: Stuart Neilson

Conference Committee

Conference Chair

Dr. Paul Sharkey, Department of Cybernetics, University of Reading, UK

International Programme Committee

Dr. David Brown, VIRART, Dept. of Manufacturing Engineering, University of Nottingham, UK.

Mr. Riadh Cammoun, Commisariat a Energie Atomique, Cedex, France.

Dr. Sylvie Collins, Dept. of Psychology, University of Leeds, UK.

Mr. Martyn Cooper, Technology & Disability Research Group, Dept. of Cybernetics, University of Reading, UK.

Prof. Paolo Dario, Scuola Superiore D. Anna, Pisa, Italy.

Dr. Alistair Edwards, Dept. Computer Science, University of York, UK.

Prof. Toshio Fukuda, Dept. of Mechano-Informatics and Systems, Nagoya University, Japan.

Prof. Ephraim Glinert, Computer Science Dept., Rensselaer Polytechnic Institute, Troy, NY, USA.

Prof. Jim Hewit, Dept. of Appl. Physics and Electronic & Manuf. Eng., University of Dundee, UK.

Dr. Javier Ibañez-Guzmán, ESA IGELEC, Saint-Nazaire, France.

Dr. Harry Murphy, Centre on Disabilities, California State University, Northridge, USA.

Dr. Stuart Neilson, Centre for Health and Disability Studies, Brunel University, London, UK.

Prof. Alan Newell, Applied Computer Studies Division, University of Dundee, UK.

Dr. Helen Petrie, Dept. of Health and Human Sciences, University of Hertfordshire, UK.

Mr. Chris Ramsden, Clinical Engineering and Design Systems, Royal Hospital for Neuro-Disability, London, UK.

Dr. David Roberts, Dept. of Cybernetics, University of Reading, UK.

Prof. David Rose, Dept. of Psychology, University of East London, UK.

Mr. Simon Rushton, Dept. of Psychology, University of Edinburgh, UK.

Dr. Paul Sharkey, Dept. of Cybernetics, University of Reading, UK.

Mr. Phil Smythe, BT, Martlesham Heath, Ipswich, UK.

Prof. Kevin Warwick, Dept. of Cybernetics, University of Reading, UK.

Dr. Paul Wilson, Virtual Reality Research Group, Dept. of Psychology, University of Leicester, UK.

Dr. Ali Zalzala, Dept. of Automatic Control and Systems Engineering, University of Sheffield, UK.

Organisation Committee

Mr. Martyn Cooper, Technology & Disability Research Group, Dept. of Cybernetics, University of Reading, UK.

Dr. Stuart Neilson, Centre for Health and Disability Studies, Brunel University, London, UK.

Dr. David Roberts, ***Conference Technical Director***, Dept. of Cybernetics, University of Reading, UK.

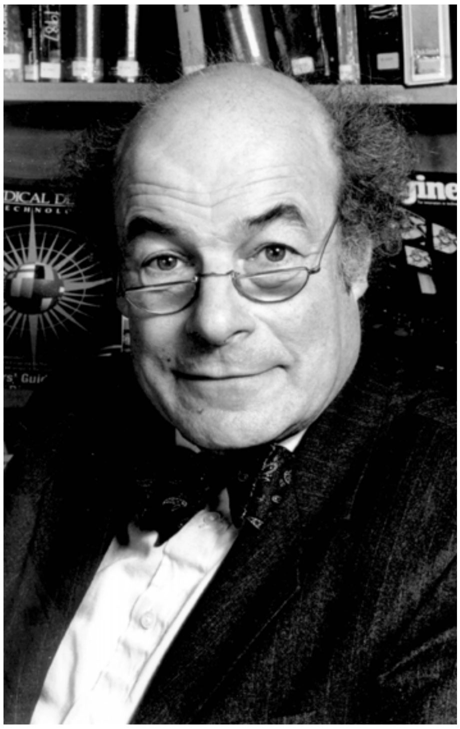
Dr. Paul Sharkey, Dept. of Cybernetics, University of Reading, UK.

Mr. Phil Smythe, BT, Martlesham Heath, Ipswich, UK.

 ISBN 0 7049 1140 X

 ISBN 0 7049 1140 X

Conference Banquet Speaker



Prof. H.S. Wolff, BSc, FIBiol, FIEE, FRSA, FBES

Professor Wolff emigrated from Germany with his family in September 1939. After school in London and Oxford, he worked for four years at the Radcliffe Infirmary at Oxford and at the Pneumoconiosis Research Unit near Cardiff (concerned with the design of medical and environmental measuring equipment) before reading Physiology and Physics at University College, London. In 1954 he graduated with First Class Honours.

He was initially employed in the Division of Human Physiology at the National Institute for Medical Research, where he specialised in the development of instrumentation suitable for field work. In 1962 he founded and became director of the Biomedical Engineering Division.

In 1971 he founded a Division of Bioengineering at the new Clinical Research Centre at Harrow, which was concerned with a wide spectrum of applications of technology to medical research and medical care, and with technology transfer to industry.

Since 1975 he has held a number of honorary appointments with the European Space Agency and until July 1991 was chairman of the Microgravity Advisory Committee, responsible for making policy for the scientific exploitation of the low gravity environment present on orbiting spacecraft. He has also served as one of the UK representatives on the European Science Foundation Space Science Committee.

In 1983 he founded the Brunel Institute for Bioengineering at Brunel University, Uxbridge, Middlesex. The Institute is financially totally self-supporting and holds contracts for work in space research, medical instrumentation and technology for the improvement of the quality of life for the elderly (Tools for Living). In connection with the latter activity, Professor Wolff was given the Harding Award for 1989, awarded alternately by Action Research and RADAR. In 1992 he was the recipient of the Edinburgh Medal, awarded in recognition of an outstanding contribution by a scientist to society. In 1987 he was made a Fellow of University College, London, and in 1989 a vice-president of the College of Occupational Therapists. In 1993 he received an Honorary Doctorate from the Open University and in 1993 was made a Fellow of the Institution of Electrical Engineers. In 1994 he was made a Fellow of the Biological Engineering Society, and in 1995 received an Honorary Doctorate from De Montford University, Leicester.

He has always been interested in the scientific and technical education of the young and lectures widely. He is perhaps best know by the public for a number of television series, such as Young Scientist of the Year, the Great Egg Race and Great Experiments Which Changed the World.

His personal interests range from the invention of high technology devices to the widespread and sensible application of technology to the problems of the elderly and disabled. He regards the communication of enthusiasm for science and engineering to the young as a very high priority.

The Conference banquet will be held in the Long Room of Shoppenhangers Manor, on Tuesday, 9th July, at 8 p.m.

 ISBN 0 7049 1140 X

Introduction

The purpose of this first European Conference on Disability, Virtual Reality and Associated Technologies is to provide a forum for international experts and researchers to present and review how advances in the general area of Virtual Reality can be used to assist people with Disability. The initial Call for Papers generated considerable interest, with high-quality contributions from researchers in several countries, many from beyond Europe.

The International Programme Committee have selected 30 papers for presentation at the conference, collected into 6 plenary sessions, *Communications and Language*; *Virtual and Enhanced Environments*; *Rehabilitation (I & II)*; *Visual Impairment, Ambisonics and Mobility*; and *Technology*.

Conference Addresses

The conference will be held over three days, with a Keynote Address at the start of each day.

The first Keynote Address will be given by Dr. Harry Murphy, Director of the Center on Disabilities, at California State University, Northridge (CSUN), who will review research in the fields of movement disorder, anxiety, and the treatment of injury. Dr Murphy co-ordinates two conferences on disability (Technology and Disability, and Virtual Reality and Disability) held each year in Los Angeles.

The second Keynote Speaker is Prof. David Rose, Head of the Department of Psychology, University of East London who will present an overview of how virtual reality relates to current research in neurological rehabilitation. Previously a Reader at Goldsmiths' College, University of London, Prof. Rose's principal research interest is in the recovery of function following brain damage. He is currently investigating how environment enrichment through the use of VR may be used for rehabilitation.

The Keynote Address on the final day of the conference will be given by Mr. Arthur Zwern, President of General Reality Company, who will introduce a novel system to enable visually impaired computer users to navigate about a computer screen using a head-mounted display interface. He will also discuss subject trials of the technology. Mr. Zwern has previously worked at Hughes Aircraft Corporation, developing some of the earliest full field-of-view 3D environments for battlefield simulation, and at KLA Instruments in the area of laser technology, before founding Voice Innovation and, in 1994, General Reality Company. He is the inventor of the Virtual Computer Monitor.

The Opening Address to the Conference will be given by Kevin Warwick, Professor of Cybernetics, Head of the Department of Cybernetics, University of Reading.

Workshop

In addition to the plenary presentations, there will be a parallel exhibition and workshop with ongoing access to virtual environments, demonstrations of some of the applications presented at the conference, access to Web browsers, and demonstrations from companies and other organisations of ongoing research related to the general subject area of the conference.

In addition to the parallel workshop, there will be an art exhibition, by disabled artists.

 ISBN 0 7049 1140 X

Acknowledgements

I would like to thank the Programme Committee, for their input to the conference format and focus, and to their commitment to the review process, the authors of all the papers submitted to the conference, the Organisation Committee, and the students who have helped out over the three days of the conference. I also wish to thank Harry Murphy and CSUN for continuing support for this conference.

Finally, I would like to especially thank BT, particularly Phil Smythe, for providing the initial sponsorship without which this Conference would surely have floundered.

On behalf of ECDVRAT I would like to welcome all delegates to the Conference. I sincerely hope that delegates find the conference and workshop to be of great interest. ECDVRAT welcomes any feedback that you can provide to enable us to improve for the next time round !

Paul Sharkey

Artwork and Conference Layout by *skelp*, adapted from original ideas of Eric Phipps and David Angus (ECDVRAT Logo) and from original clipart from the CorelDraw Clipart Library (ECDVRAT '96 Graphic). Eric Phipps and David Angus (both of Project DISCOVER) and *skelp* may be contacted through ECDVRAT.

 ISBN 0 7049 1140 X

Conference Sponsors

The principal sponsors of ECDVRAT '96 are BT, the Institution of Electrical Engineers, and the Department of Cybernetics, University of Reading.

The following companies have sponsored delegates with disabilities: BT; Department of Cybernetics, University of Reading; Silicon Graphics Limited; Virtual Presence; and the Holiday Inn, Maidenhead/Winsdor.[1]

The following companies and organisations provided equipment and/or demonstrations for the parallel workshop: BT; Department of Cybernetics, University of Reading; Department of Psychology, University of Reading; Silicon Graphics Limited; Virtual Presence; Division plc; Acting Up; Project DISCOVER; and Advanced Visual Presentations.[1]

Additional help in publicising the conference has been gratefully received from the Mechatronics Forum of the Institution of Mechanical Engineers, the Institution of Physics and Engineering in Medicine and Biology, the European Institute for Design and Disability and the National Rehabilitation Board (Ireland).

[1] Correct at time of press, 10 June 1996.

Proc. 1st Euro. Conf. Disability, Virtual Reality & Assoc. Tech., Maidenhead, UK, 1996
ISBN 0 7049 1140 X

ECDVRAT '96
Keynote Addresses

Day 1: 8 July, 1996

Virtual reality and persons with disabilities

Harry J. Murphy

CENTER ON DISABILITIES,
California State University, Northridge,
Northridge, California 91330 - 8340, USA.

Day 2: 9 July, 1996

Virtual reality in rehabilitation following traumatic brain injury

F. D. Rose

Department of Psychology,
University of East London,
London E15 4LZ, UK.

Day 3: 10 July, 1996

Virtual computer monitor for visually-impaired users

Arthur L. Zwern

General Reality Company,
124 Race Street,
San Jose, California 95126, USA.

Virtual reality and persons with disabilities

Harry J. Murphy

CENTER ON DISABILITIES
California State University, Northridge
Northridge, California 91330 - 8340
Phone: 818 885 2578
FAX: 818 885 4929

HMURPHY@CSUN.EDU

ABSTRACT

There are many exciting developments in the field of virtual reality. This presentation will use slides and videotapes to demonstrate three important applications of the technologies of virtual reality: one with people with movement disorders, another with people who experience height anxiety, and a third that deals with the treatment of injuries on the battlefield. More people in the field of disability need to become involved in the potential of virtual reality. Conferences such as this one play an important role in stimulating ideas that may result in more applications.

Keywords: overview, movement disorder, anxiety, injury, Center on Disabilities (CSUN)

1. INTRODUCTION

The CENTER ON DISABILITIES at California State University, Northridge (CSUN) has three major functions: (1) provide leadership to the Office of Disabled Student Services at the university, where 700 students with disabilities are assisted with educational support services such as counseling and tutoring; provide training to students with disabilities in assistive technologies which will assist them in their university studies, prepare them for employment, and help them live independently, (2) conduct major international conferences in the area of technology and disability, and (3) engage in a variety of grants and contracts such as our current training program, “Leadership and Technology Management,” (LTM).

As many of you know, the CENTER ON DISABILITIES conducts the largest, annual, international conference of its kind in the world, “Technology and Persons with Disabilities,” each March in Los Angeles. This conference deals with many issues in the area of technology and disability. We try to keep a focus on new and emerging technologies to bring these to the attention of our field, and to advocate for their use among people with disabilities. This conference brings more than 2,000 people each year to Los Angeles from almost every state and 25 foreign countries. There are about 150 exhibitors, and 300 speakers at this conference.

In 1991, the conference keynote speaker, Ted Saenger, President and CEO of Pacific Telesis, the large phone company on the west coast of the United States, spoke on “Virtuality.” In broad terms, Mr. Saenger addressed the issue of virtual environments including the virtual corporation and the virtual office, and the power of telecommuting for individuals with disabilities. He also whetted our appetite for virtual reality: three dimensional, interactive, computer generated worlds. It was clear that virtual reality had great potential for training, as well as other applications, to assist people with disabilities.

2. THE VIRTUAL REALITY FIELD BECOMES INTERESTED IN DISABILITY; THE DISABILITY FIELD BECOMES INTERESTED IN VIRTUAL REALITY

At about the same time that Ted Saenger spoke to us on “Virtuality,” we had contacted Jaron Lanier, who is credited with coining the term, “Virtual Reality.” Lanier at that time was chief scientist with VPL in Redwood City, California. We booked Lanier as keynote speaker for our 1992 conference, “Technology and Persons with Disabilities,” and invited him to bring virtual reality equipment (input glove, head mounted display, hardware and software) to the conference to demonstrate virtual reality to our participants.

 ISBN 0 7049 1140 X

Lanier did so to the great interest of the field of technology and disability. It become clear to us that we had identified a major developing technology. We assumed the obligation to follow this technology and keep its potential before the field.

To this end, we developed a separate conference devoted entirely to virtual reality and disability. This conference has been conducted in 1993, 1994, and 1995 in San Francisco. Keynote speakers have included Brenda Premo, Director, California State Department of Rehabilitation, Ray Bradbury, the famous science-fiction writer, and once again, Jaron Lanier. Attendees from the virtual reality field, and the disability field, have had the opportunity to mingle, discuss applications, and share ideas of what the future might hold for persons with disabilities.

The objectives of the separate conference,"Virtual Reality and Persons with Disabilities," are:

(1) engage the virtual reality field in disability issues; engage the disability field in the new technologies of virtual reality

(2) encourage new applications and provide a forum for dissemination of information

(3) engage the international community

3. PARTICIPATION OF THE INTERNATIONAL COMMUNITY

The CENTER ON DISABILITIES is very active in the international community. Participants from about 25 foreign countries participate in the conference, "Technology and Persons with Disabilities," and participants from about a dozen countries participated in our most recent "Virtual Reality and Persons with Disabilities" conference. I have keynoted conferences in six or seven countries and have provided technical assistance in Sweden and England to those wishing to develop "Virtual Reality and Person with Disabilities" conferences of their own. We strongly encourage our friends around the world to become more engaged in assistive technology (including virtual reality) field through conferences, research, and demonstration projects.

4. INTERESTING DEVELOPMENTS IN VIRTUAL REALITY AND DISABILITY

The major part of this presentation will be conducted through the showing of videotapes of leading applications. Several that are worthy of mention include:

(1) Suzanne Weghorst of the University of Washington. Dr. Weghorst has worked with people with movement disorders, such as Parkinson's Disease. She has developed a computer-generated world that trains a person to walk in relation to virtual objects. As you will see on the videotape, the training effect results in dramatic improvement.

(2) Ralph Lamson of Kaiser Hospital. Dr. Lamson has worked with people who fear heights. His computer-generated worlds provide training in incremental increases in distances. He has demonstrated reduction of fear through blood pressure and heart rate measures.

(3) Richard Satava of Walter Reed Hospital. Dr. Satava is a Colonel in the U. S. Army and a surgeon. His work includes a project in telepresence and will be shown in the videotape, "Modern Medical Battlefield." Through television and robotics, paramedics on site, and a surgeon at a remote location, can communicate and intervene in order to treat a wounded soldier at a battlefield location.

5. SUMMARY

There are many exciting developments in the field of virtual reality. Applications in the field of disability are at an early stage. Most of the applications have been developed with "off the shelf" technologies, and minimal specialized programming. More people in the field of disability need to become involved in the potential of virtual reality. Conferences such as this one play an important role in stimulating ideas that may result in more applications.

Virtual reality in rehabilitation following traumatic brain injury

F. D. Rose

Department of Psychology,
University of East London,
London E15 4LZ, UK

F.D.ROSE@UEL.AC.UK

ABSTRACT

The potential use of virtual reality (VR) in neurological rehabilitation has frequently been discussed. This paper relates current thinking on the subject to the clinically defined concepts of impairment, disability and handicap. It is concluded that VR has a contribution to make to reducing all three, as well as in the initial assessment of the consequences of traumatic brain injury.

Keywords: traumatic brain injury, assessment, impairment, disability, handicap

1. THE TECHNOLOGY OF VIRTUAL REALITY

> "Imagine a wrap-around television with three dimensional programmes, including three dimensional sound, and solid objects that you can pick up and manipulate, even feel with your fingers and your hands. Imagine immersing yourself in an artificial world and actively exploring it. Imagine that you are the creator as well as the consumer of your artificial experience, with the power to use a gesture or word to remold the world you see and hear and feel." (Rheingold, 1991)

More prosaically VR might be described as a computer technology which allows us to create detailed three dimensional representations of particular real life or imaginary situations, which can be examined and manipulated and within which one can move around.

The visual, auditory and tactile aspects of the computer generated virtual environment are delivered to the subject through visual display units and speakers within the helmet (head mounted display) and through heat and pressure emitting devices in "data gloves" or a "body suit" which looks a bit like a space suit. It was with good reason that Stewart (1992) described someone fully attired for VR as resembling "a mime in scuba gear"! These sensory experiences, in turn, are dependent upon the subject's movements within the environment which are relayed back to the computer from the helmet sensor, the hand held joy stick or other control device and the data gloves or body suit. For example, if a subject in a virtual room looks to the left or right, or up or down, the computer will detect these movements through the helmet sensor and alter accordingly the visual images relayed to the subject via the helmet display units. Similarly, if the subject touches an object in a VR environment the computer will detect this and deliver the sensation of tactile pressure through the pressure devices in the gloves.

Its most vital characteristics are "presence" , the feeling of being immersed in the computer generated world, - and that it is interactive. Within the virtual world created by the computer every response the "user" makes has a consequence to which he/she must adapt in terms both of mental processing and behaviour. What VR allows us to do is to temporarily isolate a person from his/her normal sensory environment and substitute for it an artificial computer generated environment built to the precise specifications of the programmer.

This sort of VR is at the forefront of research. Some aspects of it, for example the data gloves, require considerable further development (Durlach and Mavor, 1995). Financially it is at the very top end of the VR market, costing well in excess of £150,000. Consequently, its applications to real world problems are somewhat limited at present. However, VR technology is available at more affordable prices. A good quality immersive system based on the head mounted display but without the data gloves or body suit can be purchased for £65,000.

In addition to the immersive variety so far described VR also comes in a so-called "non-immersive" form which is very much cheaper still. Here the visual aspects of the computer generated VR environment are presented not through helmet mounted displays but on a conventional computer monitor or projected onto a screen. The subject controls his/her movement within the environment by means of a joy stick or other control device. In non-immersive VR it is as if the

 ISBN 0 7049 1140 X

subject is looking at the immersive VR environment, previously described, through a window (the computer monitor or screen). Of course as the size of this window is increased, from a conventional computer monitor, to a slide projection screen, to a series of screens which surround the subject (Browning, Cruz-Neira, Sandin et al, 1993), so the subject's sense of immersion can be increased. Although never conveying the sense of complete immersion in the environment associated with the more expensive immersive systems, non-immersive VR is, in other respects, comparable.

Clearly the potential uses of VR extend far beyond the computer games so often associated with it (see Durlach and Mavor, chapter 12). Already VR has been applied to the training of surgeons, pilots and drivers where training in a computer generated virtual world avoids the danger, expense and problems of monitoring and control associated with training in the real life situation. Similarly VR has been used by engineers, architects and designers to visualise real life structures (eg oil rigs, buildings, machinery) prior to actually building them, and by scientists to visualise systems difficult to illustrate in other ways (eg molecular interactions, effects of reduced gravity in various situations etc).

One area with very considerable potential for VR applications is the area of neurological rehabilitation (Rose and Johnson, 1994; Rose et al 1996; Rizzo and Buckwalter, 1995; Pugnetti et al, 1995). However, if those involved in the development of VR are to be of help to the clinicians and therapists working in this area of rehabilitation it is important that they familiarise themselves with the clinicians' terminology and the detail of the problems they seek to address.

2. TRAUMATIC BRAIN INJURY – THE SILENT EPIDEMIC

Traumatic brain damage has been labelled "the silent epidemic" (Klein, 1982). Frankowski, Annegers and Whitman (1985), in a review of seven major reports on the incidence of traumatic brain injury in the United States, arrived at an average incidence figure of approximately 250 per 100,000 of the population. This figure is in good agreement with statistics from Australia and the United Kingdom. Road accidents are the biggest single cause, accounting for between a third and a half of all such injuries. Other major causes are assaults and falls. Males are twice as likely to suffer a traumatic brain injury as females. Since the peak incidence falls in the 15 - 24 age range, the consequent problems are likely to be very long term and to become even longer term as medical science becomes ever more successful in increasing survival rates. The need for effective rehabilitative strategies is obvious. Moreover, these strategies must take account of practical considerations such as increasing pressure on staff time and the predominant age group within this category of patients. It is also crucial that there is clear agreement about the precise aims of whatever rehabilitative strategies are adopted.

The consequences of damage to the brain are usually described in terms of three concepts, impairments, disabilities and handicaps. These are defined as follows:

An impairment is:

> " - - - any loss or abnormality of psychological, physiological or anatomical structure or function." (World Health Organisation, 1980, p.27)

Disability refers to the effects of an impairment on a person's ability to perform an activity in a normal manner. The definition of this term is:

> " - - - any restriction or lack (resulting from an impairment) of ability to perform an activity in the manner or within the range considered normal for a human being." (World Health Organisation, 1980, p.29)

The final term, handicap, is defined as:

> " - - - a disadvantage for a given individual, resulting from an impairment or a disability, that limits or prevents the fulfilment of a role that is normal (depending on age, sex and social and cultural factors) for the individual." (World Health Organisation, 1980, p.29)

The term "impairment simply labels the effect of the injury on the brain and its function. The term "disability" assesses the impairment due to the brain injury in terms of its effects on what would be considered a normal profile of activities for a fit person. Finally, the term "handicap" places the disability within the context of that particular person's previous abilities, expectations and aspirations.

Whilst the model is not universally accepted (Johnston, 1996) the terms "impairment", "disability", and "handicap" define a progression of consequences of traumatic brain injury which link the initial injury with eventual outcome and, most importantly, identify targets for rehabilitation strategies.

VR has a potential role in rehabilitation of all three consequences of brain damage. However, it is important to recognise the differing levels of objectives in the three cases. For example, with an impairment the primary objective of VR rehabilitation may be to bring about clearly defined changes in brain structure or function; with a disability the objective in using VR may be to directly facilitate the learning process; and with a handicap the objective may be to use VR in a prosthetic manner.

However, before any rehabilitation can begin the patient's condition must carefully assessed. VR has a valuable role here also.

3. VIRTUAL REALITY IN ASSESSMENT OF REHABILITATION NEEDS

The first step in any successful rehabilitation programme is the accurate and comprehensive assessment of the patient's current abilities. Whilst the methodology for assessing sensory and motor capacities is well developed there is a continuing debate about how best to assess cognitive functions such as attention, memory and reasoning and how what is measured under these headings relates to practical skills in real life settings. This is the issue of the "ecological validity" of measures. The 1970s and 1980s saw a reaction against laboratory based measures of cognitive function which had been developed within the mainstream of experimental psychology. Increasingly these were seen as too narrow and artificial to give an accurate guide to cognitive function in real life situations and, in consequence, several tests of "everyday" cognitive function have been developed. However, in their turn these so-called "ecologically valid" measures have been criticised for a lack of rigorous control of the test situation (Banaji and Crowder, 1989).

Andrews et al (1995) have suggested that a possible solution to the problem lies in measuring cognitive function within a VR environment. VR allows the measurement of cognitive function to be made in the context of interaction with a realistic everyday environment without sacrificing the opportunity to maintain strict control over every aspect of the test situation. For example, a distractible patient's ability to carry out simple procedures in a kitchen can be tested without there being any danger. Since it is under the control of the computer programmer, by definition, VR allows precise and detailed analysis of the environment within which cognitive function is assessed and to which the subject's responses are being measured. Moreover, since interaction with a VR environment can be made contingent on a wide range of motor responses it is possible to measure cognitive function in an everyday situation in people whose motor disabilities restricts movement in the real world. Similarly, particular aspects of the sensory array which are presented to the patient can be artificially enhanced to help overcome partial sensory loss (Middleton, 1992).

In other words VR allows us to place a neurological patient in a variety of precisely controlled simulated environments which are entirely safe and which are intentionally programmed in order to avoid, as far as possible, the patient being at a disadvantage as a consequence of sensory or motor impairments. Having created this "level playing field" we can simply observe how the patient behaves - how active, how interested, how distractible -as well as seeking to investigate particular cognitive functions.

The work referred to above (Andrews et al, 1995) was concerned with using VR to measure memory. VR is also currently being used to develop measures of spatial memory and of contralateral neglect although results have not yet been published.

In providing ecologically valid measures of cognitive function VR will provide a valuable resource for those involved in cognitive rehabilitation. However, VR has the potential to contribute to a more fundamental type of assessment of traumatic brain injury. Frequently, clinicians are confronted with patients whose level of responsiveness is so reduced that it is difficult to arrive at any reasonable estimate of their residual abilities. An extreme instance is patients who are in a persistent vegetative state (PVS). In 1993 the High Court granted permission to those caring for Hillsborough victim, Tony Bland, to remove his nasogastric feeding tube and thus allow him to die. Since then there has been a great deal of debate about the ethical and legal problems raised by PVS patients (Jennett, 1993; Andrews, 1993) and understandable public concern about the adequacy of the methods available to assess the true cognitive function of such patients. As Murphy (1995) has observed, neither brain scanning (CT/MRI) nor electroencephalography (EEG) can reliably predict or detect PVS and the label is ultimately applied to patients on the basis of "clinical judgement". The concern of both the general public and the clinicians is that the failure to show "consistent psychologically meaningful responses" (Jennett, 1993) to a sensory stimulus, which is characteristic of PVS, might be overcome if only the right stimulus or stimulus combinations could be found.

Fully immersive VR (the headset, body suit and data gloves), in combination with suitable input/output devices to interface between the patient's behavioural and physiological responses (Knapp and Lusted, 1992), could be of great help here. VR would allow the patient to be exposed to a sensory world of a complexity which it is impossible to deliver in any other way. Sensory stimuli could be delivered singly, or in combination, without context or in meaningful and familiar contexts (the patient's own home or workplace could easily be recreated), and over prolonged periods of time. Moreover, responses to stimulation could be sought, not only in overt behaviour, but in subtle movements and a range of physiological responses. Certainly VR has the potential to improve on existing assessments of sensory responsiveness (Freeman, 1993) in maximising the chance of identifying the right combination of stimuli and minimising the chance of missing a meaningful response. Of course, if this application of VR were to be successfully developed it could also be used to deliver sensory stimulation for therapeutic purposes (Wilson and McMillan, 1993).

4. VIRTUAL REALITY IN REHABILITATION OF IMPAIRMENT

As will be apparent from the definitions given above, the term impairment is used to refer primarily to compromised anatomy and physiology and its immediate functional consequences. Can VR be used to reduce impairments? Although as yet there is little direct evidence that it can, a case can certainly be made that it should.

A primary effect of most forms of neurological insult is a reduction in cerebral arousal - activation. This combines with other common neuropsychological impairments, for example in attention, memory and motivation, to result in significantly reduced levels of interaction between the patient and his/her environment whether at home or on the ward. Coexisting sensory and motor impairments, such as hemiplegia, typically restrict interaction still further.

Clinicians agree that this is undesirable and that environmental interaction is vital to the rehabilitation process. Moreover their clinical judgement is supported by an extensive scientific literature. Neuroplasticity, the brain's capacity to modify its structure and function in response to experience gained from interaction with the environment is no longer in doubt (Rose and Johnson, 1996). Animal studies have shown that increased levels of environmental interaction result in a more highly developed and more efficient brain (Renner and Rosenzweig, 1987). Brain changes attributable to increased interaction include a greater amount of cerebral cortex (the part of the brain primarily associated with cognitive function), more profuse connections between cortical neurons, increased activity in glial cells and a higher metabolic rate within the cerebral cortex. Increased environmental interaction has also been shown to enhance functional recovery following many types of brain damage in animals and, in particular following damage to the cerebral cortex (Rose, 1988; Will and Kelche, 1992).

Conventional therapies in neurological rehabilitation, such as physiotherapy, occupational therapy and speech therapy, involve increased levels of interaction, of course. However, research has shown that, for example, stroke patients typically spend only 30 to 60 minutes per day in formal therapy (Tinson, 1989). Consequently, there are lengthy periods in the patient's day when levels of environmental interaction are quite low.

Here, again, VR can be of help to the therapist. VR provides a powerful means of increasing levels of environmental interaction. Importantly it is generally a compelling experience, and largely inescapable, unlike more conventional computer based cognitive rehabilitation programmes (Bradley, Welch and Skilbeck, 1993). Moreover, as noted above, since interaction with a VR environment can be made contingent upon whatever motor capacity the patient has, and also take account of sensory impairments, this technology is eminently well suited to this therapeutic interaction.

The therapeutic role of environmental interaction in the brain damaged rat is not yet fully understood. Moreover, there is always a question mark over the extent to which data derived from rats can be extrapolated to humans. Nevertheless, it is possible that interaction with a virtual environment might have direct effects on the brains of neurological patients, increasing their efficiency and, consequently, maximising their functional output. The possibility of success, given that VR is a non-invasive and low risk strategy, would certainly indicate further investigation.

Indeed the argument may be even more compelling than that. Just as the animal studies indicate that environmental enrichment causes thickening of the cerebral cortex of the brain, so they also show that the opposite of enrichment, environmental impoverishment, causes cortical thinning (i.e. actually damages the brain). Often, in discussing traumatic brain damage a distinction is drawn between different injuries. The first injury is due to the impact between brain and windscreen, road, or whatever. The second injury, which is to an extent avoidable, involves further damage to the brain as a result of complications arising from the first injury, for example, a sudden drop in blood pressure which damages brain tissue by starving it of oxygen, or the swelling of the brain which damages brain tissue by crushing it. The third injury, degenerative damage, occurs as nerve pathways severed by the initial impact degenerate causing secondary damage elsewhere in the brain. However, if as a result of brain damage patients become less interactive (their environments effectively become impoverished) this may represent yet another injury, the fourth injury. VR therapy, then, may be the means not only of enhancing function but also preventing further damage.

As this particular application of VR to rehabilitation is developed it will be necessary to investigate the effects of exposure to VR on both nervous system structure and function and on behaviour in order to establish whether interaction with a virtual environment in neurologically impaired humans can be equated with changes associated with interaction with the real environment in brain lesioned animals. Already there have been reports of psychophysiological changes during interaction with a VR environment (Pugnetti, et al, 1994) and Decety, et al (1994) have carried out PET scans following exposure to VR. However, as yet nervous system changes accompanying interaction with VR environments is a largely unexplored area.

5. VIRTUAL REALITY IN THE REHABILITATION OF DISABILITY

The remedy for a failure to perform a normal activity in a normal manner (i.e. a disability) would most obviously seem to lie in the training process. As noted above, VR has proved to be a valuable training aid where training in real life situations would be impractical because, for example, it would be dangerous or unduly expensive. In the specific context of neurological rehabilitation VR also has great potential where training in real life situations, although not totally impractical, is made difficult because of the patient's sensory, motor and cognitive disabilities. Because the virtual training situation has been constructed in every last detail to the computer programmer's specification, certain aspects or categories of sensory stimuli can be accentuated to offset partial sensory impairment. The salience of stimuli and the links between them can also be emphasised to offset some of the effects of cognitive impairments. Once again movement within the training situation can be precisely geared to whatever motor abilities the patient has. There are also other benefits. In terms of staff resources it is clearly less time consuming to train a patient in a controlled and danger free VR environment than in a real life environment which is much less predictable and possibly fraught with danger for both the patient and other people.

A variation on the VR theme which has particular value in a training context is known as "augmented reality". By placing half silvered optical surfaces within the head mounted display, positioned in front of each eye, it is possible for the subject to retain a view of the real world but to have superimposed upon it virtual images. For example, a trainee surgeon could have a virtual image of the vascular system superimposed upon the area of the real body to guide his/her incision. In a neurological rehabilitation context, for example, linguistic labels could be superimposed upon real objects in the environments of patients suffering from object agnosia.

VR is currently being developed as a training aid in several neurological rehabilitation contexts. Emmett (1994), using the knowledge that despite their difficulty in walking Parkinson's patients do step over objects placed in their paths, presented visual obstacles via a head mounted display to achieve normal gait. VR has also been used to develop everyday living skills for children with severe learning disabilities (Brown and Wilson, 1995; Brown, Stewart and Wilson, 1995). This system consists of three programmes, a virtual house which includes an interactive kitchen, a virtual city for developing traffic sense, and a virtual supermarket to train subjects to choose and pay for goods. Currently, it is being used in ten special schools. A somewhat similar endeavour is Mowafty and Pollack's (1995) "Train to Travel". This project was devised to enable people with cognitive impairments to use public transport. Following training on basic skills, including recognition of landmarks, the students are immersed in a simulation of an actual fixed bus route for as many rides as they need. This form of training eliminates the need for a teacher to accompany them on real trips, which can take up to fifteen hours to reach the same level of competence. In addition to these published examples (not an exhaustive list) there are ongoing attempts to use VR in training patients to overcome impairments in attention, incidental and spatial memory, visuospatial function and to correct contralateral neglect.

The arguments for applying VR specifically to the training of patients with traumatic brain injuries have been rehearsed on several occasions (see, for example, Rizzo and Buckwalter, 1995: Rose et al, 1996). Without doubt VR embodies many of the characteristics of an ideal training medium (Darrow, 1995) and, in this respect, has significant advantages over traditional computer based cognitive rehabilitation formats (Bradley et al, 1993). However, as yet there appear to be no published empirical evidence of the efficacy of its use in this context (Darrow, 1995). It is to be hoped that this situation will have been remedied before the next European Conference on Disability, Virtual Reality and Associated Technologies.

Of course central to any claims of success for these training procedures is the demonstration that what is learned in VR transfers to real life situations. Whilst some have made such claims (Standen and Cromby, 1995; Wilson, 1993) others have not found any significant transfer (Foreman et al - personal communication; Kozac et al, 1993). For a fuller review of evidence on transfer from the virtual to the real world, see Rizzo and Buckwalter (1995). Certainly further systematic research is needed on this transfer of training issue.

6. VIRTUAL REALITY IN THE REHABILITATION OF HANDICAP

The term handicap refers to the disadvantage for a given individual resulting from impairment or disability. Alternatively handicap might be described a pattern of difficulty experienced by a person as a consequence of the juxtaposition of a particular pattern of impairments and disabilities, on the one hand, and, on the other, that person's lifestyle and aspirations. If everything possible is already being done to reduce impairments and disabilities and if life style and aspirations are not to be compromised more than necessary, the use of prosthetic devices is indicated. Examples include, at the simplest level, notebooks to help overcome memory problems and, at a more technological level, devices for utilising the eye movements of paralysed patients to operate keyboards. VR also has a potential contribution to make here. Indeed the addition of VR to the existing technologies employed in prosthetics promises to revolutionise the lives of the disabled. However, progress depends upon progress not only in VR technology but also in the development of input devices which interface between whatever response repertoire the patient has (Knapp and Lusted, 1992) and the virtual world, and developments in robotics which allow the patient's actions in the virtual world to be translated into actions in the real world.

Of all the uses of VR in neurological rehabilitation this is the one which perhaps requires most development. But it is also one of the most exciting. Ultimately one can imagine a situation in which the disabled might be able to carry out tasks in their real world environments by operating within a linked virtual version of it. Recent work by Simsarian and Fahl (1995) investigates this possibility. Lasko-Harvill (1993) has claimed that: "In VR the distinction between people with and without disabilities disappears." Exaggeration for the sake of emphasis, perhaps. However, the combined resources of VR and robotics promises to empower the disabled to an extent undreamt of even a few years ago. However, there are others at this conference much better qualified to review work on the role of VR in prosthetics.

7. VIRTUAL REALITY AND CLINICAL ACCEPTANCE

Remarkable as it is VR is a technology which does seem to attract rather more than its fair share of hyperbole. Discussing developments in medicine, nanotechnology and VR, Fruchterman asserts that: "We are on the threshold of the Age of Magic - - - ." (Fruchterman, 1992, p.15) Still more expansive is a quotation from John Perry Barlow:

> "I - - am in cyberspace, a universe churned up from computer code, then fed into my eyes by a set of goggles through whose twin video screens I see this new world. All that remains of my corporeal self is a glowing, golden hand floating before me like Macbeth's dagger. I point my finger and drift down its length to the bookshelf on the office wall -"
>
> Barlow (quoted by Marcus, 1992)

If those who advocate the use of VR in rehabilitation are to be taken seriously by clinicians it is perhaps advisable to curb the temptation to wax lyrical in this way. The realities of day to day life on a traumatic brain injury rehabilitation unit do not sit happily with such, as yet, unsupported promise. We should continue to be mindful of the observation made in the National Academy of Sciences report (Durlach and Mavor, 1995), that so far for VR the "excitement to accomplishment ratio" remains high.

Clinicians will have other concerns as well. For example they may see cost and technical complexity as a barrier to the development of VR therapy. There is also, as with all new treatments, an ethical question. As medical interventions go a VR based therapy is unlikely to be particularly contentious. It is non - invasive and must be seen as a low risk strategy. However, there is evidence that exposure to VR, particularly immersive VR, can have side effects. These include visual disturbances and motion sickness.

Mon-Williams et al (1993) reported that normal healthy subjects experienced transient reduced binocular vision after wearing a head mounted display for just ten minutes. They argued that this was due to compromised visual experience resulting from the generation of 3-D visual space from 2-D images. However, subsequent research (Rushton et al, 1994) found that 'new-generation' head-sets, such as the Visette 2000 HMD, did not lead to changes in binocular function even after thirty minutes use. Nor did these authors find any significant changes across a range of visual performance measures. Nonetheless, Rushton et al (1994) caution that these findings do not apply to stereoscopic displays, and evaluations of this type of display still need to be carried out. Related to the question of visual symptoms is that of nausea following use of a head mounted display. In immersive VR environments the display devices utilise a number of oculomotor systems. The incongruity between visual and vestibular motion cues can produce sickness symptoms. This so-called "simulator sickness" produces similar symptoms to motion sickness, including disorientation, sweating, nausea, headache, and general discomfort. Regan (1995) found that 61% of subjects (n=150) reported symptoms at some stage during a 20 minute immersion and 10 minute post-immersion period. However, Kolasinki (1995) has found considerable variability in the extent to which subjects in VR do suffer from this condition. No significant side effects have been reported with non immersive VR.

If VR is to be used clinically it is important that its effects on bodily systems be thoroughly investigated (Eberhart and Kizakevich, 1993) but, on the basis of research so far, side effects do not appear to represent a serious barrier to the use of VR in neurological rehabilitation. However, it is important to remain vigilant. For example, Middleton (1992) warns that immersion in VR can be quite disorienting for the hearing impaired or deaf and great caution seems advisable at this stage in using VR with patients displaying psychiatric symptoms. As we develop VR as a therapy it is probable that we will discover still more exclusion criteria. Certainly they should not deter us - any more than cost and technical complexity should deter us - from seeking to exploit the very considerable potential for VR in brain damage rehabilitation.

Although a relatively young technology the very great potential of VR to improve the lives of those with traumatic brain injury, and other types of brain damage, is not in dispute. However, in order to fulfil this potential we must invest great effort in research, development and evaluation. To this end we need to forge sound working relationships with our clinical colleagues and take their perspective and their aims as our starting point. We have just entered the second half of the "Decade of the Brain" (Goldstein, 1990). It is not unrealistic to hope that by the end of the decade VR therapy will be part of the clinician's armoury in tackling the problems of traumatic brain injury.

8. REFERENCES

Andrews, K. (1993). Should PVS patients be treated? *Neuropsych. Rehab.*, **3**, pp109-119.

Andrews, T.K., Rose, F.D., Leadbetter, A.G., Attree, E.A. & Painter, J. (1995). The use of virtual reality in the assessment of cognitive ability. In: *The European Context for Assistive Technology: Proceedings of the 2nd TIDE Congress* (I. Placencia Porrero & R. Puig de la Bellacasa, Eds), IOS Press: Amsterdam pp 276-279.

Banaji, M.J. & Crowder, R.C. (1989). The bankruptcy of everyday memory. *Am. Psychol.*, **44**, pp1185-1193.

Bradley, V.A., Welch,J.L. and Skilbeck, C.E. (1993). *Cognitive Retraining Using Microcomputers*. Lawrence Erlbaum Associates, Hillsdale, USA.

Brown, D. & Wilson, J. (1995). LIVE: Learning in virtual environments. *Ability*, **15**, pp24-25.

Brown, D.J., Stewart, D.S. and Wilson, J.R. (1995). Ethical pathways to virtual learning. In: Murphy, H.J. (ed), *Proc: Third Ann. Intnl. Conf. "Virtual Reality and Persons With Disabilities"* San Francisco.

Browning, D.R., Cruz-Neira, C., Sandin, D.J. & DeFanti, T.A. (1993). CAVE Automatic virtual environment: Projection-based virtual environments and disability. In: Murphy, H.J. (ed), *Proc: First Ann. Intnl. Conf. "Virtual Reality and Persons With Disabilities"* San Francisco, pp17-22.

Darrow, M.S. (1995), Virtual reality's increasing potential for meeting needs of persons with disabilities: What about cognitive impairments? In: Murphy, H.J. (ed), *Proc: Third Ann. Intnl. Conf. "Virtual Reality and Persons With Disabilities"* San Francisco.

Decety, J., Perani, D., Jeannerod, M., Bettinardi, V., Tadary, B., Woods, R., Mazziotta, J.C. & Fazio, F. (1994). Mapping motor representations with positron emission tomography. *Nature*, **371**:600-602.

Durlach, N. I. and Mavor, A. S. (Eds.) (1995). *Virtual Reality. Scientific and Technical Challenges.* National Academy Press, Washington, D.C.

Eberhart, R.C. & Kizakevich, P.N. (1993). Determining physiological effects of using VR equipment. In: Murphy, H.J. (ed), *Proc: First Ann. Intnl. Conf. "Virtual Reality and Persons With Disabilities"* San Francisco, pp47-49.

Emmett, A. (1994). Virtual reality helps steady the gait of Parkinson's patients. *Computer Graphics World*, **17**, 11, pp17-18.

Frankowski, R.F., Annegers, J.F. and Whitman, S. (1985). Epidemiological and descriptive studies. Part 1: The descriptive epidemiology of head trauma in the United States. In *Central Nervous System Trauma Status Report.* (D.P.Becker and J.T.Povlishock, Eds.) National Institutes of Health, USA, pp 33-43.

Freeman, E.A. (1993). The clinical assessment of coma. *Neuropsych. Rehab.*, **3**, pp139-147.

Fruchterman, J.R. (1992) The age of magic. In H. Murphy, (ed.) *Proc. of Virtual Reality and Persons with Disabilities*, Los Angeles, pp 15-20.

Goldstein, M., (1990) The decade of the brain, *Neurology*, **40**, p321.

Jennett, B. (1993). Vegetative survival: the medical and ethical dilemmas. *Neuropsych. Rehab.*, **3**, 99-108.

Johnston, M. (1996). Models of disability. *The Psychologist.*, **9**, 205-210.

Klein, F.C. (1982). Silent epidemic: head injuries difficult to diagnose, get rising attention, *Wall Street Journal*, 24th November.

Kolasinski, E.M. (1995). Simulator sickness in virtual environments. *Tech. Report 1027: U.S. Army Res. Inst. for the Behav. and Soc. Sci.*

Kozac, J.J., Hancock, P.A., Arthur, E.J. & Chrysler, S.T. (1993). Transfer of training from virtual reality. *Ergonomics*, **36**, 7, pp777-784.

Knapp, B. & Lusted, H.S. (1992). Biocontrollers: A direct link from the nervous system to the computer. In: *Proc., Medicine meets Virtual Reality*, University of California at La Jolla, San Diego.

Lasko-Harvill, A. (1993). User interface devices for virtual reality as technology for people with disabilities. In: Murphy, H.J. (ed), *Proc.: First Ann. Intnl. Conf. "Virtual Reality and Persons With Disabilities"* San Francisco, pp59-65.

Marcus, S. (1992) Virtual realities: from the concrete to the barely imaginable. In H. Murphy, (ed.) *Proc. of Virtual Reality and Persons with Disabilities,* Los Angeles, pp 31-35.

Middleton, T. (1992) Matching virtual reality solutions to special needs. In H. Murphy, (ed.) *Proc. of Virtual Reality and Persons with Disabilities*, Los Angeles, pp 37-45.

Mon-Williams, M.A., Wann, J.P., Rushton, S.K. & Ackerley, R. (1993). Real problems with virtual worlds. *Ophthalmic and Physiological Optics*, **13**:435-436.

Mowafty, L. & Pollack, J. (1995). Train to travel. *Ability*, **15**:18-20.

Murphy, L. (1995), Cognition and the persistent vegetative state. *The Psychologist*, 8, 445-447.

Pugnetti, L., Mendozza, L., Cattaneo, A.M., Guzetti, S., Mezzetti, C., Coglati, C., Dell'Abate, E., Brancotti, A. & Venanzi, D. (1994). Psychophysiological activity during virtual reality (VR) testing. *Journal of Psychophysiology*, **8**(4):361-362.

Pugnetti, L., Mendozzi, L., Motta, A., Cattaneo, A. and Barbieri, E. (1995). Evaluation and retraining of adults' cognitive impairments: which role for virtual reality technology? *Computers in Biology and Medicine*, 25, 213-227.

Regan, C. (1995) An investigation into nausea and other side-effects of head-coupled immersive virtual reality. *Virtual Reality*, 1, 1, 17-32.

Renner, M.J. & Rosenzweig, M.R. (1987). *Enriched and impoverished environments: Effects on brain and behaviour*. Springer-Verlag: New York.

Rheingold, H. (1991). *Virtual Reality*. Secker & Warburg: London.

Rizzo, A.A. and Buckwalter, J.G. (1995). Theoretical and practical issues for the use of virtual reality in the cognitive rehabilitation of persons with acquired brain injuries - an update. In: Murphy, H.J. (ed), *Proc.: Third Ann. Intl Conf. "Virtual Reality and Persons With Disabilities"* San Francisco.

Rose, F.D. (1988). Environmental enrichment and recovery of function following brain damage in the rat. *Med. Sci. Res.*, **16**, pp257-263.

Rose, F.D. & Johnson, D.A. (1994). Virtual reality in brain damage rehabilitation. *Med. Sci. Res.*, **22,** p82.

Rose, F.D. and Johnson, D.A. (1996). Brains, injuries and outcome. In *Brain Injury and After*. (F.D.Rose and D.A.Johnson, Eds) Wileys, Chichester, pp 1-20.

Rose, F.D., Johnson, D.A., Attree, E.A., Leadbetter, A,G. and Andrews, T.K. (1996). Virtual reality in neurological rehabilitation, *Br. J. Ther. Rehab*. **3**, pp223-228.

Rushton, S., Mon-Williams, M. & Wann, J. (1994). Binocular vision in a bi-ocular world.: New generation head-mounted displays avoid causing visual deficit. *Displays*, **15**, pp255-260.

Simsarian, K.T. and Fahl, L.E. (1995). Using virtual and augmented reality to control an assistive mobile robot. In: Murphy, H.J. (ed), *Proc.: Third Ann. Intl. Conf. "Virtual Reality and Persons With Disabilities"* San Francisco.

Standen, P.J. and Cromby, J.J. (1995) Can students with developmental disability use virtual reality to learn skills which will transfer to the real world? In: Murphy, H.J. (ed), *Proc.: Third Ann. Intnl Conf. "Virtual Reality and Persons With Disabilities"* San Francisco.

Stewart, D. (1992). Through the looking glass into an artificial world - via computer simulation. *Smithsonian*, **21,** 10, pp36-45.

Tinson, D.J. (1989). How stroke patients spend their days: An observational study of the treatment regime offered to patients with movement disorders in hospitals following stroke. *International Disabil. Stud.*, **11**, pp45-49.

Will, B. & Kelche, C. (1992). In *Recovery from brain damage: Reflections and directions*. (Rose,

F.D. & Johnson, D.A., eds). Plenum Press: New York.

Wilson, P. (1993) Nearly There. *Spec. Children*, **68**, pp28-30.

Wilson, S.L. and McMillan, T. (1993). A review of the evidence for the effectiveness of sensory stimulation treatment for coma and vegetative states, *Neuropsych. Rehab.*, **3**, pp149-160.

World Health Organisation (1980). *International Classification of Impairments, Disabilities and Handicaps*. Geneva: World Health Organisation.

This paper is based, in part, on a paper previously published in the British Journal of Therapy and Rehabilitation and on chapter 1 of "Brian Injury and After. Towards Improved Outcome" (Rose,F.D & Johnson,D.A., eds. Wiley, 1996).

Virtual computer monitor for visually-impaired users

Arthur L. Zwern[1] and Gregory L. Goodrich[2]

[1]General Reality Company
124 Race Street
San Jose, California 95126 USA
(408)289-8340

[2]USDVA Western Blind Rehabilitation Center
3801 Miranda Avenue
Palo Alto, CA 94304 USA

[1]*Arthur@genreality.com*, [2]*Goodrich@roses.stanford.edu*

ABSTRACT

Conventional computer display products for the visually impaired are limited by the amount of enlarged imagery that can be displayed at any one time, and by awkward methods for navigating about the scene. This paper describes a prototype system designed to address these problems by providing a head-mounted display interface which allows the user to position a cursor anywhere on an enlarged virtual page by turning to view that location in space. Also discussed are human subject trials underway at SRI International and the Veterans Administration under a US Department of Health and Human Service research grant, and technical issues requiring further investigation.

Keywords: low vision, adaptive devices, computer access, head-mounted displays, screen enlargers

1. BACKGROUND: THE FIELD NAVIGATION PROBLEM

Among the visually-impaired population, the most common approach to computer access is specialized software that magnifies the image on the computer monitor. This is because simpler solutions such as moving closer to the monitor, using a larger monitor, adding a screen magnifier, or using a spectacle-mounted telescopic system provide either limited magnification or a very limited viewing field.

These computer display magnification solutions operate by magnifying the original image of the computer screen to a "virtual screen" whose size is much larger than the physical monitor. The visually-impaired user then operates the computer by using a mouse, joystick, or cursor keys to control which portion of the virtual screen is shown on the physical monitor at any one time.

Unfortunately, there are two basic shortcomings to this approach. The first problem is spatial orientation, in that it is difficult to determine where on the page one's view is directed at any given time. At 15x magnification, a 15" monitor can only display about 1% of a standard 8.5"x11" page, making most computer work essentially impossible. The second problem is dynamic control, in that all of the various control schemes for navigating about the page are cumbersome, confusing, and slow. Together, these problems were termed the "field navigation" problem in the US National Advisory Eye Council's 1994-98 National Plan, in which the Low Vision and its Rehabilitation Panel identified "text navigation" as a particularly promising opportunity for new technologies (Legge et al, 1994).

2. VIRTUAL REALITY TECHNOLOGY APPLIED TO FIELD NAVIGATION

In response, the Virtual Computer Monitor (VCM) was conceived by General Reality, and a breadboard first constructed in mid-1994. This breadboard received an award as "Most Innovative New Device" at CSUN's Virtual Reality & Persons With Disabilities Conference in August 1995. Under a US National Eye Institute sponsored Phase I Small Business Innovation & Research project initiated during 1995, improved "proof-of-concept" prototypes were constructed from commercially-available virtual reality components.

As shown conceptually in Figure 1, the VCM operates by combining a head-mounted display (HMD), head tracker, and screen enlarger software to fix the enlarged virtual screen of data in space and scan the user's line of sight across the data, instead of scanning the virtual screen across the display device as done in conventional screen enlargement systems. By incorporating a cursor at a fixed position in the user's visual field, interaction with computer data is then possible by turning one's head to the desired insertion point and clicking a mouse button.

 ISBN 0 7049 1140 X

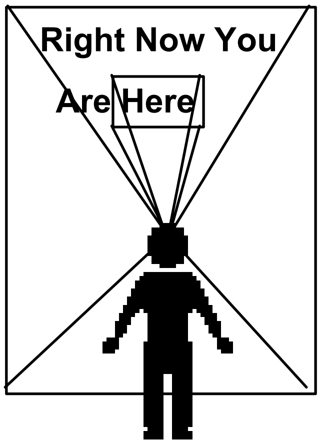


Figure 1: *VCM Concept*

The VCM approach represents an entirely new methodology for providing visually-impaired individuals with enhanced access to computer data and other visual content. The research team as well as others in the disabilities field hypothesize that this methodology can provide significant value for various populations of computer access-impaired users as a result of the following properties of the VCM approach:

1) An HMD-based display system can provide significantly greater instantaneous display field-of-view than a large conventional computer monitor, at lower system cost, and in a portable configuration.
2) An HMD-based display system can provide an unlimited field-of-regard, accessible merely by turning one's head. This should provide a more intuitive interface, since turning one's head instead of using a mouse provides natural proprioceptive feedback and frees the hands for input instead of navigation.
3) The VCM's ability to fix a computer control cursor at a user's preferred retinal locus can provide a significant enabling advantage to users with large central scotomata or smaller distributed scotomata, as well as to low vision quadriplegic users.
4) An HMD-based display system can be adjusted to any desired apparent display focus distance independently of instantaneous field-of-view, which may be more comfortable for some users and may help prevent the incidence of myopia caused by extended close viewing of conventional monitors.
5) HMD-based displays block undesirable external stimuli and remain fixed with respect to the user's visual field regardless of intentional or uncontrolled head motions, which can enable computer access for individuals with non-optically related computer visualization impairments.

The overall goal of the VCM effort is to demonstrate the benefits of these conceptual advantages sufficiently to justify commercial introduction of the system.

3. PHASE I PROTOTYPE DEVELOPMENT AND CONFIGURATION

3.1 Target Platform Scope

The Phase I prototype development resulted in VCM systems designed and constructed for both Macintosh and PC platforms. The completed systems function similarly, and are shown in block diagram form in Figure 2. The Macintosh system is shown photographically in Figure 3, which demonstrates the portability of the VCM system.

Proc. 1st Euro. Conf. Disability, Virtual Reality & Assoc. Tech., Maidenhead, UK, 1996

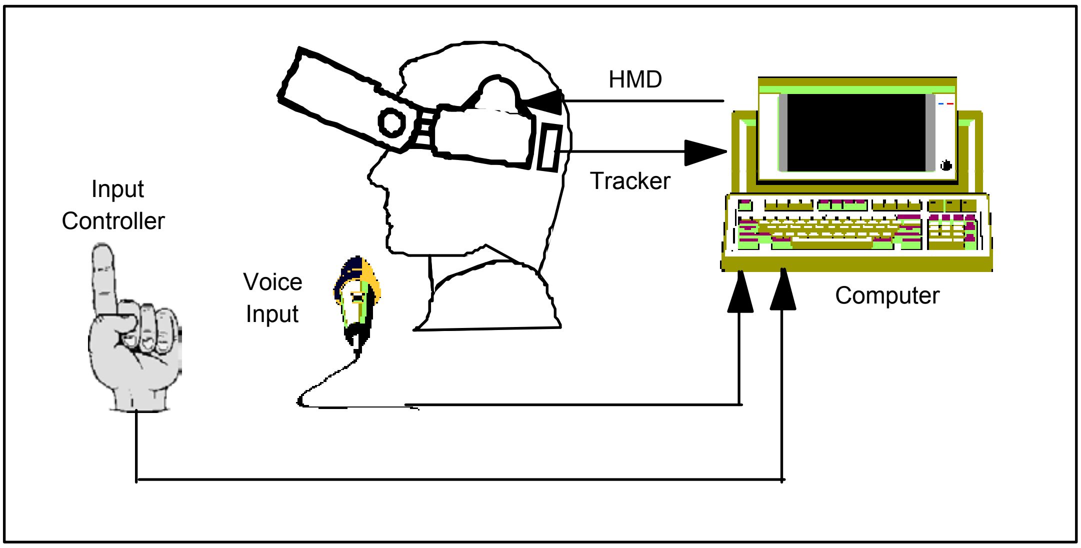


Figure 2: *Virtual Computer Monitor Block Diagram*

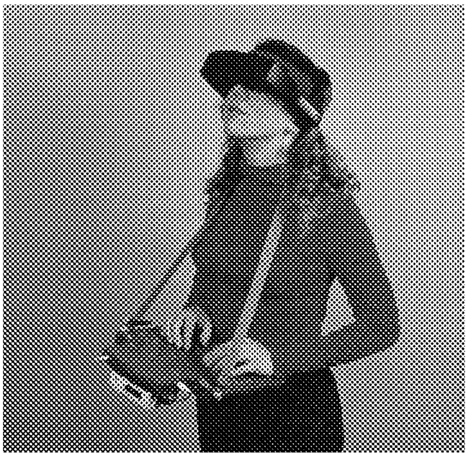


Figure 3: *Virtual Computer Monitor Macintosh Prototype*

3.2 Tracking System Design

In an ideal VCM system, the head tracker senses every rotational head movement with perfect accuracy, and the viewport location is changed instantaneously in response. The result would be an enlarged document that appears to float at a fixed position in front of the user, and which remains fixed in space as the user turns to look at various portions of it. Unfortunately, neither perfect accuracy nor instantaneous response can be achieved in a real system. The resulting imperfections and delays can cause the document to appear to jitter or move in various ways unintended by the user. This can be frustrating, and is believed to be a major cause of vertigo, dizziness, and other short-term symptoms often termed “simulator sickness”.

To prevent simulator sickness issues, a design goal of 40msec maximum lag between head movement and updated display video was set. Recent work (Azuma, 1994) has shown that total system lags of this magnitude are required for augmented reality applications (in which virtual objects are visually overlaid on physical objects using a see-through HMD), and it was postulated that a similar performance requirement applies to the VCM application. Additional criteria for the tracking system included low cost and portability, to support near-term commercialization. An analysis of available tracking technologies was performed, and is summarized in Table 1. Based on this analysis, prototypes were designed and constructed using two alternative technologies, magnetic/gravitational and gyroscopic.

Table 1: *Available Tracking Technologies*

Technology	Benefits	Drawbacks
AC/DC Electromagnetic	• Low Latency • Familiarity	• Bulky Transmitter • High Cost • Metallic Interference
EM Compass + Gravitational	• Small & Sourceless • Low Cost/Available	• Ambient Jitter (Floor) • Settling Time
Solid-State Gyroscope and/or Accelerometer	• Small & Sourceless • Very Low Latency • Potential Low Cost	• Drift • Dynamic Range • Commercial Availability
Optical	• Low Latency • High Accuracy • Low Interference	• Bulky Transmitter • Occlusion Problem • Cost
Ultrasonic	• Low Latency • High Accuracy	• Line-of-sight • Bulky Electronics • Interference
Image-Based Analysis	• Least Encumbering • Most Flexible	• Compute-Intensive • Availability/Cost • Occlusion Problem

3.2.1 Magnetic/Gravitational Tracking Results. This device operates by sensing the earth's magnetic and gravitational fields, and reporting changes in the sensed data that result from orientation changes of the tracker with respect to those fields. The highest performance magnetic/gravitational tracker available commercially was used.

Upon testing, it was found that the magnetic/gravitational tracking approach provided insufficient performance for the VCM application. The primary problem was jitter, wherein the imagery jumps randomly by a small amount one or more times per second. In addition, the interface device used to combine tracker output with mouse output added latency to the image updates, which caused the movement of the viewpoint to lag head movements. As a result, selecting objects on the screen was found to be difficult, since overlaying the cursor over the imagery often required several attempts. The only feasible approach to solving the jitter problem was determined to be averaging data over several samples prior to output, which would have exacerbated the lag problem.

3.2.2 Gyroscopic Tracking Results. The gyroscopic approach uses two tiny solid-state gyroscopes, each of which senses accelerations about one rotational axis. In theory, such devices can provide extremely rapid response to head motions (below 10msec), with essentially no ambient jitter. However, limitations to the gyroscopic approach include drift (which results in the imagery within the VCM moving steadily in one direction even when the user's head is still) and dynamic range (wherein very slow or very fast orientation changes are not properly tracked).

Two gyroscope suppliers were considered. The first, VR Systems, is based in the United Kingdom, and markets a $1,700 gyroscopic tracker for use as a cursor control device for the motor-impaired. The second, Gyration, is based in Mountain View, CA, and markets a $129 gyroscopic mouse used for stand-up computer-based presentations. The Gyration unit was selected for Phase I testing as it is much less costly and available more quickly in the US. It was incorporated in the VCM by means of sorting through six Gyropoint mice from Gyration to find the two units exhibiting the lowest ambient drift. The selected mice were then modified to serve as trackers and mounted on head-mounted displays. This implementation was found to provide extremely rapid, responsive tracking with insignificant jitter, and was used in the human subject trials. However, periodic manual recalibration to manage the drift is still necessary.

For future implementations, the project team plans to eliminate the drift issue by developing a hybrid tracker using a combination of gyroscopes or accelerometers for rapid response and magnetometers to periodically correct drift.

3.3 Display System Design

The baseline prototypes utilize a standard CyberEye Model CE-100 HMD. This device is designed for professional virtual reality and data display applications, and provides 420x230 display pixels over a 22.5 by 16.9 degree field-of-view, focused at near-infinity. This results in individual RGB pixels which subtend about 3.2 arc-minutes as seen by the user. The display uses one 0.7" diagonal full-color active matrix liquid crystal display (AMLCD) for each eye, with optics that provide the equivalent of a 19" viewing monitor.

3.4 Software Design

3.4.1 Pointer Control Software. The pointer control software is a device driver that accepts input from the head tracker and mouse, and interprets that data to control the screen pointer. Conceptually, this software must support gross positioning on the magnified page using head tracker input, fine positioning using mouse input, and selection of screen objects based on input of mouse "clicks".

The pointer control software problem was simplified on the Macintosh platform due to the Apple Desktop Bus (ADB), which supports daisy-chaining of multiple input peripherals. Since the Gyropoint connects directly to the ADB, no special software was required to combine tracker and mouse inputs on the Macintosh. Instead, the standard Macintosh mouse driver software was used.

On the PC platform, providing for dual-input control devices was more challenging since the PC architecture does not include an analog of the ADB. Therefore, combining the data from two control devices to control the screen pointer was found to require either an external interface device or a customized controller driver capable of combining inputs from the mouse port and the serial port. Initially, the external interface device approach was attempted and ruled out due to lag, while development of a customized driver was ruled out due to budgetary constraints. In response to this barrier, a dialog was initiated with Microsoft, who provided a software driver capable of providing this functionality. This driver has been found adequate for human subject trials.

For future implementations of the VCM, the project team intends to incorporate the following features in the tracking driver:

1) Drift Correction: The first processing stage of the software will use output from the magnetic/gravitational tracker to correct drift in the gyroscopic tracker.
2) Prediction: To maximize tracking performance, predictive algorithms such as Kalman filtering will be applied to the drift-corrected tracking output. Such algorithms have been demonstrated to provide major reductions in perceived lag, with corresponding reductions in simulator sickness symptoms (Emura and Tachi, 1994; Foxlin, 1996).
3) Field-of-Regard Selection: This software will modify tracking sensitivity based on user input, which provides the user with the ability to manage the overall field-of-regard in the VCM system. This is required in order to limit the amount of head rotation to a comfortable range.

3.4.2 Screen Magnifier. The second software component required for the VCM is a screen magnifier capable of magnifying application output and delivering the appropriate portion of the magnified image to the display buffer. Commercially-available screen magnifiers were used to limit project cost and to allow side-by-side comparison between the VCM and conventional screen enlargement approaches using the same user interface software.

On the Macintosh platform, the only commercially available screen enlarger is inLARGE, by Berkeley Systems (Berkeley, CA). The latest (unreleased) version was used to ensure compatibility with late-model Macintosh systems and with the VGA/NTSC converters selected for the project.

On the PC platform, a large number of commercially available screen magnifiers exist, including LP-DOS (VisionWare Software, Boston MA), Vista (Telesensory, Inc., Mountain View, CA), Magic (Microsystems Software, Inc., Framingham, MA), Zoomtext (Ai Squared, Manchester Ctr., VT), and Magnum GT (Artic Technologies International, Inc.). Each was tested to evaluate performance within the VCM regime. In the case of LP-DOS, an unreleased new version was obtained. This version provided faster screen updating, which reduced overall perceived tracking lag.

While all of the PC platform enlargers function in the VCM regime, the Macintosh inLARGE product contains a beneficial feature which could not be found in any Windows product. This feature maintains the cursor in the center of the viewport even when the user tries to move beyond the edge of the magnified desktop. In the Windows enlargers, moving to the edge of the desktop unfreezes the cursor from the center of the viewport and moves it to the edge of the desktop. This is suboptimal, but was unavoidable for initial human subject trials.

While no Windows product provides this constant centering feature, LP-DOS does provide it when used under the DOS operating system. In addition, LP-DOS provides the ability to maintain the screen pointer at locations other than the center of the screen. It is believed that this feature can be important for users with central vision obstructions such as might occur in macular degeneration. However, evaluation of the benefit of non-central screen pointer positions has not been performed due to the lack of suitable DOS applications. A future release of LP-DOS for Windows will be available in late 1996, and is expected to provide both constant centering and selectable pointer positions in the Windows environment.

4. PHASE I RESULTS FEEDBACK

4.1 Informal Results

The VCM prototypes have been demonstrated at three virtual reality conferences, where several hundred normally-sighted individuals used the system to navigate about a typical computer desktop after only one sentence of instruction. While no formal data was taken, it is clear that the interface functions intuitively, since productive interaction with the desktop folders was instantaneously achieved even among those unfamiliar with virtual reality concepts. This is supported by the observation that a very high percentage of trade show attendees immediately navigated to the "games" folder, where they booted "Wolfenstein 3D".

4.2 Safety Evaluation

Long-term physiological effects of VCM use are a major concern that must be addressed as early as possible in the product development cycle, and are being evaluated by SRI International, one of the world's leading research institutions in the area of HMD safety. The safety evaluation is being performed in three stages:

1) Compile Research On HMD Safety: SRI has updated its database, focusing on physiological effects of HMD use. This search included a review of physiological effects of liquid crystal display (LCD) and other visual display systems, physiological effects of inadequate ergonomic design of HMDs, and visual stimuli that can induce epilepsy.
2) Evaluate Subjective User Reactions: SRI is conducting a series of tests on user acceptance and ease of adjustment of HMD designs for head size, interpupillary distance, and vertical phoria (differences in the vertical position of the two eye's images). In these tests, subjects don a VCM, adjust the HMD for their head and eyes, and are asked to comment on 1) the comfort of the HMD after adjustment, 2) any difficulties they had in making adjustments, and 3) whether they were able to see clear, single, fused images prior to making adjustments for the eyes, 4) whether they were able to see single, fused images after making any necessary adjustments for interpupillary distance or vertical phoria, and 5) whether they experienced any vertigo or eye, face, neck, headache, or bodily discomfort.
3) Establish Guidelines For Extended Use Of HMDs: SRI has developed and is performing a set of standard visual and vestibular tests aimed at determining whether an upper limit should be imposed on use of the VCM. In these tests, subjects access the computer while wearing the VCM for periods between 15 minutes and 2 hours. The access tasks consist of a simple game involving 2-dimensional localization and identification of objects, repositioning those objects in valid places within the scene, and cancellation of occasional dialog windows that pop up when non-valid positions are selected. A scoring system is kept to encourage active participation of the subjects.

Tests are administered before and after use to measure visual modulation transfer function (detection threshold criterion), visual acuity, near and far phorias (vertical and horizontal), near point of accommodation, near point of convergence, and stereo disparity sensitivity. These vision tests are used to determine whether the subject has: 1) developed eye strain, 2) experienced a loss in either far or near visual acuity, 3) had a change in ocular alignment or 4) a reduction in depth perception. Routine vestibular tests of balance and locomotion are also administered to subjects. The entire test procedure is repeated using a conventional screen enlarger to enable detection of differential effects between the VCM and conventional monitors.

While the SRI safety evaluation is not complete, no results indicating fundamental safety problems in the VCM approach have been identified. A minority of users have noted vertigo, dizziness, and analogous symptoms known to occur in virtual reality simulations when the delay between head motion and display regeneration is appreciable. These results are attributed to tracking errors and artifacts in the prototype VCM system used in the prototype, and solutions have been identified.

Interestingly, a growing body of research suggests that extended close work such as reading or computer use may be a major cause of myopia (short-sightedness), which affects nearly 50% of pre-high-school children in the US (Seachrist, 1995). Given these findings, it is logical to assume that users with visual impairments unrelated to emmetropy (accurate eye focal length) may develop myopia as a result of working extremely close to conventional computer monitors. This has not been researched, but if confirmed, it may be the case that HMD use can prevent this problem. This is because the HMD can be focused at any arbitrary distance from the user, which can optically eliminate extended close viewing. If this is found true, it may be the case that for some individuals and some tasks, HMDs may actually prove safer than conventional displays.

4.3 Performance Evaluation

The Phase I performance evaluation effort is using one of the Western Blind Rehabilitation Center's low vision computer training workstations. These workstation are used on a regular basis to provide computer access training to 30-35 legally blind veterans annually out of the center's 180 annual students. As such, these workstations are outfitted with "the latest" available low vision access tools, including a wide range of screen enlargement software, large monitors, and other specialized hardware.

As of this date, experimentation has begun using professional staff and students of the Western Blind Rehabilitation Center as subjects. The vision of these subjects ranges from fully-sighted to legally blind. Each subject is given an overview presentation regarding the virtual monitor concept and the purpose of the study, then provided with a VCM training and practice session. While this is sufficient to allow use of the system, it does not overcome potential bias against the VCM caused by familiarity with only the conventional enlarger solutions. Each subject is then tested for computer access capability on the VCM, and on standard (13"-14") and large (20"-23") monitors, using identical software with all display devices. Magnification is adjusted to provide the same apparent text size on all monitor devices to control for the acuity variable.

While the evaluation at the VA Blind Rehabilitation Center is not complete, every user tested has provided strongly positive feedback about the VCM's promise, and praised the virtual desktop concept. Several stated that "it is fun to use". Most users provided suggestions and requests for further development, with individual requests including wider FOV, monocular operation, monochrome operation, increased contrast, reduced brightness, disabling tracking during typing, accommodation for smaller head sizes, increased resolution, and larger icon sizes. As some of these suggestions conflict, all will be evaluated further and implemented as necessary during the Phase II effort.

5. SERENDIPITOUS DISCOVERIES

The project team strongly believes that one of the major benefits of applying leading edge technologies such as virtual reality to a driving problem such as low-vision computer access is serendipity, meaning unexpected discoveries. Several such serendipitous results were found during the Phase I effort, and will be more fully explored during Phase II.

5.1 Real-Time Video Magnification

The first serendipitous result occurred via realization that the VCM can be modified in a straightforward manner to magnify live video, and present it in a virtual monitor approach. This approach was conceived in response to the problems visually-impaired students face in classroom lecture environments such as mathematics courses.

As shown in Figure 4, a VCM for real-time video use requires addition of a camera and video capture board to the baseline VCM. By positioning the camera to view the entire blackboard (or overhead projector screen), an image of the scene is captured into memory, and can be enlarged and displayed using the VCM. The student can then easily turn to look around the enlarged virtual blackboard, thus facilitating visual access. A second benefit can be achieved by incorporating a command which captures the image for later viewing, thus allowing the student to concentrate on learning instead of note taking.

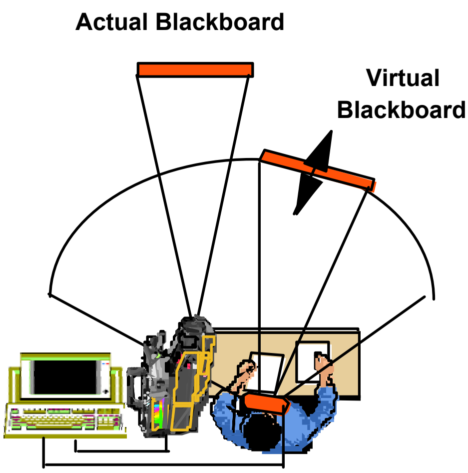


Figure 4: *Real-Time Video Magnification Application*

It is noted that this approach is expected to provide superior access performance for real-time images than alternative devices such as the Low Vision Enhancement System (LVES) developed at Johns Hopkins University. This is because the LVES system mounts the camera on the user's head, so any jitter in the user's head is magnified to the same degree as the viewed image (just as is the case when using binoculars). In contrast, the VCM configured for live video uses a fixed camera, so jitter in the user's head position is unmagnified.

5.2 New VCM Applications

A key serendipitous result achieved during the Phase I effort was discovery of several new classes of computer access impairments that can be addressed through variants of the VCM. These include Dystonia, Parkinson's Disease, Cerebral Palsy, Autism, Attention Deficit Disorder, Dyslexia, and mobility impairments. In some of these cases, it is believed that benefits can be achieved simply because the HMD blocks undesirable external distractions, forcing the user to focus on the computer-based task at hand. In other cases, the benefits are achieved because wearing an HMD causes the display to remain in view despite uncontrolled head motions.

Exploring these benefits was outside the scope of the Phase I effort, however, one human subject was tested outside of the formal study, and the following case summary was provided by the clinicians involved:

> "15 years ago, Richard had a stroke resulting in Dystonia. The resulting head jerks and twists left him unable to continue in his work as a professional proofreader and narrator. He could not read because of the involuntary head jerks which would jerk his gaze away from the text. By the time the muscle contraction subsided and he had relocated his reading place, it would happen again.
>
> Over the years, many potential solutions were examined and tested without success. Various drugs were used to control the involuntary muscle movements, some of which had side effects worse than the Dystonia. Botulism was even injected into the muscles to partially paralyze them. Many avenues of technology were explored. The CyberEye head-mounted display was a last ditch attempt to meet his need. Because the display moves with his head jerks whenever they occur, it was reasoned he could maintain a constant gaze on the screen.
>
> Initially, we were concerned about possible torsional problems with the headset. Would it cause discomfort or shift position during his involuntary head jerks? Our evaluation revealed this was not a problem. Screen enlargement software was necessary because of a slight amount of "text blur" which still occurred due to the limited resolution of current-generation head-mounted displays. There were some initial compatibility problems between the screen enlargement software and word processing software in "panning" mode, but these were resolved.
>
> Using a computer with screen enlargement software and the VR headset, a scanner, fax/modem, a modified mouse with a large cursor and recording equipment, Richard is able once again to pursue a job he loves. For the first time in 15 years, he is able to read paragraph after paragraph of text without stopping and to narrate non-stop!
>
> Although it was never designed to help a person with Richard's condition, the CyberEye proved to be the missing link in his long odyssey to rehabilitation. What for many is used as a gaming device, became for Richard, a life changing device. A solution which more than meets the eye."

[Note: Researchers Rita Howells and Garry Bowman work for the Bureau of Blind Services, Illinois Dept. of Rehabilitation Services, State of Illinois.]

6. PHASE II RESEARCH PLAN

The VCM research to date can be considered exploratory research, as it leveraged commercially available components to construct and subjectively evaluate the VCM concept. In contrast, the planned Phase II research and development effort will be a focused technology development effort with a clear goal: design a VCM system meeting a specific set of user requirements.

The planned Phase II effort will initiate by answering key open questions regarding technical requirements for an optimized VCM product, especially in the area of display brightness/contrast/FOV/color depth. The results of the Phase I and initial Phase II work will then be used to develop an optimized HMD/tracker/enlarger combination designed specifically for the VCM application. In addition, the prototype VCM systems will be enhanced to improve performance in the baseline low-vision computer access application, and add capabilities for new applications such as classroom lecture access. Finally, this optimized prototype will be evaluated in an extensive series of user studies to validate the VCM system and improvements.

7. SUMMARY AND CONCLUSIONS

The project team believes the Phase I results to date are valuable and important for the following reasons:

1) Extensive progress towards determination of user requirements for a commercially-viable VCM system design has been achieved. Most of the detailed requirements for an optimized Phase II prototype have been generated.
2) The first VCM functioning in the PC environment has been designed, constructed, debugged, and demonstrated.
3) Professional computer trainers for the visually-impaired (including visually-impaired trainers) have used the prototype VCM system, and strongly praised the approach.
4) The existing body of HMD safety and performance research has been extended, and the first hypothesis that HMDs can actually be safer than conventional monitors has been generated.
5) Additional benefits of the VCM for unexpected populations have been discovered.
6) Significant momentum has been generated for further development and commercialization throughout the technology and disabilities field. This momentum includes technical support from industry leaders ranging from Apple and Microsoft to screen enlarger developers, plus an extensive list of clinicians, resellers, and consultants interested in patient evaluations following release of a pre-production version to beta test.

8. REFERENCES

R. Azuma and G. Bishop (1994), Improving Static and Dynamic Registration in an Optical See-Through HMD, *Proceedings of SIGGRAPH '94*,.

S. Emura and S. Tachi (1994) Compensation of Time Lag Between Actual and Virtual Spaces by Multi-Sensor Integration, *Proceedings of 1994 IEEE International Conference on Sensor Fusion and Integration for Intelligent Systems (MFI'94)*, pp. 463-469.

E. Foxlin (1996) Inertial Head-Tracker Sensor Fusion by a Complementary Separate-Bias Kalman Filter, *Proceedings of IEEE VRAIS '96, 1996, IEEE Press #PR07295*, pp. 185-194.

G Legge, D Pelli, et. al. (1994) Report of the Low Vision and its Rehabilitation Panel, *Vision Research - A National Plan 1994-1998. US National Advisory Eye Council,,* pg 313.

L. Seachrist (1994) Growing In and Out of Focus, *Science News* **148**, *Nov. 11, 1995* pp. 318-319.

ECDVRAT '96
Conference Sessions

Day 1: 8 July, 1996

Morning: Communications and Language
Afternoon: Virtual and Enhanced Environments

Day 2: 9 July, 1996

Morning: Rehabilitation I
Afternoon: Technology

Day 3: 10 July, 1996

Morning: Visual Impairment, Ambisonics, and Mobility
Afternoon: Rehabilitation II

Communications and Language

Day 1: 8 July, 1996, Morning

Recognition of sign language gestures using neural networks

Peter Vamplew

Department of Computer Science, University of Tasmania
GPO Box 252C, Hobart, Tasmania 7001, Australia

vamplew@cs.utas.edu.au

ABSTRACT

This paper describes the structure and performance of the SLARTI sign language recognition system developed at the University of Tasmania. SLARTI uses a modular architecture consisting of multiple feature-recognition neural networks and a nearest-neighbour classifier to recognise Australian sign language (Auslan) hand gestures.

Keywords: sign language, hand gestures, communication aid

1. INTRODUCTION

Sign languages such as BSL and Auslan (Australian sign language) are the primary form of communication between members of the Deaf community. However these languages are not widely known outside of these communities, and hence a communications barrier can exist between Deaf and hearing people. The hand tracking technologies developed for VR enable the possibility of creating portable devices which can convert sign language to speech, as an approach to overcoming these difficulties. The SLARTI system is a prototype of such a device, based on Auslan.

2. SYSTEM DESIGN

2.1 Input Hardware

In computer recognition of spoken language, speech data is captured using a microphone connected to an analog-to-digital converter. Similarly a data-capturing device is also required in order to recognise sign language; in this case measuring the position and movement of the signer's hands. Two broad categories of input hardware have been used for recognition of hand gestures – glove-based devices such as those used by Kramer et al (1989) and Fels et al (1993), and camera-based systems as used by Holden (1993). The latter approach has some benefits, particularly as it does not require specialised hardware, but this is offset by the complexity of the computer vision problems faced in extracting the necessary data about the hands from a visual image. Therefore for this research glove-based input was used, as this allowed the research effort to be focused on the area of sign recognition rather than that of data capturing.

The specific input devices used in developing SLARTI were a CyberGlove, and a Polhemus IsoTrak. The CyberGlove measures the degree of flexing of the various joints of the hand and wrist. The version of the CyberGlove used for this research provides 18 sensors. The Polhemus allows tracking of the spatial position and orientation of the hand with respect to a fixed electro-magnetic source. Using only a single glove restricts the system to the recognition of one-handed signs (and hence eliminates the possibility of recognising the Auslan manual alphabet which is two-handed), but it is envisaged that the techniques used in developing the system could be extended to two-handed signs if appropriate input hardware was available.

2.2 System Architecture

Linguistic analysis of sign language has revealed that signs can be described in terms of four basic manual features, which may be modified in meaning by more subtle factors such as body language and facial expression (see for example Johnston 1989). The handshape defines the configuration of the joints of the hand. Orientation specifies the direction the hand and fingers are pointing, whilst the place of articulation is the location of the hand relative to the body. The most complex feature is motion, which consists of a change over time of any combination of the other three features (although for this research only changes in location have been considered).

The task of transforming a stream of input data directly into a classification of the sign being performed is an extremely difficult one. Instead the approach taken within SLARTI was to initially process the input data so as to

 ISBN 0 7049 1140 X

produce a description of this sequence in terms of the four features discussed above. The sign can then be classified on the basis of this feature vector. The SLARTI system consists of four separate feature-extraction neural networks, each trained specifically to recognise one of the features of the sign. The feature vector produced by these networks is then used to perform the overall classification of the input sequence, as shown in Figure 1.

This approach of decomposing the problem and applying a modular structure of networks has a number of benefits. First, as demonstrated on the task of speech-recognition by Waibel et al (1989), it allows the use of several smaller networks rather than one massive network and thereby reduces the amount of training time and data required. It may also result in superior classification accuracy. Second, it produces a system which can more easily be extended. The features recognised by the feature-extraction networks are expressive enough to describe an extremely large number of signs, not all of which may be recognised by the final classifier. If the vocabulary of the system is to be extended then only the final classifier will require modification. This greatly reduces the costs involved in performing such expansion of the system, and makes it practical to tailor the vocabulary of the system to a particular user.

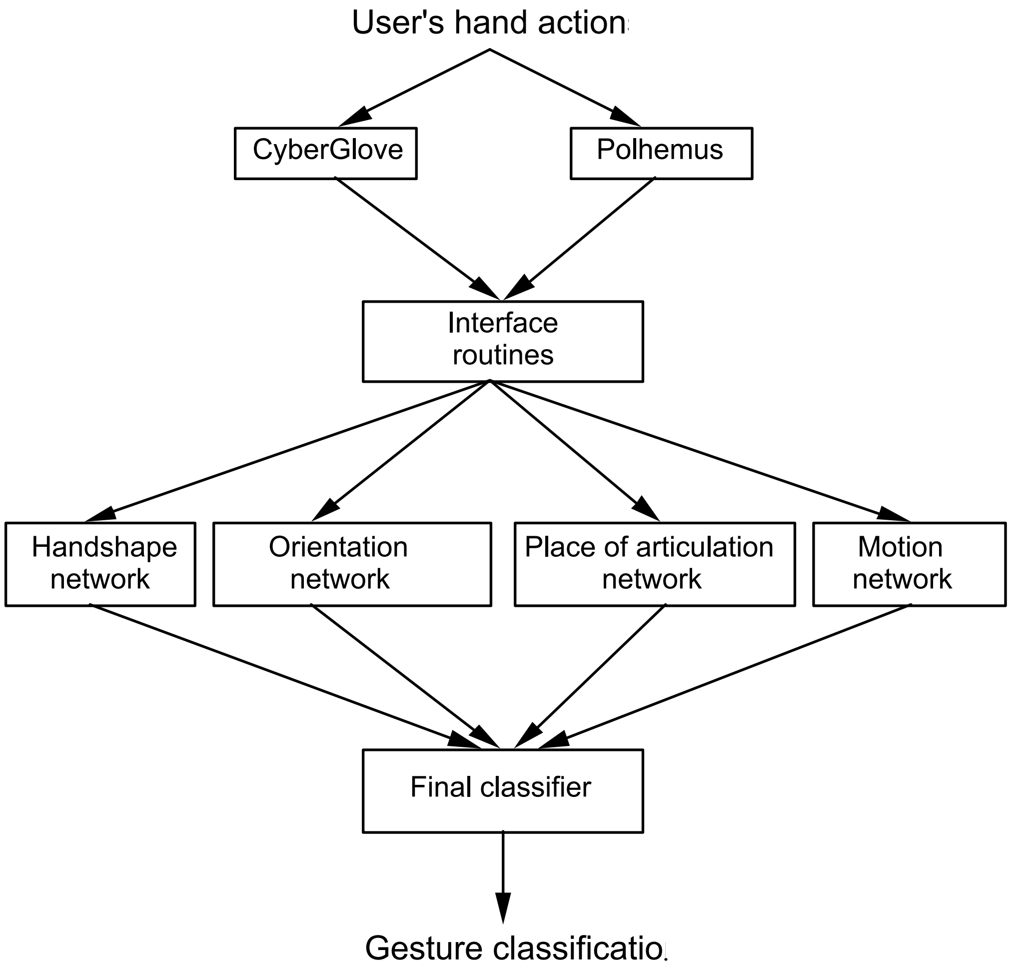


Figure 1. *The modular architecture of the SLARTI system*

3. FEATURE EXTRACTION NETWORKS

3.1 Data Gathering and Training Methodology

All of the feature-extraction networks were trained on examples gathered from 7 signers (which will be referred to as the registered signers), and tested on both fresh examples from the same signers and examples from 3 other signers (the unregistered signers), so as to assess the possibility of creating a signer-independent system. A fully-connected feed-forward architecture with a single hidden layer was used for all four networks and backpropagation without momentum was used as the training algorithm. All input data were scaled to lie in the range -1 to 1. The results reported are the average of results over 25 trials from different starting weights.

3.2 Handshape Recognition Network

 ISBN 0 7049 1140 X

Johnston (1989) identified 30 different primary handshapes used in Auslan, which can be further subdivided into 61 variant handshapes, although for the purposes of classifying signs it is only necessary to be able to distinguish between the primary handshapes. For each of the registered signers, 4 examples of each of the 61 variant handshapes was gathered for use in a training set. A further example of each handshape was gathered from all 10 users to constitute the 2 test sets. Prior to gathering the handshapes each user was asked to perform a simple calibration routine consisting of several handshapes chosen to measure the range of movement of the user's joints. By calibrating the handshape data relative to these extremities it was hoped to improve the network's ability to generalise to the unregistered users.

Networks were trained on both the calibrated and uncalibrated data. In both cases the networks had 18 inputs, 40 hidden nodes and 30 output nodes (this will be denoted as an 18:40:30 architecture), and were trained for 1,000,000 pattern presentations with a learning rate of 0.2 The results reported in Table 1, show that although the calibration process slightly reduced performance on the registered test set, it had a larger beneficial effect on the unregistered test set, and therefore this calibration was incorporated into the final system.

Table 1. *Mean classification accuracy of networks trained using raw and calibrated versions of the handshape data sets*

	Training set	Registered test set	Unreg. test set
Raw data	97.9	96.6	87.9
Calibrated data	98.0	96.2	89.9

3.2 Orientation Recognition Network

The orientation of the hand can be described in terms of two orthogonal directions – the facing of the palm, and the direction in which the hand is pointing. If we consider only six possible directions (up, down, left, right, towards the signer, away from the signer) then there are 15 different orientations used in Auslan (in fact some signs involve directions such as 'left and up', but such small distinctions are never the sole difference between signs).

The input to the network consisted of the 3 orientation values from the Polhemus sensor, and also calibrated values for the 2 wrist sensors on the CyberGlove. The latter was required as the positioning of the Polhemus mount on the CyberGlove was above the wrist for some users, meaning that the orientation values were affected by the degree to which the wrist was flexed (early trials conducted without any wrist data performed poorly). The orientation values returned by the Polhemus were cyclical in nature (ranging from 0 to 255, and then back to 0). To avoid the problems caused by this discontinuity in the input data the network was presented with the sine and cosine of the orientation values rather than the raw values. Therefore the networks had an 8:14:15 topology.

The results of training these networks are reported in Table 2. These show that the overall accuracy is only moderate. However if these mistakes are broken down in terms of the misclassification of the component directions of the orientation, then it can be seen that the majority of errors consist of confusing adjacent directions. These mistakes are less likely to be important in distinguishing between signs than if the network were to confuse opposite directions.

Table 2. *Mean percentage accuracy obtained by different encodings on the hand orientation data, broken down by error of the two component directions*

Directions	Training set	Reg. test set	Unreg. test set
Both correct	94.8	90.4	89.1
One correct, one adjacent	2.5	5.0	5.9
Both adjacent	2.7	4.5	4.9
One correct, one opposite	0.0	0.0	0.0
One adjacent, one opposite	0.0	0.1	0.0
Both opposite	0.0	0.0	0.0

3.3 Location Recognition Network

Auslan signs can occur in three main groups of locations – neutral space (the space in front of the signer's body), primary locations (those on or near the body or head) and secondary locations (on or near the hands). In order to recognise secondary locations it is necessary to track the position of both hands, which would require a second Polhemus sensor. Therefore the SLART system considers only the 16 primary locations as well as neutral space. Any signs made using the subordinate hand as a base were regarded as being performed in neutral space.

The location and orientation features affect each other, and so the training and test data for the location network was gathered at the same time as the orientation data. For the same reason it was necessary to provide the network with the orientation and wrist flex values as inputs, in addition to the 3 Polhemus location values. Hence the networks developed had an 11:19:19 architecture. They were trained for 1,000,000 pattern presentations at a learning rate of 0.1. As with the handshape data, a calibration routine was used to measure the extremes of the input values for each user. In this case each signer made 5 gestures which measured the maximum extent of movement of their hand in each direction. Networks were trained using both the calibrated and uncalibrated data.

Table 3. *Mean classification accuracy of networks trained on the raw and calibrated versions of the hand location data*

	Training set	Registered test set	Unreg. test set
Raw data	71.7	67.7	64.5
Calibrated data	80.5	74.7	68.4

Two conclusions can be drawn from the results in Table 3. First the calibration routine was extremely beneficial, increasing performance on both the registered and unregistered test sets. Second, the location network achieved a much lower level of accuracy than any of the other feature-extraction networks. This is due primarily to the tracking technology used. The Polhemus measures the position of the glove relative to a fixed source, which for these experiments was placed on a wooden desk behind and to the left of the signer. Ideally the input to the system would be the position of the hand relative to the signer's body, rather than relative to the Polhemus source. In the current system any change in the positioning of the body relative to the source will affect the accuracy of the system, particularly with regards to the closely-spaced locations on the signer's head. This is one area in which a visually-based tracking system would have an advantage as it would allow more direct measurement of the hand position relative to the body.

3.4 Motion Recognition Network

Motion is the feature for which it is most difficult to enumerate the complete range of possible categories used within Auslan, as many signs involve 'tracing' motions which indicate the shape of an object, and hence are unique to that sign. For this research only the 13 most commonly used motions were classified, consisting of simple movement of the hand in either direction along the three primary spatial axes, back-and-forth motions along the same axes, circling motions aligned with the axes and stationary.

Motion differs from the other features in that it is inherently temporal in nature. Two approaches were taken to dealing with this aspect of the problem. The first was to use a recurrent network with 3 inputs per time frame, feeding into a layer of 30 recurrently-interconnected nodes (13 of these were output nodes, the remainder served to store the network's internal state). The input values were the difference in location from the previous time-step. This recurrent network was trained using the backpropagation-through-time algorithm with a learning rate of 0.05.

The second approach was to pre-process the input sequence to extract features for presentation to a standard feed-forward network. For each of the three location values the change at each time step was summed over the entire sequence (giving a measure of the difference between the initial and final position), as was the absolute value of the changes. The sum of the absolute value of the change in velocity at each time step was also presented to the network, as was the length of the input sequence, giving a total of 8 inputs. The network architecture used was 8:8:13, and this was trained for 750,000 pattern presentations at a learning rate of 0.05.

Table 4 compares the results obtained by the two network architectures. It can be seen that the non-recurrent network fared much better, slightly outperforming the recurrent network on the training data but giving a significant improvement in generalisation to the test sets. Therefore a non-recurrent network was used in the final system.

Table 4. *Mean classification accuracy of recurrent and non-recurrent networks on the hand motion data*

	Training set	Registered test set	Unreg. test set
Recurrent net	89.7	78.6	63.4
Non-recurrent net	93.5	91.6	75.7

4. CLASSIFICATION OF SIGNS

Once all of the feature-extraction networks had been trained, the best network for each feature was selected for inclusion in the final system (as determined by performance on the registered test set). Table 5 summarises the performance of these networks.

Table 5. *Summary of the performance of the best network for each feature on the training set and test set for the registered and unregistered signers*

	Training set	Registered test set	Unreg. test set
Handshape	98.0	97.4	89.5
Orientation	94.5	91.6	89.2
Location	80.9	76.4	69.0
Motion	93.7	92.3	76.9

Each signer was asked to perform 52 signs selected from Auslan to form SLARTI's initial vocabulary. Unfortunately due to age-related failure of the CyberGlove it was only possible to gather test sets from 4 of the 7 registered signers, although training sets were gathered from all 7. Test sets were also gathered from the 3 unregistered signers.

The 52 signs were randomly divided into 13 sequences of 4 signs which were performed by each signer, manually indicating the start and end of each sequence via a switch held in the non-signing hand. The signs were segmented at these points, and the input sequence was processed by the feature-extraction nets. The handshape, orientation and location features were found for both the start and end of the sequence, whilst the motion feature was extracted for the entire sequence. Hence each sign was described by a vector of 7 features which were then used to perform the final classification. A neural network was not used for this final classifier for two reasons. First the size of the resultant network (139 inputs, 52 outputs) would require an extremely large number of training examples in order to achieve a suitable level of generalisability. Second, this approach would mean retraining this large network any time that changes were made to the system vocabulary. For this reason other pattern classification techniques were preferred.

The first method used was the nearest-neighbour lookup algorithm. Four variants of this simple algorithm were used. One difference was in the nature of the examples considered by the lookup – in one version the examples from the training sets were used, whilst the second version used instead the definitions of the signs as derived from the Auslan dictionary. The second difference was in the nature of the distance measure used. In the simple distance measure (SDM) all categories of a feature were considered equidistant from each other. A heuristic distance measure (HDM) was also tested, which was derived by examination of the confusion matrices of the feature-extraction networks on the training examples. This heuristic aimed to account for the systematic errors introduced by the feature networks, by weighting these errors less heavily.

The results of these variants of the nearest neighbour lookup for each signer are reported in Table 6.

From Table 6 it can be seen that using the simple distance measure the lookup algorithm using the training examples easily outperforms that using the sign definitions. However the heuristic distance measure successfully captures the extra information present in the training examples, as it enables equal or better performance to be obtained using only the sign definitions. This is extremely useful as it allows the vocabulary to be extended without the need to gather examples of the new signs.

Table 6. *Classification accuracy of the nearest neighbour lookup algorithm on complete signs from each signer*

Signer	Definitions (SDM)	Definitions (HDM)	Training set (SDM)	Training set (HDM)
A	88.5	94.2	92.3	94.2
C	71.2	92.3	100.0	100.0
E	71.2	96.2	67.3	90.4
F	86.5	94.2	86.5	88.5
Reg. signers (mean)	79.4	94.2	86.5	93.3
D	67.3	82.7	75.0	86.5
H	65.4	88.5	76.9	75.0
K	71.2	84.6	84.6	82.7
Unreg. signers (mean)	68.0	85.3	78.8	81.4

The second classification algorithm trialed was the C4.5 inductive learning system developed by Quinlan (1992). C4.5 builds a decision tree on the basis of training examples, which can subsequently be pruned to obtain a smaller tree. The process of generating the decision tree is extremely fast in comparison to neural networks, meaning that creating a new decision tree every time the vocabulary was extended is a viable proposition.

Table 7. *Classification accuracy of the C4.5 algorithm on complete signs from each signer*

Signer	Standard unpruned	Standard pruned	Subset unpruned	Subset pruned
Tree size	649	397	140	133
Training examples	92.3	88.2	96.2	95.9
A	86.5	84.6	90.4	90.4
C	96.8	92.3	98.1	98.1
E	73.1	71.2	55.8	55.8
F	78.8	78.8	76.5	76.5
Reg. signers (mean)	83.8	81.7	80.2	80.2
D	63.5	63.5	65.4	67.3
H	63.5	61.5	65.4	65.4
K	71.2	69.2	78.8	78.8
Unreg. signers (mean)	66.1	64.7	69.9	70.5

Table 7 reports results for C4.5 using both the pruned and unpruned versions of the tree, and bot with and without the subsetting option (this option allows each node in the decision tree to incorporate multiple values of an attribute). The results obtained by C4.5 are generally below those obtained by applying the nearest neighbours lookup algorithm to the same training examples, even if only the simple distance measure is used. In particular the nearest neighbours technique generalises much better to the unregistered signers.

5. CONCLUSION

SLARTI is capable of classifying Auslan signs with an accuracy of around 94% on the signers used in training, and about 85% for other signers. The modular design of the system allows for future enhancement of the system both in terms of expanding its vocabulary, and in improving the recognition accuracy. The major area in which accuracy could be improved is in the classification of sign location where the performance could be improved by the addition of extra position tracking sensors.

Currently the hardware used is not portable enough to be used in the real-world as a communications device, but it could be applied as a teaching aid for people learning Auslan. The techniques developed are not specific to Auslan, and so the system could easily be adapted to other sign languages or for other gesture recognition systems (for example, as part of a VR interface or for robotic control).

6. REFERENCES

S Fels and G Hinton (1993), Glove-Talk: A Neural Network Interface Between a Data-Glove and a Speech Synthesiser, *IEEE Transactions on Neural Networks*, **4**, *1*, pp. 2-8

E Holden (1993), Current Status of the Sign Motion Understanding System, Technical Report 93/7, Department of Computer Science, University of Western Australia,

T Johnston (1989), Auslan: The Sign Language of the Australian Deaf Community, PhD thesis, Department of Linguistics, University of Sydney

J Kramer and L Leifer (1989), The Talking Glove: A Speaking Aid for Nonvocal Deaf and Deaf-Blind Individuals, *RESNA 12th Annual Conference*, New Orleans, Louisiana

J Quinlan (1992), *C4.5: Programs for Machine Learning*, Morgan Kaufmann

A Waibel, H Sawai and K Shikano (1989), Modularity and Scaling in Large Phonemic Neural Networks, *IEEE Transactions on Acoustics, Speech and Signal Processing*, **37**,*12*

Tackling isolation and the expression of emotion in a virtual medium

D. J. Roberts[1], C. Wood[2] and A. Gibbens[2]

[1]Department of Cybernetics, University of Reading,
Whiteknights, Reading, UK.

[2]Project DISCOVER,
London, UK.

[1]*cybdjr@cyber.reading.ac.uk*

ABSTRACT

This paper discusses how virtual reality may be used as a medium for expressive communication and thus aid in tackling the problems of isolation faced by many people with disabilities. The principle causes and effects of isolation are explained, along with the strengths of the virtual medium that make it applicable to this problem area. The concept of a virtual meeting place is introduced, where users can overcome the physical problems of mobility and communication. The more esoteric problem of communicating emotion is addressed and a number of ways in which this may be tackled in virtual reality are proposed.

Keywords: disabilities, emotion, isolation, expression, awareness, multi-user

1. INTRODUCTION

Project DISCOVER is group of people with differing abilities who have been meeting over the past three years, to discuss ways in which virtual reality can help them overcome some of the problems they face. After analysing several application areas, the group has concentrated on tackling the underlying problems of isolation. Isolation may arise from a lack of mobility, the attitudes of the able bodied, or a physical or psychological inability to express emotion. The group has developed a number of concepts that show how the medium of virtual reality may be used to overcome some of these underlying problems. The problems of mobility may be partly overcome in a multi-user environment. This environment, along with the custom representation of users, could tackle shyness and help educate able bodied people. Multi-user environments are already being developed by many researchers. The emphasis of the DISCOVER group is to use virtual reality as a medium for the expression of emotion. A multi-user environment, that promotes the meeting of friends and allows them to build relationships through emotional dialogue, is not as fantastic as it sounds. Most of the building blocks are already in place, but are separated by the disciplines of their creators. By bringing together people of differing abilities with relevant backgrounds, Project DISCOVER hopes to provide the impetus to bring these blocks together.

1.1 Structure of the paper

The DISCOVER group is introduced in section 1 which summarises their aims, and both the problems they have been, and are now addressing. The emphasis of the groups current work is tackling the problems of isolation through representing emotion in a virtual medium. Section 2 discusses the principle causes and effects of isolation and proposes how these causes may be temporarily reduced within a virtual environment. The concept of a virtual meeting place where users can overcome the physical problems of mobility and communication is introduced in section 3. Section 4 addresses the more esoteric problems of communicating emotion and suggests a number of ways in which this may be tackled in virtual reality. Section 5 discusses how levels of emotions may be monitored provided the kind of emotion is known.

2. PROJECT DISCOVER

Project DISCOVER was founded three years ago with the aim of investigating the technology of virtual reality as a medium for expression for people with disabilities. The DISCOVER group consists of around ten active voluntary

 ISBN 0 7049 1140 X

members of various abilities and backgrounds. The various life experiences of the group members provides a strong backbone for discussing the issues faced by those with disabilities. A strong emphasis has been made to allow the people the project was meant to benefit, to not only steer the work but to drive it. The meetings combine working structure with a social emphasise which is aimed at bringing people together to share experiences. Although academic involvement has been present from an early stage, this has been primarily in a informal advisory capacity and to act as a gateway between the group and the developers of the technology.

DISCOVER defined three categories of human experience that would gain useful insight though the medium of virtual reality: basic everyday, emotional and educational. Everyday experience encompasses daily living functions including coping with the physical environment, picking up a kettle, opening a door, et cetera. Emotional experiences are those caused indirectly by the disability such as feelings of isolation, relationships and sexual needs. Education targets the attitudes of the able bodied through the promotion of awareness, and those with disabilities through addressing the lack of self esteem, often associated with disabilities, which must be overcome to achieve empowerment, control and independence.

The group came up with a number of possible application areas that they felt could be of benefit to those with disabilities. Four of these areas: Emotion artistic; Training of the able bodied; Construction assessment and Everyday living, where analysed by a separate sub-group for a period of around six months. By the end of this period it had become apparent that the area of greatest interest and potential was the representation of emotion as an aid to reducing the feelings isolation.

3. WHY USE VIRTUAL REALITY TO TACKLE ISOLATION

This section looks at the causes of isolation and the effects that it has on individuals. It then continues by suggesting the attributes of virtual reality that make it applicable for tackling some of these problems.

3.1 Isolation

One of the greatest problems faced by people with disabilities is that of isolation. This problem arises from several factors including the attitudes of other people, a lack of self confidence and difficulties associated with mobility, access and communication.

The attitudes of other people can make those with disabilities reluctant to communicate with them. Some people may be dismissive, wrongly thinking that people that do not adhere to what they consider to be the norm, to be of less value. At the other extreme and often with the best intentions, some people may treat anyone with a noticeable disability with patronising attention. Both these behaviours come from a basic misunderstanding of disabilities which itself stems from a lack of education and contact.

Many people suffer from a lack of self confidence or shyness. This is usually brought about by a lack of self esteem which in turn stems from the attitudes of others. People may often dwell on what they perceive as their own shortcomings which in turn increases their nervousness. This clearly becomes a recursive problem spiralling the individual into isolation. Living within an institution can often make a person become over reliant on the help of others and so lowers confidence and self esteem (Gwalter, 1995).

There may be many reasons why a persons might find travel difficult. Obvious examples include impaired sight or hearing, confinement to a wheel chair, ill health, shyness or perhaps a lack of energy. Furthermore, even if someone achieves his or her destination, there is often no guarantee that they will be able to obtain access or that the facilities will be available for them to comfortably participate in group activities. This leaves many people reluctant to leave the comfort and security of their own home and thus leads them to a solitary life style. Even though there might be others living in the same building, it is often more difficult to keep on good terms with people that one see all the time, especially when the building begins to give the impression of imprisonment.

Some disabilities may severely impair communication. The most obvious category being those that affect the speech. Communication may also be impaired by a lack of control in body movement. Where someone has both difficulty speaking and in controlling facial muscles, that person will find it harder to communicate emotion. This is of particular importance because it hinders the building of personal relationships at any level.

3.2 Applicability of Virtual Reality

Virtual reality has many features that make it suitable for aiding the fight against isolation. Most importantly it offers the potential for the cost effective creation and support of tailor made environments in which people can meet without leaving their own home. Although this might sound futuristic it should be noted that multi-user worlds containing many thousands of participants are already being developed by the military. Research at the Universities of Reading and

Nottingham is developing a virtual environment that may support large numbers of users connected over the Internet or Telephone lines. This project is sponsored by BT and Silicon Graphics and so has the potential to lead to low to medium cost systems on the market.

Virtual reality allows people to be placed in synthetic environments and thus almost any role may be played or scenario explored. Many every day experiences, for example not being able to get through a door, can be reflected with reasonable accuracy and effect. Although virtual worlds are computer generated and, as yet, lack some sensual stimuli such as inertia, taste and smell, they can offer a much greater feeling of 'presence' than watching a film. In practice the mind quickly adapts and the user stops worrying that the experience is not real. This feeling of presence is reinforced because the virtual experience is interactive, that is, the state of the world is affected by the user. As user actions will be in response to the world they perceive, it can be seen that a closed-loop effect will emerge where the experience of the user will depend upon their own actions.

Within a virtual environment, the user may choose an embodiment to represent the way in which they want to be perceived. People that are uncomfortable about communicating with those with disabilities might be able leave their prejudices behind when confronted with a neutral embodiment. This could then educate them that people are the same underneath.

Embodiment may allow people to escape from feelings of self conscious. There is no reason that a person in a wheel chair may wish to be represented that way in a virtual environment. Perhaps a small person might find they can communicate in an entirely different way if they take on a persona that resembles a body builder. This might seem negative and suggest that it would encourage an individual to rely on an “imaginary” self image. This has analogies with the film, The Mask, where the timid hero can only be brave when he wears a mask of special powers that transforms his body and character. One of the morals of this film is, however, that the hero begins to realise this and analyse why he relied on an image. Using a persona in a virtual medium might give someone initial confidence and after persistent use might help them realise that what they look like should not effect who they are.

Conversely, a person might well be proud of how they look and in such cases a reasonably accurate representation would be chosen. Furthermore, it might be interesting for an able bodied person to masquerade as someone with disabilities to find if others treat them any differently.

4. OVERCOMING PHYSICAL CONSTRAINTS IN A VIRTUAL MEETING PLACE

Imagine a virtual meeting place that may be entered by anyone with a cable connection to their home and a personal computer. How would such a meeting place be structured ? A number of issues have to be tackled that would allow users to: find the people you wish to meet; find a good place to interact; and find suitable ways to interact.

4.1 Finding people

Just as in the real world, not everyone will want to talk to each other, especially at the same time. There must therefore be ways in which you as a participant may locate those with which you wish to converse or interact. Once you find them you might want somewhere to go that will suit the kind of interaction in which you wish to partake. If you want a personal conversation then you do not want to stand in a crowd. Also analogous to the real world, it is often easier to get on with people if you can join them in some activity. For example you might wish to play a ball game or go to the theatre. This suggests the concept of a city, but in a real city it is hard to find new friends and it takes a long time to get somewhere.

The concept developed by the DISCOVER group is a maze where people can wander around and find others of a similar vane. The questions that springs to mind is how do you find someone in a maze and how do you know what they are like. The answer is that the maze can be made up of sectors or neighbourhoods that reflect interests and personalities. For instance, some people might congregate around the football pitch which would be located in the sports sector. Others might wait in an ornate garden, in the park sector. Unlike a city the maze would be hierarchical. At the top level, sectors would be represented as miniature three-dimensional icons. Any of these icons can be reached in under a minute and when entered would expand into the sectors. This is a well used concept in virtual environments where icons are used as portals between worlds.

Once inside a sector, how do you find individual people ? If you had pre-arranged the meeting then you could have pre-arranged the meeting spot. Alternatively you might know where certain people can often be found. What about if you wish to find someone new ? The answer is that people can advertise. In the real world people use clothes to divide themselves into sub-cultures. This can be expended in the virtual environment to entire appearances. Perhaps a cat lover may take on the embodiment of a cat. The problem with physical embodiments is that you must be quite close to see them. Increasing there size would not be practical as everyone would do the same and the environment would simply change scale. Furthermore, someone's appearance is not the best indication of his or her mood. The DISCOVER group propose that sound or music is a good way to attract initial interest. Once initial interest has been established, physical appearance, body posture and gesture are useful to gauge a more detailed impression of an individual. Music is strongly connected to emotion and so an individual could choose music that represented the way they felt. This is a natural thing to do and as such would require little or no training. An immediate problem that springs to mind is that this would create a cacophony of confused sound in the environment. This is however not necessarily true. Some virtual reality systems, for example dVS, incorporate sound that may be positioned at any point in a three dimensional space. Such a virtual sound source fades naturally as you move away from it in the environment.

4.2 Finding Somewhere to Go

When you find the people you wish to be with you may well want to go somewhere. You could perhaps just take a stroll around the sector or perhaps leave it and go to another. The maze is hierarchical and so each sector may itself hold portals to related meeting places. Perhaps one portal might lead to foyer from which any one of a number of theatres might be reached. Alternatively you might wish to take your companions to a personal place that may tailored to the kind of interaction you wish to take place or might just represent the kind of person you are. Such an environment is very important as an aid for getting to know people. We all feel more comfortable in our own home and its an important part of social interaction to occasionally act as the host. People typically decorate their home to reflect their self image. This is more difficult to do if you are severely disabled. As this is your own place you might as well carry it around with you as a turtle carries her shell. Imagine being a magician and able to conjure up a doorway into your own home no matter where you are.

4.3 Aids to Interaction

Many people may just wish to talk but that can be done in the real world with a telephone. Virtual reality opens up many more possibilities for group interaction. Perhaps you might wish to go as a group of spectators to a ball game or perhaps you would like to participate in the game. In the project under development at Reading and Nottingham, the first test application is a virtual arena in which the audience can watch the game and hold conversations with the people sat next to them in the stadium. Analogous to some television shows, members of the audience may even be invited down to participate.

Such group activities, although important, diverge from the work of the DISCOVER group and the emphasis of this paper. We are more interested in interaction scenarios that will encourage the communication of emotion.

5. COMMUNICATION OF EMOTION

The DISCOVER group has spent considerable time discussing the issues that the members find important. Of these issues those of difficulties in building relationships, isolation and loneliness where thought paramount. These arise from numerous causes including lack of self confidence; lack of opportunities for self expression; and the lack of control over ones life (Wood et. al, 1995). It is felt by the group that the medium of virtual reality has strong potential for addressing the issues by providing an environment, that when entered, partly alleviates these underlying causes.

5.1 Why use Virtual Reality to Express Emotion

So why do we want to communicate emotion, can we not do that in the real world ? The answer to that question is that some people can more easily than others. Just shyness or lack of self esteem can prevent someone from saying how they feel. More fundamentally, emotion is usually conveyed in the tone and volume of the voice and by facial expressions. Not all people have full control of these. In addition, people may wish to be able to communicate with others whom it is often impractical to regularly visit. Furthermore, the times when we need to convey emotion are usually not predictable and so long journeys are out of the question.

Virtual reality is a potentially rich medium that opens up many new ways of representing emotion. We are however not starting from scratch. Many lessons can be learnt both from observing the way the body exhibits emotion and ways in which artists and musicians have expressed emotion through the ages. The advantage with virtual reality over conventional artistic mediums is that it combines their qualities. Virtual reality can have the colour of paintings, the three dimensions of sculpture, the complexity and essence of music, and the dynamic nature of a performance artist.

5.2 The Human Body

The human body conveys as much information about emotion as the voice. Three aspects of the human form are of particular importance in judging emotion: the expression on the face; the movement and position of hands and arms; and the body posture. Many artists have used recreations of the human form to depict emotion. Egon Schiele, for example, makes strong use on the face, hands and arms in self portraits that provide an emotive insight into his state of mind.

The human face contains many muscles and measuring and modelling it in real time requires considerable computational resources. One system controlled a virtual face representing the games character Super Mario (????). To control this representation, measuring joints were attached between main facial muscle groups and a light wait exoskeleton. This allowed the virtual head to give fairly realistic face movement when talking.

Fairly complex human figures are already used in some virtual reality applications. Most of these applications are presently supported by a human modelling system called Jack (Badler, et al, 1995). The movement of these figures may be completely controlled through external input such as tracking devices, be fully or partially autonomous. Fully autonomous figures are controlled solely by a computer program and do not represent the actions of a user. Partially autonomous figures may be commanded by the user to move to a goal posture or gesture. Natural movement is then achieved by either undergoing a pre-defined set of transitions or through inverse kinametric modelling. It can be seen how a set of expressions could be found to depict given emotions and off line interpolation used to provide natural movement between them.

Modelling natural movement however requires considerable computation restricting such simulation to expensive computers. The required computation may be reduced by only allowing pre-recorded movements, or by reducing the resolution of the figure in terms of the number of moving joints.

Other research has shown (Waldrop, 1995) that reasonable interaction may be achieved by the user, controlling through tracking devices, the head, arms, hands and orientation of the torso. This work did not however address the communication of emotion.

5.3 Colour

Colour is a particularly powerful tool in the representation of emotion. Just using basic colours such as red, blue and yellow, provides a good indication of basic emotion. A prototype system at Reading represented a user as a box with three simple facial expressions mapped to the above colours. Although primitive, this provided an obvious representation of happy, sad or neutral moods. Subtler blends of colour could depict a wider set of emotions as well as there level. For example, Klimt's painting of the daughter of the King who was locked in the tower to protect her innocence, but is then seduced by the devil disguise as an offering of gold, uses a colouring of the cheeks that leaves little doubt to her emotion.

5.4 Sound

Speech is, for the majority of people, the primary form of communication. Emotion is conveyed in speech through its pitch, volume, speed and content. For people with full control of their vocal functions, the most effective method of communication through sound is simply to communicate their voice. Not all people are however able to convey emotion in their voice. Systems are already available that convert text to speech. These could be adapted to change pitch and volume with emotion.

Other sounds can evoke given emotions. It is thought that the brain has developed to produce emotional queues for the cortex that are triggered by given sounds and pitches (Peterson, 1996). The probable reason for this is that we must instinctively respond to given sounds. For example, anger or perhaps fear must be instilled by sounds of aggression, and a parent must be given incentive to respond to cries from his or her baby. Music is one of the most powerful mediums for the expression of emotion. This is probably because it combines these instincts with rhythms that resonate with the natural clocks in the body.

As mentioned earlier, it is relatively trivial to match music to given moods. For example, Vivaldi's Spring quartet gives a feeling of well being, Mahler's Resurrection instils strides of powerful moving emotion, whereas Ravel's Ballero changes the mood to romance and the making of love. What is more difficult is dynamically following the level of mood or even changing the mood entirely. Although a composer might be able to express such changes, it would be difficult to automate the creation such evocative music on a computer in real time.

Both colour and sound are good mediums for the expression of emotion. When combined the effect can become much greater. This effect can be seen at any night club or musical show. Kandinsky related the sound of orchestral instruments to colour, for example, green represented the sound of a violin. A prime example of the power of such combinations are the concerts by the rock band Pink Floyd. In these concerts the combination of music, icons, video and light controlled by music sensitive computer programs, brings powerful feelings of emotion, even to those of us that would not usually listen to the music alone. Pink Floyd have shown how changing mood through lighting and colour may be achieved through computer in the real world. It can be seen how this effect could be even more powerful in a virtual environment where the computer software can change any aspect of the environment it chooses.

There is no reason why colour may only be used on the face and body, or why sound must come from the mouth or an instrument the figure is holding. Virtual reality provides the potential for a persons emotions to be dynamically depicted in their clothing and even their surroundings.

5.5 Clothes

People often wear clothes that match the way they feel. In the real world, it is impractical to change our clothes with our mood. A user could build up a wardrobe of virtual clothes just as they would in the real world. In stead of choosing between each garment before changing they could catalogue outfits to map certain emotions. For example, a black shirt would probably represent a sombre mood, whereas a colourful tee-shirt would represent a more relaxed frame of mind. While in the virtual environment, the underlying system would then swap and change the clothing of the user to match their mood. Unlike in the real world, these clothes could be given predefined behaviours to represent the level of the mood they depict. For example, the tee-shirt could get brighter as the user becomes happier, or a dress might get shorter as the user becomes more ... well that can be left to the imagination.

5.6 Aura

It is sometimes possible to subliminally sense emotion emanating from a person in the real world. This could be represented in the virtual world be surrounding the figure with a transparent aura of colour, sound or a combination of the two. Although this is less natural than the methods discussed above, it does offer a simple and effective method of communicating emotion. The DISCOVER group has investigated the use of colour surrounding a figure. Mary Gyes pictures (Fig. 1) shows an easily recognisable iconic representations of a happy and an angry person surrounded by a supporting coloured auras. One can imagine the effect of animating such a figure in a three dimensional virtual environment.

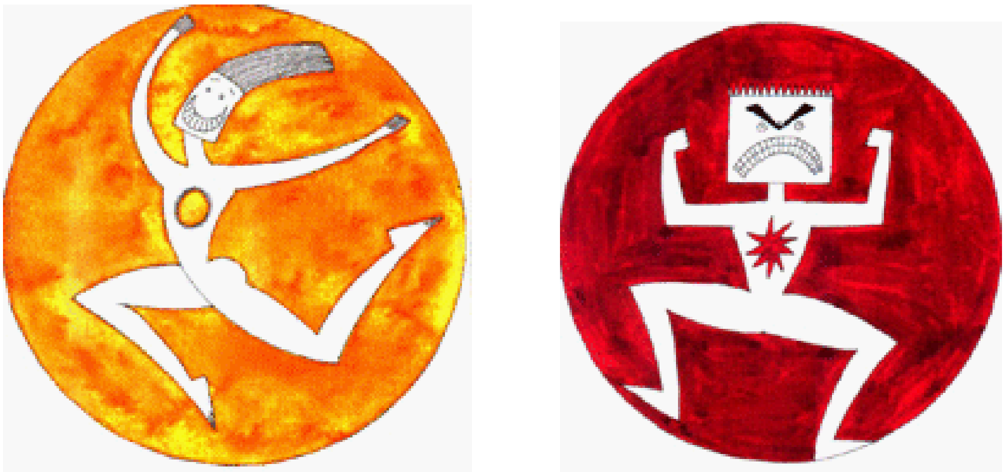


Figure 1. *Happy and angry icons*

5.7 Dynamic emotional environments

The previous section introduced the concept of interacting in a custom personal space within the environment. The esoteric of this environment could be chosen in advance to represent or encourage a given mood. In virtual reality any object may be given behaviour and may interact with the users. This suggests the concept of the esoteric of the space dynamically changing with the mood of those that occupy it. At the most natural level, the music and level of lighting could change.

Taking a leap away from what we are used to, why not have the surrounding environment subtly changing its own essence. For example, the colour of the walls and furniture of a room could change. In a landscape environment, sun could represent a feeling of well-being. If the conversation starts to anger the participants, the sky could cloud over. Extremes of emotion could drive clouds racing across the sky or even evoke thunder. Claire Woods picture of the Cloud Room (Fig. 2) combines the concept of a physical structure surrounded by clouds to give a picture of tranquillity. Many readers will dismiss this as a fantastic exercise, but there is little doubt that such phenomena would be emotive. It must be remembered that the virtual representation of a human is limited in the underlying emotion it can instil in others. Using the environment could act as an amplification of impressionism that could make up for these deficiencies.

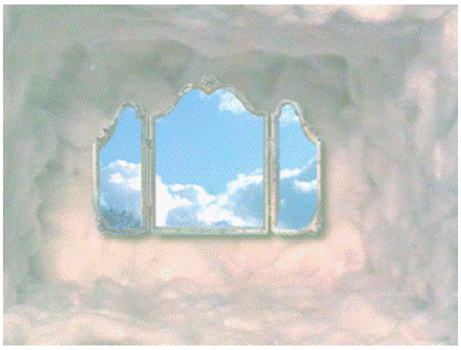


Figure 2. *The Cloud Room*

Members of the DISCOVER group have produced portraits that clearly depict given emotions in the scene and have discussed ways in which these scenes may be enhanced in an interactive virtual environment. For example, Eric Phipps depicts his feelings of isolation as a baby devoid of human contact and left in a colourless, machinery packed factory. In the virtual environment, the machinery would relentlessly whirr and just before another user comes into sight of the baby he will hear it crying. If the user decides to approach the baby, the human contact would slow the machinery and the colour of the environment would lighten. Were the user to interact in some way with the baby it would stop crying. In a similar picture, Eric depicts isolation as a child on the top of a skyscraper being watched by distant bystanders (Fig. 3).

Figure 3. *Isolation of a baby*

Michael Solomon depicts isolation in a portrait that places himself in a wheelchair at the center of an empty stage in a packed theatre (Fig. 4). This picture of isolation would quickly change if members of the audiance joined him on the stage. In another painting Micheal represents fear and perhaps feelings of a lack of control as being pushed away from his companion and down the side of a mountain (Fig. 5). In a dynamic virtual environment, the slope of this mountain could perhaps represent the sum of Michael's anxiety and the anger of his companion.

Figure 4. *Isolation of a stage*

Figure 5. *Anxiety of a hill*

6. MEASURING EMOTION

We have discussed how emotion may be both statically and dynamically represented within a virtual environment. What is left is the question of providing the system with an indication of the emotion of the users. Furthermore, would a user necessarily want their emotion to be reflected automatically or would they want some control. Considerable research has been undertaken by psychologists in an attempt to measure emotion. Ekman used three psychophysiological measures: heart rate, finger temperature; and skin conductivity (Ekman et al, 1990). These give a good measure of the level of emotion with finger temperature and skin conductivity also helping to distinguish between fear and anger.

Although varying levels of emotion can be monitored, it is in general difficult to distinguish between given emotional states. A current undergraduate project at Reading is developing a prototype system that when told which emotion the subject is undergoing will measure its level and represent this with a simple iconic face in a virtual environment. A set of buttons will allow the user to select the emotion they are feeling and the above measures will be taken to determine the level.

7. CONCLUSION

The DISCOVER project has brought together people of different abilities, background and experience to discuss ways in which virtual reality could help those with disabilities. From a number of application areas, the group has chosen to concentrate on tackling the problems of isolation through the expression of emotion in a virtual medium. The concept of a virtual meeting place was proposed, where people could overcome the problems mobility, access and self consciousness. Many people find it particularly hard to express emotions. This may be brought about by psychological problems such as shyness, or physical problems such as voice and muscle control. The proposal of a virtual meeting place was developed into a novel concept for facilitating the expression of emotion with others who might be separated by a considerable physical distance. A system supporting such a concept could potentially tackle many of the underlying causes of isolation. Although no such system exists, it was shown that all the building blocks where either already available or under development. It was shown how colour, sound, music, body posture and the environment itself could represent the emotions of the participants. Finally, methods for a system capable of ascertaining a persons emotion was investigated and the principles of a prototype emotion monitoring system introduced.

Acknowledgements: The authors would like to thank Phil Clark of BT; Roger Allison of Silicon Graphics; and Ambitron Ltd. for providing loan of, or access to, equipment. Thanks to Phil Smithe of BT; and Paul Sharkey, Paul

Sandoz and Martin Cooper of Reading University's Department of Cybernetics for facilitating a workshop that provided the impetus for this conference. Thanks to Michelle McIntosh and Andy Richardson of the Department of Cybernetics for their input on the measuring of emotions; Thanks also to John Ladle and Linda Bunn of Acting Up; and to Mr and Mrs Wood of the International Cerebral Palsy Foundation. Finally thanks goes to the other members of DISCOVER: David Angus, Toby Brandon, Malini Chib, Chris Gilliard, Mary Gye, Ian Gwalter, Eric Phipps, Sarah Potts, Steven Ocane, Michael Solomans, and Stuart Turner.

8. REFERENCES

N Badler, D Metaxas, B Webber and M Steedman (1995), The centre for human modelling and simulation, *Presence*, Vol. 4. No 1.

I Gwalter, (1995), Independent Living", *Ability*.

P Peterson (1996), Music and the Mind, *Channel 4 TV documentary*, May.

D Snowdon and A West (1994) The AVIARY system: A prototype implementation, *6th ERCIM Workshop*, June, Stockholm.

M Waldrop, S Pratt, R McGhee, S Falby and M Zyda (1995), Real-time upper body articulation of humans in a networked interactive virtual environment, *Proc. 1st Conf. Simulation and Interaction in Virtual Environments*, Iowa, p.p. 210.

Wood, C., Gwalter, I., Chib, M., Roberts, D., Wood, M., 1995), Communicating Emotion Using Virtual Reality, *The Royal Society*, June, London.

ECDVRAT '96

Virtual and Enhanced Environments

Day 1: 8 July, 1996, Afternoon

Applications of virtual reality technology to wheelchair remote steering systems

Robert W. Gundersen[1], Stephen J. Smith[2], and Ben A. Abbott[3]

Center for Self-Organizing Intelligent Systems,
College of Engineering, Utah State University
UMC 4120, Logan, UT 84322-4120, USA

[1]*snowvax@cc.usu.edu*, [2]*slngp@cc.usu.edu*, [3]*abbott@sol.ece.usu.edu*

ABSTRACT

The Center for Self-Organizing and Intelligent Systems at Utah State University has been engaged in a two year project to investigate the application of virtual reality and associated technologies as a means for assisting the disabled to steer and control motorized wheelchairs. There have already been several interesting investigations aimed at steering virtual wheelchairs in virtual, computer generated, environments. This paper, however, reports on how this technology may be used to assist, or even completely take over, the task of steering and navigating a real wheel chair in real environments. The basic objective is to arrive at affordable and effective systems that can be used to improve the independence and quality of life of the disabled.

Keywords: wheelchair control, path planning, fuzzy logic, ultrasonic sonar, obstacle avoidance

1. INTRODUCTION

There has recently been an encouraging increase in attention directed toward uses of virtual reality (VR) technology for applications other than commercial video-games and similar forms of self-entertainment. Among the more promising are applications to the fields of medical technology, architectural design, telecommunications and, significantly, products and services designed to help and improve the quality of life for the severely disabled. As examples of the latter, researchers for several organizations have successfully demonstrated the use of VR as a tool for training new users to safely and efficiently operate wheelchairs (for example, Inman et al., 1994) while others have focused on the application of VR as a unique opportunity for architectural designers to experience first hand the problems and difficulties facing the disabled in poorly designed homes and other buildings (Swan et al., 1994).

The Center for Self-Organizing and Intelligent Systems (CSOIS) at Utah State University has been involved since the Spring of 1994 in a project designed to enhance the wheelchair mobility of the severely disabled and the aged. Prior to that time, CSOIS had been actively involved with NASA and other international space agencies, working to develop telepresence and fully autonomous intelligent micro-robotic vehicle navigation and control systems for planned Mars exploration missions (McJunkin et al., 1994). The wheelchair project originated with a suggestion by Dr. Bruce Murray, ex-director of NASA's Jet Propulsion Laboratory, to investigate the transfer of this technology to useful non-space applications.

The sections to follow will describe the technology, results, and possible uses of the CSOIS wheelchair research project.

2. VIRTUAL PRESENCE WHEELCHAIR CONTROL

2.1 Background.

Stereo telepresence control plays an important fall-back role in navigating both the Russian Marsokhod and the NASA-JPL Rocky Rover micro-robotic planetary exploration vehicles. In the event that the vehicles encounter an obstacle or some other terrain feature which baffles the primary autonomous navigation and control systems, dual on-board CCD cameras can be used to send stereo images back to earth-based remote operators, who are then expected to send return signals and manually steer the vehicles around the difficulty. Since CSOIS engineers had participated in the development of this technology, it was natural to select telepresence as the first approach to remote wheelchair control. The idea was simply to provide a remote operator with a joystick, or some other appropriate hand controller, with which to drive a motorized wheelchair from a stereo telepresence image. The image was to be sensed by cameras on-board the wheelchair and transmitted to a computer monitor for display at the operator's work station.

It rapidly becomes obvious, however, that steering a wheelchair by telepresence presents a number of challenges over and above those of steering a small, stable, and relatively slow moving planetary exploration vehicle. Not the least of these problems is the need to take into consideration that the operator of a wheelchair will likely not be as highly trained and practiced as a space scientist or engineer. As a result, the approach was broadened to include a more general sensory environment. Stereo sight and sound, for example, allows an operator to instantly determine not only that the wheelchair is located on a sidewalk, but possibly out in the street and in the path of oncoming traffic. Speakers on the wheelchair permit the operator to ask questions of the occupant or bystanders. On-board electronic inclinometers can be slaved to a gimbaled operator platform to duplicate the orientation of a wheelchair in imminent danger of tipping. In general, the CSOIS idea was to immerse the remote operator in a sensory environment so close to the true environment that the operator could be considered *virtually* present in the wheelchair. To distinguish this approach from the purely visual concept of telepresence control, and to acknowledge the contribution of virtual reality (VR) technology, the investigators agreed that the new technology should be called *virtual presence* (VP) wheelchair control.

2.2 *User Input and Cost Considerations.*

In an investigation of this type, it is easy for engineers to get carried away with the technological developments and to forget the real objectives of the research - i.e., to enhance the quality of life for the disabled. Fortunately, CSOIS has had the significant advantage of being co-located at Utah State University with the National Center for Persons with Disabilities (CPD), which has provided regular access to the real customers for this project, the disabled users themselves. The investigators have also made a point of regular attendance and presentation of results at several Conferences on Technology for Persons with Disabilities, sponsored and organized by Dr. Harry Murphy and his California State University, Northridge, Center on Disabilities (Powell et al, 1994; Smith and Gundersen, 1995; Smith et al, 1995). Feedback obtained in this manner directly from the disabled users of wheelchairs has been instrumental in shaping the goals and objectives of the project.

Cost has been a constant concern throughout the project. Obviously, the objective of improving the wheelchair users quality of life will not have been achieved if the end result is unaffordable. For this reason the center has set the goal of limiting the cost of retrofitting a motorized wheelchair with a remote steering system to no more than twenty percent of the wheelchair cost.

2.3 *Systems Description.*

The CSOIS virtual presence wheelchair system consists of two functional subsystems, the on-board subsystem and the (remote) control station subsystem, shown as Figures 1 and 2 respectively. Each subsystem is controlled by its own processor, with communication between processors provided via an RS-232 serial radio frequency (RF) link. The central processor for the onboard system is an Onset TT8™ micro-controller. The TT8 has a Motorola 68332 processor and peripheral devices such as A/D, UART, Timers/Counters, and digital I/O ports, all on a single 2"x3" board. The primary function of the on-board controller is to receive control commands from the remote control system and convert them to digital signals which drive the wheelchair motor control system. For the current CSOIS design, remote system control commands are received by a Proxlink™ RF modem transceiver, with a range of 1000 ft. (Modems with ranges in excess of this distance are readily available, but, in the USA, must be operated under a license issued by the Federal Communications Commission (FCC).) A program running on the TT8 converts this data to numerical drive commands. The numerical drive commands are then fed to a MAXIM MAX500 D/A converter to generate analog voltages that drive the wheelchair motor controller. One of the design goals of the CSOIS system has been to arrive at a configuration which can be readily added on, or retrofitted, to existing commercially available motorized wheelchairs. CSOIS has been using an ARROW-XT™ wheelchair as a prototype. With the design described above, it has not been necessary to make any modifications to the ARROW motor microcontroller.

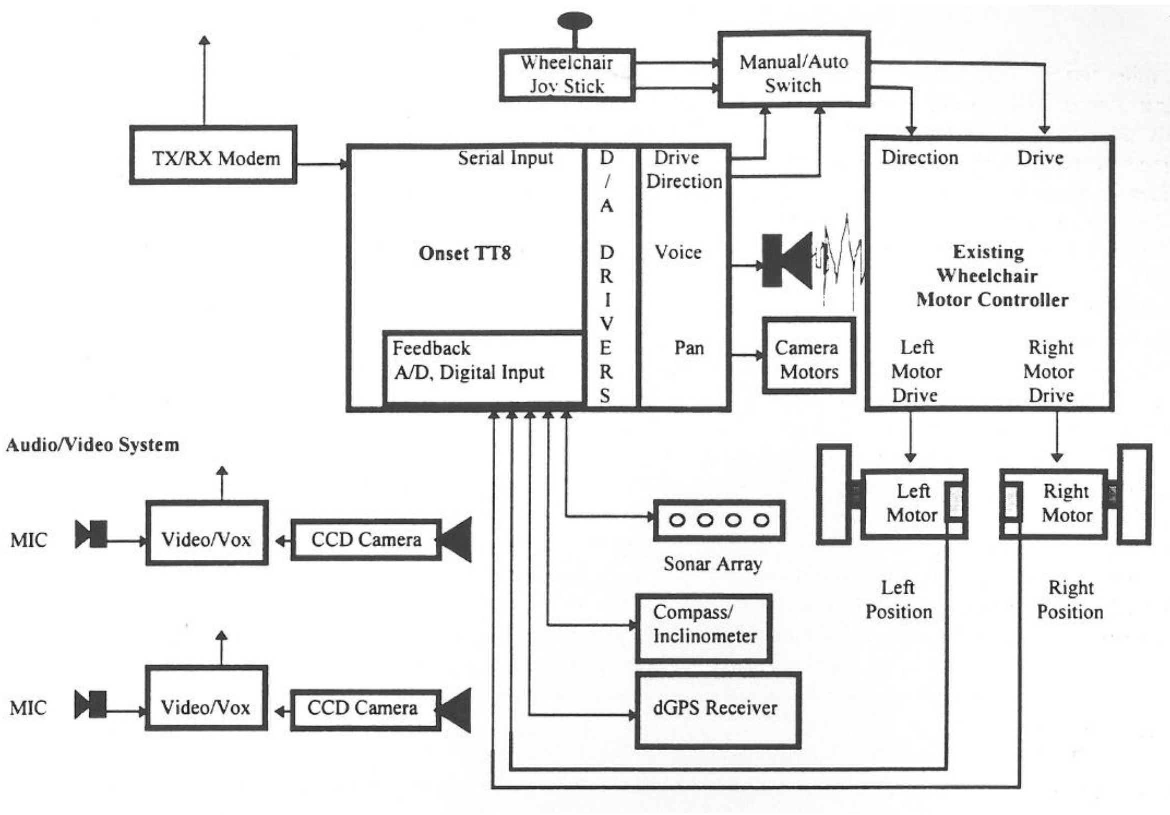


Figure 1. *Onboard System Diagram.*

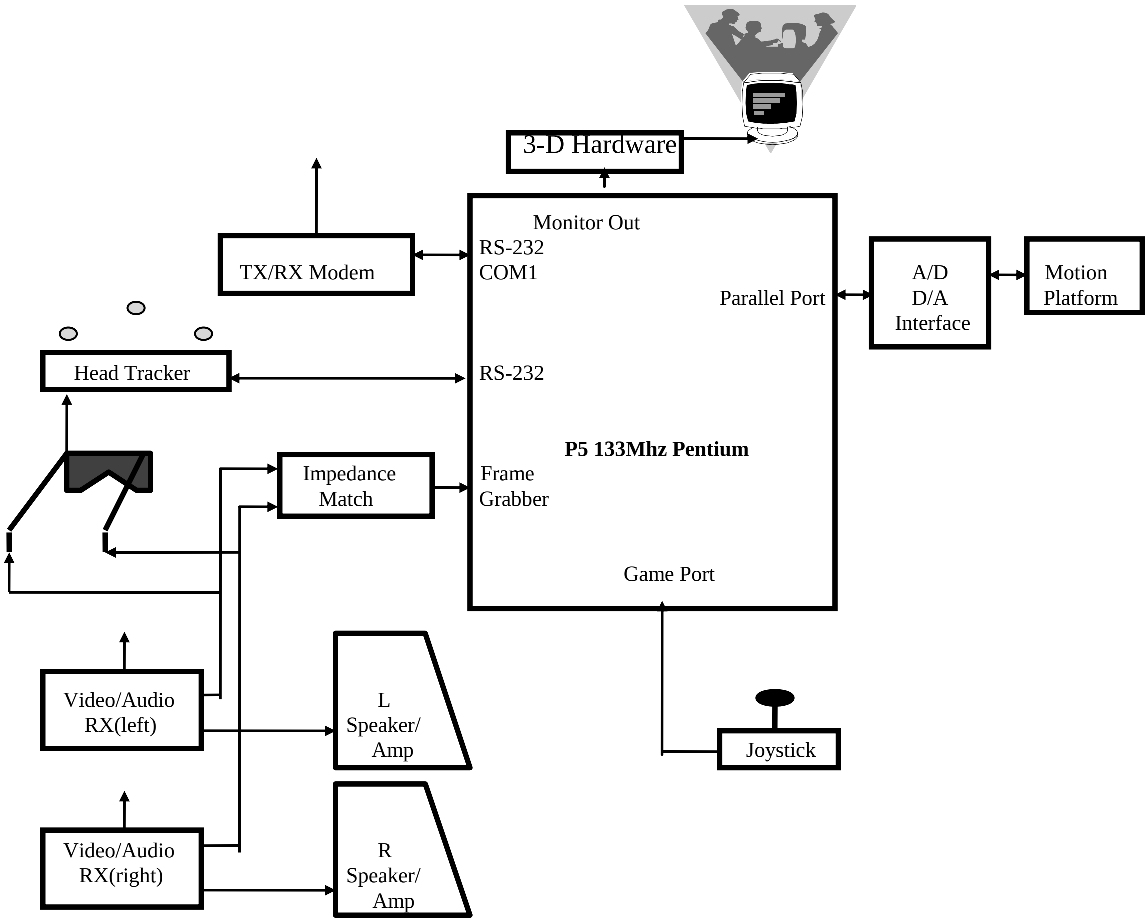


Figure 2. *Remote Control System Diagram.*

The other half of the onboard system is the visual feedback system, which consists of two CCD cameras mounted on a rigid, vernier adjustable, stereo mounting system. This mounting system allows for precise adjustment of the spacing of the cameras and their angle with respect to each other. Both BC-710 color and Pulnix TM-7CN black and white CCD cameras have been used by CSOIS. Two Pelco WLV video transmitters transmit the NTSC signal from the cameras to the remote control system. The range of these transmitters is 1000 ft. In earlier experiments, only one transmitter was used and a video switching system was employed to allow each camera to transmit alternately. To achieve stereo vision, the video signal then had to be de-multiplexed at the remote control system before viewing. While this approach lowered the cost of the system by one video transmitter, the received picture suffered from jitter, sync problems, and a slower frame rate of 15 frames per second. The two transmitter design gives jitter free video at 30 frames per second. The video transmitters include an audio transmission channel. By connecting an amplified microphone to this input, stereo audio signals can be sent to the remote control system.

2.4 The Remote Control System.

Processing at the remote control system end is accomplished using a 133 Mhz Multimedia equipped Pentium workstation. VIRTUAL i-O i-glasses!™ are used for stereo presentation of visual and audio feedback. The remote wheelchair driver currently uses a joystick type controller to input driving commands, with a program running on the Pentium computer converting the output commands to a digital signal for transmission by a RS-232 RF modem to the wheelchair on-board system. Video and audio signals from the wheelchair are received by Pelco WLV receivers and either passed directly to the i-glasses!, or to a frame grabber inside the Pentium if display on a computer monitor, rather than the i-glasses!, is desired. In the latter case, CSOIS has been using a Crystal EYES™ viewing system for stereo visual presentation. The audio signal can be fed to the computer's stereo speaker system.

Figure 3 shows a remote operator, wearing a VIRTUAL i-O headset, and controlling the wheelchair in Figure 4, which is equipped with the on-board system described in section *2.3.* Note that the headset is supplied with a head-tracking unit, which can be slaved to a gimbaled camera mounting system. Thus, the operator is not constrained to "see" only in the forward direction of the wheelchair, but can look around at will.

Figure 3. *VP operator wearing VIRTUAL i-O i-glasses! ™.*

Figure 4. *Arrow-XT ™ configured for VP Control.*

2.5 Uses of Virtual Presence Control.

Perhaps the most likely use of virtual presence wheelchair control is provided by the example of an individual confined to a wheelchair and showing symptoms of the relatively early stages of Alzheimer's disease. This individual may normally be perfectly able to go for a lone outing around the block or down the street and back, without an attendant. However, there is always the risk that the individual may become confused one day, and unable to remember the way home. If the wheelchair is equipped with virtual presence steering, it would only be necessary for the occupant, or possibly a passerby, to press an alarm button to obtain the attention of a remote operator at the base station. Immediately upon donning the virtual reality glasses, the operator will be able to see, hear, and look around in order to determine the location of the wheelchair and take action to either steer it out of harm's way, or possibly all the way back to the operator's station.

Another use suggested for virtual presence control would be as a "helper" system for either the severely disabled, or for new or unskilled users. Such users may, for example, be able to move around a hospital, retirement, or nursing home by themselves for the most part, but could experience difficulties and frustrations if required to enter a room through a narrow door, or enter a crowded elevator. If the wheelchair were equipped with a virtual presence control system, it would only be necessary for the occupant to signal a remote operator to take over, who could then return control back to the occupant once the difficulty was negotiated.

3. FULLY AUTONOMOUS NAVIGATION AND CONTROL

3.1 Background.

Stereo telepresence control has one serious shortcoming with regard to its use on planetary exploration vehicles - transmission time. For example, an estimated forty minutes would have to go by in order to transmit one stereo image from Mars back to Earth, and then relay the operators steering command back to the vehicle. At that rate, an entire mission could be spent trying to get around one rock. Consequently, stereo telepresence is planned for use only as an emergency backup, while primary control is usually left to navigation using an on-board fully autonomous system. CSOIS was actively involved in the development of autonomous systems for use on the NASA-JPL Rocky Rover series of micro-robotic planetary exploration vehicles (McJunkin et al., 1994; Madsen and Gundersen, 1994; Gundersen et al., 1995), so it was natural to look into the question of whether this, or a slightly modified technology could be usefully transferred to the wheelchair steering and navigation problem.

Two possible approaches to fully autonomous wheelchair systems are described in this part of the paper, both based upon the concept of global steering by path programming and local steering via an autonomous obstacle detection and avoidance scheme. The two approaches differ only in the methodology used for path programming - operator-based planning from computer imaging on the one hand, versus a strictly computer-based optimal dynamic programming scheme on the other.

3.2 Operator-Based Path Programming.

Both approaches to global path planning assume the availability of prior maps, floor plans, or reasonably accurate architectural information from which a two or three-dimensional computer model of a building's interior can be constructed. These approaches are more intended for moving about inside a home, a building, or relatively small outdoor area such as a garden, since it is more likely that the necessary prior map information will be available, or easily generated, for the more limited regions. Also, it will shortly be seen that the second, optimal dynamic programming scheme has some very desirable features. However, these features require that the position or location of the wheelchair must be accurately tracked while the chair is en route to its goal. This type of information can be easily obtained using differential global satellite positioning system (dGPS) technology. Unfortunately, dGPS would be prohibitively expensive for this application, and only useful out of doors anyway. An alternate approach is required and will be described in section 3.4.1 below.

The idea of operator-based path programming is straight forward enough provided some form of information is available, such as builder's plans or blueprints, which can be used to build either a two or three-dimensional computer imaging model of the relevant area. For example, Utah State University has computerized most of its buildings and grounds blueprints, which can in turn be used with a commercial program such as AutoCAD☺ to build a two-dimensional computer image of the area, with monitor pixel locations registered one-to-one with the dimensions of the blueprint. Similarly, the blueprints furnish sufficient information to allow a program such as Performer☺ to create a three-dimensional model suitable for viewing by two-dimensional projection techniques, or through virtual reality optical viewing devices such as i-glasses! or Crystal EYES. A computer operator then simply traces the desired path from the wheelchairs present position to the desired location. This data is stored in the form of a sequence of path commands for execution by the wheelchair computer station, and transmitted to the wheelchair on command of the occupant. In reality, little seems to be gained by going to the extra software and hardware expense of three-dimensional rendering of blueprint data, unless the terrain is outdoors where one path might be chosen over another because of steepness or some other terrain feature. Of course, there may be similar indoor features, such as narrow doorways and difficult corners to negotiate which could affect path selection, in which case the expense may be justifiable.

3.3 Computer-Based Path Programming

The objective of the path planning could be as simple as proceeding from the present location of the wheelchair to the desired location in the shortest possible time, or it could be complicated by requiring the chair to travel a route which avoids use of certain congested hallways or wheelchair unfriendly architectural features, such as too narrow doorways or sharp turns. These criteria and preferences can be selected by the occupant on a pull-down menu residing on a small computer monitor on-board the vehicle. Path planning works backward from the specified goal to find paths satisfying the selected path criteria from *all* relevant starting points, including the present position of the wheelchair, to the goal. These paths are optimal, relative to the selected path criteria, and obtained using the method of incremental dynamic programming (Dietterich and Flann, 1995).

If there are no unexpected obstacles to completing the computed path, the wheel chair will start from the present position and follow that particular computed route to the goal. If this is the case, one could say that there has been a lot of wasted computation; i.e., in order to obtain alternate paths from all of the other possible starting points to the goal. However, if someone has blocked off a hallway with furniture, and the computed route fails, then the other computations have not been wasted. The wheelchair obstacle detection system, or possibly the occupant, now tells the computer that it has a *new* starting point, it's current location, and off they go.

What is nice about this scheme, is that once an unexpected obstacle to wheelchair movement is discovered, it is no longer unknown. Indeed, that information can now be added to the building and grounds model used by all of the wheelchairs equipped with the system. The next time any wheelchair path is planned, that obstacle will automatically be taken into consideration. If all of the wheelchairs are able to communicate with each other, say by radio frequency modems, these corrections could even be made while the chairs were *en route* to their respective goals.

One problem with implementing this scheme would appear to be the computational load it adds to the TT8 on-board computer. A similar, but even more computationally intensive scheme, is already being used by CSOIS on another project, however, and has been programmed for and is operating on the TT8. The real problem with implementing the scheme arises from its demands on knowing the instantaneous position of the wheelchair.

3.4 Systems Description.

At present, the path planning systems have been implemented with the master controller located at a remote computing center, using the same communication system as described earlier to relay on-board measurements to the control station, and control commands back to the wheelchair wheel controllers. Future plans call for simply downloading the path plan to the on-board TT8, so that the wheelchair will operate autonomously until a new route or path plan is required. There are some major differences in on-board sensors and instrumentation, however. The next two sections discuss these modifications and their functions.

3.4.1 Position Tracking. Regardless of whether the computer-based or operator-based method is used, the path, once planned, consists of a locus of points on a map of the area, with the coordinates of each point given relative to a convenient coordinate system. Given that the present position and angular orientation of the wheelchair is known in that coordinate system, a vector command can be computed and sent to the controller on-board the wheelchair, which instructs it to precede at a certain angle for a certain distance in order to arrive at its next programmed position. By using this one-step-ahead protocol, it is possible to minimize the cumulative effect of position error. Nevertheless, a reasonably accurate method for tracking the present position of the wheelchair is required. For example, it may be necessary to make a turn and enter a room through a narrow door. The positioning system must be accurate enough to avoid the obvious consequences of missing the doorway. In terms of planning itself, knowing the expected position error allows the path planner to accept or reject various available paths.

As mentioned earlier, dGPS systems offer one way of tracking position. Currently available dGPS systems can track within a few centimeters. Unfortunately, systems with this kind of accuracy are still prohibitively expensive, and are generally suitable only for outdoor applications. Consequently, it is necessary to rely upon some form of improved dead reckoning. In the case of a wheelchair traveling on a level surface, this can usually be accomplished with satisfactory accuracy by counting wheel revolutions, together with the use of an electronic compass to obtain instantaneous heading information. Center investigators have interfaced two Hewlett Packard (HP) HEDS-6310 optical encoders to the Onset TT8 for relative angular measurements, mounting one encoder on the left and one on the right wheel drive motors. Assuming the heading command is executed first so that travel is in a straight line, the average of these two counts provides an accurate distance measurement, once again assuming a level and reasonably smooth surface. True heading information is obtained from a Precision Navigation Vector-2x electronic compass that has been interfaced with the TT8. If even greater accuracy is required, it is possible, at the expense of some additional instrumentation, to update the absolute position of the wheelchair at a few way sites common to most of the possible route choices. For example, infra-red checkpoints could be established at each end of a hallway.

3.4.2 Obstacle Detection and Avoidance. Because the planned path is obtained from "historical" data, such as maps etc., it is always possible to encounter unexpected obstacles in the way of the wheelchair. Thus, an effective path planning system has to include some method for detecting and reacting to their presence. Since the method is to be autonomous, coping with the obstacles has to be done automatically. This is not a new problem. Obstacle detection and avoidance is also a system design requirement for planetary exploration vehicles and an area where CSOIS has concentrated a considerable amount of time and interest (Madsen and Gundersen, 1994; McJunkin et al., 1994; Gundersen et al., 1995). It may not be necessary to transfer all of this technology to the wheelchair application, however, since both the type of obstacles and the conditions are likely to be quite different.

For example, it has already been assumed that the autonomous path planning approaches of this paper are most feasible for moving occupied wheelchairs around buildings or limited outdoor areas. Thus the wheelchair will be spending a lot of its time in hallways or sidewalks, where an unexpected object will more than likely be in the form of another wheelchair, a person, or some similar dynamic object. In such cases, all that should be necessary is to detect the presence of the obstacle, stop, and wait for the obstacle to move, or be moved. Just in case, the TT8 can be used to time the duration of the wait, and if excessive, a call for assistance and the present position of the wheelchair could be issued via the central computer. CSOIS investigators have found that mounting a single inexpensive Polaroid Ultrasonic Sonar sensor with a 30 degree cone of ultrasound searching directly in front of the wheelchair, works quite well. Students walking down the hallways are usually out of the path before the wheelchair even comes to a stop. Some of the more preoccupied faculty, on the other hand, appear to need obstacle detection and avoidance devices more than the wheelchair.

3.4.3 Fuzzy Obstacle Detection and Avoidance. Of course it is possible to imagine applications where the simple obstacle detection and avoidance scheme of the preceding section is insufficient. For example, it may be necessary for the wheelchair to pass through a lobby with furniture and decorative plants. These obstacles would not show up in building blueprints, and could be moved around regularly. This possibility and similar situations could alone justify the extra expense of an autonomous obstacle avoidance system. However, there is another use for such a system which may be of even more value. That is, it could be used to reduce path planning accuracy requirements. For example, even with careful and conservative planning to take in possible position and orientation errors, it is still possible that the chair could arrive at a narrow doorway too far off its mark. Sensing that a collision was imminent, an obstacle detection system could take over and successfully steer through the door.

Two approaches to obstacle detection and avoidance have been investigated by CSOIS. In this section we shall describe a fuzzy logic scheme, which is based on a similar system developed by CSOIS (with the encouragement of NASA-JPL) for possible secondary mission uses of the JPL Rocky Rover class of micro-robotic planetary exploration vehicles (McJunkin et al., 1994). The objective of that project was to use the intelligent systems technologies of neural networks and/or fuzzy inference machines to arrive at a fully autonomous navigation and control scheme, without making any changes whatsoever in the already established Rocky Rover sensor, computer, or electronic hardware systems. Sensors available on the Rocky Rover include an array of five striping (planar) lasers and two CCD cameras, all mounted at suitable angles to provide a "view" of the scene in front of the vehicle. On a perfectly flat surface, the light appears as a straight line. However, any object interrupting the flat surface causes a discontinuity, or a jump, to occur at a point representative of the distance to that object. By aiming the cameras at known non-coincident angles, elementary trigonometry can be used to determine the approximate height of the object, using the distance from where the pixel should have been to where the interrupting object had displaced it.

The CSOIS approach was to use only the pixels corresponding to the intersection of *three* scan lines of the CCD image with each of the five laser lines appearing in the image, *a total of only 15 active data points*. Figure 5 illustrates how this minimal information can be translated into an approximate, or *fuzzy*, map of the terrain immediately in front of the vehicle. For the wheelchair application, the three lines were chosen to correspond to distances of one, two, and three chair lengths in front of the sensors. As shown, each intersection point serves as the center of an ellipsoid. A fuzzy membership function is then used to color the ellipsoid so as to provide a measure of the "hazard" presented by an obstacle, with a null membership value indicating the absence of an obstacle altogether and a membership of unity indicating a certainly impassable object. Using the traffic light analogy, these membership values have been mapped into a color continuum ranging from green ("0") to amber ("0.5") to red ("1"). It will be noted that the rectangular figure in the center of the approximate image is surrounded by ellipsoids, twenty six in all. As the vehicle moves forward, the three rows of ellipsoids in front of the vehicle are allowed to shift backward one length. In effect the fuzzy display "remembers" what was in front of the vehicle.

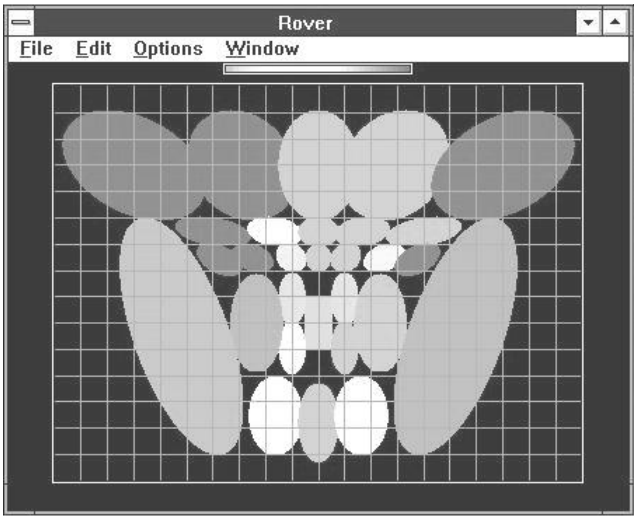


Figure 5. *A fuzzy approximate map.*

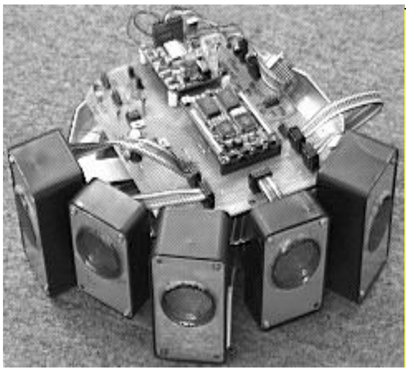
Figure 6. *Actual scene in front of chair.*

The fuzzy display is not directly used in the actual system. Instead, a human operator trains on the fuzzy display until expert in detecting and avoiding obstacles in the wheelchairs path. The expert operator then devises a set of rules of the form, " If the goal is straight ahead, and If all of the ellipsoids directly ahead are nearly green; Then steer straight ahead". In this manner, a rule-base was built with twenty six fuzzy input variables corresponding to the twenty six ellipsoids surrounding the vehicle, three fuzzy input variables corresponding to desired heading, and three fuzzy output steering commands; "steer to the left", "steer to the right", "straight ahead" and "straight back". The scheme has been used to successfully steer through many examples of cluttered and complex room arrangements. It is rather slow, but that seems not to be a problem, since it would only be needed when an unexpected obstacle is encountered. It is accurate, and is able to successfully negotiate its way through narrow doorways and around tight corners.

The laser system approach does have two problems associated with it. The first is obviously cost. The second problem arises from concerns that the laser array may be alarming to some individuals in the vicinity, or at least annoying. There is also a perceived safety factor, even though the power level of the lasers is within the safe region. For these reasons, CSOIS has been investigating a second approach, which uses ultrasonic sonar sensors instead of lasers.

3.4.4 Ultrasonic Sonar Obstacle Detection and Avoidance. Although development of an ultrasonic alternative to the laser sensing array has not been entirely finalized at the time of this writing, results to date suggest that it should serve as a more than satisfactory and significantly less expensive replacement for the laser system. The current concept calls for an array of five Polaroid 6500 series sonar ranging modules to be arranged in a pattern such as that shown in Figure 7 (a). Each sonar emits a 30° cone of ultrasonic 50 kHz sound, measuring distance to first objects in the cone with an error of less than 2 cm at distances from 0.26 to 10.7 meters away. The amount of energy return is proportional to the object surface area tangential to the sonic arc (Maslin, 1983). Figure 7 (b) shows one way to display the output of such an array for the scene of Figure 7 (c), as seen from a camera mounted and looking straight ahead of the wheelchair. Just as in the case of the fuzzy display of laser-sensed data, a human operator uses the sonar display to detect and avoid obstacles lying in the path of a wheelchair controlled by the operator. After becoming "expert", the operator then devises a set of rules for a fuzzy inference machine, which then is used to replace the operator.

Proc. 1st Euro. Conf. Disability, Virtual Reality & Assoc. Tech., Maidenhead, UK, 1996
©1996 ECDVRAT and University of Reading, UK ; ISBN 0 7049 1140 X



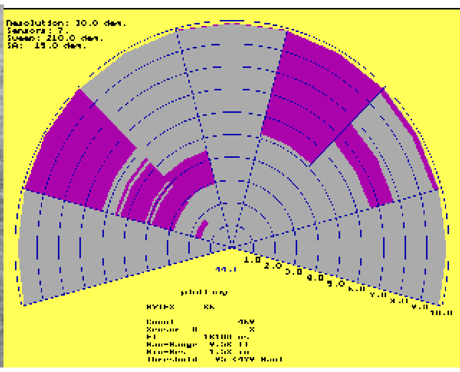


Figure 7 (a). *Sonar array* **Figure 7 (b).** *Sonar map* **Figure 7 (c).** *View from wheelchair*

In addition to being less costly, the sonar system is also much faster than the laser system, since is not necessary to proacess the laser images obtained by the CCD cameras (which can be eliminated, along with the striping lasers). Finally, most humans will be completely unaware of the sonic energy emitted by the sensors. A higher frequency than 50kHz may have to be used, however, if pets or guide dogs are in the area.

3.5 Uses of an Autonomous Wheelchair Controller.

As discussed above, an autonomously controlled system of the type described in this section would serve quite a different purpose than the previously described “emergency” virtual presence system. It would allow severely disabled individuals a degree of independence and freedom not otherwise attainable, allowing them to move around the home, hospital, or care center without constant supervision.

An interesting possibility which has been suggested is to use an autonomously controlled wheelchair during the training process; for example, by providing frustration-free mobility during “off-training” hours, or by serving as a back-up takeover in case the trainee becomes overly tired. Further, demands on the human occupant could be reduced by designing the system to augment human control of the wheelchair. In this sense, it may be possible to design the autonomous system to gradually relinquish control in favor of the trainee, as the individuals skill and confidence increases.

It may be valuable to combine the skills of a human operator and an autonomous system. The occupant could assume responsibility for “steering” the chair by pointing it in the desired direction, relying on the autonomous system to keep it on the desired path and to avoid obstacles. With such a scheme, a wheelchair user could, for example, relax and enjoy window shopping in a mall, without having to be constantly on guard for other persons or objects getting in the way. In another scenario, the autonomous system might be used only to stay on lookout for dangerous situations such as curbs or stairways, automatically stopping the wheelchair or simply sounding a warning alarm to alert the occupant to the hazard.

4. CONCLUSIONS

In this paper we have reported on the results to date of a two year investigation at Utah State University into the uses of virtual reality and its associated technologies for assisting the disabled in steering and controlling wheelchairs. There have been several interesting investigations of the use of virtual reality for steering virtual wheelchairs through virtual, computer generated, environments. The purpose of the USU investigation has concentrated instead on using the technologies associated with virtual reality to operate a real wheelchair in real world environments. The results to date have been quite encouraging and satisfactory. In particular, it appears that systems can be realized that are both useful and cost effective.

Acknowledgements: The authors would like to acknowledge the support of the Marriner S. Eccles Foundation and the State of Utah Centers of Excellence Program for their support and encouragement of this research.

6. REFERENCES

T G Dietterich and N S Flann (1995), Explanation-based learning and reinforcement learning: A unified view. *Proceedings of the International Machine Learning Conference*, Snowbird, UT.

R W Gundersen, G Olsen, and G Powell (1995), A fuzzy-inference scheme for fully autonomous control of Rover class planetary exploration vehicles, *Proceedings of the 46th International Astronautical Conf.*, Oslo, Norway.

D Inman, J Peaks, K Loge, and V Chen (1994), Teaching orthopedically impaired children to drive motorized wheelchairs in virtual reality, *Proceedings of the 2nd Annual International Conference on Virtual Reality and Persons with disabilities*, San Francisco, CA.

D Madsen and R W Gundersen (1994), An autonomous fuzzy logic Rover navigation scheme, *Proceedings of the 2nd Joint Conf. On Information Sciences*, Durham, N.C.

G D Maslin (1983), A simple ultrasonic ranging system, *Presented at the 102nd Convention of the Audio Engineering Society*, Cincinnati, Ohio.

T McJunkin, G Powell, and R W Gundersen (1994), A unique approach to advanced obstacle detection and avoidance, *2nd International Symp. On Missions, Technologies, and Design of planetary mobile vehicles*, Moscow-St. Petersburg, Russia.

G Powell, T McJunkin, and R W Gundersen (1994), Autonomous wheelchair control by an approximate virtual reality system, *Proceedings of the 2nd Annual International Conference on virtual reality and persons with disabilities*, San Francisco, CA.

S J Smith and R W Gundersen (1995), Telepresence wheelchair control for severely disabled persons, *Proceedings of the 10th Annual International Conference on Technology and Persons with Disabilities* ,Los Angeles, CA.

S J Smith, R W Gundersen, and B A Abbott (1995), Virtual Presence and autonomous wheelchair control: An update, *Proceedings of the 3rd Annual International Conference on virtual reality and persons with disabilities*, San Francisco, CA.

J Swan, D Stredney, W Carlson, and B Blostein (1994), The determination of wheelchair user proficiency and environmental accessibility through virtual simulation, *Proceedings of the 2nd Annual International Conference on Virtual Reality and Persons with disabilities*, San Francisco, CA.

Integrating augmented reality with home systems

J.C. Hammond, P. M. Sharkey, G.T. Foster

Department of Cybernetics, University of Reading
Whiteknights, Reading, RG6 6AY, UK

J.Hammond@HSOnline.co.uk, P.M.Sharkey@reading.ac.uk, cybgtf@cyber.rdg.ac.uk

ABSTRACT

Augmented Reality systems overlay computer generated information onto a user's natural senses. Where this additional information is visual, the information is overlaid on the user's natural visual field of view through a head mounted (or "head-up") display device. Integrated Home Systems provides a network that links every electrical device in the home which provides to a user both control and data transparency across the network. This paper discusses the integration of AR techniques with a real world infrastructure and demonstrates this through the integration of an visual AR prototype with an existing Home Systems technology test rig. The integration of these two technologies provides the basis of a specialised information/control device that allows the control and diagnostics of Home Systems devices from the basic action of looking at them.[1]

Keywords: augmented reality, home systems, control devices

1. INTRODUCTION

Virtual Reality has been the subject of intense development over recent years with much of the technology being driven by the military and entertainment's (computer games) industries. As a result of advances in headset technology, associated model and graphics handling techniques are now available that are opening up new application areas. Augmented Reality (AR) is one such area. Augmented reality has been previously investigated in the military and automobile industries, primarily through the development of the Head-Up Display (HUD), where visual information regarding the current dynamic state of the vehicle (e.g. and aeroplane, helicopter or car) is displayed to the pilot by projection of this data directly onto the windscreen of the vehicle, or (more expensively) onto the visor of a pilot's helmet. The principal advantage this type of display offers is the provision to the pilot of essential information (airspeed, altitude, attitude, etc.), whilst allowing the pilot to maintain visual coverage of the airspace through which he is flying. The European Commission research project *Prometheus* has taken this concept one stage further, by investigating how a computer generated scene may be overlaid unto the real scene to enhance the user's performance. Specifically, by instrumenting a car with cameras, other sensors and real-time image processing computers, the scene in front of a driver may be enhanced by overlaying outlines of the road and any obstacles directly onto the windscreen, even when such obstacles are obscured by fog or in night driving conditions.

This paper presents an adaptation of this concept with the specific application area of disability, and details the development of a prototype visual AR system. The central idea is to provide a person with (non-visual) disability additional visual information that may assist them on a day-to-day basis. Thus, a person with mobility difficulty may be able to look about an environment and, though the additional information provided by the AR display, ascertain what objects within their field of view may be influenced. By coupling such an information provider to some form of control one can easily envisage the user visually selecting a functional object and controlling the object by simply looking at it. Home Systems provides the ability to associate function with common objects in a typical home, such as doors, windows, lights, alarm and safety systems, and so on. We show: how a simple, inexpensive, head mounted display may be adapted to allow a user to view a typical room, where the 3D structure of the room is overlaid in wire-frame onto the actual scene; how this overlay can update in real-time as the user looks around; how functional items within the room can be identified and highlighted through the display as the user looks at them; and, finally, how the function may be accessed in a simple intuitive way.

[1] This project was developed as a final year project by the first author for the degree of BSc in Cybernetics and Control Engineering, from the Universit of Reading, UK

 ISBN 0 7047 1140 X

1.1. *Augmented Reality*

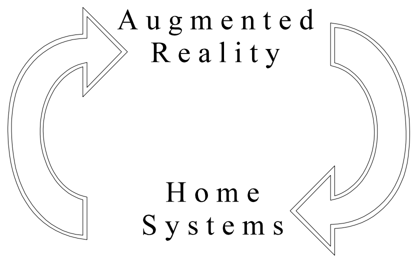
A visual AR system overlays three-dimensional computer images over a user's normal field of view through the use of a 3D head-up display. AR is different from VR because the model world that it creates is not an abstracted 'virtual' world but a close 'mirror' of the actual environment. The ideas of modelling closely correlated mirror worlds can be seen in Gelernter (1992). Virtual Reality systems are currently at a state where the computing power needed to generate a believable and useful environment is still not widely available. However, the graphical display requirements for AR systems are reduced since the amount of computer graphics can be relatively small compared to the field of view but still maintain useful information.

The close correlation with actuality makes the AR system particularly useful for real world Human Computer Interface (HCI). In order for any AR system to operate effectively it must have complete information about the environment in which it is operating. Further, to make the system a useful tool it must be able to alter the state of an object in actuality and reflect that change in the computer model. Hence a method is required to enable measurement, control and feedback between the computer model and the real world.

1.2. *Home Systems*

For the domestic environment a technological infrastructure capable of supporting an AR system exists in the Home Systems (HS) control network. HS provides a network that links every electrical device in the home, this provides both control and data transparency across the network. Integration of AR with HS opens up the possibility of modelling the domestic environment and providing a novel method for interaction within the home. A multitude of scenarios are possible, from opening a window simply by looking at it, to changing the heating control system from examining a radiator. The implications of such technologies for rehabilitation can be found in Cooper (1996). HS and AR are therefore mutually supportive technologies.

- Information infrastructure.
- Actual devices to control.
- Link to "actuality".
- Network model.



- Human Computer Interface.
- Extends HS location model.
- Novel alternate control method.

1.3. *Structure of Paper*

The paper is structured as follows. In Section 2 we review Home Systems and introduce a number of projects associated with this work. Section 3 discusses the hardware development of the Augmented Reality System and prototype head-mounted display, while Section 4 presents an integration of the Augmented Reality System within a Home Systems test rig set up within the HS Research Laboratory. This report concludes with a wider discussion on how such a prototype AR system may be further developed.

2. HOME SYSTEMS TECHNOLOGY

2.1. *Home Systems Overview*

The Home Systems (HS) specification, which defines networked control systems for the domestic environment, was developed by the ESPRIT Home Systems Consortium (Project 2431) and the integrated Interactive Home (Project 5448). The specification is upheld by the European Home Systems Association (EHSA). HS supports networked communications on a variety of media including twisted pair and power line. Nodes are based upon the common 8051 family of microprocessors and the HS system is an 8 bit addressed token passing network.

In HS terms, devices connected to the network fall into the following two categories:

- *Complex Device* (CiD): The CiD also has no system 'intelligence', however it does support token passing and is therefore able to broadcast data onto the network. An example CiD may be a door locking device that signals back onto the network whether or not it has successfully completed its task.
- *Feature Controller* (FC). A feature controller is a device that contains 'system intelligence', i.e. in the application layer of the device it is able to process and act upon information from network devices. The nature of the system information processing capabilities of a feature controller is entirely up to the designer. An

example FC might be a security system feature controller. This controller is able to gather information from sensors and issue controlling actions. Also it would be able to communicate with other feature controllers where appropriate. For instance it might alert a communications feature controller to phone for help in an emergency situation.

2.2. *The HS-ADEPT Project*

The HS-ADEPT Project (Home Systems - Access of Disabled and Elderly People to this Technology), funded by the European Commission TIDE Initiative, is co-ordinated by the Department of Cybernetics, University of Reading, see Ferreira et al (1995). The aim of the project has been to develop emerging Home Systems technologies to provide the technological infrastructure to allow independent living for disabled and elderly people. The work at Reading has concentrated on the system integration and control issues relating to networked technologies. Project developments are currently being installed in the homes of 3 disabled people in the UK and a rehabilitation centre in Portugal, where they will be subject to independent real world trials.

2.3. *User Interface Feature Controller*

2.3.1. Structure and Model. The major problem when attempting to control any networked system is defining the physical location of nodes around a network. The problem manifests itself when trying to control a specific device across the network as each similar device actually looks the same. So choosing which particular light you want to turn on from a large list presents several control problems. This problem is made more acute with concepts such as 'plug and play' where devices can actively be moved around a home. The solution adopted by the HS-ADEPT project was the development of a User Interface Feature Controller (UIFC) that holds local information about control facilities available in a certain location.

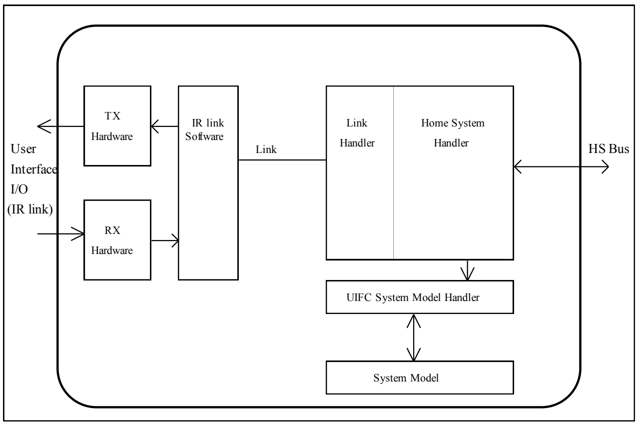


Figure 1. *Schematic of the User Interface Feature Controller*

Fig. 1 shows the schematic of the UIFC. The philosophy behind the UIFC is very simple: it works on the premise that devices are most likely to be operated in the immediate vicinity of the user. The user is tagged to a specific location through an Infra Red (IR) communication link, the relatively short range of IR communications effectively locating the user to the nearest transceiver. In practice, a UIFC is placed in each room of the house so control actions are defined on a room by room basis. The IR communications are handled by a separate microprocessor board and the validated codes are passed to the UIFC node via an RS232 link.

The UIFC maintains a local model of devices that are accessible from that point. The model is initiated by placing the UIFC in a 'Learning Mode' and enrolling the necessary devices. So, for a particular room, the UIFC may learn the local devices such as the lights, doors and windows. However, there is no constraint as to which devices are learnt in this way. So if global control actions are required, such as the security and safety system or the front door unit, they can simply be learnt as well. The System Model Handler then keeps a dynamic model of devices relevant to it's location.

In a typical control scenario the IR link receives a message, for instance to turn on the light. This message is parity checked and, if valid, is then passed on to the UIFC node. The model then attempts to find the local light and operate it. Feedback is passed back through the bi-directional IR link as to whether the control action was successful. The advantages of using the local FC concept are a direct consequence of the UIFC not having to hold a detailed representation of the entire Home System network to which it is attached:

- Software for the UIFC is less complex as the UIFC has less devices to address. This reduces the size of the dynamic model stored, the amount of ROM and RAM space required and the complexity of the software.

- Reduced software complexity produces an increase in execution speed.
- A reduction in the size of the dynamic model means that a control action requested by the user can be implemented faster as there is less time required to search the model for matching network information..
- For the UIFC to maintain an up-to-date dynamic model it must be informed of changes to devices on the HS network. If every UIFC is required to keep track of every device on the HS network this generates a high amount of bus traffic. With every UIFC having to monitor only a subset of the devices on the HS network there is a considerable reduction in bus traffic and, as a consequence, a reduction in the software activity within the UIFC.
- Local control will allow only the activation of devices which are local to the user. The user has visual feedback as confirmation of his action and the safety associated with that feedback.

2.3.2. IR Link. The IR link was designed to give a generic message structure that allows the construction of very cheap and simple User Interface devices. Effectively this means that control devices can be pre-programmed – the command to control a device is the same whichever room you are in – and the UIFC translates the requested action to the appropriate device. The IR link is bi-directional although unidirectional communications are adequate for normal operation – the return message largely ignored. However the return message does allow feedback to the user. For example, the Apple Newton based control devices developed at the University of Reading exhibit several useful features as a result of the bi-directional communications, which include:

- context specific menuing (for instance if the front door bell rings the door entry menu is automatically displayed);
- alarm display; and
- network diagnostics.

The full set of codes for objects can be found in the *Home Systems Specification Command Tables* in *Home Systems Specification Release 1.2*. (1995).

3. THE AUGMENTED REALITY SYSTEM

A visual augmented reality system comprises three major elements: computing, optics, and sensors. The computer stores a reasonably accurate 3D wire-frame model of the environment within which the user is traversing. Head tracking sensors can establish the location of the user within the environment and the direction in which a user is looking (this assumes that the user's eyes look along a 'Cyclopean' or gaze direction). The computer model estimates the field of view of the actual environment and presents this information, via Liquid Crystal Displays (LCDs) back to the user through dedicated optics. The optics (LCD screens, half-silvered mirrors and lenses) ensure that the wire-frame model coincides with the real world view. The hardware implementation of a prototype AR system is detailed below.

3.1. Overview

The AR system is shown schematically in Fig. 2, which details the information transfer between different devices within the system. The AR system is based around an application developed on a PowerMac 8100AV computer which controls the interaction between the head tracker, the graphics display devices and the IR link to the HS network. The primary link of the system is that between the PowerMac and the headset as well as the receiver for the head tracker. The direction of the arrows in Fig. 2 denotes the direction of communications in the current system – the link between the PowerMac and Home Systems is unidirectional, while the communications between the PowerMac and head tracker are bi-directional. For the integration to the Home Systems network (see Section 4) the PowerMac utilises an IR link for serial communications.

The display is produced using a pair of LCD television screens with 512 x 768 pixels. The screens are mounted on the headset in an assemblage of optics consisting of lenses and half-silvered mirrors, thus allowing the images generated by the screens to be overlaid onto the users visual field.

The position and orientation tracking of the headset is performed by a Flock of Birds™ (FOB) Tracker (from Ascension Technology Corporation). This has a range of c. 2.5m from its transmitter. The FOB works by transmitting

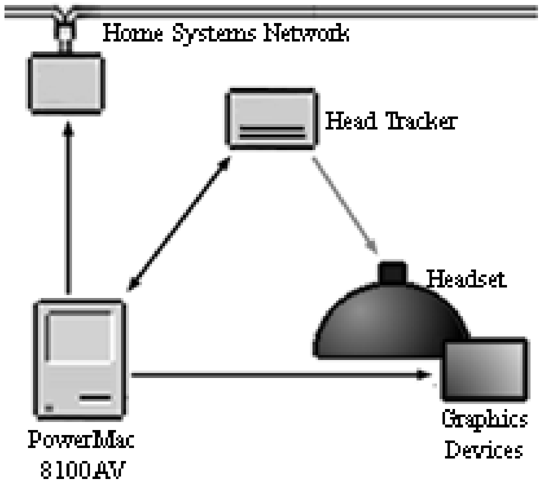


Figure 2. *System communication schematic*

a pulsed DC magnetic field which is received at the headset and all six degrees of freedom are determined by the measured field characteristics. The head tracker is capable of up to 144 measurements a second and has a translational accuracy of approximately 2mm with an angular accuracy of 30 arc-seconds. The sample rate of the FOB exceeds the required update rate for any display device by a factor of 8, and the resolution is approximately equal to that of the human eye when the distances for this application are considered. Due to the geometry of the AR headset prototype, one pixel represents approximately 50 arc-seconds of the user's field of view. As a result, there is little degradation in the quality of the display attributed to the accuracy of the head tracker.

The software on the PowerMac is an object orientated application that performs all calculations on the model of the environment. These allow the graphics to be updated given the position of the user in the environment. The image is generally a wire-frame representation of the environment so that the user is able to clearly see the outside world without obstruction. This also cuts down on the time it takes to render the graphics and therefore the latency of the system with degrading the information conveyed.

3.1. Optical Hardware

The optical hardware consists of half-silvered mirrors and lenses mounted in front of the user's eyes allowing each eye to simultaneously view both the outside world and the images on the LCD displays. The configuration of the assembly can be seen in Fig. 3. This shows one half of the optics on the right side of the headset. For each eye the optical components consist of a half-silvered mirror mounted at 45 degrees that allows the user to see through the headset and to simultaneously see the computer generated images. The images are focused, via a lens, to a suitable distance in front of the user. Without the lenses the user would not be able to focus on both the graphics and the environment at the same but would have to continually re-focus between the two. The main problems to be solved for the optical hardware were:

- *Graphics field of view:* the field of view should be as large as possible without degrading the quality of the graphics.
- *Brightness of the virtual images:* the brightness should be sufficient to allow the user to clearly see the virtual images under normal lighting or include the provision to alter the brightness and contrast to suit the conditions.
- *Focusing of the virtual images:* the images should be set at a comfortable distance in front of the user, probably between 3-5 metres.
- *Alignment of the virtual images:* the images should be aligned accurately enough to allow the user to easily associate the virtual images with their physical counterparts. They should also be accurate enough to prevent incorrect interaction with adjacent devices, for example, two windows physically close in an environment should have separate images with no overlap, thus preventing one window to be activated when the intention of the user was to activate the other. This covers physical alignment of the optics as well as factors affecting distortions in the images, for example aberrations caused by the lenses.

The half-silvered mirrors were chosen to have a 50:50 reflect-transmit ratio so that they give equal luminance to the virtual and real images. Any further adjustments are made by the brightness and contrast of the LCD displays. The mirrors are mounted 30mm from the users eyes covering the entire visual field not obstructed by the assemblage. This means that the real world is seen with uniform clarity .

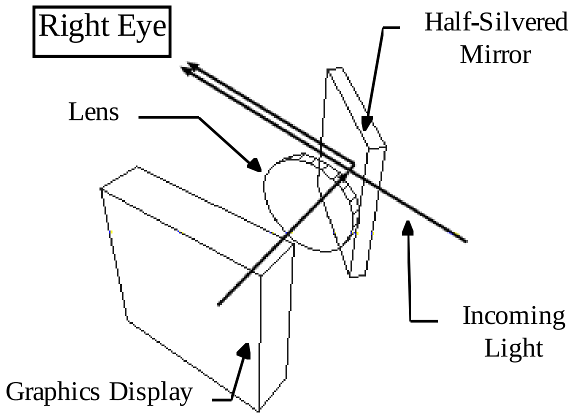


Figure 3. *Basic Optical Configuration (Right Hand Side)*

Each lens is mounted 40 mm from the mirrors and perpendicular to the user's line of sight. The lenses are plano-convex with diameters of 50mm and a focal lengths 125mm. Hence graphics screens placed at a distance equal to the focal length appear to be at infinity, allowing the user to see these combined with the real world at anything beyond approximately 3m. The size of the lenses mean the virtual image covers 36.32 degrees of the users visual field. The eyes normal field of view being approximately ±100° horizontally and ±60° vertically but most of the activity the eye concentrates on is focused on the fovea. This combined with the fact that most of the images displayed will probably be at distances of 3m or more means 36.32 degrees is sufficient for the application. Because each lens is less than 73 mm from the eyes no light is lost from the graphics displays and when viewing the images the effects of aberrations are minimal. These do exist but only at the extremes of the images away from the concentration of the fovea.

3.2. Display Software

The software written specifically for the application is based on the XYZ Fixed Angles transformation given by J. Vince (1995) to position the user in the virtual environment. This uses a single homogeneous equation to rotate and translate the environment before the graphics for each eye are mapped on to a two dimensional virtual plane (one for each eye) given the focal length of the lens. For speed these calculations manipulate integer values wherever possible and therefore do introduce slight errors in the transformation although in vital areas such as the measurement of angles long values are used. The errors noted in this transformation are negligibly small but it is the intention to eradicate these at a later date by scaling up all values to be significantly large integers, including those currently being represented by longs and only return the values to their correct size immediately before they are mapped to the display.

The model of the environment is stored as a series of shapes in a linked list. This list is constant once the model has been defined. A second identical list is used to store the transformed model after every movement of the user. The model is then clipped using a simple algorithm determining all those objects behind the user. More advanced clipping is not required as the graphics are only wire-frame and do not need such common 3D procedures such as back face removal. The simplicity of the operations cuts down on the latency of the system allowing the addition of more specific procedures for object selection.

The most application specific procedure detects whether the user is facing an object. Each object with functionality is assigned one or more facets, defined by a series of four 3D co-ordinates. If the vector projected from the user straight through the centre of the virtual screen intersects this space then the user can activate the object. The procedure works by projecting a vector in the model given the users position and orientation, three points are taken on the plane of the facet and a vector normal to this is calculated. After simple checking to ensure that the two normals are not parallel indicating that the users line of sight never intersects the plane, equation (1) shows how the point of intersection is ascertained.

There are a series of restrictions for the model that make the equations for the positioning of the point of intersection simpler and quicker. The four points of the detection facet must all be in the same plane and the facet must be a rectangle. This is to check whether the point of intersection is within the facet the vector between the point of intersection and each corner of the facet is found and the angles between these and the adjacent side of the facet is measured. If all four are less than 90 degrees then the point of intersection is deemed to be within the facet. These

©1996 ECDVRAT and University of Reading, UK ; ISBN 0 7047 1140 X

restrictions are not serious disadvantages since this only requires that all devices be represented by a flat rectangle. The distance of the user from the point of intersection is h given by equation (5). The procedure only activates the nearest of the facets indeterminate of which object it is on. In this way only one object can be selected and acted upon. Later additions could be to assign different actions to different facets on the same object, for example enabling a door be locked from one side only.

$$\mathbf{n}_x (u_x - x_p) + \mathbf{n}_y (u_y - y_p) + \mathbf{n}_z (u_z - z_p) = 0 \tag{1}$$

$$u_x = x + h \sin(\phi_{yaw}) \tag{2}$$

$$u_y = y + h \sin(-\phi_{pitch}) \tag{3}$$

$$u_z = z + h \cos(\phi_{yaw}) \tag{4}$$

$$h = \frac{\text{(N.Point-N.Position)}}{\text{(N.Unit)}^{-1}} \tag{5}$$

where

$\mathbf{n}_x$, $\mathbf{n}_y$ and $\mathbf{n}_z$ are vectors normal to the plane,

$\mathbf{u} = [u_x, u_y, u_z]^T$ is the point of intersection,

and $\mathbf{p} = [x_p, y_p, z_p]^T$ is a point on the plane.

These calculations can easily be modified to allow greater versatility by defining the facet by only three points. This constrains the facet to a single plane whereas four points does not. The principle for ascertaining the relative position of the point of intersection is similar but will not work on the 90 degree calculation rather some other specific to triangles. This unfortunately requires more storage area for the facets as well as a slower more complicated algorithm not required for this application.

3.3. Head Tracker and Calibration

The Flock of Birds position and orientation tracker runs directly out of the PowerMac modem port using the RS232 protocol. On issuing specific ASCII characters to the FOB it will return x, y, z, roll, pitch and yaw information of the receiver relative to the transmitter. The returned information is encoded in 12 bytes, 2 for each measurement which have to be decoded by a series of bit transformations as described in the reference documentation of the FOB. Unfortunately since the FOB only measures the position of the receiver relative to the transmitter an accurate position of the transmitter relative to the base co-ordinates of the model must be known.

The acuity of the human eye is relatively good at 40 seconds of an arc (or more simply the eye can distinguish between two points 1.5 mm apart 10m away). This means that the accuracy of the transmitter needs to be as exact as possible so that the alignment of virtual model and real environment is as close as possible. This is a problem known as static registration and is common in AR systems, mostly unseen in VR due to the lack of frame of reference for the user in a completely immersive display. Static and dynamic registration (the appearance of the virtual image to lag behind the real environment when the user is moving) are being researched by a number of groups including Azuma and Bishop (1994).

To place the head tracker accurately without the need to physically measure the position of the transmitter before the model is initially configured a feedback procedure was implemented. This procedure was designed to automatically calculate the orientation and position of the transmitter and in the process take into account the position of the receiver unit relative to the eyes of the user. The procedure also removes any errors in the positioning of the receiver on the top of the headset and should be specific to each user that configures it. The procedure works by prompting the user to align a cross they are able to see in their display to certain points in the environment of known position. The first point is a cross set in front of a mirror and is used to find the initial offset angles of the tracking equipment relative to the co-ordinate frame of the room. By aligning the cross in the users virtual display with the cross in front of the mirror and its reflection it is guaranteed that as long as the mirror is mounted flat on the wall the user will be on a line perpendicular to the wall and therefore the results provided by the tracker are offsets. After aligning with two more crosses with significant x, y and z displacements it is possible to calculate the position of the eyes of the user and therefore the relative position of the tracker transmitter. The model of the environment is configured from these calculations. This procedure provides a satisfactory static registration for the integration to HS and its control.

3.4. Model

The physical model used by the AR system is provided by the device information held in the Home Systems devices. The principle function of the AR system is the accurate display this model of the environment. The current system relies on an installation procedure for the definition of this model and is determined by physically measuring the environment and recording its characteristics in a pre-defined format enabling a dynamic interpretation by the application. Further developments will have such information available directly from the network.

The format for the environment model contains a variety of information, some unique to that model whilst other symbols represent objects contained within the application. The application is written in such a manner that allows any shape to be generated as long as it is definable by a series of straight lines. A shape may simply be for visual reference only but by the inclusion of specific characters the object is no longer assigned the attributes of a simple shape but takes on the functionality of a window, door, light or any of the objects in the applications hierarchy. By doing this a set of co-ordinates in three dimensional space might be allocated the functions of a window, for example with position and security (locked or unlocked) states. The only remaining set of information required is the 'Detection Frame' defined as the area of the object that when looked at, the user may act upon the object.

The information contained within the application for each object is currently constant although at a later stage this paper discusses the increased benefits of a completely dynamic model obtained from sources other than the user. The application stores data in each of the objects in the hierarchy on the possible states the object can be in as well as handles to menu descriptions that ultimately provide the user with their options; handles to audio sound-bites that provide an additional feedback to any changes in the state of an object and the necessary codes to interface at a basic level to the Home Systems network.

4. INTEGRATING AUGMENTED REALITY AND HOME SYSTEMS.

4.1. The Test System

The test system as shown in Fig. 4 has been implemented in the Distributed Systems Research laboratory with effectively a single room scenario. This employs one UIFC that holds the network control model for a variety of devices including a door actuator, window actuator, lights and a security system.

Presently the AR model has been synchronised by manually entering the devices present in the room. The various devices are highlighted by placing a wire frame representation over the object with basic on/off control actions implemented using a simple switch. Proportional commands such as dimmer intensity can be set through the keyboard. Continuing development will enable the physical model to be automatically loaded from the UIFC model into the AR model.

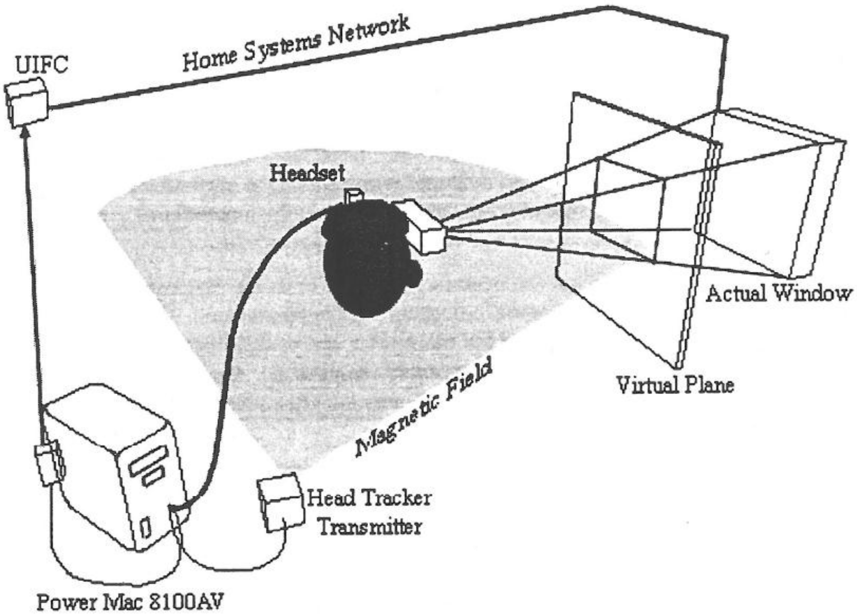


Figure 4. *The test installation*

4.1.1. Accuracy. The system at the moment is not as accurate as was originally intended due to the combination of errors in various separate areas. These are the errors introduced in modelling the environment, inaccuracies in the

Proc. 1st Euro. Conf. Disability, Virtual Reality & Assoc. Tech., Maidenhead, UK, 1996
©1996 ECDVRAT and University of Reading, UK ; ISBN 0 7047 1140 X

calibration calculations as well as errors introduced in the transformation of the model. On their own these are negligibly small but combined they introduce registration errors of up to 10 - 12 inches. All these errors can be reduced with significant work but the system does perform the necessary tasks and the registration error can be avoided with intelligent modelling of the environment, for example, not modelling the edges of an object but using a symbol or hot spot overlaid onto the device to indicate functionality. There are also possible errors inherent in the design of the headset due to the proximity of the head tracker receiver to the metal frame of the optics assembly as well as the weight incorporated to balance the weight on the front of the headset. This could be easily be solved by changing the construction materials but the effect of the metal is currently unknown without removing the compounding errors from other parts of the system.

4.1.2. Comfort. The current headset is a prototype and is not in its smallest possible form. This means that the users head and neck are subject to a substantial weight. This makes wearing the headset intolerable for prolonged periods of time and means that the device would not be viable without serious ergonomic design to reduce weight and increase comfort.

4.1.3. Control. The system is designed to only pick the nearest of any overlapping devices and therefore to control any obscured devices the user must physically move so that the obstruction is removed. For the persons this system is intended this may not be a viable option but it would not be too difficult to re-engineer the software to allow, with specific commands, the user to toggle between different devices intersecting their line of sight. The system in its present state utilises the press of a mouse button to input user action and this means that the entire system is tied to the proximity of the PowerMac. With further redesign of this input device right up to speech recognition capabilities provided by the PowerMac this restraint will be lifted.

4.1.4. Display. It was intended that the wearer would be able to easily see both virtual and real images concurrently. In specific environments this is severely compromised due to ambient light almost drowning out the virtual images. Unfortunately the contrast of the current screens is fixed and only the brightness is variable. Shielding the graphics displays does increase the quality of the display and reduce this problem but the graphics still maintain a constant level and whereas looking to a dark corner of the room provides good conditions for observing the images viewing a bright window does not.

4.2. The User Perspective

At the time of writing there have been no extensive trials of the system covering a wide range of users. First impressions suggest that the system has potential applications for people with quadriplegia or other high level paralysis who are without visual impairment and are able to look around a room. However preliminary findings have raised a number of interesting points.

The optical display need not be high resolution. If a method of highlighting devices by ‘hot spots’ is employed a simplified headset could be employed.

The user feedback does not have to be optical. All the AR systems to date rely on an optical representation of the room. However this information could be provided in an audible form so a headset would not have the display devices just a small earpiece. This could be used to help a visually impaired person control devices or orient themselves in an unknown environment.

To initiate control actions most actuation devices are appropriate. Here any form of switch could be used from suck & puff switches to voice actuation. The only unsuitable switches are those that cause a gross head movement such as head switches.

Factors such as head tremor affect the system. Tremor is a major problem for the AR system, filtering of the tracking data along with broad tolerances on device locations goes some way to alleviate the problem.

4.3. Further Work

The power of integrating the technologies of Augmented Reality and Home Systems has been shown to be to the mutual benefit of information sharing, this is still at a limited level and highlights the need for the parallel development of the technologies. In particular the ability to download the model of the environment directly from the UIFC will be a significant advance, this direct feedback link makes true synchronisation of the AR model and the actual environment a reality.

The first stage would simply be to recreate the model from the devices on the network who would supply physical dimensions as well as detection frames and information determining the identity of each device. Increasing the dynamic nature of the objects in the AR system and removing the descriptions of a window, door etc. to the network there is the advantage of being able to slot new and previously undefined devices onto the system that carry their own dimensions, states, activation codes and handles without the need for re-engineering the hierarchy of the AR system.

Surpassing this the AR system has the potential of displaying almost any visual data the HS network can potentially store about itself. The discussion of this paper to this point has assumed the existence of a single model of the environment which could be described as the 'User level'. The network level error detection capabilities allow errors to be pin pointed to specific pieces of hardware. A model of the physical connections and apparatus on the network could be downloaded to the user under the heading to 'Engineer level'. With this model the user would be able to view the state of the network as well as any available statistics from individual components suitably superimposed onto the corresponding device. This would require that the AR system would hold links to various models of the network held within the UIFC that could be download as requested and updated with any change in state of the environment.

5. CONCLUSIONS

2.4. Extensions to specification to enable AR applications

While HS is able to offer the technological infrastructure and control capabilities that enables real world interaction with Augmented Reality systems, there is, as yet, no physical installation information (i.e. no local model) carried by a particular device. As a result, the AR model must be built up from physically measuring up a room. The next steps in the development process is to produce a number of devices that carry their physical dimensions and installation details. This level of detail may then be stored in the model so that when a person wearing the AR headset enters a room the local model downloads all the necessary physical data for the AR system to build its own environmental model.

As a proof of concept the results from this project have clearly shown the advantages of the integration of AR systems with existing control technologies. For commercial systems to be viable there are some technological advances to be made both in display, tracking and network technologies.

- *Users.* The AR system represents a different way on interacting with the environment, using the system as a control device has considerable potential for a broad range of users. The effectiveness of the dynamic AR model and display depends upon suitable actuation methods.

- *Display*. As the AR system only requires a relatively small amount of information to be overlaid on a users normal field of view there is the potential to use very small cheap LCD units to place 'hot spots' to represent devices. However it is clear that for practical purposes a smaller and cheaper head mounted display is required.

- *Home Systems.* Here HS is demonstrated as an ideal technological infrastructure for complimentary AR modelling. However for real Mirror World type developments physical locations must be provided in node data structures and protocol for the interaction of the HS model and AR model must be developed.

Acknowledgements. The authors would like to thank Wayne Rudge for work on the UIFC; Darren Wenn for hardware development of the IR link; Martyn Cooper and HS-ADEPT for the use of the Demonstration System; and John Wann for ue of the Head Tracker.

6. REFERENCES

R. Azuma & G. Bishop (1994) Improving Static and Dynamic Registration in an Optical See-Through HMD *Proceedings of SIGGRAPH '94* (Orlando, FL, 24-29 July 1994) *Computer Graphics*, Annual Conference Series.

M. Cooper & D. A. Keating. (1996) Implications of the emerging home systems technologies for rehabilitation. *Med. End. Phys*. vol. 18, pp176-180. April 1996.

J. M. Ferreira, J. L. Ramalho, M. Cooper, D. A. Keating. (1995) Smart Homes and the HS-ADPET Project: Applications and User Interfaces. *Proc. ECART3 Conference*, Lisbon, October 1995. pp341-343.

D. Gelernter (1992). *Mirror Worlds*, Oxford University Press, Oxford.

J. Vince (1995). *Virtual Reality Systems.* Addison-Wesley, London.

Home Systems Specification Release 1.2. (1995) Technical Report, Department of Cybernetics, University of Reading.

Using virtual reality in the adaptation of environments for disabled people

R C Davies[1] and J Eriksson[2]

[1,2] Division of Work Environments, University of Lund,
Ole Römers Väg 1, S221 00 Lund, SWEDEN.

[1]Change@Work, University of Lund,
Box 118, S221 00 Lund, SWEDEN.

[1] *roy.davies@amt.lth.se*, [2] *joakim.eriksson@amt.lth.se*

ABSTRACT

This paper describes work on a computer aided design and visualisation toolkit for the adaptation of homes and workplaces for disabled people. The basis for this work comes from a number of case studies which used a prototype planning tool based on ordinary 3D modelling and drawing packages. These case studies highlighted the need for a 3D object library containing office furniture, mobility aids, building construction elements and so forth. Three further prototype tools have also been developed: a user-friendly design and visualisation tool for non-computists based on a 3D graphics API; a mannequin modeller; and a VR based design and visualisation tool.

Keywords: physical disability, adaptations, visualisation, mannequins, 3D object library, Quickdraw 3D.

1. INTRODUCTION

People who have a disability or who have become disabled often require alterations to existing environments to accommodate their needs, or may wish to design their own work places or homes with their specific needs in mind. Traditionally, the design process is performed by architects - professionals who presumably know best what the disabled person wants.

This situation prompted the question: But what about us, can't we have a say in how the environment shall look? This question was raised, of course, by the people with the disabilities themselves and their occupational therapists. Now consider the following statement:

Pictures can convey graphical information more efficiently than words. By the same reasoning, three dimensional visualisation of a potential environment can, for most people, be preferable than textual descriptions or even architectural plans (Wagner, 1994). The former is not usually enough to build up a proper feeling for the environment and the later can be hard to interpret by laypeople. 3D visualisation also promotes communication and is less open to misinterpretation.

Put these together, and one comes naturally to the idea of using computerised three dimensional visualisation as a tool to help include the people who want the alterations, and others affected, in the design process. This not only allows the disabled people to incorporate their own ideas, but ensures that they feel more at home in the environment they have helped to plan. However, there is a complication: The tools that architects use for visualisation - CAD packages - are designed for computer literate professionals. Simpler tools need to be developed that can be used by the disabled people, occupational therapists and other, not necessarily computer literate, people.

Furthermore, if an alteration is made, or an environment built that requires further adjustment and modification due to unforeseen problems, then money is wasted. With more effective visualisation tools at the start, this inefficiency can be minimised.

These points form the basis for the work described in this paper.

2. BACKGROUND

When alterations to an environment are carried out, architects, perhaps in consultation with occupational therapists, engineers and others, discuss what is needed with the help of two dimensional plans and drawings. These plans and drawings can be hard to understand, are open to misinterpretation and hinder the *'how would it be if we put this, here...'* approach. Furthermore, testing usability factors can be extremely difficult, and relies heavily on the experience of the designers with disability related problems. Mistakes can easily be made, and features required for a particular person's needs may be omitted.

By using a 3D design and visualisation toolkit, many people, not just the design professionals, can work together in a medium that is not so easily misunderstood, using an iterative planning process where ideas can be visualised, tested for suitability and re-thought until all are happy with the final decision.

This planning toolkit could be used in the following ways (Fig. 1). A group can work together around a screen, sharing ideas and viewing interesting configurations. When a viable environment has been decided upon, those who actually will be using the environment can 'walk through' it and see whether it meets their criteria. Finally, the computer can be told to test the environment automatically against ergonomic heuristics using the anthropometric measurements of the people involved.

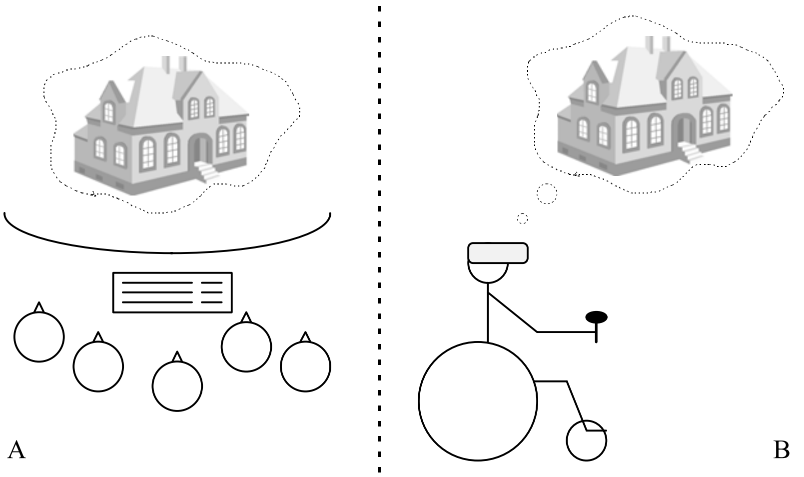


Figure 1. *Alternative Design Strategies, A: Group work, B: Single Person Visualisation*

This idea for a planning tool has been developed into a prototype using conventional modelling tools and has been tested in a series of case studies.

3. THE CASE STUDIES AND PLANNING TOOL PROTOTYPE

Six case studies of real-life planning situations were performed between 1990 and 1993, investigating the usefulness and effectiveness of a computer based planning tool that would enable all people responsible, or affected, to participate on equal terms. The tool was meant to support the planner (who is usually an occupational therapist) in designing and evaluating multiple alternatives at an early stage, and in making improvements throughout the planning process. As well as to be used during planning sessions, in order to visualise suggestions, and in making instant changes, thus enhancing communication and participation among the people involved (Eriksson *et al*, 1995; Eriksson and Johansson, 1996).

The planning tool was prototyped with a set of commercially available software packages in order to conceptualise possible features of a planning tool, and to evaluate the usefulness and efficiency of such features. The prototype involved 3D-modelling and multimedia software based on a Macintosh computer (Apple Inc.). The programs most frequently used were Swivel 3D Professional, Modelshop II, MacroModel, MacroMind Three-D and MacroMind Director (MacroMedia Inc.). Fig. 2 shows schematically the use of the different programs.

Proc. 1st Euro. Conf. Disability, Virtual Reality & Assoc. Tech., Maidenhead, UK, 1996
©1996 ECDVRAT and University of Reading, UK ; ISBN 0 7049 1140 X

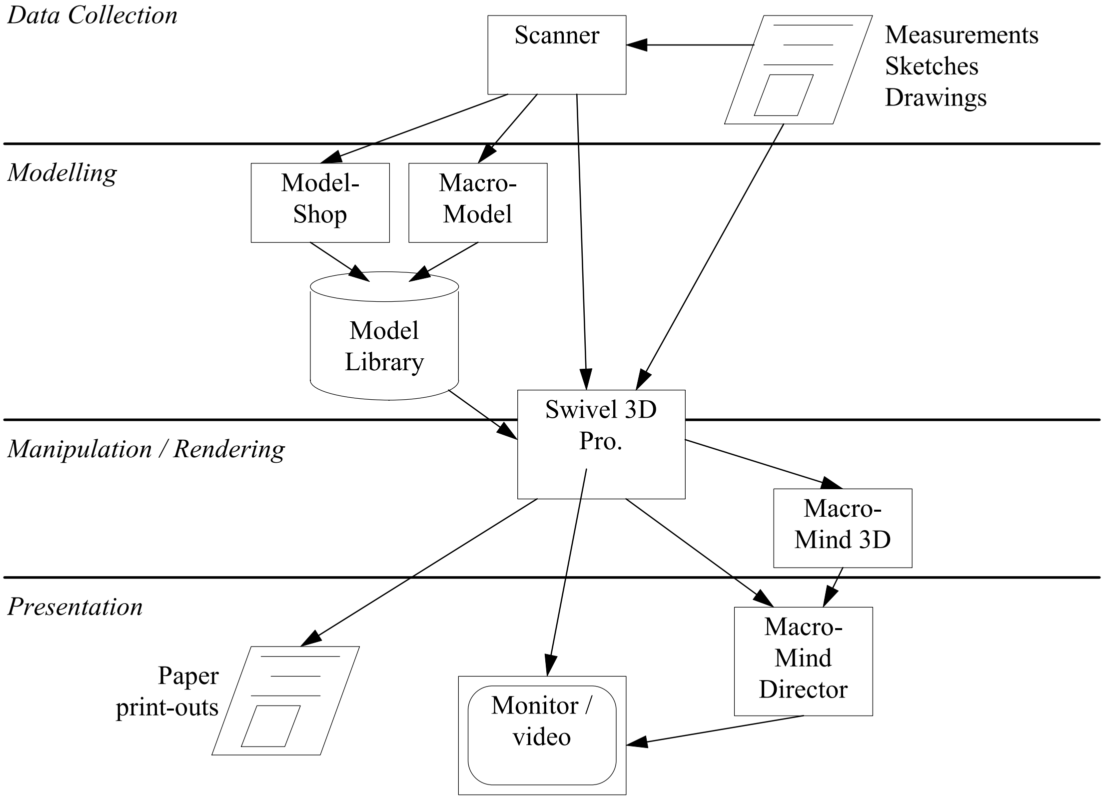


Figure 2. *A summary of the prototype compilation used in the case studies.*

The case studies were evaluated with observations, interviews and questionnaires. In the last three cases, all participating people received a uniform set of questionnaire and interview questions. Table 1 summarises the results of the questionnaire from case studies four to Six (Eriksson and Johansson, 1996). If the subject answered "yes" to the first question, then the options were A="It was possible to visualise a suggestion"; B="It was possible to test clearance for objects/people"; C="It was possible to test reach"; D="Others".

Table 1. *Summary of results for case studies 4 to 6.*

		Useful in the case	Supported understanding	Supported discussions	Encouraged participation	Useful tool in the future
Case 4	Subject	yes; B,C	yes	yes	yes	yes
	Employer	yes; A,B,C,D[1]	yes	yes	yes	yes
	Therapist	yes; C,D[2]	yes	yes	yes	yes
Case 5	Relative	yes; A,B	yes	yes	yes	yes
	Therapist, clin.	yes; B	yes	yes	no	yes
	Therapist, distr.	yes; B,C	uncertain	yes	uncertain	yes
	Engineer	yes; B	uncertain	yes	uncertain	yes
	Home help	yes; A,B,C,D[3]	yes	uncertain	yes	yes
Case 6	Subject	no	no	uncertain	no	yes
	Relative	no	uncertain	no	no	uncertain
	Therapist, clin.	yes; B	uncertain	yes	yes	yes
	Therapist, distr.	uncertain	uncertain	uncertain	yes	yes
	Engineer	no	no	no	no	yes

[1] "Print-outs supported my presentation for the other employees."
[2] "One could see the workplace from an arbitrary viewpoint"
[3] "Provided several alternatives."

4. SUMMARY OF THE CASE STUDIES, THE WAY FORWARD

The questionnaire answers indicated that laypeople seemed to appreciate the tool mainly as a visualisation aid, while the therapists and engineers emphasised the ability to evaluate a solution's functionality. However, the opportunity to interact with the model and to evaluate clearance and accessibility, seemed equally appreciated.

A frequently mentioned advantage was that mistakes could be more easily avoided. All therapists agreed that it is worth spending extra time, if it pays off in better communication and if flaws can be discovered at the planning stage. Some participants remarked that suggestions can be over-elaborated, and that details that are not possible to display graphically risk being forgotten. Concern was also expressed that the tool itself could become too much in focus, thus diverting attention away from the actual planning issues. The level of detail and realism was generally considered sufficient, and one therapist thought that it was better with a coarse animation rather than a video recording for instance, since it may then be easier to focus on the issue of concern, that is, accessibility. In general, the answers indicated that people who were professionally experienced about these kinds of adaptations, required fewer details and less realism than those with no previous experience.

The case studies also indicated that a planning tool, such as the one prototyped, can be useful to a planning group in supporting understanding and active participation amongst all kinds of participants. It also makes it possible for a professional planner to make designs of future environments and evaluate the functionality with high accuracy. Dealing with various kinds of physical impairments, it is important that the human models can be adapted to an individual's size and physical abilities.

The planning tool can support an iterative planning process: Initially, rough models of several alternatives can be presented in order to start discussions. When a certain solution has been agreed upon, the design can continuously, and in finer detail, be improved based on discussions in subsequent sessions. However, to take advantage of such an iterative process, it is important that all the people concerned can attend throughout the whole process to a greater extent than they do today. For homes, it may be important to include construction engineers, nursing and home-service personnel. For workplaces, colleagues, employers and assistants should be represented.

During the sessions, it was apparent that the planning tool was used not only to show pre-manufactured images or animations, but also to interactively view and manipulate the 3-D models. We believe that this is essential in order to support active participation.

With the prototype used, operations such as manipulating viewpoints, rearranging furniture, adjusting postures, etc., were carried out quite easily and directly, but several improvements needed to be made to simplify operations and increase the level of interactivity. Future development should also include frequently requested tools such as measurement control and collision detection.

At the conclusion of the case studies, it was apparent that a toolkit for the design and visualisation of home and work environments should be composed of a number of facets:

1. A library of 3D objects to be used as a basis for environment design.
2. A user-friendly design and visualisation programme.
3. Kinematic and dynamic mannequins based on anthropometric data for added realism and testing of new environments.

Furthermore, a new and exciting tool was becoming available that could be used as a complement to the above: Virtual Reality. All these tools are essential and interwoven (Fig 3), together combining to form a complete product.

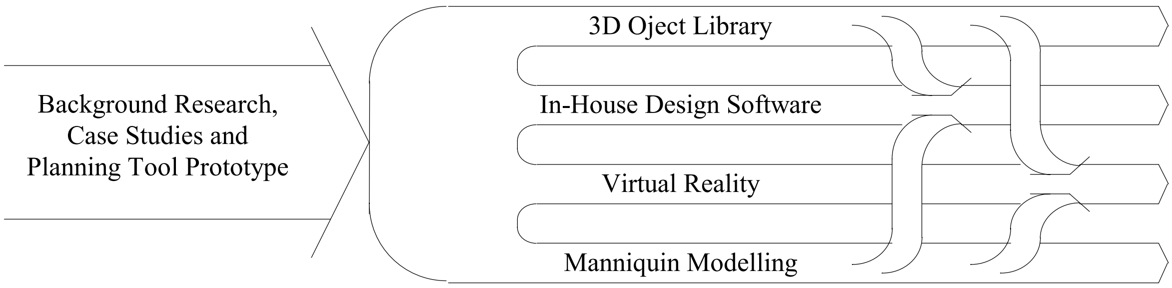


Figure 3. *The interplay between the different aspects of the design and visualisation system.*

These four aspects and their roles in the design and visualisation tool are discussed in the following pages.

Proc. 1st Euro. Conf. Disability, Virtual Reality & Assoc. Tech., Maidenhead, UK, 1996

5. THE NEED FOR AN OBJECT LIBRARY

The first aspect to be considered is the 3D object library. Such a library must be based on a standard file format, consisting of objects such as construction elements, furniture, office equipment, transportation aids, etc. This may concern thousands of articles, and yet one can never expect that the library will cover more than a fraction of all such items available on the market. In the immediate future, it will probably be necessary to concentrate on a limited set of environments such as office-sites, kitchens and bathrooms. To efficiently build a large model library, therefore, it must also be possible to exchange files between different 3-D modelling and CAD programs, but unfortunately, there are presently no generally accepted 3-D graphics formats (beside traditional CAD standards, such as DXF). In the future, one may hope for the breakthrough of modern standards, such as VRML and 3DMF. Maintaining and distributing an extensive object library will probably require a support organisation, which could also be responsible for technical support and in training the occupational therapists.

6. A PLANNING TOOL USING QUICKDRAW 3D

The object library, however, is of no use alone. It needs to be complemented by a design and visualisation tool that is both easy to use and sufficiently powerful to help in the task of environment design and visualisation. CAD packages that are available are too complex for ordinary people, whilst most PC based 3D modelling programmes are either too slow or not adequate for visualisation.

An in-house planning tool prototype is therefore being developed (Eriksson *et al*, 1996), dubbed 'Magrathea' (Fig. 4), which exclusively supports the tasks a common user, for example an occupational therapist, would work with most frequently:

1. fetching construction elements from an object library, such as wall-, door-, and window-modules, and assembling them into one or many rooms;
2. fetching furniture, or other equipment of interest, and creating different interior arrangements;
3. testing ergonomic aspects such as reach, clearance, accessibility, etc., with help of mannequins and, if applicable, transportation aids; and
4. showing and interacting with different suggestions during planning sessions.

In order to provide a cheap and simple program, Magrathea operates in an open environment, where features that are more peripheral to a common users' usual work are supported by separate programs or software modules. For instance, a therapist would rarely have the time and skills to custom-design models, or prepare sophisticated presentations, such as animated sequences and photo-realistic renderings. Hence, the program is backed-up by the 3D object library.

Magrathea utilises QuickDraw 3D (Apple Computer Inc.), which is a 3D graphics technology that provides an open, cross-platform file-format and API (Application Programming Interface) (Apple Computer Inc, 1995). The file format, 3DMF (3D MetaFile), can retain data that constitutes a complete 3D-scene with various kinds of objects and attributes, such as cameras, lights, transformations, geometries, textures, shaders, etc. It also possible to handle custom defined objects and attributes. The API calls an extension module of the Macintosh Operating System (Apple Computer Inc.), which supports file handling, object manipulation, rendering, hardware acceleration, and so forth. (it is reported to also be supported under Windows in the future).

Magrathea features the following:

– interactive rendering with optional hardware acceleration;
– key-controlled camera navigation (walkthrough), and support for multiple views;
– direct exportation and importation of 3D-objects in a standard format (3DMF);
– manipulation tools for moving, rotating, and linking objects;
– browser window for object searching, and assessing a world's hierarchical structure;
– object info window for alpha-numerical control of various parameters and attributes; and
– a tool for measuring distances in a world.

 ISBN 0 7049 1140 X

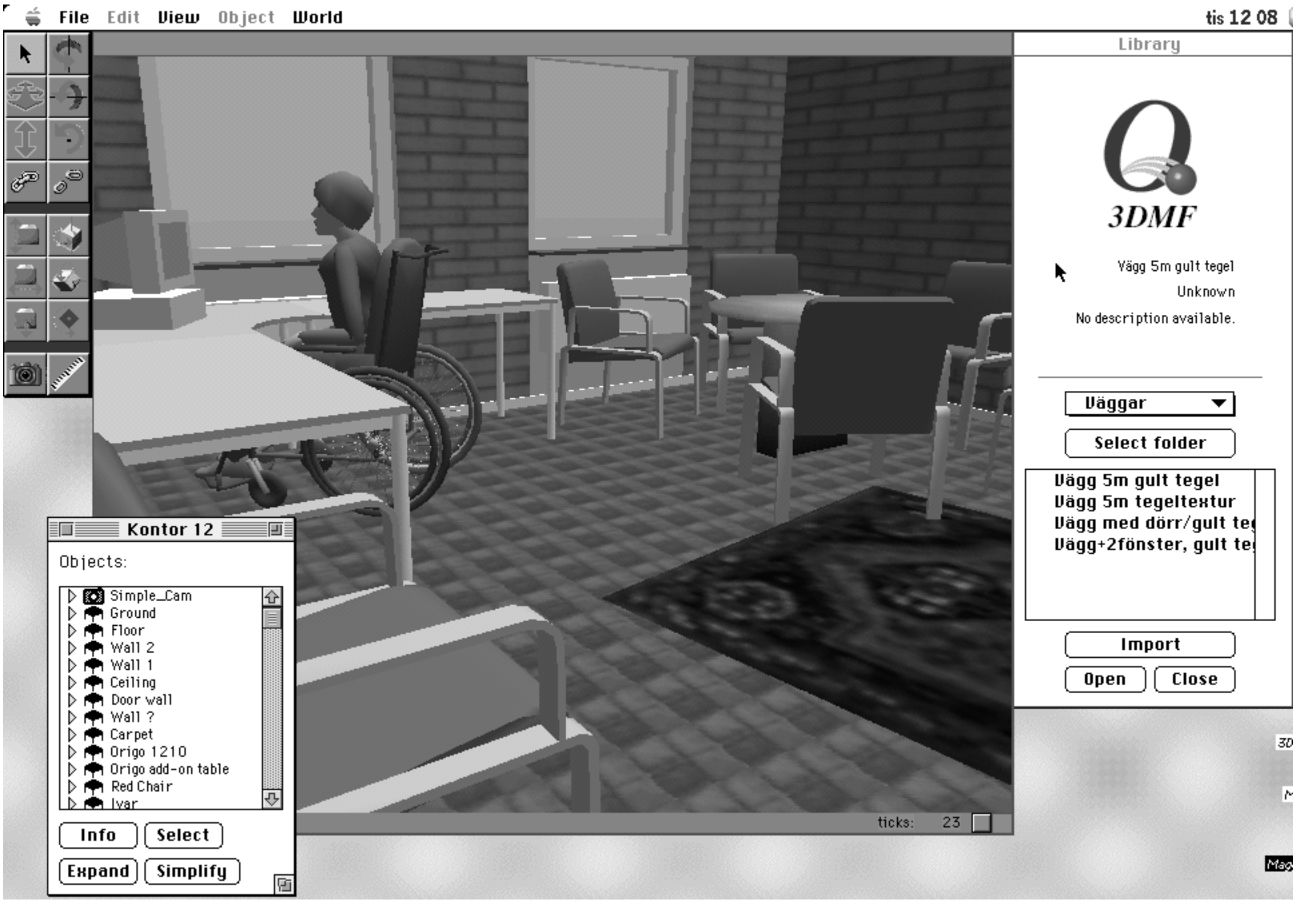


Figure 4. *A screen-dump of the Magrathea interface.*

7. MANNEQUINS

Mannequins (models of humans) are a common aid in ergonomic evaluation which provide designers and architects, for example, with an opportunity to assess factors such as postural comfort, accessibility, clearance, reach and vision early in the design process. Mannequins can also form the basis for automated motion and biomechanical calculation or simulation. Unfortunately, mannequins have traditionally been available only in expensive, expert-oriented CAD-systems, in incompatible formats, making it impossible to compare results and exchange models between different systems.

The inclusion of mannequins, therefore, in the environment design and visualisation toolkit, is vital to improve realism and help in the assessment of suitability for the disabled person. An open mannequin-design should support:

- simplicity in creating, modifying and interacting with mannequins;
- comparability of results between different applications;
- portability of mannequins between different file-formats and applications; and
- flexibility for both developers and end-users in modifying the mannequins' characteristics, complexity and behaviour.

In order to build both static and dynamic mannequins, parameters of interest are:

- stature and body mass;
- anthropometric body dimensions and their proportions to stature;
- dimensions of body segments;
- joint offsets/link lengths;
- body segments' mass and centre of mass locations;
- range of joint motion; and
- moment of inertia.

Unfortunately, there does not exist a unified or complete set of data of a sufficiently large population upon which to base such a mannequin modeller. This means that assumptions, estimations and extrapolations need to be made, all of which compromise the accuracy.

Proc. 1st Euro. Conf. Disability, Virtual Reality & Assoc. Tech., Maidenhead, UK, 1996
©1996 ECDVRAT and University of Reading, UK ; ISBN 0 7049 1140 X

In this project, a mannequin tool has been designed (Eriksson, 1994) that allows static mannequins to be constructed, based on anthropometric data from a North-European population (Jürgens *et al.*, 1990). Plans have also been made to allow mannequins to be constructed in an object-oriented way using detail levels ranging from overall body dimensions to individual segment sizes. One begins, for example, by choosing the sex, height and weight of the mannequin. This then gives a mannequin which can be progressively, and in ever-finer detail, modified to suit the individual to which the environment is being tested.

8. WHAT DOES VIRTUAL REALITY HAVE TO OFFER?

Virtual Reality is fast becoming another tool that can be added to the PC programmer's arsenal. In the area of visualisation, it offers the ability to actually climb into environments, listen to them, perhaps even touch objects, thus providing a better sense of space than other options (Kalawsky, 1993). However, in the past, VR has been exclusive to those with large powerful computers, and bank accounts to match. Nowadays, with the increase in personal computer power and diminishing costs, VR has finally reached the point where ordinary PCs can be used (such as those currently used for word processing and other tasks) and the software can be distributed to and used by ordinary people, not just computer gurus.

But what are the benefits of VR in the area of environment design and visualisation? Obviously, VR offers the opportunity to use full immersion hardware. From the case studies, it was seen that this would help in evaluating a potential environment, although desktop VR would be more suitable in the development phase when several people must work together.

Realism is another crucial feature. With collision detection, patterns, textures, object behaviours and immediate feedback from movements, an environment can be made to seem more realistic. Furthermore, collision detection and behaviours can be utilised in environment evaluation - testing for suitability, tightness of fit for a wheelchair, reach capabilities, mannequin movement and so forth. When it comes to displaying the results of a design, a VR system can offer the ability for automatic walkthroughs, animation and still picture display. These are all important, particularly if some of the design members are remotely located and do not have the correct software or hardware - design alternatives can be pre-packaged in a stand-alone application, sent on video, or photographed and printed out. For those design members that are computerised, an environment can be constructed on a central computer, made available via network, (using perhaps VRML), then viewed, modified and commented on without requiring all the members to be in one place at one time. Even if there is no network, a model can be sent on a disk.

Nevertheless, due to the decision to use only PC based systems, there will be limitations placed on how realistic the environment can be designed to look and still operate at an acceptable speed. This, however, is considered to be only a temporary problem due to the current trend in increasing personal computer speed.

9. A PLANNING TOOL USING VIRTUAL REALITY

In order to explore the use of VR in the planning tool, a PC based system has been purchased (Superscape™) which allows the creation of 3D worlds, selection of objects from an object library, behaviours, interface design and many other features of interest to this project that are not currently planned for the Quickdraw 3D-based planning tool.

At the time of writing this paper, work on the VR part of the planning tool had not yet begun, nevertheless, a number of design directives have been formulated. The VR tool must:

- be compatible with the object library and mannequins;
- allow construction of environments, decorating, placement of furniture etc;
- allow free visualisation from any direction, including walkthroughs;
- allow users to use special VR input and output devices (6 dof mice, joystick, datagloves, HMD, 3D shutterglasses etc);
- allow for pre-recorded walkthroughs; and
- allow models to be published via VRML over the internet.

Furthermore, in designing the interface, it is important to ensure user-friendliness and that the features desired by the users are in fact included (Preece *et al*, 1994). As much as can be gleaned from the case studies will be incorporated into the design, and once the prototype is sufficiently ready, user evaluation will take place.

Development will also continue on the Quickdraw 3D-based planning tool, the aim being to learn what might be possible using the VR system as a prototyping tool, with desirable features ported to the in-house programme if the users deem them suitable. The VR system will also provide possibilities for full immersion and behaviours. One can, therefore, imagine both systems developing side-by-side, twisting together to eventually form a single planning and visualisation tool.

10. CONCLUSION

This work is still in its prototype phase, nevertheless, early indications are good. The case studies have allowed a plan for the future to be formulated based on real user's needs and ideas. A cross-platform 3D object library is planned with arrangements being made to set up a development and support company. In addition, an in-house developed design and visualisation tool has been built which allows placement of objects from the library and 3D interactive visualisation. The work on realistic mannequins has produced exciting results, allowing ergonomic considerations to be taken into account in the evaluation of an environment.

Finally, the plans for the use of a Virtual Reality based software development package are made and a prototype is being constructed.

Once the prototypes of these four tools are complete, the next wave of evaluations can commence involving more real users and providing a new spring-board from which to leap.

Acknowledgements: We would like to thank the Swedish Work Environment Fund and The Swedish Handicap Institute for their support during the earlier stages of this project, as well as Gerd Johansson for supervision.

11. REFERENCES

Apple Computer Inc. (1995), *3D Graphics Programming with QuickDraw 3D*, Addison Wesley, New York.

J Eriksson, M Lövstrand, M Nilsson, R Olsson, M Vågbäck (1996), Magrathea, A Prototype for Interactive Planning of Interior Adaptations, *To be submitted.*

J Eriksson, G I Johansson (1996), Adaptation of workplaces and homes for disabled people using Computer Aided Design, *International Journal of Industrial Ergonomics*, **17**, *2*, pp 153-162.

J Eriksson, G I Johansson, K R Akselsson (1995), A Planning Tool Prototype for Environment Adaptation, *IEEE Trans. on Rehab. Engineering*, **3**, *3*, pp 283-287.

J Eriksson (1994), The design of a computer based human model and biomechanical algorithm, *Computer Aided Planning of Environment Adaptations for Physically Disabled*, Licentiate Thesis, Lund, Sweden.

H W Jürgens, I A Aune, U Pieper (1990), International data on anthropometry, *Occupational and Health series*, **65**, International Labour Office, Geneva.

R S Kalawsky (1993), *The Science of Virtual Reality and Virtual Environments*, Addison-Wesley, Cambridge.

J Preece, Y Rogers, H Sharp, D Benyon, S Holland, T Carey (1994), *Human-Computer Interaction*, Addison-Wesley, Wokingham, England.

E Wagner (1994), *System Interface Design, a broader perspective,* Studentlitteratur, Lund, Sweden.

AN EMERGENT METHODOLOGY FOR THE DESIGN, DEVELOPMENT AND IMPLEMENTATION OF VIRTUAL LEARNING ENVIRONMENTS.

D.J. Brown[1] and D.S. Stewart[2].

[1] VIRART, Department of Manufacturing Engineering and Operations Management, University Park, Nottingham, NG7 2RD.

[2] The Shepherd School, Harvey Road, Off Beechdale Lane, Aspley, Nottingham.

[1] epzdjb@epn1.maneng.nott.ac.uk

ABSTRACT

As any research field graduates from infancy to youth (or it may be argued survives in this case) there is a need to structure and order the methods you have used for getting answers to questions. This should allow us rationalise these methods and select the ones that have given good and practical answers. Successful methods can then be embedded within an overall methodology for the design, development and implementation of the Learning In Virtual Environments (LIVE) program. This program consists of virtual environments (VE) to teach basic life, communicational and personal and social skills, developed in conjunction with the staff and pupils at the Shepherd School.

This overall methodology will be presented in a series of five stages that exert a contingent influence on the final nature of any virtual learning environment. Stage one seeks to embed any development in this field in contemporary educational theory. Stage two looks at the role of parents, teachers, pupils and caregivers in the development of these VE. Stage three considers the role of testing in refining any developments we make, whilst stage four considers the ethical implementation of these learning aids in the classroom. Finally, stage five looks at the development of guidelines for the optimal design of virtual learning environments for the various types of learning.

Keywords: virtual environments, methodology, life skills, communication.

INTRODUCTION

The LIVE project was established five years ago, in partnership with the Shepherd school, to investigate whether VE had a role to play in special education. Where as the Shepherd school used drama, three dimensional art, solar-visualisation rooms and experiential techniques to stimulate and educate their students, virtual reality systems offered a platform for independent movement and autonomous action in a rich and varied set of electronically generated environments, all accessible in a classroom.

The teaching curriculum at the Shepherd school is based around the development of very practical life skills that will afford students some degree of independent living later on in life. Indeed one of the first applications suggested was the production of VE to teach the meaning of Makaton Symbols. These are a series of two dimensional icons, and used in conjunction with British Sign Language act as a communicational system by students with severe learning and language disabilities.

Why a VR system approach in this case? Several reasons were put forward. First, some students are so disabled that the may only ever learn the meaning of ten symbols in their entire school life. Even if it were proven that the VE method didn't encourage the students to learn the meaning of more symbols or at a faster rate, they still provided a stimulating and alternative teaching method, especially important for both teachers and pupils in the light of the former sentence. Second, a classroom setting can be extremely limited in both sufficient objects to represent the meaning of each Makaton symbol, and then with a sufficient range of each object to ensure that students do not over specify what each symbol can represent. Using VE, a whole world of experience can be brought into the classroom, with a range of examples for each symbol, a range of settings and modes of interactions that may not be possible otherwise. Other traditional teaching methods were also deemed flawed, for example a picture card may be used to teach the meaning of the symbol for car, and the use of a two dimensional abstract representation to teach the meaning of another two dimensional abstract icon was questioned.

As we have found subsequently there is usually more than one reason to develop a virtual learning environment. It is hardly ever the only solution however, and should be used in conjunction with other methods, adding to a battery of educative tools to stimulate students with severe learning difficulties. And virtual experience should never be used as an excuse for not providing real world experience, but more rather to fill in experience gaps, when it is not possible for students to gain regular first hand experience due to physical disability, logistical constraints and well meaning but over protecting parents and teachers.

Since this first application was developed we have developed a range of VE to teach basic life experiences (City, House, Supermarket, Skiing), interactive sound environments for students within the Autistic Continuum, reaction worlds for profound and multiply disabled children who can only operate switches and finally environments to teach health education to students with moderate learning difficulties.

It has always be our aim to test and evaluate the learning environments we have produced, but initially this testing phase was very much on an ad-hoc basis, including an initial evaluation of the usability of the available input devices and a study to determine whether students could recognise and articulate the meaning of objects they could do so in the real world. This then gave way to a much more systematic approach, borne out of a developing partnership with the Department of Learning Disabilities, based within the Queens Medical Centre also at the University of Nottingham. Since then we have sought to determine the transferability of skills learnt in a VE to the real world, whether the use of VE can encourage self directed activity and the testing and design of appropriate navigation and input devices for students with a huge range of physical and cognitive abilities.

But it is the five factors that drive the creation of any new virtual learning environment (and indeed it may be argued the program itself), that have emerged into a methodology to describe a potential role of VE in special education that will be the main subject of this paper.

STAGE1: Embedding the conceptualisation and design of virtual learning environments in contemporary education theory.

New technologies are often introduced first into mainstream education, only later to be adapted for special education use. In this instance however, the role of VE in special education appears at least as pertinent in a special needs setting as it does elsewhere. There exists a strong mapping between the attributes of VE and the goals of what is considered good contemporary special educational practices.

First, their use may encourage **self directed activity.** Many students with learning difficulties experience so little control over things other children take for granted that they may assume they are going to play a passive role all the time (Simms, 1994). VE can stimulate a child's curiosity and to this end environments have been built that allow a student to enter an underground station and take a tube train to the destination of their choice or to find out what happens when an empty kettle is boiled. Bruner (1968), Vygotsky (1978) and Piaget (see wood 1988) have emphasised the importance of self directed activity in their theories. For students who have limited opportunity to do so in real life, VE can offer a rich and varied set of opportunities to initiate self directed activity in a safe arena.

Second, given that their use might promote self directed activity in a group of students who are said to be deficient in this area, we must by good design build VE that are **motivational** and inspiring to use. Stuart and Thomas (1993) believe that whereas the age of television has bred passive disengaged students with short attention spans, the use of VE may be able to captivate students and foster their active involvement in their own education. Our 'develop-test-refine' methodology will employ empirical research to establish the optimal design of VE for the various types of learning, one of the criterion of which being whether students are motivated to use them.

Third, the **role of play** is given high importance in developmental theories of education. Students with disabilities may often be protected for longer than others (Shakespeare, 1995), often not gaining experiences necessary for further development. Vygotsky (1978) emphasised the importance of play in liberating children from constraints, whilst Bruner (1968) describes how play allows the systematic uncoupling of means and ends. Students can embody themselves in other characters and play 'lets pretend'. Brenda Laurel (1991) points toward the strong identification that players can feel with artificial characters in a computer database as an example of the human capacity for *mimesis*, to which Aristotle attributed the soul changing power of drama. Our team has adapted the virtual city to teach road safety; credit points are given for good action, for instance using a pedestrian crossing whilst credits are lost for crossing the road at an unmarked junction. The aim is to gain credits to stay alive in the game.

Fourth, many computer learning systems rely on abstracted symbol systems such as English or Mathematics. VE have their own **natural semantics** (Bricken, 1991); the qualities of objects can be discovered by direct interaction with them. Children with learning difficulties are commonly termed 'concrete thinkers', learning best through direct through practical experience. Our virtual house looks and operates very much like a real house, with a kitchen in which you can learn to make a simple snack. This form of learning, independent of abstracted symbol systems, is ideally suited to the abilities of students with severe learning disabilities.

fifth, desktop VR platforms offer a **shared public experience**. Both student and facilitator can share and discuss the environment. Bruner (1968) has drawn attention to the social context out of which skills develop. The importance of the role of instruction has been developed by Vygotsky (1978), in his concept of the 'zone of proximal development' defined as the distance between a child's actual developmental level as determined by independent problem solving and the higher level of 'potential development as determined through problem solving under adult guidance or in collaboration with more capable peers'. This is particularly pertinent to our environments used to encourage language development, where the student is encouraged to respond to a virtual manakin and articulate the meaning of symbols encountered to a teacher, carer or more capable peer.

Sixth, VE can act as a **great equaliser of physical abilities.** Provided the student can operate simple input and navigation devices (joysticks, switches, touchscreens), they can move through and interact with environments they may be restricted from doing so in real life, and gain experience and learn skills that their physical disabilities may prevent them from doing so in real life. This of course depends on good interface design an issue already extensively investigated by our team (Hall, 1993 and Crozier, 1996). People with disabilities already face a huge range of barriers to social and educational interaction in real life and we do not want to compound these by poor interface design and create a new set of barriers to interaction with VE (Vanderheiden *et al*, 1992).

And **seventh**, VE can provide a **safe space** in which to practice skills that are dangerous and risky to do so in real life. Problems can be encountered and consequences demonstrated without exposing the student to any real danger. A pilot project is now underway in Nottingham to develop virtual learning scenarios to teach health education to students with moderate learning difficulties. In this way a student might be approached by a stranger in the virtual city, and we can teach appropriate behaviour in response to such a situation in an non-abstract and safe environment.

STAGE 2: The role of teachers, parents, caregivers and students in the development of virtual learning environments.

It is a central policy of VIRART that all virtual learning environments should follow a Develop-Test-Refine process. It is the pupils, teachers, parents and caregivers who should play a major role in the developmental stage of any virtual learning project. They suggest the original application, define what it should look like, its functionality, multimedia facilities and modes of access.

By tapping into this knowledge resource we can ensure that any application developed will be relevant and useable by the group of students to which it is aimed. This is especially important in the field of disability where resources are already stretched and any new innovation should bring maximum benefit, especially when one considers that to provide such a resource in a special education classroom may drain resources from other projects.

The role of teachers and parents in this developmental stage also extends to the platforms on which the virtual learning experience will be displayed. When we first approached the Shepherd School their only reservation about the project was centred around media speculation of the hypothesised health and safety risks associated with the use of head mounted displays. For this reason all of our developments take place on desktop equipment. These platforms also have the advantage of being affordable by special schools, an important issue to consider in the developmental stage of any virtual learning project. It would be unwise to place any student with a learning disability in a headmounted display until the hypothesised physiological, psychological and visual side effects have been rigorously established. The study VIRART has been commissioned to undertake by the Health and Safety Executive (HSE) has so far established:

* Some participants, in some VR systems, with some VE, do report various side effects.

* The most common reported side effects are dizziness, nausea, 'disorientation', and a group of symptoms that might be categorised as visual fatigue.

* For most participants, any side effects wear off quite quickly after VR participation finishes.

Even when the results are extensive and conclusive the teachers at the Shepherd School have pointed out that these issues are likely to be much more complex for students with learning difficulties.

Teachers, parents and caregivers will also have a role to play in the formulation and writing of any manuals or tutoring systems describing how to set up and use the LIVE system to maximum effect. Again this is important if the virtual learning environments are to be used in the home as well as the classroom or rehabilitation centre.

STAGE 3: The role of testing in the refinement of virtual learning environments.

Along side developmental issues, testing plays a major role in shaping and defining virtual learning environments. A brief description will now follow of the major tests we have carried out in the LIVE program and the subsequent effects they have had.

Test1: Can students recognise and articulate the meaning of Makaton symbols they knew in real life?
This initial test was carried out using a virtual warehouse containing a number of different objects, a series of randomly flashed Makaton symbols, representing the meaning of each of these objects and eight randomly selected students, some of which had previous computer experience. In summary, our findings were:

* Two of the students experienced no trouble in examining each symbol in turn and then quickly

and accurately identifying the object that each randomly generated Makaton symbol represented.

* Three of the students appeared to recognise objects and handsigns but failed to match the correct object with the randomly flashed Makaton symbol.

* The other three students correctly matched some objects with the correct Makaton symbol.

From this initial test the team was encouraged to continue with its program to teach Makaton symbols, but to place the objects in future in their correct contextual setting and use other clues (such as scanned images and sound) to help students to accurately identify virtual objects. The Derbyshire Assessment Program was also adopted as a method of monitoring the ability to students to recognise and articulate the meaning of grouped sets of words.

Test2: Can the use of virtual learning environments encourage self directed activity in students with learning difficulties?
Self directed activity as we have discussed is one of the areas students with learning difficulties are said to be deficient in. The team decided to investigate whether students were using the virtual learning environments in this way or in a more conventionally didactic manner.

For this test we used the Makaton environments and recorded the students autonomous moves, such as pointing and the use of the joystick and mouse against instructions issued by the teacher. What was found was that where as in the first instance the teacher was mainly directing the students as to what to do and where to go next, these directives began to fall away sharply and be replaced by the students own autonomous moves.

This finding led us to believe that as with any learning system there will be an initial learning curve, but that the use of this system can encourage self directed activity in students with severe learning difficulties, allowing them to direct their own course of events, something that they might not expect to do so much in the real life.

Test3: Can students generalise basic skills they have learnt in virtual environments to the real world?
This study investigated whether basic shopping skills learnt in the virtual supermarket could be transferred to the real world. A control and experimental group, both containing around ten students of similar real world shopping experience, was used. A baseline assessment tested their ability at a local supermarket to select items using a picture based shopping list. There was no significant difference in the two groups at this stage, both in the number of correct items picked up first time and the total time elapsed for the shopping experience.

There then followed a period of virtual training. The experimental group each experienced between three and ten sessions using the virtual supermarket to select a series of goods from the virtual aisles using a picture based shopping list. The control group were also allowed to use a series of virtual learning environments, however they were not allowed to use the virtual supermarket.

All of the students then revisited the real supermarket. Again challenged to select a variety of goods, analysis of the subsequent data revealed that not only did the experimental group have a significantly faster overall shopping time but they were also more accurate, selecting the correct item off the shelf first time on a more frequent basis than the control group.

This is a very encouraging result, proving that it is possible to use VE to teach basic life skills to students with severe learning disabilities. It is not planned that the virtual supermarket be used as a replacement for real world experience, but more rather as a support, used on a frequent or every day basis to build skill in between the times when students can visit a real supermarket in school time or supported by their parents.

Test4: Appropriate input devices for students with motor skills disabilities.
People with disabilities already experience a whole range of access issues in their lives and we do not want to create an extra set of barriers when creating virtual learning environments. For this reason we needed to review the currently available navigation devices (spaceball, spacemouse, joystick, keyboard) and input devices (mouse, touchscreen, keyboard, switch) and assess their usability within VE by students with a wide range of motor skills abilities.

Our first analysis (Hall, 1993), a population stereotype study, looked at the usability of the currently available navigation devices and revealed that a joystick which allowed just two simultaneous degrees of freedom created least frustration in students when attempting to navigate from one virtual position to another.

The latest study (Crozier, 1996) investigated the usability of input devices and makes recommendations for the operational improvement of touchscreens within VE and software to support the use of more specialist devices, such as switches and soundbeam, to allow access to VE by students with the most severe motor skills disabilities.

Test 5: Can students with severe learning disabilities transfer spatial skills they have learnt in a virtual environment to the real world?
This current and on going experiment has used a virtual model of the Shepherd School. There are tremendous training and rehabilitation implications if it is possible to use VE to teach spatial and navigational skills to this group of students. What is planned is a virtual treasure hunt game where students receive clues as to where the treasure is buried in the virtual school. The students will then be challenged to find the treasure in real life, thus indicating whether routes can be learnt in a VE and related to an existing building.

STAGE 4: The implementation of the virtual learning environments in a classroom setting.

The ethics involved in the placement of virtual learning environments are important drives in the way we shape and form them. Issues of how they are placed in the curriculum, whether they drain resources from other important projects, whether they unnecessarily raise expectations or simply are used as a convenient excuse for real world experience should all be considered at the out set of any project. There are many ethical issues which programmers must consider when designing educational VE for students with severe learning disabilities. Three central issues are:

* the need to include appropriate content.

* the need to provide context and setting.
* their own power in constructing users' virtual experiences.

Each of these three issues will now be discussed in more detail.

Debates about the appropriate content of VE gain additional force when they are intended for educational use by students with severe learning disabilities, who are typically considered more vulnerable and open to suggestion than others. Desktop VE have an obvious resemblance to television, but because they promote active engagement rather than passive viewing (Stuart & Thomas, 1991) their potential influence is greater. One study by Calvert & Tan (1994) looked at the impact upon levels of arousal and aggressive thoughts in 36 non-learning disabled young adults of either watching or playing a violent desktop VE game, and found that those playing the game showed higher levels of arousal and more aggressive thoughts than those just watching. The content of educational VE must be carefully chosen to make sure not only that they are appropriate to the task, but also to ensure that no "hidden agendas" are inadvertently introduced.

Important though they are, debates about content are not unique to VE and do not go to the root of most people's concern about their educational use. What is more specific to VE is their potential to provide experiences devoid of their usual social context. Actions in VE do not have the real social costs associated with their real world analogues and Whalley (1993) argues that as a result continual use of VE might foster infantile thinking and feelings of omnipotence, especially in those whose ability to negotiate their place within the socially shared physical world is already impaired.

Programmers can address this concern by building explicit moral direction into their educational VE. This can be done in a number of ways. First, programmers can attempt to model in VE the negative consequences of anti-social actions in the real world. Second, at crucial points in the program they can add explicit references to the real world beyond the computer screen: by embedding video sequences, or explicit instructions to the user to consider the issues raised and discuss them with someone else. Third, programmers can consider limiting access to some parts of the program with parental-locking mechanisms, ensuring that parents or school staff must authorise and monitor their use. This option should be reserved for exceptional cases - for example, to prevent students from using a photo-realistic sex education VE merely for titillation.

These recommendations to programmers emphasise the responsibility of teachers and parents in providing both context and setting for students' exposure to media images, to assist them in arriving at appropriate interpretations. It is this that is perhaps the most fundamental ethical requirement for the educational use of VE.

Finally, we must consider the massive power of programmers to influence the virtual experiences of users. The ethical responsibility of VE programmers can be compared to that of town planners or architects in creating spaces which are more conducive to some kinds of activities than others. This issue is especially relevant when the intended users are students with severe learning disabilities, who are unlikely to have much involvement in the design process. VE for students with learning disabilities may also have therapeutic uses, in which case the ethical responsibilities of the programmer are analogous to those of a doctor or therapist. In all cases, programmers need to be aware that their creations are not value-free - as Boal (1995) observes, "technology is congealed ideology". Development of VE in close consultation with the intended end users or their representatives, parents or teachers -even devolving to them important decisions on content and structure - is the best way to ensure that programmers' power is deployed in ethically acceptable ways.

STAGE 5: The optimal design of virtual learning environments.

Empirical testing has aimed to establish the educational veracity of virtual learning environments and this has led us to believe that we could employ a similar process to help determine the optimal design of VE for the different 'types' of learning. This could help us to produce guidelines on the optimal construction of VE using current technology by investigating issues concerning complexity of virtual learning environments and the ensuing speeds of access, the use of scanned images, video sequences, embedded sequences and contextural three dimensional sound.

VE are now being planned and constructed composed of varying degrees of the above factors, to investigate how their relative presence in an environment might influence the transference of basic life and spatial skills to the real world. Also of interest is the dichotomy between the use of embedded sequences and scaffolding techniques in VE that, on the one hand offer total freedom and autonomy, whilst on the other seek to have control structures embedded within them so that important lessons will be learnt. For example, within the virtual kitchen it will only be possible to boil the kettle after filling it with cold water, or experiencing a limited degree of freedom whilst navigating to ensure that the student won't become lost or disorientated and thus experience a degree of frustration.

The use of three dimensional sound may be of help in virtual navigation exercises but perhaps a distraction when attempting to focus in on and learn selection and payment skills in a supermarket. If we try to replicate the buzzing and random complexity of the real world in the virtual city, it may look very realistic, but a student's ability to move through it and interact with it at a speed where they could generalise the lessons learnt there might be considerably hindered.

The development of guidelines for the construction of virtual learning environments will be a significant influence on the way be plan, develop and use them in special schools.

CONCLUSION

VIRART, the Shepherd School and the Department of Learning Disabilities at the University of Nottingham have described an emergent methodology to define a possible role for virtual learning environments in special education. As the LIVE project matures it has become clear that the VE we produce for use by students with a range of learning and motor skills difficulties are influenced by a set of five factors that act together to drive and shape not only individual learning environments but also the program as a whole.

A battery of tests has also been gathered and exploited to enable us to evaluate the methods which we use to define the LIVE program. It is this structured methodology that may help us to not only turn ideas into working programs in special schools today in an efficient manner but also share our findings with other groups working on defining the role of VE in special education and provide a framework for extending the use of VE into other areas, notably mainstream education and training.

REFERENCES

Boal, I. (1995) A flow of monsters: Luddism and virtual technologies. In "Resisting the Virtual Life: the culture and politics of information" Brook, J. & Boal, I. (eds) San Francisco, City Lights

Bricken M. (1991) Virtual reality learning environments: potential and challenges. Computer Graphics **25** (3).

Bruner, J.S. (1968) Processes of cognitive growth: infancy USA, Clark University Press.

Calvert, S. & Tan, S.L. (1994) Impact of virtual reality on young adults physiological arousal and aggressive thoughts: interaction versus observation. Journal of Applied Developmental Psychology 15,1 125-139.

Crozier, J. (1996) Evaluation of navigation and input devices for use within virtual learning environments by students with severe learning difficulties. Undergraduate Thesis, Department of Manufacturing Engineering, University of Nottingham.

Hall, J. (1993) Explorations of population expectations and stereotypes with relevance to design. Undergraduate Thesis, Department of Manufacturing Engineering, University of Nottingham.

Laurel, B. (1991) Computers as theater. Menlo Park, California: Addison Wesley.

Piaget, J.S. (1950) The psychology of intelligence. London, Routledge and Kegan Paul.

Shakespeare, R. (1975) The psychology of Handicap. Methuen, London.

Simms, D. (1994) Multimedia camp empowers disabled kids. IEEE Computer Graphics and Applications, January, 13-14.

Stuart, R. & Thomas, J.C. (1991) The implications of education in cyberspace. Multimedia Review, 2(2) 17-27

Vanderheiden, G.C., Mendenhall, J., and Andersen, T., (1992) Access issues related to VR for people with disabilities. Trace Reprint Series.

Vygotsky, L.S. (1978) Mind in society: the development of higher mental processes. Cambridge, Mass, Harvard University Press

Whalley, L.J. (1993) Ethical issues in the application of virtual reality to the treatment of mental disorders. In "Virtual Reality Systems" Earnshaw, R.A., Gigante, M.A. & Jones, H. (eds) London, Academic Press.

Virtual partners in cyberspace - a proposal

Magnus Magnusson

University of Karlstad
Handikapp & Språk/ Utbildningsvetenskap
S-651 88 KARLSTAD, SWEDEN

Magnus_Magnusson@tp44.bad.se

ABSTRACT

People with Aphasia have great use for computer support in their rehabilitation, aided by different types of software. Most of the software consists of sets of exercises with a clear language content. However, in this paper there is a presentation of ideas around a more general communication situation where a computer-simulated environment could offer people with Aphasia a broader experience of interaction and communication. Examples are given from work with videotelephony and different types of simulation software.

Keywords: simulation, videotelephony, therapy, aphasia, telematics

1. INTRODUCTION

Simulated environments have appeared in computer programs since the first commercial computer systems were introduced on the market. In the early days, text-based adventure-games like the legendary ZORK gave users the illusion of being in a fantasy-world inside the computer, where experiences were manifold although limited to text. Another famous non-game application was the misunderstood advisory system Eliza which gave the user the illusion of having a dialogue with the machine.

Today, the computer industry offers us simulated environments, where it is possible to control a computer program with movements of the body, speech and even the electrical currents in your muscles, myoelectrical impulses. It is also possible to experience the contents of the program through all your sensory organs. In other words, you interact with the computer and the program in much the same way as you interact with your physical environment or even with other people. Oral, visual and even tactile experiences can be simulated through computer applications.

However, we have a long way to go before the market can offer programs of a more total-immersive type which might be expected from my over-simplisistic lines above. It is possible to use datagloves, headsets and even data-suits and interact with the computer which has been shown by pioneers like Jaron Lanier and others. It is also possible to interact as an artist with your own or anyone else's piece of art - music or painting or whatever - through the interface of a camera and a microphone which has been shown by Myron Krueger and others (Krueger, 1991; Newby, 1993). However, interacting through taste, smell and the proprioceptive systems still are lacking outside the highly specialized gas sensor systems of the laboratories (Sundgren, 1992; Teil, 1992) . Tactile interaction is also very limited. However, tactile screens are in common use.

It is indeed interesting to see that the world(s) on the other side of the screen are so tempting to artists and people in general working in the cultural field. Cultural innovators like Brenda Laurel, Jaron Lanier, David Rokeby, Rafael Lozano-Hemmer, John Vincent and others have enrichened the concept of art by opening similar doors of not only perception but also of active interaction with art itself through experiments with VR. The goal seems to be the enlargening of what is humanly possible to perceive and do. The end result becomes a simulation, either of reality itself or of the artist's perception of reality or even of the alternative realities of the spectator(s).

The general tendency in the computer industry today seems to be the fact that computers become more and more invisible or transparent. More and more functions are included into the computers or the (system) program. The latest piece of information about this tendency which reached my ears was the piece of news telling me that voice control will be a standard part of the new version of the WARP operating system.

2. VR AND SPECIAL NEEDS - GENERAL REMARKS

In a similar way, developers and researchers have found the field of VR rewarding and tempting when trying to develop new ways for people with special needs to interact in a better way with other people and the world in general. This conference is the fifth one in the world which concentrates exclusively on VR and disability and if we would count the total number of presentations and papers dealing with VR and special needs or disability we would probably find at least five hundred papers or presentations from the last five years. This is quite a number of papers and of ideas and it is a good measure of the amount of human creativity which has been spent in this area. I also have to add that most of these presentations have been made at some of the conferences created by the real pioneer in this field, Dr Harry Murphy Director and founder of the Centre on Disabilities, CSUN, California.

Very briefly I have looked over this impressive list of presentations and most of them seem to concentrate on the needs of people with hearing problems, motor dysfunctions, visual disabilities and in a few cases people with cognitive problems and I just want to give a few examples. In the last case, the problem of learning and understanding has been a main focus (Brown & al, 1995; Standen & al, 1995). In the field of visual disability some very exciting ideas have been presented for instance regarding the ability to construct spatial correlates, just using auditive stimuli (Cohen, 1995). In the case of motor disabilities, the main work seems to concentrate on developing more or less body-independent control-mechanisms or interfaces(Henry & al, 1992; Knapp & al, 1992; Knapp, 1995; Riedel, 1993). A good review of the VR-field, mainly from a medical perspective has been compiled by Dr Walter Greenleaf, another pioneer (Greenleaf, 1995). An interesting example of VR as a diagnostical and analytical tool can be found in the concept of holography (Szymanski, 1995).

3.VR AND COMMUNICATION DISABILITIES

However, very few applications have concentrated on the needs of speech or language impaired people. Apart from my own tentative work in the field of visual remote communication and simulation software for people with Aphasia (Magnusson & al, 1995) I think that most of the work in the language disability field and VR comes from UK, notably Dundee (Brophy-Arnott & al, 1992; Waller, 1993; Cairns & al, 1992), Edinburgh and London (Roy & al, 1993). Reviews of the general situation regarding computer application for people with Aphasia have been published in the journal Aphasiology several times (Loverso, 1992 and several others). However, it is important to remember when we talk about speech and language impairment that the field of AAC includes the needs of severely motor disabled persons with communication disabilities and there you will find many ideas regarding VR-based communication. The examples are many and well known.

An interesting detail is the fact that although there has existed a semi-annual American journal for about ten years called CUSH/Computer Users in Speech and Hearing, not too many of the otherwise interesting articles in that journal concentrate on applications for people with Aphasia.

4. APHASIA AND COMPUTER APPLICATIONS

This presentation, however, will concentrate on the needs of the language disabled people with Aphasia. The concept of Aphasia is very complex and, depending on your theoretical standpoint regarding the definition of the syndrome, there are several ways of defining Aphasia. A basic definition would say that Aphasia is a problem where the Aphasic person might have trouble in producing or understanding/ decoding language. Treatment for Aphasia often consists of training strategies to retrieve or recreate language structures of different types. This means that training or therapy is very important to the Aphasic person. If we want to create good training instruments for Aphasics we have to consider many different aspects of communication.

Since a language deficity like Aphasia means that you probably have difficulties in your ability to handle any type of symbol you will have difficulties in handling any computer interface since all computer interfaces are heavily abstract and symbol oriented, either using letters or icons. Another difficulty in the general handling of computers is that a common program consists of series of commands which could be difficult to remember correctly. In the old DOS-days it was difficult for someone with Aphasia (or Dyslexia) to spell the complicated commands correctly. Today, the icon-based interface might be spatially confusing since too many open windows or desktops tend to confuse a user so that in the end you really don't know where you are at the moment, since information seems to be hidden in many different graphical codes all over the screen.Few researchers have done any real serious research into the results of having to use complex interfaces and what complexity really consists of. Some work has been done regarding the use of multimedia (Cairns & al, 1992). This field is certainly open to some basic research.

Many computer based programs for Aphasia training have been developed over the years. So far, most of the programs consist of basic exercises where the user of the program encounters a set of questions or trials, mostly based on text and pictures. The training situation, however, is far from the real and basic communication situation and I believe that if we are to create optimal training software for people with language disabilities, then we have to recreate the general communication situation itself. And to do this we have to define that situation and all its different parts so that in the end we will have a simulated reality. In this simulation we will have to consider pragmatics and body-language and also social behaviour and also the biological aspect of communication.

5. VR AND APHASIA

The optimal situation would be the (dyadic) situation which would offer most input and output channels to the participants, probably the basic situation where two persons are close two each other in the same room and within touching distance of each other. When I talk about the basic communication situation I simply mean such a situation which is as much a product of your body as of your personality. A basic morpheme in the language of life - to put it a little bit high-sounding. In other words, if it would be possible to create a computer-based situation of that sort, then it would be possible to give a computer-based experience not only of parts of the communication process but also of the totality of the communication process with speech, eye-contact, movements, touches etc as natural parts of the computer use.

At this point, some listener or reader might make the remark “why bother with simulated reality when we have the real thing?” Unfortunately, as many therapists, relatives and Aphasics themselves know to well, the “real thing” meaning other people to talk to and interact with are not always available so the simulation might become the support, helping you to find your way into the so called real world, at your own comvenience.

So far, the world of Aphasia therapy has seen very little of this advanced type of program. I will give a short history of the development of computer-based Aphasia therapy leading up to more and more reality-like software and then give a few suggestions as to the next steps of development.

The very first computer programs were used in Aphasia therapy and consisted of text-based practices in word-finding, filling in sentences, finding opposites etc. In the mid fifties the mainframe based Plato system was used in a few trials in the USA but the first real programs were developed for personal computers in the middle of the seventies, in several countries including Sweden and UK and others. During the eighties real program packages containing sets of exercises were developed and several programs became commercially available. There also were developed programs not specifically for therapist-supported training but also for self-training and for use as personal communication aids. Still the programs were concentrated on linguistic training and the pragmatics of communication was not so evident in the software.

In the early nineties, software development had reached a level where it was possible to offer advanced multimedia solutions and program developers in the Aphasia field started to build scenarios for the programs where the communication situation was more interesting than just the isolated language test or practice itself. In an average program package it was possible to find exercises for memory training, word-order, syntax, semantics, basic mathematics, spatial orientation, phonological training etc. This type of program is available in most European countries as well as in the US, Australia, Japan etc.

There has also been a controversy of dialectic proportions regarding the use of computers for people with Aphasia. Many people have agreed that the computer has all the potentialities to give people with Aphasia many new possibilities and we all have seen the flexibility of training software. On the other hand, the language based structure of most computer interfaces seems to make independent computer handling very difficult for people with Aphasia and since computers are an integrated part of modern society, it seems as if computers would make society much more difficult to cope with for people with Aphasia. Besides, it has been difficult to create any really new computer ideas for Aphasics since the therapeutic software mostly is translated directly from the paper-version. In other words, developing new and creative computer solutions for people with Aphasia has been considered a very difficult task (Ahlsen, 1995; Waller, 1995).

However, the first time anyone officially expressed the new idea of a whole communication situation in the form of a computer program was the time when the Scottish speech therapist Alison Crerar coined the term “microworld” (Crerar, 1991), thereby meaning that the training situation with the computer should be likened to the sensation of experiencing a special environment or “world” where the Aphasic person would be able to experience communication as a form of interaction with the software. Several programs during the nineties have tried to offer this possibility by creating rooms, landscapes, buildings, countries etc in the computer where it is possible to “wander” around and discover different things and where the training content lies in the fact that you have to identify verbally the things you encounter.

I will show a few examples of software of this type from Sweden and the other Nordic countries. This type of program gives the user the feeling that he/she is moving around in an environment where it is possible to use and learn from the things that you encounter in this environment, shortly that it is a question of interacting in a broader sense of the word than in the older types of programs.

This type of program is in reality a simulation and simulations of processes, stories, adventures, journeys etc. They have become very common during the last five years. We only need to think about the family of most popular games - the Sim-family games where Sim-earth might be the most wellknown example.

The simulations for people with Aphasia have so far only covered manipulation of objects and language tools and experiencing a special environment. No software has yet offered the interaction with a simulated person other than in very primitive form. And all the programs are limited to mouse- or keyboard control. What would be the next step would be to offer the Aphasic person the possibility to interact with the program in a more direct form.

6. SUGGESTIONS AND IDEAS

A simple example would be to make programs were it would be possible to use a touch screen. thereby creating the feling that it is possible to manipulate the objects and tools on the screen directly without any intermediate medium or change of perspective. Few trials have been made yet, making use of this technology. In the same way it would be natural to offer control solutions where it is possible to talk to the computer or to make gestures to the computer which would take in the gestures through a camera connected to the computer.

We have a few examples where the computer has helped a person with Aphasia and traumatic brain injury to express the feelings related to the communication difficulties through computer-based art (Addison, 1995). We also have examples where computer software has been created to guide and support someone with Aphasia constantly through situations and environments, once familiar and taken for granted.

To sum it up, we have good programs, good experiences of the use of computers by people with Aphasia, good knowledge of what Aphasia therapy is ablout. However, we have insufficient control tools to make it possible to inetract in a broader sense with the computer and the software. We also do not have a lot of software where it is possible to train the interaction with another person, either a real person or a simulated person.

Therefore I would like to suggest two main ways to spend research and developmental money for people with Aphasia:

1. The development of easier ways to access and use networks like the InterNet or direct contact between two individuals, in other words, meeting real partners electronically from a distance. If you call the InterNet using WWW today and make a search using the word Aphasia you will probably find about 100 hits including the presentation of a few American researches, some bibliographies from scientific journals and the homepage of the heavy metal rock group Aphasia (!). Accessing the InterNet is still a little bit cmbersome and to move around in the network yo have to be good at reading and writing as well as endowed with quite a lot of patience - not to speak about all the knowledge about hardware and software which is necessary.

Another way of using standard technology to access people and sites of iformation over a network would be to use video-telephony, either over InterNet (Cu-See-me) or over ISDN or narrowband telephone network. My own research is concentrating on the use of ISDN-based videotelephony as a tool to give therapy to people with Aphasia over a distance. How transparent is the videotelephone technology for users with Aphasia? Will the participants experience a videotelephone-meeting in the same way as a “real” physical meeting?

Very valuable views on general design problems regarding the accessability of remote information systems have been presented in the work at British Telecom (Smythe, 1993). Within our small community of researchers into the accessability problems of different networks there seems to be consensus on the need to make sites on InterNet as accessible as possible (Serflek & al, 1995). To paraphrase Krueger’s words, a good program has to be like a place where you experience things and discover new aspects of the physical world. When you become accustomed to your computer and the program you have to trust the technology and feel that you sort of belong together (Treviranus, 1995).

2.The development of a totally new type of software where it would be possible to meet a simulated communication partner within the computer. This type of software would have to be controlled through, camera, microphone, touch-screen beside the usual mouse/keyboard platform. The software would also have to contain some sort of voice recognition and pen-writing function. In other words, a platform containing such a type of software will be so power demanding that the personal computer of today will probably be too limited. The large computers from film-making would probably be a more promising platform. And sofar, no concrete development has been made with such an idea for people with language disabilities, at least not to my knowledge.

This type of software would certainly offer a totally new possibility of language training for people with Aphasia as well as for therapists. Apart from being slightly controversial as an idea where we really have to be careful about the definition of “human communication” it is fascinating. Imagining meeting a virtual discussion partner on the screen where the software would adapt to your reactions and you might be able to choose from different personalities to interact with. In Sweden and Denmark we have started discussions to see what such a joint effort would demand and lead to. We would be happy to welcome other interested discussion partners.

7. CONCLUSION

To sum it up, the software with the working name “The Virtual Partner” would have to offer the user the possibilities to

- train a dysfunction
- test an ability
- get information
- get a role model
- get a way to express the inexpressable

My final and basic hypothesis would then be:

> If you create a computer simulation that would function as close as possible to the situation where two people sit in the same room communicating with each other using all output and input channels available, then you have created the optimal software for the Aphasic person.

If you agree - let’s do it!

8. REFERENCES

Ahlsén Elisabeth (1994) Aphasia as a Challenge to AAC In: Brodin, J. & Björck-Åkesson, E. (Eds) *Methodological Issues in Research in Augmentative and Alternative Communication, proceedings from the 3rd ISAAC Research Symposium in AAC,* Kerkrade, Oct 14-15, 1994, ppg 34-41

Addison Rita K. (1995) Virtual Reality: A New Way To Design Neuroperceptual Tools In: Murphy, Harry (Ed): *Proceedings from 3rd conference on Virtual Reality and Persons with Disabilities,* (1995), University of Illinois, Chicago

Brophy-Arnott M.B. Newell Alan F. Arnott J.L. Condie D. (1992) A Survey of the Communication-Impaired Population of Tayside *European J of Dis of Comm* 27(1992)2 ppg 159-174

Brown David Stewart David (1995) The Learning in Virtual Environments (Live) Program In: Murphy, Harry (Ed): *Proceedings from 3rd conference on Virtual Reality and Persons with Disabilities,* (1995), Univ. of Nottingham

Cohen Michael (1995) Audio Windows for Synchronous and Asynchronous Conferencing In: Murphy, Harry (Ed): *Proceedings from 3rd conference on Virtual Reality and Persons with Disabilities,* (1995), Human Interface Lab, Aizu

Crerar, Alison (1992) Adventures in the Microworld *Ability* (1992)5 pp 5-7

Greenleaf Walter (1995) Applications of Virtual Reality Technology to Therapy and Rehabilitation In: Murphy, Harry (Ed): *Proceedings from 3rd conference on Virtual Reality and Persons with Disabilities,* (1995), pp University of Illinois, Chicago

Henry, Tyson R., Yeatts, Andrew K, Hudson, Scott E., Myers Feiner, Brad A. Steven (1992) *A Nose Gesture Interface Device: Extending Virtual Realities,* Presence 1(1992)2, pp 258-261

Knapp, R. Benjamin, Lusted, Hugh S. (1992) Biocontrollers for the Physically Disabled: A Direct Link from Nervous System to Computer CSUN/California State University, In H. Murphy (Ed): *Proceedings of the Seventh Annual Conference: “Technology and Persons with Disabilities”*, pp 291-294

Knapp Benjamin (1995) The Use of BioElectric Signals for Navigating a Three Dimensional Computer Environment In: Murphy, Harry (Ed): *Proceedings from 3rd conference on Virtual Reality and Persons with Disabilities,* (1995), pp San José

Krueger, Myron W. (1991) *Artificial Reality II* Addison/Wesley Publishing Co, New York, 290 pp

Loverso F.L. Prescott T. E. Selinger M. Microcomputer Treatment Applications in Aphasiology *Aphasiology*, 6(1992)2, ppg 155-164

Magnusson Magnus Gunnilstam Olle (1995) Remote therapy for Persons with Aphasia In: Murphy, Harry (Ed): *Proceedings from 3rd conference on Virtual Reality and Persons with Disabilities,* (1995), pp University of Karlstad

Newby, Gregory B. (1993) Gesture Recognition Using Statistical Similarity, CSUN/California State University, In H. Murphy (Ed): *Proceedings of the first Annual Conference: "Virtual Reality and Disability"*

Riedel, O (1993) VIRUSI - Virtal User Interface - Icon- orientierte Benutzerschnitt-stelle für VR-Applikationen, In: *Virtual reality ´93, Anwendungen und trends,* pp 227 - 244

Roy, David M., Panayi, Marilyn, Harwin, William, Fawcus, Robert (1993) The Enhancement of Interaction For People With Severe Speech and Physical Impairment Through Computer Recognition of Gesture and Manipulation, In: *Proceedings from Virtual Reality and Persons with Disabilities,* San Fransisco June 17-18, 1993

Serflek Christopher Treviranus Jutta (1995) Accessible VRML: Shouldn't Virtual Ramps Be Easier to Build? In: Murphy, Harry (Ed): *Proceedings from 3rd conference on Virtual Reality and Persons with Disabilities,* (1995), pp Utah Univ.

Smythe, Phil (1993) Use of Networked VR to Assist People with Special Needs, In: *Proceedings from Virtual Reality and Persons with Disabilities,* San Fransisco June 17-18, 1993

Standen P.J. Cromby J.J. (1995) Can Students with Developmental Disability Use Virtual Reality to Learn Skills which Will Transfer to the Real World In: Murphy, Harry (Ed): *Proceedings from 3rd conference on Virtual Reality and Persons with Disabilities,* (1995), pp, University of Nottingham

Sundgren, Hans, Winquist, Fredrik, Lundström, Ingemar (1992) Artificial Olfactory System Based on Field Effect Devices Montpellier, *Proceedings of the conference Interface to Real & Virtual Worlds,* pp 463-472

Szymanski Axel (1995) Virtual Holography in Diagnosis and Therapy of Sensorimotor Disturbances In: Murphy, Harry (Ed): *Proceedings from 3rd conference on Virtual Reality and Persons with Disabilities,* (1995), pp Utah Univ.

Teil, Geneviève (1992) L'ordinateur peut-il assister le gout? Montpellier, Proceedings of the conference Interface to Real & Virtual Worlds, pp 439-450

Treviranus Jutta (1994) Mastering Alternative Computer Access: The Role of Understanding, Trust, and Automaticity Assistive Technology (1994)6, ppg 26-41

Waller Annalu (1994) Dysphasia: The Cinderella of the AAC World In: Brodin, J. & Björck-Åkesson, E. (Eds) Methodological Issues in Research in Augmentative and Alternative Communication, Proc from the 3rd ISAAC Research Symposium in AAC, Kerkrade, Oct 14-15, 1994, ppg 41-49

ECDVRAT '96

Rehabilitation I

Day 2: 9 July, 1996, Morning

Using virtual reality environments to aid spatial awareness in disabled children

Danaë Stanton, Paul Wilson and Nigel Foreman

Virtual Reality Research Group, Department of Psychology, University of Leicester,
University Road, Leicester, UK.

des1@le.ac.uk, pnw1@le.ac.uk, for@le.ac.uk

ABSTRACT

Several spatial tasks were presented to subjects in computer simulated environments to ascertain whether spatial skills could be trained and enhanced using this medium. Studies carried out at Leicester University have demonstrated that exploration of a simulated building by disabled and able bodied subjects allows them to acquire considerable information about the spatial layout of that specific building. The present studies extended these earlier findings. In Experiment 1, transfer of spatial skills between different virtual environments was investigated. The results confirmed that the skills disabled children acquired using computer simulated environments improved with exposure to successive environments. To eliminate the possibility that learning was non-specific, Experiment 2 compared 3-D exploration and 2-D (control) exploration, finding the former to be superior. Thus the interactivity and three-dimensionality of virtual environments seem to be crucial to spatial learning. Further research is being carried out to establish the nature and extent of the improvement in spatial skills of physically disabled children after intensive exploration of complex virtual environments and thus the value and limitations of VR as a training tool for children with mobility problems.

Keywords: simulated environments, spatial cognition, 3-D and 2-D tasks, physical disability, children

1. INTRODUCTION

Children with mobility problems are limited in their ability to explore their surroundings independently. Those who use a wheelchair may find some places inaccessible, and when guided by welfare assistants or helpers they may take repetitive routes over which they have no control. Children using walking aids may find it tiring and uncomfortable to walk long distances and they may be unable to explore freely for reasons of physical safety.

Limited independent exploration leaves children with poor spatial knowledge of environments, even familiar environments such as their schools (Foreman et al, 1989), and difficulty in coping with novel environments. They may have problems forming effective cognitive spatial maps, which impairs their ability to interpret secondary sources of spatial information such as cartographic maps, and to make detours or take short cuts when moving about in built environments. This poor spatial knowledge may have educational implications, but it also creates a feeling of disempowerment, adversely affecting the individual's confidence in public places (Foreman et al, 1995.) Thus, if a means can be found to improve spatial knowledge in disabled children, by providing independent exploratory experience from an early age, this could have a substantial impact on their overall quality of life, both as pupils and later as adults.

Computer-simulated environments appear to provide an ideal solution to this problem. Children can explore computer simulated environments independently using appropriate interface devices that are tailored to their particular skills and disability. They can explore freely and safely. Such exploration should increase their confidence in subsequent real life exploration, particularly if they can visit an environment in simulated form before encountering the real one. Moreover, spatial disorientation is not confined to disabled individuals; able bodied people who possess a poor sense of direction may also benefit from virtual exploration. Kozlowski and Bryant (1977) concluded that for people to show a good sense of direction it was necessary for them (a) to make a conscious effort to orientate themselves, and (b) to provide them with repeated exposure to the test environment. Clearly, both of these requirements can be met with ease and safety using virtual environments.

Studies carried out at Leicester University in the past have investigated the transfer of spatial knowledge from a virtual to a real environment, establishing that both able-bodied and physically disabled children can acquire substantial spatial knowledge from virtual exploration alone. Able-bodied 18-year old students from a sixth form college were tested in a first study. They had never visited the Psychology building before. Half of them explored a simulation of the building and half explored the real building. Both groups then completed a battery of spatial tests, which included the use of a pointing device to indicate landmarks (such as emergency exits and fire alarms, to which their attention had been drawn during exploration) which they could not see from the testing location. Half of each group was tested “in” the virtual test room (i.e., on computer) and half in the real test room. They were also asked to estimate distances between various objects and make plan and side elevation drawings of the building. It was found that all subjects were able to draw accurate maps and point to objects with greater accuracy than a control group who made reasoned guesses (Wilson 1993; Wilson and Foreman, 1993).

In a second experiment, ten physically disabled children explored the simulation of the Psychology building and were then given similar tasks to those described above. Exploration was encouraged in the form of a game in which the children had to locate fire alarms and fire hoses and activate them, thus unlocking a virtual fire door, from which they could “escape” from the building. The children were then asked to indicate, using the pointing device and by describing routes, where they thought items of fire equipment were located in the real building. Finally, they were asked to escort the experimenter to the real fire equipment and fire exit. Subjects were generally accurate in their angle estimations and could describe shortest routes to objects. They had clearly acquired a good deal of spatial information. For example, they could readily tell the experimenters which objects they expected to find behind doors as they were taken through the real building (Wilson 1993; Wilson and Foreman, 1993).

Similar results have been reported from the United States Army Research Institute in Orlando, where Regian and colleagues (Regian et al, 1992) evaluated how well able bodied subjects could learn to navigate through a virtual office building. Subjects trained in the virtual building learned better than those in the control group, who were shown a series of photographs of the building, and nearly as well as those who explored the real building. The researchers found that subjects learned both spatial-procedural tasks and spatial-navigational skills within a virtual environment. However Kozak et al (1993) failed to obtain transfer from a virtual environment to a real environment for a pick-and-place task. It is clearly crucial to address the issue of which tasks and skills do, and do not, transfer between virtual and real environments. To date, with a few exceptions (Azar, 1996), little work has been conducted to address such psychological issues concerned with human behaviour in virtual environments and to determine how behaviours in virtual environments may differ from behaviours in the real world.

The present studies extended the earlier work in Leicester, examining whether computer simulated environments could be used to train spatial skills generally, i.e., whether spatial-perceptual abilities per se are enhanced with repeated virtual testing. In Experiment 1 we examined whether skills acquired by children using simulated environments would transfer to others, experienced subsequently.

2. EXPERIMENT 1

2.1 Method

2.1.1 Subjects. These were 8 physically disabled children, 4 boys and 4 girls, having a mean age of 11.88 years. Children selected for the study were those having substantial mobility problems. School staff were asked to rate the mobility of each child on a scale of 0: independently mobile, to 10: completely immobile. The average rating was 4.7. Their conditions ranged from cerebral palsy to a heart condition. Every child was able to use a computer keyboard.

2.1.2 Design. Each subject explored three computer-simulated environments at fortnightly intervals. Each session consisted of an exploration phase followed by a number of tests of spatial knowledge. A different environment was explored in each session but the spatial tests remained the same. The independent variables were the layout of the environments and the positions of six target objects. The dependent variables were the subjects’ angle estimations (made using two different types of pointing device), the time taken to find a specified object, and the quality of maps that subjects were able to draw of the experimental environment.

2.1.3 Apparatus. The experimental environments were created using Superscape Virtual Reality Toolkit and were presented on an Intel Pentium 90 with SVGA graphics, displayed on a 14 in. monitor. All three environments consisted of three rooms joined by a T-shaped corridor. Each room was subdivided into smaller sections using two or three walls of ceiling height. In the first environment each room was coloured differently: one burgundy, one pale green and one blue, such that they were clearly discriminable. In the corridor leading out of the central (pale green) room was a START sign. Movements of the viewpoint were controlled using keyboard arrow keys. The viewpoint was 18000 units in height and the walls were 45000 units high. Six objects were placed in the environment, two in each room. These objects were: a piano, an animated shape, a flag, a decorated shield, a camera, and a rotating globe. In order to measure the subjects’ estimates of the positions of objects in the environment, three testing viewpoints were programmed, one in each room. At each viewpoint subjects could pan 360 deg. around the Y-axis, direction being recorded in 5 degree steps. These viewpoints were used to record accuracy of pointing using the screen cross-sights. The second pointing device was a hand held pointing device. This was made from a flat, circular piece of board coloured blue and red. A full circle was marked on the board, segmented in 5 degree steps and a moveable arrow, rotating about the centre of the circle, was used by subjects to indicate chosen directions/angles. When using the hand pointing device the same three testing positions were presented on the screen as were used for the computer pointing estimates. A schematic plan view of the environment (minus the objects) was generated on A4 paper for an object placement task, in which subjects had to place small crosses to indicate where they thought the objects were located.

The second virtual test environment was coloured differently. One room was pink, one turquoise and one orange, and the starting point was moved from the centre room to the right hand (orange) room. Six objects were placed in the environment, two in each room, in the same positions as those occupied by objects in the previous experiment. These objects were: a tank, a postbox, a traffic cone, a “no smoking” sign, a torch and an animated robot. In the third environment, the rooms were moved so that the room that was originally on the right hand side of the corridor was now on the left, the original left hand room being moved to the centre position and the original central room being to the right hand side. Each room was again coloured differently, one yellow, one dark green and one stone. The starting point was placed in the left hand (green) room. Six objects were placed in the environment, two in each room. These objects were: a rotating fairground wheel, a star, a map of the U.K., a car, a clock and an animated insect. Three testing viewpoints were again programmed; however they were positioned differently from those in the previous two environments. A stopwatch was used to measure the time taken to find a chosen object before and after exploration.

2.1.4 Procedure. Subjects were tested individually. At the beginning of each session the experimenter demonstrated a tour through the environment, beginning at the starting point and visiting each room in turn before returning to the starting point. Subjects were told to explore the environment, find the six objects, and try to remember both where the objects were located and the layout of the envivonment. They were made aware that their memory would later be tested. The subject was then given a demonstration of the type of pointing task that they would be asked to complete following exploration.

Before commencing general exploration, the subject was first asked to locate one object, picked randomly by the experimenter, as fast as possible. Latency to find the object was recorded and the experimenter then reset the viewpoint to the starting point. (On test days 2 and 3, the selected test object was always located in a different position from the one used in the previous session, and was always an object that occupied a location distant from the starting point rather than an adjacent location.) The subject was then asked to explore the environment for as long as necessary for them to feel confident about being able to carry out the tests. The experimenter monitored the subject’s exploration to ensure that all six objects had been encountered. When the subject indicated that they felt confident about their familiarity with the environmental layout, the experimenter used the keyboard keys to “transport” them from the starting point to each of the testing points in turn, where they were asked to point toward all six objects in turn, either using the cross sights in the centre of the computer screen (half of the subjects) or the hand held pointer (the remaining half.). The three testing points were then retested in the same order. The subject was then given the same test but using the alternative pointing device (screen-based, or hand-held.) Each subject used the pointing devices in the same order for all three test environments. Note that since subjects were “led” from the starting point by the experimenter to each of the testing positions they could monitor their route through the environment. None of the objects was visible from any of the testing positions.

In the second environment the testing positions were in the same place as in the first, but they were visited in a different order. When the pointing task was completed the experimenter took the subject back to the starting point and asked them to find once again, as quickly as possible, the object that they had been asked to locate before exploration. The experimenter recorded the time taken. Next, the subject was given a blank sheet of A4 paper and asked to draw an outline of the test environment complete with all of the objects correctly positioned. Finally, a plan outline of the environment was presented to the subject who was asked to mark the positions of all objects as accurately as possible.

2.2 Results

A repeated measures analysis of variance (ANOVA) was carried out on median angle estimation error scores using the computer pointing device. A significant main effect was found, $F(2,14) = 6.56$, $p < 0.05$, reflecting significant differences among sessions. A Newman Keuls test showed that session 3 error scores were lower than those of session 1 and 2, but that scores for the latter sessions did not differ (Fig.1.) For equivalent scores using the hand-held device, no significant effects were obtained, $F(2,12) = 0.83$, $p > 0.05$.

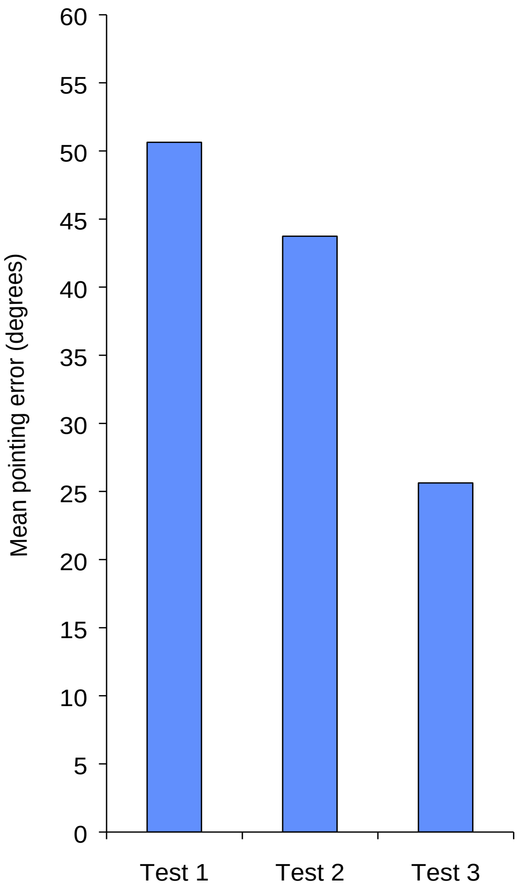


Figure 1. *Group error scores for angle estimations using the computer pointing device.*

The time taken to find a given object before and after exploration is shown in figure 2. One tailed t-tests were used to assess the significance of differences. There was a significant drop in the latency before exploration $t(7) = 2.64$, $p < 0.05$ and also after exploration $t(7) = 1.98$, $p < 0.05$. Error scores for the mapping tasks were computed as the difference in degrees between the actual object direction and the childrens' indicated direction. These were averaged across objects for each map. A paired samples t-test was conducted on the mean error scores for objects positioned on an outline map. The mean error scores for tests 1 and 3 were 55.7 (SD=32.8) and 30.3 (SD=21.0), respectively; a difference that was found to be significant, $t(7) = 3.02$, $p < 0.05$. (Fig.4.) There was no significant effect for the hand drawn maps, $t(6) = 1.63$, $p > 0.05$.

2.3 Discussion

The results show that on several criteria, childrens' performance on spatial tasks improves with repeated experience of virtual environments. Improvement was evident when measured using the computer screen as a pointing device, but no improvement was seen when the hand pointing device was used. The latter device requires a fairly complex transformation from large-scale locomotor space to a small, artificial device, which proved difficult for the children to use. From their verbal reports, they appeared to have problems relating the direction of the pointer to real world locations. The ANOVA on the median error scores for angle estimations using computer pointing shows that error scores reduced significantly over trials, though post hoc analysis revealed that improvement was evident only by test 3 (which differed significantly from both tests 1 and 2.) Little improvement occurred between the first 2 test sessions, which did not differ significantly. Childrens' positioning of objects on an outline map improved significantly between test 1 and 3, though their hand drawn maps did not show any significant change. The latter measure is subject to great variability and is a less powerful measure of cognitive map formation. The average time the children required to find a selected object was greatly reduced after exploration, indicating that they had learned specific routes. However there was a general practice effect involved, inasmuch as children were faster both before *and* after exploration on the last session when compared with the first.

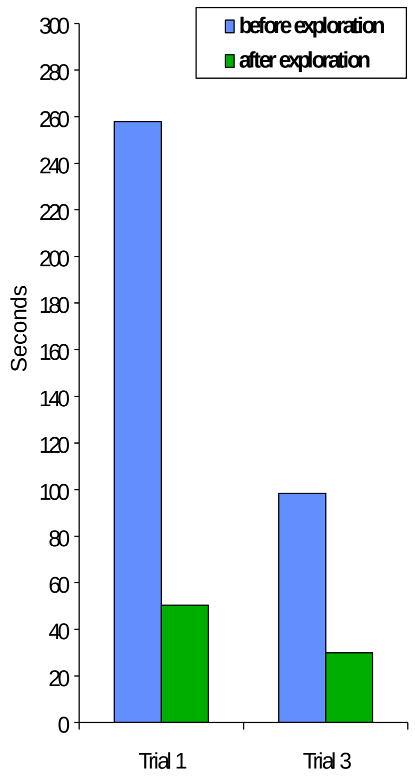


Figure 2. *Time taken to find a given object.*

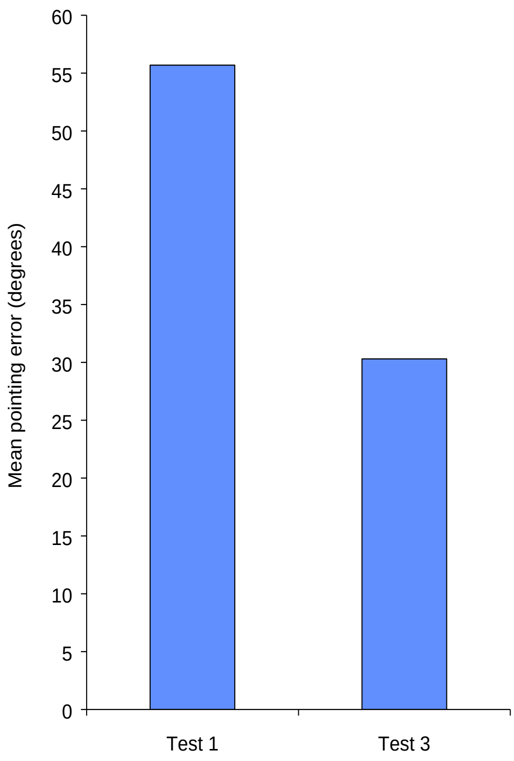


Figure 3. *Mean error scores for objects positioned on an outline map.*

No significant correlation was found between mobility and motor skill, but it is perhaps unrealistic to draw conclusions from a relatively small sample of subjects. A possible confounding factor in the experimental design was that the order in which subjects explored the environments was not counterbalanced. Although there may be differences between environments in terms of their navigability, this is unlikely to have influenced the results since each environment in the present study was constructed from the same basic elements (i.e., the same three rooms and the same corridor.)

The results of this study are very encouraging since diminishing error scores are likely to reflect children's improved use of some form of internal representation (cognitive map) which is necessary for the solution of the tasks. The fact that significant improvement only emerged in test 3 suggests that several exposures are needed for learning to take place. Nevertheless, having shown that childrens' spatial abilities improved after exploring virtual environments, it was necessary to establish that this improvement was due to the interactivity and three-dimensionality of the virtual environments and not to other non-specific factors such as improved confidence and familiarity with the experimenter, or computer and keyboard familiarity**.** This was examined in Experiment 2 by comparing 3-D exploration with 2-D (control) exploration. Navigational measures were supplemented by other spatial tasks, measuring various aspects of visuo-spatial skill, in order to test the generality of cognitive spatial improvement occuring after 3-D training.

3. EXPERIMENT 2

3.1 Method

3.1.1 Subjects. There were 24 subjects, 3 girls and 21 boys, having a mean age of 10.38 years. The subjects were selected on the same criteria as Experiment 1.

3.1.2 Design. A pretest consisting of a battery of three spatial tasks (see below) was followed by four 30 min. sessions of interaction with either two dimensional or three dimensional computer graphics in a "game" format. Subjects were then retested on the same battery of spatial tasks as before. Note that this design ensured that children in both experimental and control conditions spent an equal amount of time with the experimenter and gained an equal degree of familiarity with the keyboard.

3.1.3 Apparatus. The virtual environments used in Experiment 1 were reused in one of the spatial tests in this study, though here they were presented in counterbalanced order to avoid effects due to differential navigability. In addition the *Money Standardized Road-Map Test of Directional Sense* was used and an adapted version of the *Shepard and Metzler Mental Rotation Test*. The Money Road-Map Test requires subjects to follow a route through a stylised street map, and make judgments at 32 turning points as to whether a turn is to the right or to the left. (The difficulty in the task is that subjects have to make these judgments when "travelling" in different directions on the page.) They are not allowed to rotate the paper. The adapted Shepard and Metzler Test consists of thirty two pictures of a geometric shape which is composed of seven cubes. The shapes were created using Superscape Virtual Reality Toolkit and were printed in black and white. One shape was created and eight of the 32 pictures displayed this shape in different orientations, rotated in the (depth) z-plane by various angles about its centre point. Another eight pictures showed the shape rotated through various angles in the (picture) y-plane about its centre point. The original shape was then reflected 180 degrees and eight of the pictures displayed this shape at various degrees of (depth) z-plane rotation, and the final eight pictures displayed this shape at various degrees of (picture) y-plane rotation.

3.1.4 Procedure. Each child attended six test sessions. In the first session the child was given the Money Road-Map and the modified Shepard and Metzler Mental Rotation tests. Before taking the Money test, subjects were asked to indicate first their right hand and then their left ear. (All subjects could do this easily.) The map was then placed before the subject, who was told to imagine that they were following the path shown on the map and at each turn to say whether they would be making a left or a right turn. Subjects then followed a short practice route, the experimenter correcting any mistakes that they made. The full test then followed, during which no feedback or assistance was given.

For the Shepard and Metzler test, each subject was shown examples of the original shape rotating on the computer screen and was then tested on the thirty two pictures. They were provided with a reference picture of the original shape, and they were shown the thirty two pictures one at a time. The subject had to say, for each picture, whether it was the *same* shape as the reference or a *different* shape, the different shape being the reflected shape. Subjects were shown eight depth pictures (four the same, four different), eight picture plane pictures (four the same, four different), and sixteen of the depth and picture plane shuffled (eight the same, eight different). The pictures were shuffled between subjects.

At this point the child explored one of the novel environments used in Experiment 1 and completed the same spatial tests (finding an object as quickly as possible, estimating angles from given positions and drawing maps). Half of the children were assigned to the 2-D and half to the 3-D groups. The next four sessions consisted of individual 30 min. sessions of exposure to either two dimensional, or three dimensional environments which the child explored with the experimenter. The two dimensional environments were selected from the popular games market and consisted of non-violent platform and adventure games. The three dimensional environments were created using Superscape Virtual Reality Toolkit and consisted of large scale and small scale environments (such as a leisure centre or an office) in which the child could explore and interact with objects. The exploration was presented in the form of a game by asking the child to try and find objects and interact with them.

Finally**,** in the sixth session the children carried out the same tests as in the first session (the Money test, the Shepard and Metzler Test, exploration of a novel virtual environment) and was tested in the same way as previously. Note that the novel environment was different from the one explored in the first session.

3.2 Results

A two factor mixed ANOVA was carried out on the median error scores for computer pointing for 2-D and 3-D groups. An interaction was found between the 2-D/3-D factor and test sessions, $F(1, 22) = 6.34$, $p < 0.05$. A one-tailed t-test between two dimensional and three dimensional exposure on the post test revealed a significant effect $t(22) = 2.02$, $p < 0.05$. A related t-test on the two dimensional group from pre- to posttest revealed no significant difference ($p > 0.05$.). However a related t-test on the three dimensional group from pre- to posttest did reveal a significant difference $t(11) = 2.28$, $p < 0.05$. (Fig.4) The time taken before and after exploration at pretest and posttest in both conditions are illustrated in figures 5 and 6. Although the scores for the 3-D group, after exploration, were falling in the right direction, they failed to reach significance when analysed using a two factor mixed ANOVA, $F(1,22) = 2.41$, $p > 0.05$. A two factor mixed ANOVA carried out on the error scores for objects placed on the outline map, and on the hand drawn map revealed no significant differences from pretest to posttest or by condition. Test scores from both the Money and the Shepard and Metzler tests, analysed using a two factor mixed ANOVA, failed to reveal improvement from pretest to posttest.

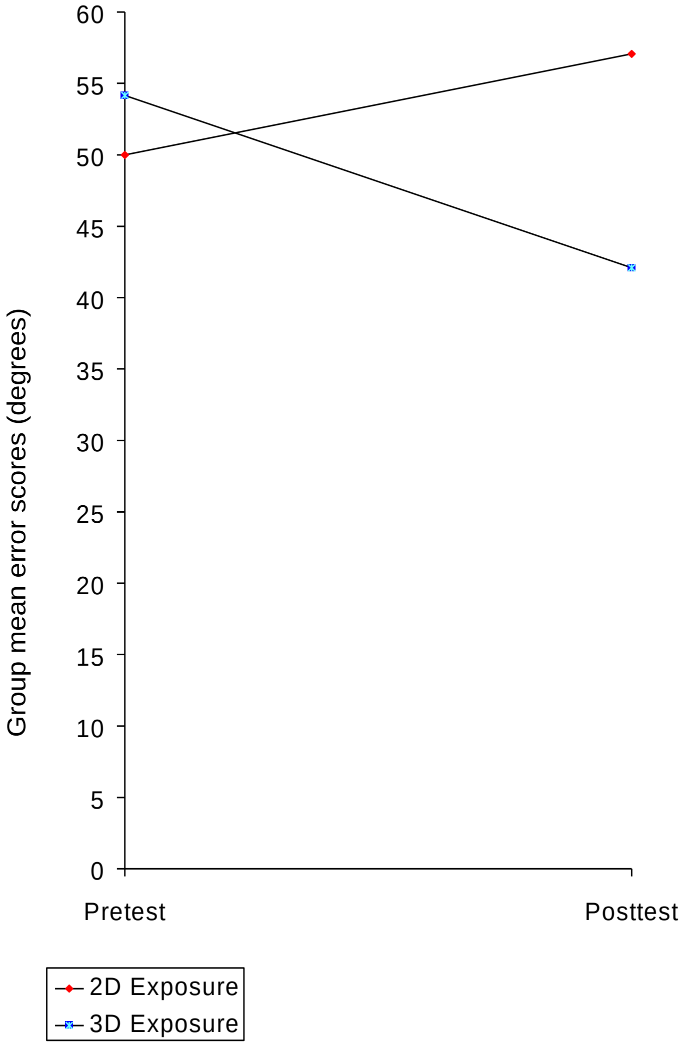
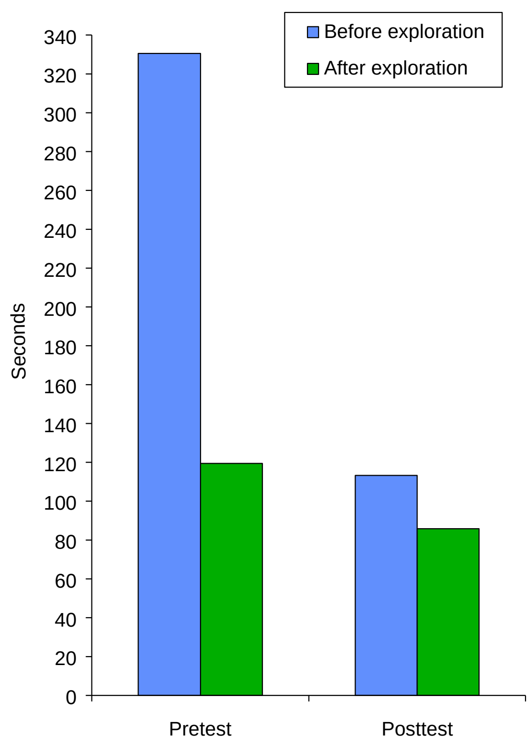


Figure 4. *Two dimensional vs three dimensional interaction for angle estimations using the computer pointing device.*

3.3 Discussion.

The results show that childrens' scores on navigational spatial tasks improve when they use 3-D training but not 2-D training. This was particularly clear from the task which we have found in this and other studies to be the most effective measure of spatial skill, namely the computer pointing task. The data support the notion that improvement in spatial skill seen in this and previous experiments is specifically due to the unique features of simulated environments created using 3-D graphics, and which provide both three dimensionality and real-time interactivity. The interaction reveals a difference between the performance of the 2-D and 3-D groups between pretest and posttest. While the rise in 2-D scores (i.e., a slight worstening in performance) between pretest to posttest was surprising, the significant drop in the scores of 3-D subjects (i.e., significantly improved scores) reflects their improved spatial performance. The reduced time taken to find a given object by 3-D subjects further reinforces this conclusion, suggesting that children in this group had learned routes and acquired navigational information of a kind that would be beneficial in an equivalent real environment. For the 2-D condition the posttest score (after exploration) was actually higher than the corresponding pretest score. This could not be due to differences in the navigability of the pretest and posttest environments, as the order in which they were used was carefully counterbalanced.

Measures of competence in drawing maps and placing objects on outline maps failed to reveal any improvements in the posttest phase, though these measures are subject to considerable variability, largely due to difficulties in choosing appropriate scoring criteria. We are currently investigating alternative criteria. The Money Road-Map and the Shepard and Metzler tests also failed to reveal improvement, after 2-D or 3-D experience, and thus we conclude that spatial skills learned in a virtual environment are fairly specific. There is apparently no generalisation to non-navigational skills, which may reflect a quite different set of underlying visuo-spatial cognitive abilities.



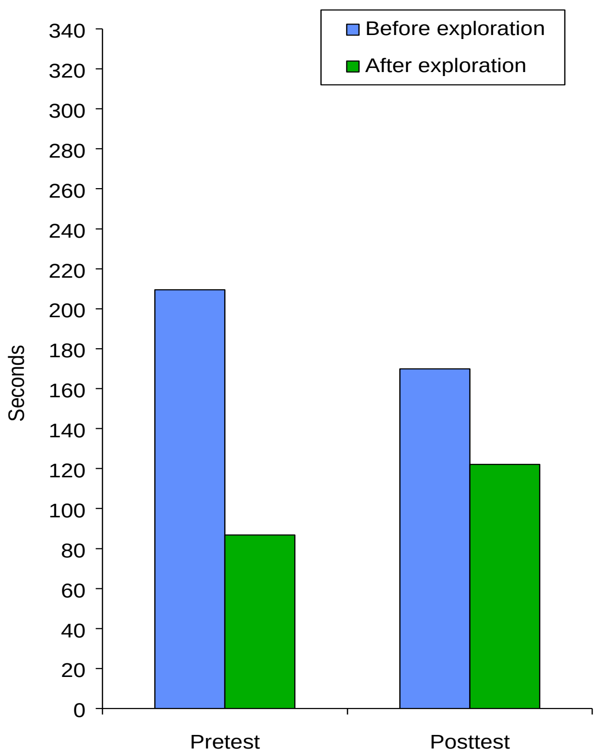


Figure 5. *Time taken to find a given object for the 3D group.*

Figure 6. *Time taken to find a given object for the 2D group*

If this design were replicated, the 3-D environments used in the four 30 min. sessions could all be created as large scale environments. In the present experiment in which subjects explored large and small scale environments, it was found that when the subject was in a small scale space (just one room) little exploration took place. The subject spent the majority of the time purely interacting with objects. The problems children are most likely to encounter in reality are more likely to occur in larger scale environments where goals can not be viewed from one point in the environment.

4. GENERAL DISCUSSION

It is evident from the present studies that there is a transfer of spatial skill from one virtual environment to another and that this improvement in spatial ability is primarily due to the interactivity and three dimensionality of virtual environments. These results have important implications for disabled children in that virtual reality is proving to be a potential training medium for skills that could enhance quality of life for many disabled children. The software used in the present study is affordable by many schools. We have shown that subjects acquire information from virtual exploration which enables them to understand the spatial layout of an environment to a considerable degree. However, it is important to discover the optimal ways of presenting virtual information, and the optimal modes of interaction. We are currently investigating alternative input devices, and screen, projection and head-immersion modes of presentation, using a variety of cognitive spatial (detour and short-cut) tests, in order to assess what aspects of spatial skill are successfully acquired when subjects navigate in virtual worlds, and whether alternative modes of presentation and interactivity may be used to provide more comprehensive training.

Acknowledgements: The authors gratefully acknowledge financial support from Action Research and BT.

5. REFERENCES

B. Azar (1996) Training is enhanced by Virtual Reality. *Monitor of the American Psychological Association*, **27** (March), pp 24.

N P Foreman, C Orencas, E Nicholas, P Morton, and M Gell (1989) Spatial awareness in seven to eleven year-old physically handicapped children in mainstream schools. *European Journal of Special Needs Education*, **4**, pp. 171-179.

N P Foreman, P Wilson, and D Stanton (1995) Virtual reality applications: Spatial competence and personal confidence. *7th TIDE Bridge Phase Workshop*, Brussels 7 July.

J J Kozac, P A Hancock, E J Arthur and S T Chrysler (1993), Transfer of training from virtual reality, *Ergonomics*, **36**, 7, pp. 777-784.

L T Kozlowski and K J Bryant (1977), Sense of Direction, Spatial Orientation and Cognitive Maps, *Journal of Experimental Psychology: Human Perception and Performance*, **3**, 4, pp.590-598.

J Metzler and R N Shepard (1982), Transformational studies of the Internal Representation of Three-Dimensional Objects. In *Mental Images and their Transformations* (Shepard and Cooper), The MIT press, London, pp. 25-72.

J Money (1965), *A Standardized Road-Map of Direction Sense*, The John Hopkins Press, Baltimore.

J W Regian, W L Shebilske and J M Monk (1992), Virtual Reality: An Instructional Medium for Visuo-Spatial Tasks, *Journal of Communication*, **42**, 4, pp.136-149.

P Wilson (1993), Nearly There. *Special Children*, **68**, pp. 28-30.

P Wilson and N Foreman (1993) Transfer of information from virtual to real space: Implications for people with physical disability. *Eurographics Technical Report Series. ISSN*.1017-4656, pp.21-25.

Successful transfer to the real world of skills practised in a virtual environment by students with severe learning difficulties

J.J. Cromby, P.J. Standen, J. Newman & H. Tasker

Department of Learning Disabilities,
University of Nottingham Medical School
Queens Medical Centre, Nottingham NG7 2UH

John.Cromby@nottingham.ac.uk

ABSTRACT

Nineteen students with severe learning difficulties aged between 14 and 19 years completed a shopping task in a real supermarket before 9 students, the experimental group, had twice-weekly sessions carrying out a similar task in a virtual supermarket. The remaining 10 students formed the control group, matched with the experimental group for ability, age and sex. They had the same number of sessions using other virtual environments. Although there was no difference between the two groups at baseline, on repeating the task in the real supermarket the experimental group were significantly faster and more accurate than the control group.

Keywords: transfer, generalisation, training for independent living

1. INTRODUCTION

1.1 Overview

Before virtual environments can be widely used as an educational medium it must be demonstrated that skills practised in them can transfer successfully to the real world. This is a particular issue when the intended users are students with severe learning difficulties (in other countries, mental handicap or mental retardation), for whom generalisation of skills from training situations to the real world is notoriously difficult. This paper describes a pilot study showing successful transfer to the real world of shopping skills practised in a virtual supermarket by a group of students with severe learning difficulties.

1.2 Background

With the continuing expansion of “care in the community” there is an urgent need to train young people with learning difficulties in the skills required for independent living, yet the training methods currently in common use are problematic. On the one hand, if skills are taught in a training situation initial learning may not transfer easily to novel situations (Miller 1992). Assessment of social skills training programmes for people with learning disabilities shows that skills and knowledge learnt in training are unlikely to generalise beyond the teaching situation and that even when they do they are rarely maintained for very long after training has ceased (Robertson et al, 1984; Davies & Rogers, 1985). This failure to generalise has been attributed to both the design of the intervention (Stokes & Baer, 1977; Ward & Gow, 1982) and to the individual’s perceived status within the group varying from setting to setting (Selman & Jaquette, 1977).

Alternatively, training can be attempted in situ. However, this is both time consuming and staff intensive and hence prohibitively expensive. It also involves an element of risk, and in some instances the consequences of making a mistake may be so great that training in the real world is simply not an option. For example, use of a light industrial workshop by people with learning disabilities would require a grasp of basic health and safety principles before any in situ training could even be attempted.

Virtual environments may provide a solution to this problem. With a combination of careful design and programmatic use directed at clear and obtainable learning objectives they can be made to mimic closely the salient or critical features of the real world, making it more likely that skills will be transferred successfully. Where risk is a factor they permit the user to learn by making mistakes without suffering the real consequences of their errors. They may also have a particular pertinance in the training of people with learning disabilities because they promote self-directed activity, decrease reliance on language or other abstract symbol systems, and minimise the effects of physical impairments (Cromby et al, in press).

For virtual environments to be of any use in training, learning obviously has to generalise to similar experiences in physical reality (Bricken, 1991). An unsuccessful attempt to demonstrate the generalisation of skills learnt in a virtual environment (Kozak et al., 1993) attributed the failure to the lack of veridacity of the virtual environment and the over-simplistic task (moving cans to target locations). However, Wilson (1993) describes successful attempts to teach children with physical disabilities the location of fire exits and appliances using a virtual model of the real building in which they were tested.

The present study examined whether shopping skills practised in a virtual supermarket by students with severe learning difficulties would transfer to a shopping task in a real supermarket. This task was chosen because of its high ecological validity and the relative ease with which it could be quantified. Using a supermarket is precisely the kind of activity which people with learning disabilities must be able to accomplish if they are ever to acquire any meaningful degree of independence.

2. METHOD

2.1 Selection of participants

The research took place at a school for students with severe mental retardation aged between three and 19 years. School staff nominated students to take part in the research if they met the following criteria:

1. They had sufficient motor skills and visual ability to be able to use the computer terminal and joystick
2. They were sufficiently able to carry out a real shopping trip with minimal staff support
3. They had used virtual environments on at least three previous occasions
4. The parents or carers gave consent for their child to take part in the research.

This resulted in a sample of 21 students aged between 14 and 19 years whose teachers then completed the classroom version of the Vineland Adaptive Behaviour Scale (VABS). The students were then assigned to either the experimental or control group so that the two groups were matched on age, gender and Vineland score. Their parents were asked to complete a short questionnaire detailing how much their child accompanies them and helps them with shopping. There were no differences between the two groups on these measures.

2.2 Design

Students' baseline performance on a shopping task in a real supermarket was compared with their performance after an eleven week interval during which the experimental group had twice-weekly sessions using a virtual supermarket while the control group spent this time using other virtual environments.

2.3 Procedure

2.3.1 The Real Shopping Task. Students were taken to the local supermarket and individually given a shopping list of four items which they were to find, put in their trolley and take to the checkout. The list was one of six made from both miniaturised colour reproductions of packets of commonly available consumer goods such as washing powder and breakfast cereal and actual labels cut from other items such as tinned goods, which were then mounted under clear film onto pieces of A4-sized card.

Their performance was monitored by the accompanying adult who maintained a distance from the student and recorded the following measures:

1. Total time: Time taken from passing through the turnstile into the shopping area until stopping at the checkout with their shopping.
2. Total number of items: All items placed in the trolley even if they were later put back onto the shelves before reaching the checkout.
3. Total items at end: Number of items in the trolley on arrival at checkout.
4. Total items correct: Number of items from the shopping list in the trolley on arrival at the checkout.

2.3.2 The Virtual Shopping Task. This consisted of a two-aisle store viewed on the monitor of an IBM compatible computer with movement through the store achieved by use of a joystick and selection of items by a mouse. Shelves

Proc. 1st Euro. Conf. Disability, Virtual Reality & Assoc. Tech., Maidenhead, UK, 1996
©1996 ECDVRAT and University of Reading, UK ; ISBN 0 7049 1140 X

were filled with a representative selection of goods found in the local supermarket. There were five different lay-outs of goods on the shelves which the students were presented with at random each time they started a session. Students were given a list as described above and faced with the same task. Each session was initiated by entering the student's identifying number and the number of the shopping list used. The same set of measures that the staff collected on the real shopping trip were then obtained from the computer's record or by a researcher. In addition the researcher recorded the amount of time the students spent looking at the screen as a proportion of the total time spent in the virtual supermarket.

2.3.3 Other Virtual Environments. While the experimental group used the virtual supermarket the control group were free to use any of the other virtual worlds which included a virtual city, a virtual house and a ski-slope. In the days before the final assessment the control group also played a game with the shopping lists to ensure that they were not disadvantaged in familiarity with them when they returned for the second real shopping trip.

3. RESULTS

3.1 Participants

Two students in the experimental group spent very little time looking at the screen and there was no improvement in the time it took them to complete the virtual shopping task. They were therefore excluded from further analyses. The characteristics of the students who completed the trial are shown in Table I below.

Table 1. *Characteristics of participants*

	Mean Age	Gender	Mean VABS
Experimental Group	15.84	5M / 4F	5.40 (SD 1.66)
Control Group	15.67	6M / 4F	5.48 (SD 1.72)

3.2 Analysis

Data were analysed using SPSS for Windows.

3.2.1 Total Time to complete the task. Mean times taken to complete the real world shopping task for both real shopping trips are shown in Table 2 below. Neither group showed an improvement on this measure because the task had become more difficult at the second visit. The store had prepared for Christmas by changing their lay-out and increasing their stock. However, using an analysis of covariance with baseline time as a co-variate, the experimental group were significantly ($F = 7.173$, $df = 1,18$, $p<0.02$) faster than the control group at the second visit.

Table 2. *Mean Times to Complete Real Shopping Tasks*

	First Visit	Second Visit
Experimental Group	11.97 (SD 5.7)	11.59 (SD 3.1)
Control Group	11.21 (SD 6.9)	16.92 (SD 5.8)

3.2.2 Number of Items Correct. Using a Mann-Whitney U Test there was no difference between the two groups at baseline but at the second visit the experimental group had significantly ($p<0.05$) more correct items than did the control group. On the second visit there was no difference between the two groups in the number of items picked up whether correct or not (total number of items) and the experimental group had significantly ($p<0.05$) fewer items in the trolley at the checkout (total items at end).

3.2.3 Other Shopping Experience. A two way ANOVA showed that whether students accompanied their parents shopping was significantly related to final shopping time ($F = 4.70$, $df = 1,18$, $p<.05$). However, the interaction with group was not significant, i.e. the experience of accompanying parents did not differentially benefit either the experimental or the control group.

4. DISCUSSION

In this study performance on a real shopping task benefitted from the experience of shopping in a virtual supermarket. Neither group showed an improvement from baseline, but on the second visit just two weeks before Christmas the task was significantly harder. The supermarket had increased their stock from around 5,000 to around 7,500 lines, and distracting displays of attractive toys, gifts and decorations were placed prominently at the junctions of aisles. Faced with this more difficult shopping task on the second occasion the experimental group completed the task in less time than did the control group.

Students in the experimental group also selected more correct items than did the control group. This cannot be explained in terms of the students improving their chances of a correct choice by simply picking up more items since the experimental group had significantly fewer items in their trollies on arrival at the checkout.

These results were achieved using a very simple virtual environment bearing a limited resemblance to the real supermarket. It is likely that in its simplicity the virtual environment succesfully abstracted the essential features of supermarkets (packaging, aisles, shelves, checkout desk) enabling the students to recognise them in the real supermarket. The question remains of how much detail should be built into virtual environments to increase the likelihood that learning will generalise to other settings. The virtual environment needs to have sufficient detail for the learner to be able to practice skills needed in the real world (visual search, navigation) and to be recognisable as a representation of the real world. On the other hand too much detail may prevent the learner from extracting the salient features necessary for the task to be learnt. Further research may be able to identify the optimal levels of detail to facilitate both successful skill acquisition and generalisation.

Shopping in a supermarket involves a variety of skills: memorising items from the list, visually searching for them, remembering the route taken so far and the locations of items already passed. This is all made easier if the shopper has some understanding of the categorisation of goods such that finding coffee on the shelves makes it more likely that tea bags will be located nearby. In this study it is possible that the advantage that the experimental group had over the control group was mediated by a more general increase in familiarity with the task, or the fact that their sessions with the virtual supermarket were more structured than those on the other virtual environments. Research into changes in the specific sub-skills of shopping is currently underway, but measures need to be developed to assess any change in more generalised skills such as autonomy and decision making which may be produced by the experience of learning in virtual environments.

Finally, there is the the question of maintenance: ideally, skills must not only transfer to the real world but should also endure for some significant period. This raises many other issues, since regardless of the quality of the training method skills will only be maintained if there are real opportunities to practice them. A more interesting question may be how much learning experience in a virtual environment is necessary before the learner can move into the real world and begin learning there? Future research must identify critical features of appropriate educational virtual environments for students with learning difficulties, and also develop appropriate guidelines for their successful integration into the classroom.

5. REFERENCES

Bricken, W. (1991) Training in virtual reality, Proceedings of the 1st International Conference on Virtual Reality, London, pp 46-48

Cromby, J.J., Standen, P.J. & Brown, D.J. The potentials of virtual environments in the education and training of people with learning disabilities, in press, J. Intellectual Disab. Rsch

Davies, R. R. & Rogers, E. S. (1985) Social skills training with persons who are mentally retarded, Mental Retardation, 23,(4), pp. 186-196

Kozak, J. J., Hancock, P. A., Arthur, E. J. & Chrysler, S. T. (1993) Transfer of training from virtual reality, Ergonomics, 36, pp. 777-784.

Miller, E.C. (1992) Special experiences for exceptional students: integrating virtual reality into special education classrooms, “Art, Science and Visual Literacy: selected readings from the 24th Annual Conference of the International Visual Literacy Association” Pittsburgh PA, pp326-334

Robertson, I., Richardson, A. M. & Youngson, S. C. (1984) Social skills training with mentally handicapped people: a review, Br. J. Clin. Psychol., 23, 241-264.

Selman, R. L. & Jaquette, D. (1977) Stability and oscillation in interpersonal awareness: a clinical-developmental analysis. In C. B. Keasey (Ed.), Nebraska Symposium on Motivation, 25, pp. 261-304

Stokes, T. F. & Baer, D. (1977) An implicit technology of generalisation. J. App. Beh. Anal., 10, 349-367.

Ward, J. & Gow, L. P. (1982) Programming generalisation: a central problem area in clinical psychology, Educ. Psychol., 2, 231-248.

Wilson, P. (1993) Nearly There, Special Children, Sept., 28-30.

Providing motor impaired users with access to standard Graphical User Interface (GUI) software via eye-based interaction

Howell Owen Istance[1], Christian Spinner[2] and Peter Alan Howarth[3]

[1,2]Department of Computer Science, De Montfort University,
The Gateway, Leicester, LE1 9BH England, U.K.

[3]Department of Human Sciences, Loughborough University
Loughborough, Leicestershire, LE11 3TU England, UK

[1]*hoi@dmu.ac.uk*, [3]*p.a.howarth@lboro.ac.uk*

ABSTRACT

We have designed an on-screen keyboard, operated by eye-gaze, for use by motor-impaired users. It enables interaction with unmodified standard Graphic User Interface (GUI) software written for able-bodied users, and it is not solely designed around the need to enter text. The keyboard will adapt automatically to the application context by, for example, loading a specific set of keys designed for use with particular menus whenever a menu is displayed in the target application. Results of initial evaluation trials are presented and the implications for improvements in design are discussed.

Keywords: eye-control, visual keyboard, physically-challenged, disability, handicapped

1. INTRODUCTION

A practical way to enable people with various forms of motor impairment to work efficiently in office environments is to provide the means whereby they can use the same software tools at the workplace as their able-bodied colleagues. This requires building interaction devices which allow control over the range of software available and which ideally are usable by people who have different motor impairment.

Eye-based interaction is attractive in this context for a number of reasons. First, it offers the prospect of reducing learning time by providing a ‘natural’ means of pointing at a displayed object on-screen. Second, moving the eye is fast and positioning a pointer on a required object can be done quickly. Third and perhaps most important, users with severe degrees of motor impairment generally retain good ocular motor control, and so devices based on eye-movement may be used by a larger range of users with motor impairments than devices relying on other muscle groups.

These advantages have, however, to be offset against the known problems of eye-based interaction. These include the level of accuracy with which eye position can be measured, the degree of fine control that the user can exercise over eye movement and the need to be able to disengage eye-control whenever the user wishes to look at the screen without issuing commands. To overcome these problems, eye-based control may be combined with other input modalities, such as speech, so that eye-gaze is used for pointer positioning only and other modes of input are used to make selection actions (equivalent to those normally made by the mouse button). However, at this stage we have restricted ourselves to the use of the eyes alone for both cursor control and selection, rather than investigating combinations of different modalities. Previous work (Istance and Howarth, 1994) investigated the use of eye-gaze for emulation of a mouse to interact directly with standard Graphic User Interface (GUI) applications, and it was concluded then that an indirect ‘soft’ control device offered a means of overcoming many of the problems associated with the direct eye-based interaction approach. Software-generated virtual keyboards that are displayed on a display screen (which are synonymously termed either an “on-screen keyboard” or a “visual keyboard”) offer a solution to these problems, and then because there is no need to invoke additional input modalities, one does not have to rely upon the user having any form of motor control other than over their eyes alone.

The idea of a visual keyboard displayed on a screen and operated by some external device is certainly not new. Indeed, visual keyboards have been developed in the past for use with a variety of interaction devices, such as joysticks and mice, as well as eye-gaze control. Some of these have emulated normal keyboards, but in doing so have been restricted almost entirely to text entry. These have not allowed control over the variety of objects such as those

 ISBN 0 7049 1140 X

contained in menus and dialogue boxes found in modern direct manipulation GUI software. There are many applications, such as Web browsers, which have very limited requirements for text entry but instead require interaction with displayed documents, such as scrolling or clicking on links, as well as interaction with menu items. Other visual keyboards have been developed to control specific software applications only, for example communication and control systems, and can not be used as general purpose input devices.

2. FUNCTIONS AND DESIGNS OF EXISTING VIRTUAL KEYBOARDS

2.1 Single Application Keyboards

EyeScan (Eulenberg, 1985), BlinkWriter (Murphy et al., 1993), ERICA (Hutchinson et al., 1989; White et al., 1993) and EyeTracker (Friedman et al., 1981) are examples of applications which support eye-gaze controlled text entry.

Early text entry devices for physically-challenged users arranged letters according to their frequency of occurrence in text and allowed indirect selection by stepping through a matrix (e.g. MAVIS system in Schofield, 1981; HandiWriter in Ten Kate et al., 1980). This technique has the advantage of accommodating a large set of characters and other symbols within a single template. Demasco and McCoy (1992) proposed a virtual keyboard model which, besides scanning single characters, also supports scanning of words. This technique is unnecessarily restrictive for eye-based pointing and is better suited to devices reliant on more limited forms of motor control. Nevertheless, laying out character keys on the basis of frequency of use is an important design consideration. As selecting larger targets with an eyetracker is easier than selecting smaller targets, frequently used keys can be made larger within the keypad (static sizing) or they can be expanded temporarily when the pointer moves within the key (dynamic sizing), as well as being positioned appropriately.

Operating a visual keyboard by eye is comparable to the one-finger approach of a novice user who also delays each keystroke after the position of a desired key has been recognised. It is a highly sequential task as text must be entered character-by-character with forced delays between the ‘eyestrokes’. Consequently, text entry becomes tedious and keystroke savings through a word prediction system not only have the potential to improve speed but also to reduce errors and user fatigue. Koester and Levine (1994) examined user performance in text entry tasks with word prediction by considering the trade-off between the number of keystrokes and the additional cognitive and perceptual loads imposed by having to select from the presented word choices. They concluded that the cognitive cost of presenting a set of word choices, and explicitly selecting one, largely eliminated the performance advantage of keystroke savings. However, their results depended very much on the degree of disability and the chosen method of controlling the keyboard. For instance, there may be a different trade-off when using a mouthstick for typing on conventional keyboard than in the case of a visual keyboard, where the input and output channel is the same. Indeed, the absence of the need to switch between external keyboard and screen suggest that there will be a lower cognitive load when a visual keyboard is in use.

2.2 General Purpose Keyboard Emulators

EyeTyper, developed by Friedman et al. (1984) was an eye-gaze controlled keyboard which could be used in place of the standard keyboard of an IBM-PC. Hence, it was transparent to the host computer’s software but its architecture was based on a single keypad template which did not allow more than a very simplistic keyboard emulation.

More recently, the visual keyboard developed by Bishop and Myers (1993) at the University of Iowa supports control by an eyetracker device. The arrangement of character keys is advantageous and groups frequently-used characters towards the centre of the screen. In our experience, a discrepancy between visible cursor position and actual eye position is more likely in the screen corners, and the keys and button sizes of their visual keyboard are arranged to take this into account. However, the support for editing text is quite limited. Key combinations to highlight text, such as <*Ctrl+Shift+Right*> are not possible, and whereas characters are available for text input, they are not available for command selection (e.g. <*Alt+F*>). Thus it is not possible to make use of the rich functionality of a word processor like MS-Word, thereby loosing the advantage of services such as spell-checking.

3. DESIGNING THE VISUAL KEYBOARD

3.1 Customising the Keyboard

Given that the application for which the keyboard cannot be specified in advance, it is clearly important to provide end-user customisation, which can be provided through:

- providing suitable configurable defaults for the keyboard (e.g. dwell time settings)

Proc. 1st Euro. Conf. Disability, Virtual Reality & Assoc. Tech., Maidenhead, UK, 1996
©1996 ECDVRAT and University of Reading, UK ; ISBN 0 7049 1140 X

- providing an off-line keypad editor to enable new keypads to be created which are suited to different applications
- providing the option to create keys which invoke macros. For example, a keypad intended for use with Netscape might contain a key to display a particular page, and this would invoke a macro with the keystrokes necessary to display the URL prompt and enter the characters defining the data.
- self-adaptation of visual keyboard to the current application context (e.g. automatically loading a menu keypad)

3.2 Summary of Design Requirements

In summary, there are a number of requirements that a visual keyboard controlled by eye needs to satisfy:

- it should support both keyboard and mouse emulation
- it should support effective interaction with GUI components such as the scrollable lists, text fields and buttons found with dialogue boxes and not be solely designed around the need for text entry
- it should provide mechanisms to compensate for the inaccuracy inherent in eye-based control
- it should enable the user to customise the device to suit individual preferences concerning tasks within specific applications, but should allow the user to switch between different applications without the need for device reconfiguration.

4. OPERATION OF THE KEYBOARD

4.1 System Architecture

The eyetracker provides raw data on the positions of the left and right eye to the host machine (Istance and Howarth 1994). This data is processed completely separately from the visual keyboard. The software responsible for this uses the current gaze position to move the mouse pointer. The visual keyboard runs as an application, and is sensitive to the position of the mouse pointer. If the mouse pointer remains within the area of the key for a specified time (dwell time), a Windows event corresponding to the key is generated. In this way, it is possible to set different dwell times for different types of key. The visual keyboard is thus completely separate from the eyetracker system providing the data, and although we have concentrated on eye-control, the keyboard could perfectly well be used with other types of input device, such as a joystick or mouse.

4.2 Overall Design of the Visual Keyboard

The keyboard is comprised of three sections as shown in Figure 1. The left section contains the general keyboard system commands menu, the centre section contains the variable keypad, and the right section contains the keypad selection menu.

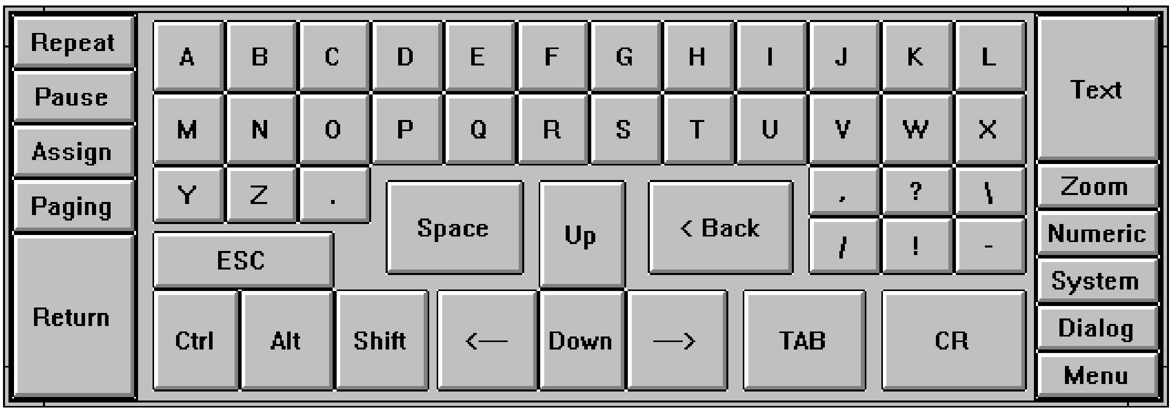


Figure 1. *Visual keyboard with text keypad loaded*

In the keypad selection menu and the system command menu, each key expands when the pointer first moves into it and reverts to its original size when the pointer moves over another key. When the cursor moves off either menu, the last key expanded remains enlarged. This makes dwelling with the eye within the key area easier without the penalty of taking up window space for keys which are not being used (the same principle as a pull-down menu). Additional keypads may be created in the keypad editor and added to the menu of standard keypads or may replace them in the menu.

4.3 Keypad Design

4.3.1 Text Keypad. The current text keypad (shown in Figure 1) incorporates two important design decisions. First, it is based on an alphabetical arrangement of characters (although this could be replaced by a Dvorak or QWERTY arrangement. Quill and Biers (1993) recommend a QWERTY arrangement, but this is rejected for this prototype because the physically-challenged user is not likely to be familiar with it. Second, the key arrangement gives prominence to the <cursor control> keys due to their relative importance not only during text entry (see also Gould et al. (1985)) but also during text editing.

4.3.2 Dialogue Keypad. The dialogue keypad shown in Figure 2 contains the keys necessary to control dialogue boxes. There are fewer keys here than in the text keypad and therefore the individual keys are larger. The main keys are the <cursor control> keys, the <next item> and the <previous item> keys (corresponding to <tab> and <shift-tab> respectively) and the <escape> key.

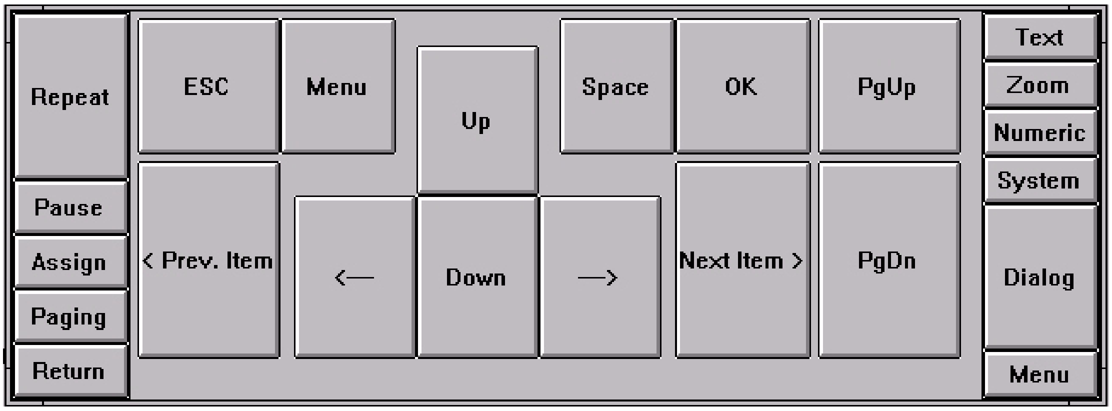


Figure 2*: Dialogue Keypad*

4.3.3 Menu Keypad. The *menu keypad* (not shown) is another context-sensitive keypad for interaction with the menu system. The main keys here are the <cursor control> keys, the <menu> key (corresponding to the <alt> key), the <escape> key and a key which contains the text of the currently highlighted menu item. Selecting this key selects the menu item (and is equivalent to pressing the <carriage return> key).

4.3.4 Zoompad. Many users may have difficulty in precisely positioning the cursor, and so a zoompad has been incorporated (Figure 3). The zoompad shown here emulates mouse commands. It incorporates the equivalent mouse command selection by means of radio buttons. The user looks at a region of the client window and after a brief dwell interval has expired, the region is copied and enlarged into the keypad. The user effects a click action within the zoom area and the event is sent to the corresponding part of the client window.

4.4 Overriding Automatic Keypad Selection

The visual keyboard automatically loads and displays keypads appropriate to the current task context as, for example, if the user accesses a menu when a dialogue box is displayed in the client application. However, this self-adaptation mechanism is not always desirable. For instance, if the first control object in the dialogue box is a text element, the user may prefer the text keypad for the initial activation. In such cases, the user can make explicit assignments using the *<Assign>* button in the left hand part of the keyboard system command menu. Subsequently, the assigned keypad will be loaded automatically when the client application context is the same.

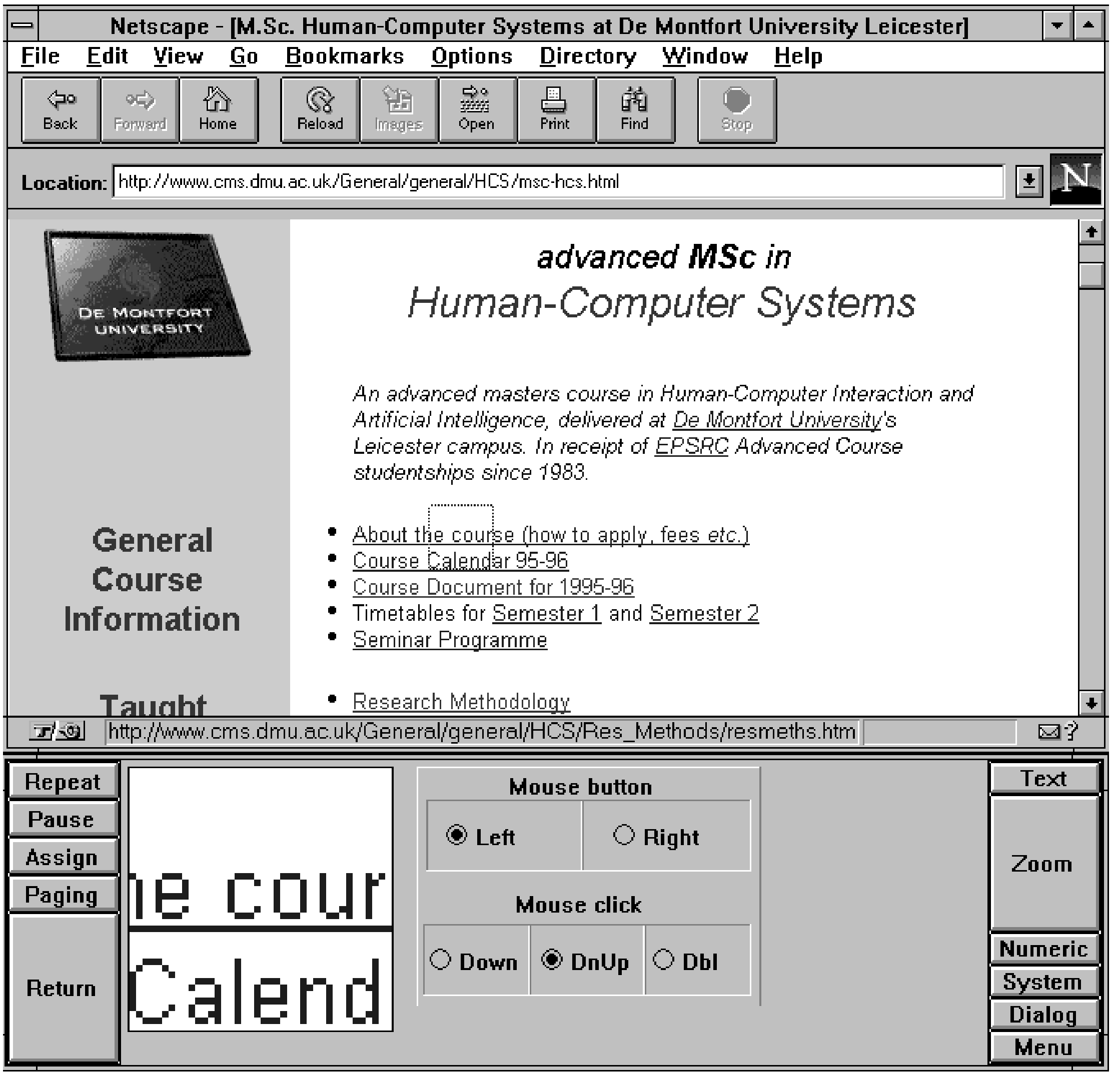


Figure 3: *Zoompad used together with Netscape*

4.5 Object Dwell Time

Dwell times for different types of keys can be set individually and these can be altered by the user to reflect personal preference. The values in table 1 show the different settings used as initial values in the evaluation trials (described in section 5). The user is warned about expiration of the dwell time by a change in the cursor. The mouse pointer is represented by a circular cursor within the keyboard and this changes to show a black spot in its centre just prior to the end of the dwell interval. If the user does not wish to select the key, they may look away at this point and selection is then inhibited.

Table 1: *Dwell time settings*

Dwell object	**Dwell time**
<*Cursor control*> keys	1000 ms
<*Enter*> key (menu keypad only)	2500 ms
System command menu	1000 ms
Keypad selection menu	500 ms
Target selection (*zoompad*)	2000 ms
Any other key	1500 ms

4.6 Window Arrangement

The window arrangement of the client application and visual keyboard is supported by three different approaches. In combination, these attempt to compensate for the fact that the on-screen keyboard has to take up a finite part of the available screen area, thereby reducing the area available for the application.

- *Paging:* In the case that the window of a client applications is partially overlapped by the visual keyboard or even requires the whole screen, the <paging> command moves the client application window so that either its top half or bottom half is displayed in the area of the screen above the keyboard window.
- *Heading:* Windows and in particular dialogue boxes which appear initially in the centre of the screen will be automatically moved to the top of the screen if there is overlap. In the case that the remaining screen space is too small, the <paging> command will be selected.
- *Bounding:* Re-sizing the visual keyboard will automatically re-size the client application window.

5. EVALUATION TRIALS

This section reports results from initial evaluation trials with a modern word processor (MS-Word 6.0). More evaluation work remains to be done, both with this application and with other types of application. The trials do, however, give a good indication of the success, or otherwise, of some of the design ideas that have been included in the keyboard. The input device was the eye-control system described previously (Istance and Howarth 1994).

5.1 Selection of Tasks

A set of five tasks, which constituted an integrated exercise, was completed by each subject:

1. Run application and open text document
2. Enter a few lines of text (an address)
3. Save the file
4. Edit the existing text
5. Require help about a specific problem

The tasks, and keystroke level actions required to execute them, are transferable to most word processors, and one objective of the evaluation trials was to judge whether the visual keyboard was effective enough to perform the specified tasks. In addition, in order to evaluate the efficiency with which the tasks were performed the effort required for the user to correct errors, or to edit text, served as a useful performance indicator.

5.2 Results and Discussion

All five (able-bodied) subjects were able to complete the tasks which were given to them. Moreover, all were able to recover from errors and consequently this usability objective of an effective visual keyboard was met in this context. The mean time spent on the completion of all tasks was 19 minutes per subject whereas the total time, including recalibration and repetition of tasks due to eyetracker problems, required on average 38 minutes. The task completion time of the whole task sequence did not vary across subjects by more than three minutes either side of the mean.

5.2.1 Text Entry and Feedback. On average, subjects were able to enter their address in about seven minutes (task 2) which corresponds to a text entry rate of only one word per minute. This is very inefficient and it was observed that most errors occurred during this task. Subjects frequently unintentionally entered a character twice and consequently the time spent on correcting those errors had a major impact on text entry efficiency. The reasons for this lay partly in the techniques used by subjects to get feedback, partly due to a lack of training, and partly due to a delay in updating cursor position.

Two subjects looked for feedback in the text document after each character was entered whereas the other subjects entered a sequence of characters before checking. When the subject's task completion times were compared, the first approach was found to be more efficient. This was because errors were recognised earlier and consequently the subjects required less cursor control keystrokes to return to the site of the error. However, one might expect reductions in key location and selection times in the second approach as the user's attention remained on the keyboard for longer periods of time. There is clearly room for improvement in the rate of text entry, however even the present rate may be acceptable if text entry is limited to a few letters, such as when entering a file name or a help item for which to search.

5.2.2 Interacting with Dialogue Boxes. During tasks 1 and 5, (see section 5.1) navigation problems occurred when interacting with dialogue boxes. In some cases, subjects mistook the control object with the input focus and so manipulated interface elements by keystrokes which were actually designated for a different object. The visual

 ISBN 0 7049 1140 X

feedback showing the current input focus was not always clear. These problems are caused by a lack of feedback when moving the input focus from one control to another. When an ordinary keyboard is used, it is possible to press a key and observe changes in the visual appearance of a control simultaneously. Using the eyes to 'press' a virtual key prevents the user from observing the effects of the command as it is taking place. However, during these tasks, unintentional key presses occurred far less frequently than with the text entry task, as one might expect given that the dialogue keypad had fewer, and therefore larger, keys.

5.2.3 Combinations of Keystrokes. The fourth task incorporated the use of shortcut keys for stepping from one word to the next in order to select and highlight text. Most subjects had no difficulties in combining two or three keys and hence were able to move the caret with shortcut keys rather than by a series of cursor control keystrokes.

5.2.4 Dwell as a Means of Activating Commands. Problems associated with interference between mouse events generated directly by the eyetracker software and similar events generated by the visual keyboard were apparent. These arose on occasions when subjects were reading a document or browsing a dialogue box. While subjects were aware that there is some form of response when dwelling too long within the visual keyboard area, the possibility of an event happening inside the client window was not apparent. Whilst it would be possible to disable all event generation in the client window by the eyetracking software, there is a case for letting the user interact directly with the client application (by the eyetracking software sending events directly to it) as well as by using the visual keyboard. The advantage lies in greater flexibility and not always having to use the keyboard on-screen. The disadvantage lies in the possibility of generating unwanted events in the client window.

The feedback provided by the change in the cursor, which was intended to warn that dwell time was about to expire, was generally misinterpreted by subjects who thought that the change signified that the key had been pressed.

5.2.5 Feedback on Modifier Key States All subjects were observed to have difficulties in remembering the current state of the modifier keys, for example, forgetting that the <*Shift*> key was locked, and this issue needs addressing. If another keypad was selected after locking the modifier key, then feedback indicating that a subsequent keystroke would lead to a key combination was lost. Errors were also made even though the keypad remained visible, but the effect of the locked key was not obvious. For example, moving the carat with the cursor control keys to correct a typing error with the <Shift> key still locked resulted in the text becoming highlighted rather than the carat simply being moved.

5.2.6 Mechanisms for Changing Keypads. All subjects were able to operate the menu-based keypad selection mechanism, although frequently this required more than one attempt. This is acceptable because the keypad selection is based on a short dwell time (500 ms). Once the desired keypad button had been "acquired" subjects could easily keep the pointer within that button because of its large size. The use of expanding, 'fish-eye' buttons has been shown to be particularly successful for this type of eye-based interaction. Furthermore, the self-adaptation by the visual keyboard by loading the appropriate keypad automatically was also successful in reducing user input.

5.4.7 User Preferences for Interaction Styles with the Visual Keyboard. In general, the most preferred means of command selection was the use of the zoompad to select the smart icons of the toolbar, followed by menu item selection. The least preferred option was using shortcut keys and accelerators. Preferences appeared to depend on the expertise with MS-Word. For instance, one subject reported frequently using shortcut keys for text editing and preferred to use the keypad supporting use of the shortcut keys. A follow-up trial where all tasks, excepting the text entry task, were performed using the zoompad showed considerable increases in speed.

6. CONCLUSIONS

This work has demonstrated how a visual keyboard controlled by eye can be used to interact with standard software produced for the able-bodied user and thus allow the physically-challenged user to benefit from the wealth of software produced for modern GUI environments. This includes being able to access the Internet using existing browsers without the need for any modification either to the browser or to the keyboard.

Furthermore, many GUI applications require precise pointing ability which is often lacking in eye-controlled systems (Istance and Howarth, 1994) and the visual keyboard is capable of assisting with these tasks. Adequate solutions for non-keyboard sensitive GUI objects have been considered problematic in the past (Shein et al., 1991, 1992). The zoompad has been shown to provide an effective solution here.

Further work is required to improve the text entry rate achievable with this keyboard. Part of the problem here lies with the eyetracking system and its associated data processing software, rather than with the keyboard itself. A major problem was the need to correct unintentional keystrokes rather than the time required to locate and press an individual key. The next phase in the development will look closely at the causes of this and will examine means of improving text input rates. In addition, means of editing and interacting with existing text using the visual keyboard will be studied more closely. At the workplace, the visual keyboard is perhaps more likely to used for editing existing documents than for original document creation.

Future work with the visual keyboard will focus on the issue of feedback and examine ways of overcoming the visual separation between the keyboard and the region of client window from which the user receives feedback about the effects of commands. Additionally, it is intended to examine how direct eye-based interaction with the client application could be integrated with the use of the visual keyboard.

The major conclusion of this project is that the visual keyboard can be considered as a valuable low-cost enhancement to the eyetracker with the capability to compensate for its limitations as a pointing device. It has demonstrated that effective interaction with standard modern software applications using eye-based interaction techniques is entirely possible. This will greatly enhance the possibilities for physically-challenged users to work with the same software products and on a more equitable basis with their able-bodied colleagues.

7. REFERENCES

Bishop, J.B. and Myers G.A. (1993) Development of an effective computer interface for persons with mobility impairment. *Proceedings of the Annual Conference on Engineering in Medicine and Biology,* **15 (3)** 1266-1267.

Demasco, P.W. and McKoy, K.F. (1992) Generating Text From Compressed Input: An Intelligent Interface for People with Severe Motor Impairments. C*ommunications of the ACM*, **35 (5)** 68-77.

Eulenberg, J.B., King, M.T. and Patterson, H. (1985) Eye-Controlled Communication. *Proceedings of Speech Technology '85 - Voice Input and Output Applications,* 210-211.

Friedman, M.B., Kiliany, G. and Dzmura, M. (1985) An Eye Gaze Controlled Keyboard. *Proceedings of the 2nd International Conference on Rehabilitation Engineering*, 446-447.

Friedman, M.B., KILIANY, G., Dzmura, M. and Anderson, D. (1981) The Eyetracker Communication System. *1st National Search for Applications to Aid the Handicapped,* John Hopkins University, 183-185.

Gould, J.D., Lewis, C. and Barnes, V. (1985) Effects of Cursor Speed on Text-Editing. *Proceedings of the CHI'85 Conference on Human Factors in Computing Systems*, 7-10.

Hutchinson, T.E., White, K.P., Martin, W.N., Reichert, K.C. and Frey, L.A. (1989) Human-Computer Interaction Using Eye-Gaze Input. *IEEE Transactions on Systems, Man, and Cybernetics,* **19 (6)** 1527-1534.

Istance, H.O. and Howarth, P.A. (1994) Keeping an Eye on your Interface: The Potential for Eye-Based Control of Graphical User Interfaces (GUI's). In Cackton, G., Draper, S.W. and Weir, G.R.S. (eds), *Proceedings of HCI'94: People and Computers IX,* 195-209.

Koester, H.H. and Levine, S.P. (1994) Modelling the Speed of Text Entry with a Word Prediction Interface. *IEEE Transactions on Rehabilitation Engineering,* **2 (3)** 177-187.

Murphy, R.A. and Basili, A. (1993) Developing the user-system interface for a communications system for ALS patients and others with severe neurological impairments. *Designing for Diversity; Proceedings of the Human Factors and Ergonomics Society,* **2,** 854-858.

Quill, L.L. and Biers, D.W. (1993) On-screen keyboards: which arrangements should be used? *Designing for Diversity; Proceedings of the Human Factors and Ergonomics Society,* **2,** 1142-1146.

Schofield, J.M. (1981) *Microcomputer-based Aids for the Disabled.* Samet, P.A. (ed.), London: Heyden.

Shein, F., Hamann, G., Brownlow, N., Treviranus, J., Milner, M. and Parnes, P. (1991) WIVIK: A Visual Keyboard for Windows 3.0. *Proceedings of the 14th Annual Conference of the Rehabilitation Engineering Society of North America (RESNA),* 160-162.

Shein, F., Treviranus, J., Hamann, G., Galvin, R., Parnes, P. and Milner, M. (1992) New directions in visual keyboards for graphical user interfaces. In Murphy, H.J.(ed.), *Proceedings of the 7th Annual Conference on Technology and Persons with Disabilities,* 465-469.

Ten Kate, J.H., Frietman, E.E.E., Stoel, F.J.M.L. and Willems, W. (1980) Eye-Controlled Communication Aids. *Medical Progress through Technology,* Springer-Verlag, **8,** 1-21.

White, K.P., Hutchinson, T.E. and Carley, J.M. (1993) Spatially Dynamic Calibration of an Eye-Tracking System. *IEEE Transactions on Systems, Man and Cybernetics,* **23 (4)** 1162-1168.

Memory processes and virtual environments: I can't remember what was there, but I can remember how I got there. Implications for people with disabilities.

E A Attree[1], B M Brooks[2], F D Rose[3], T K Andrews[4], A G Leadbetter[5] and B R Clifford[6]

Department of Psychology,
University of East London,
London, E15 4LZ, U.K.

[1]*attree@uel.ac.uk*, [2]*brooks2@uel.ac.uk*, [3]*rose2@uel.ac.uk*, [4]*tresa@uel.ac.uk*, [5]*tony@uel.ac.uk*, [6]*clifford@uel.ac.uk*

ABSTRACT

Memory deterioration is often a consequence of brain damage. Successful memory rehabilitation programmes depend upon effective methods of cognitive assessment. This paper considers the potential value of virtual reality (VR) in this context. The effect of active or passive participation in a virtual environment on specific aspects of memory was investigated. It was found that active participation enhanced memory for spatial layout whereas passive observation enhanced object memory. No differences were found for object location memory. These findings are discussed in terms of how VR may provide a means of measuring memory which combines ecological validity with experimental rigour.

Keywords: ecological validity, everyday memory, yoked-control, clinical assessment.

1. INTRODUCTION

> "VR is a computer-based technology which incorporates specialized input and output devices to allow the user to interact with and ... explore a three-dimensional virtual -or artificial - environment. In the virtual world the user can do things as routine as throwing a ball or as fantastic as flying through space. And these things can be made to occur by something as simple as a hand gesture or a nod ... (Middleton, 1992)

There are different levels, or degrees, of sensory immersion which can all be included under the headings of VR. For the sake of simplicity these are referred to as immersion at the space, object, or environment level (Larijani, 1994). Viewing a three-dimensional space, such as a tour of a virtual house before it has been built, is the most elementary level of virtual immersion, that is the space level. If the user can manipulate images or objects within that three-dimensional space then that is a deeper level of immersion, the object level. Finally, total immersion, where all references to the real world are blocked out and are replaced by substitute stimuli (visual, auditory, and tactile) provides immersion at the environment level.

These different degrees of sensory immersion offer a range of ways to provide a safe environment to enable training and rehabilitation of people with disabilities. The first conference on virtual reality and persons with disabilities was held in 1992, and in the intervening years a range of effective VR applications have been developed. These include, for example, a programme for children in special education which has been devised by the virtual reality group at Nottingham University (Brown & Wilson, 1995). This research group has developed a VR system to develop everyday living skills for children with severe learning difficulties. Applications have also been designed for adults with special needs. Mowafty & Pollack's (1995) *Train to Travel* is one such example. This project was devised to enable people with cognitive impairments to use public transport.

However, as Darrow (1995) pointed out, despite the wide range of applications for people with disabilities, the area of rehabilitation of people with specific cognitive disability is under-developed and under-researched. The potential for VR applications in the area of neurological rehabilitation has also been discussed by a number of other researchers (Rose and Johnson, 1994; Rose et al., 1996; Pugnetti et al., 1995). However, in order to develop successful rehabilitation programmes,

 ISBN 0 7049 1140 X

which address individual needs, there remains the need for effective assessment tools. The majority of existing clinically-based assessments of cognitive functioning have internal validity, but may lack both external and ecological validity.

Given that generalisation, for example across settings, is of vital importance in rehabilitation, programmes, then the generalisation of cognitive functioning assessments is also of fundamental importance.

One particular deficit, memory, is characteristic of many medical conditions - especially head injury and neurodegenerative diseases. The debate into how best to assess memory impairments has followed that of mainstream memory research. Cognitive psychologists have placed an increasing emphasis upon the measurement everyday memory - following Neisser's (1978) argument that existing laboratory measures were too narrow and artificial. However, ecologically valid measures have also been criticised, on the grounds that they lacked experimental rigour and control (Banaji and Crowder, 1989).

As previously stated, within the context of rehabilitation the assessment of everyday memory problems is important. Sunderland et al. (1983) investigated the extent to which laboratory tests could predict everyday memory. They found a selective relationship between objective and everyday memory assessments, indicating the need for ecologically valid assessment tools. Wilson (1987) also pointed out that, until the Rivermead Behavioural Memory Test (RBMT) was developed, standard memory batteries failed to identify everyday memory problems.

Andrews et al. (1995) have suggested that measuring memory within a virtual world may be a solution to the problems suggested by Banaji and Crowder. The advantage of using VR is in allowing the assessment of memory to be carried out in the context of interaction with an everyday environment without sacrificing analytic control. Furthermore, precise assessments can be made of the way in which memory and other cognitive processes operate in a multifaceted reality.

While VR environments can be designed to assess a range of memory processes, VR would seem an ideal medium within which assessments of spatial memory can be made. In keeping with the dearth of VR development in the area of cognition, few studies have been carried out to measure spatial knowledge. The acquisition of spatial knowledge is dependent upon navigational experience (Thorndyke and Hayes-Roth, 1982). Thus, passive observation, rather than active experience, may lead to the limited cognitive processing of spatial information.

VR studies in this area have found equivocal results. For example, Wilson (1993), in a series of experiments, found that both active and passive participants obtained similar levels of spatial information from virtual environments. The latter finding differed from that of Peruch, Vercher and Gauthier (1995) who found that spatial memory was superior for the active participants rather than the passive.

However, "spatial knowledge" is a somewhat broad term and the question remains as to what specific aspects of the environment a passive observer can actually encode. The present paper investigated the effects of active or passive participation in a virtual environment on spatial and object memory. It was predicted (following Thorndyke and Hayes-Roth's argument) that spatial memory would benefit from active VR participation, whereas passive observation would allow undivided attention to be focused on the objects and facilitate memory for those items. A prediction with respect to object location is uncertain, following the above assumptions, as both the active and passive conditions could contribute to enhanced object location memory.

2. STUDY 1

2.1 Method

Participants were fourteen female and sixteen male students (mean age 24 years). The study utilised a three-dimensional environment created on an IBM compatible 486dx2, using the Superscape™ modelling software package. This environment comprised of four inter-connected rooms in a house. 25 objects were situated in these rooms. Exploration of these rooms, which were displayed on a 15 SVGA inch monitor, was controlled by using a Gravis proportional joystick. In an adjoining cubicle another identical monitor was "yoked" to the original computer so that the two screens were identical at all times.

Participants were randomly allocated to the active or passive experimental condition, and tested in "yoked" pairs. The members of each pair of participants sat in adjoining cubicles in front of identical "yoked" computer screens. All participants were naive to the fact that they were "yoked". One of the pair negotiated their way through the virtual house while the other was exposed to exactly the same journey but had no control over the interaction. Both active and passive participants were asked to search the objects that they would see in the virtual environment and try to "find an umbrella which may or may not be there". Immediately each active participant had completed a route through the virtual reality rooms, the computer screens were switched off and both members of the pairs were instructed to draw the layout of the rooms in the virtual reality display (spatial memory test). Participants were then given a plan of the correct layout of the virtual reality rooms, with the location and name of one object marked in each room and asked to draw in the correct locations and write the names of any objects that they could recall from the virtual environment (object memory test). Finally, participants were given one minute to study the Rey-Osterreth Visual Memory test figure and five minutes to draw the figure from memory, in order to check whether active and passive participants differed in visual memory.

2.2 Results

Prior to the data being analysed a check was carried out to ensure that the participants did not differ in respect of visual memory scores. There was no significant difference in performance on the Rey-Osterreth Visual Memory test between the active and passive participant pairs [t(14)=0.26, p=0.80].

Spatial layout was scored on a pre-determined criterion which allocated marks according to number and shape of rooms, position of doorways and corridors, and the direction in which doors opened. Data were converted into probability scores so that spatial location and object memory could be compared directly.

It appears from Figure 1 that active participants performed better in the spatial memory task than passive participants (probabilities 0.62 and 0.43 respectively), whereas passive participants performed marginally better than active participants in the object memory task (0.29 and 0.21 respectively).

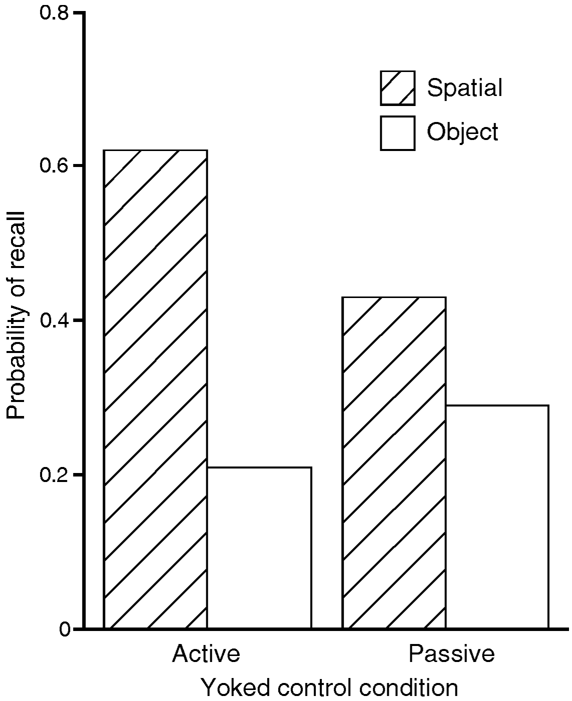


Figure 1. *Mean probability scores for spatial and object memory.*

Statistical analysis supported this interpretation of the data. A 2 x 2 repeated measures[1] analysis of variance (ANOVA) revealed a significant effect of spatial vs object memory [F(1,14)=67.04, p<0.001]. There was no significant difference in overall performance between the active and passive groups [F(1,14)=1.82, p>0.05] but, as expected, there was a significant interaction between spatial vs object memory and the active vs passive groups [F(1,14)=24.24, p<0.001][2]. Subsequent analysis, using a Bonferroni adjustment to control for the familywise error rate, demonstrated that the active group performed better in the spatial memory task (p=0.01), but the passive group's superior performance on the object memory task did not reach significance (p=0.09).

[1] The yoked control data were analysed using a 2 x 2 repeated measures ANOVA following Schweight's (1994) stipulation that "the data from a yoked design are analysed by using procedures appropriate for a within subjects design because matching subjects on a specific variable [i.e. time] relates them".

[2] A between subjects re-analysis of the yoked control condition made no difference to the overall results: a significant effect of spatial vs object memory (F[1,28]=80.95, p<0.001), no significant difference in overall performance between the active and passive groups (F[1,28]=1.18, p>0.05), and a significant interaction between spatial vs object memory and the active vs passive groups (F[1,28]=19.21, p,0.001).

A possible two marks were awarded for the correct location of each object recalled, one mark for the correct room and one for the correct location within that room. (The correct location within a room was measured by dividing each room into sextants and requiring that each object was located in the correct sextant.) Since the correct location of objects was dependent upon prior recall of objects, scores were proportionately adjusted (by dividing by the number of objects recalled) and analysed separately. This analysis showed no significant difference between the active and passive conditions [$t(14)=0.27$, $p>0.05$]. Enhanced memory for spatial layout did not, therefore, extend to remembering the correct location of objects. Neither did enhanced object memory extend to remembering object location.

3. EXPERIMENT 2

Experiment 2 was carried out to investigate whether object memory would be superior in the passive condition using a more sensitive measure of memory performance. Therefore a recognition task was used to measure object memory.

3.1 Method

Twenty-two female and eight male students (mean age 25.3 years) participated in the study. The same virtual house and objects, as used in experiment 1, were used in this experiment. Once again participants were naive as to the fact that they were "yoked", as before they were tested in yoked pairs, in adjoining rooms, with one of the pair negotiating their way through the virtual house while the other was exposed to exactly the same journey but had no control over the interaction. Participants were subsequently given a picture recognition task which comprised cards depicting the 25 objects in the virtual environment and 25 schema-relevant distractors.

3.2. Results

Participants in the active condition recognised fewer objects than participants in the passive condition (probabilities .35 and .45 respectively). Statistical analysis showed this difference to be significant [$t(14)=2.21$, $p<0.05$]. The results of Experiment 2 therefore demonstrated that object memory was better in the passive than in the active condition. Thus, using a recognition test, to provide a more sensitive measure of memory it has been found that object memory is superior in the passive condition.

4. GENERAL DISCUSSION

The main findings of the present studies were that active participants had superior recall of the spatial layout of the virtual environment, whereas passive participants had superior object memory. There was no difference between the active and passive conditions for remembering the location of objects.

The dissociation between spatial and object memory, produced by VR participation, illustrates that negotiating the virtual environment only enhanced the specific memory ability tapped by the task. Active participation via interacting with the spatial environment aided spatial memory but mot object memory. conversely, recognition memory for objects was better in the passive condition, perhaps because these participants were able to concentrate more on the task because they were not interacting with the environment. That is, unlike the active participants, they were not engaged in a divided attention task.

No differences between the groups were found in memory for object location, possibly because active and passive participation may have produced different trade-offs for the two groups. The active participants location memory was expected to be enhanced because they were actively discovering the objects and their relationships. However, the lack of enhancement for this type of memory may have been due to the fact that the active participants were carrying out a divided attention task.

Enhanced spatial memory in the active condition confirms the findings of Peruch et al (1995), who found that spatial memory was superior for their active participants relative to their passive participants. However, the present findings are contrary to those of Wilson (1993), who found no differences in spatial memory between active and passive participants. The different task requirements between these two studies may account for this discrepancy.

There are a number of implications of these findings, for the use of virtual environments for carrying out cognitive assessments. For example, this study has demonstrated how a VR setting can provide the means for the measurement of incidental memory - an aspect of memory which is part of real-life experience. To further investigate the ecological validity of VR assessment comparisons could be made between measures carried out in a real-world setting to that of the same setting constructed within VR.

Another implication of these findings is that visual experience is not the only factor in remembering aspects of a visual medium; passive viewing of an environment leads to different types of memory than that of active exploration - factors that may be important when assessing the internal validity of the test situation.

Clearly there is a need for further research in order to make use of VR technology as a means of an assessment. For example, investigating whether changing the instructions to participants to tap into intentional memory processes would remove the distinction between passive and active participation memory differences. Furthermore, other the assessment of other cognitive functions, such as attention and perception, lend themselves measurement within a virtual environment. Finally, assessments could be made of the types and intensities of environmental factors which could cause cognitive overload, and, thus, affect everyday functioning.

5. REFERENCES

M R Banaji and R C Crowder (1989), The Bankruptcy of Everyday Memory, *Am. Psychol.*, **44**, pp1185-1193.

D Brown and J Wilson (1995), LIVE: Learning in virtual environments, *Abilility*, **15,** pp24-25.

M S Darrow (1995), Virtual reality's increasing potential for meeting needs of persons with disabilities: What about cognitive impairments? In *Proc.: Third Intnl. conf. "Virtual Reality and Persons With Disabilities"*, (H J Murphy, Ed.), San Francisco.

L C Larijani (1994), *The virtual reality primer*, McGraw-Hill, New York.

T Middleton (1992), Matching Virtual Solutions to Special Needs, *Proc., Virtual Reality and Persons with Disabilities,* Los Angeles.

L Mowafty and J Pollack (1995), Train to travel, *Ability*, **15**, pp18-20.

U Neisser (1978) Memory: what are the important questions? In *Practical Aspects of Memory* (M M Gruneberg, P E Morris and R N Sykes, Eds.), Academic Press, London, pp3-24.

P Peruch, J L Vercher and G M Gauthier (1995), Acquisition of spatial knowledge through visual exploration of simulated environments, *Ecol. Psychol.*, **7,** 1, pp1-20.

L Pugnetti, L Mendozzi, A Motta, A M Cattaneo and E Barbeiri (1995), Evaluation and retraining of adult's cognitive impairments: which role for virtual reality technology?, *Computers Biol. Med.*, **25**, pp213- 227.

F D Rose and D A Johnson (1994), Virtual reality in brain damage rehabilitation, *Med. Sci. Res.*, **22**, p82.

F D Rose, D A Johnson, E A Attree, A G Leadbetter, and T K Andrews (1996), Virtual reality in neurological rehabilitation, *Br. J. Ther. Rehab.*, **3**, pp223-228.

W A Schweigert (1994), *Research methods and statistics for psychology*, California, Brooks Cole, p236.

A Sunderland, J E Harris, A D Baddeley (1983), Do laboratory tests predict everyday memory? *J. Verb. Learn. Verb. Behav.,* **22**, pp341-357.

P W Thorndyke and B Hayes-Roth (1982), Differences in spatial knowledge acquired from maps and navigation, *Cog. Psychol.*, **14**, pp560-589.

B A Wilson (1987) *Rehabilitation of memory*, Guilford Press, New York, pp89-90.

P Wilson (1993), Nearly There, *Spec. Child.*, **68**, pp28-30.

Do virtual environments promote self-directed activity? A study of students with severe learning difficulties learning Makaton sign language

P J Standen[1] & H L Low

Department of Learning Disabilities,
University of Nottingham Medical School,
Queen's Medical Centre,
Nottingham, NG7 2UH, England

[1]*mvzpjs@nottingham.ac.uk*

ABSTRACT

Eighteen students with severe learning difficulties and their teachers were videoed while using an educational virtual environment. Teachers' activity was coded into eight categories (e.g. instruction, suggestion, pointing) and the students' into three (e.g. moves in three dimensional space) and intra-rater reliability established. Significant ($p<0.0001$) decreases in rate over repeated sessions was found for all the teacher's categories with the more didactic (e.g. instruction and physical guidance) decreasing at a faster rate than suggestion and pointing. For the students' categories there was a significant increase ($p<0.05$) in rate between the first and last sessions for two of the three categories.

Keywords: self-directed activity, learning difficulties, special education, Makaton

1. INTRODUCTION

A principal claim for the value of virtual environments (VE) in education is that they encourage self-directed activity and give the student some control over the learning process (Pantelidis, 1993). Self-directed activity is given a high priority in the theories of Bruner (1968), Vygotsky (1978) and Piaget (see Wood 1988). Research on the development of perception (Held and Hein, 1963) also emphasised the essential contribution of self-directed activity.

The value of self-directed activity may be even greater for people with learning disabilities. As with all people with disabilities, children with learning difficulties have less control over things that other children take for granted and assume they will play a passive role all the time (Sims, 1994). They are often described as showing underdeveloped autonomy and self determination. This makes it difficult for them to make informed choices, for example where issues of health care are involved (Welsh Health Planning Forum, 1992). Around 20 people in every thousand have mild or moderate learning disabilities, and about three or four per thousand have severe learning disabilities and need frequent support with some aspect of daily living. With the continuing expansion of "care in the community" those currently in receipt of special education will need a set of skills to enable them to achieve some degree of independent living. The use of VE may be appropriate (Cromby et al., 1996).

Introduction of new technologies into the classroom has frequently led to predictions that they will revolutionise education. All too often these claims are not realised. According to Salem Darrow (1996), in order to avoid this happening with the introduction of VE, educators need to take a proactive planning stance in the growth of this important technology rather than the reactive stance many have taken with other educational technology developments up until now. "If educators want VE to meet learning needs especially of those students who have unusual learning needs, they must play an active role in the development of applications offering to developers their unique understanding of learning styles and good teaching practices" (Powers and Darrow, 1994).

While the media has portrayed VE as being experienced through the use of head-mounted displays, more suitable for educational use are desktop virtual environments, which can be run on ordinary computer monitors and use standard computer input devices such as a joystick, mouse, touch-screen or keyboard. They are just as able to generate a sense of "presence", i.e. the experience of feeling oneself to be present in the virtual environment and have the following advantages over the use of head-mounted displays or total immersion systems.

 ISBN 0 7049 1140 X

1. They avoid possible health and safety issues. Although currently under investigation in Nottingham, head-mounted displays have been reported to cause nausea and dizziness and produce short-term impairments in visual acuity.
2. They facilitate peer and tutor interaction due to the public nature of the display. By contrast, head-mounted displays make the user look and feel isolated. The social nature of learning and the role of instruction were emphasised by Vygotsky (1978) in his use of the concept of the “zone of proximal development”. This is defined as the distance between a child’s “actual developmental level as determined by independent problem solving” and the higher level of “potential development as determined through problem solving under adult guidance or in collaboration with more capable peers”. Bruner (1968) has also drawn attention to the social context out of which skills develop, highlighting the value of elements such as joint attention, shared activities and sensitive and responsive adults prepared to assign meaning to the child’s behaviour.
3. Desktop virtual environments can be run on an IBM-compatible computer costing less than £1,000, placing them within the budget of most schools and training centres. Although headmounted displays costing around £600 have recently become available in this country their image quality is poor. With total immersion system there are likely to be significant time lags in responding to the user’s movements and tactile interaction with the computer images is clumsy and imprecise. A sufficiently robust system of good enough quality for use in training would currently costs upwards of £25,000

The present study set out to investigate the role of VE in special education from the educator’s point of view. Specifically it looked at whether VE did promote self-directed activity, or whether teachers were using them in a more conventionally didactic manner.

2. METHOD

2.1 Participants

Eighteen teacher-student pairs took part. The students were all attending a large school for those with severe learning difficulties aged between three and nineteen years. Teachers volunteered to take part in the study after participating in a workshop on the use of virtual environments run by the research team. They were asked to nominate a student to work with who had sufficient motor skills and visual ability to be able to use the computer terminal, joystick and mouse and who had no experience of using virtual environments. Half of the students nominated were male. Seven were from the senior school and aged between 13 and 15 years, three were from the middle school aged from 11 to 13 years, five were from the junior school aged from six to eight years, and three were from the nursery and aged from three to five years. Before testing began, teachers completed the Vineland Adaptive Behaviour communication sub-domains of receptive and expressive behaviour. The expressive section assesses what the child says and includes items such as “ says at least 50 recognisable words”, “uses phrases containing a noun and a verb”. The receptive section assesses what the child understands and contains items such as “demonstrates understanding of the meaning of ‘no’ “, “follows instructions requiring an action and an object”. The students’ mean scores were 14.4 for receptive and 21.2 for expressive communication. Compared with age-group norms this indicates levels of low to adequate communicative ability.

2.2 Equipment

The VE used in this study were built to teach Makaton, a vocabulary developed for learning disabled hearing impaired adults which has since been adapted for children who have little or no expressive speech and poor comprehension. They were displayed on a desktop system, with movement controlled by a joystick and interaction by a mouse. They consisted of a series of rooms, each devoted to the learning of one symbol (e.g. telephone). Each room contained four examples of the relevant object (e.g. four different types of telephone). When the user investigated the objects they activated the appearance of the relevant written symbol, and the signing of that symbol by a mannikin at one side of the screen. For each set of four rooms, a fifth room was available where the learner could test their knowledge of the signs and symbols demonstrated in the other four rooms.

2.3 Procedure

Each pair had between four and ten twice-weekly sessions using the Makaton programme. It was not always possible to keep the same teacher with each student but 11 students remained with the same teacher throughout. The order in which the student proceeded through the programme, the number of rooms that they explored, the number of times each room was explored and the length of session (up to a maximum of 20 minutes) were left entirely to the student or teacher. Each session was recorded on videotape. Each student’s knowledge of Makaton was assessed at the beginning of the study and after their last session using 16 Makaton symbol cards which corresponded with those that were in the software. Any form of recognition of the symbol in terms of speech, signing or pointing to the physical object in the room was taken as positive.

2.4 Analysis

After repeated viewing of the tapes, teachers' activity was coded into eight categories:

1. Non 3D Instruction: any verbal instruction given by the teacher to direct child's attention back to the screen or to make a movement which was not through three-dimensional space, e.g. move the arrow to the phone, click the mouse, look at the screen, press this button.
2. Non 3D Suggestion: any verbal suggestion (can be in the form of question) given by the teacher to make a movement which was not through three-dimensional space. Unlike point one above, it does not tell the student how to accomplish a task but simply prompts the student, e.g. let's see what the man does, would you like to turn the pages of this book? what do you think the doors can do? shall we have a look at another room?
3. 3D Instruction: any verbal instruction given by the teacher to make a movement through three-dimensional space, e.g. find the pencils, move to the right, press this button on the joystick to bring yourself up.
4. 3D Suggestion: any verbal suggestion (can be in the form of a question) by the teacher to make a movement through three-dimensional space, e.g. would you like to move closer? Would you like to see what else is in the room?
5. Pointing: this is when the teacher points to the screen to direct the student's attention or to instruct the student to move the arrow from one place to another. It often accompanies other behaviours (e.g. a question) but will be scored as a separate category.
6. Physical Guidance: teacher physically guides the student in using the joystick, mouse or keyboard, e.g. puts hand over student's hand to move mouse. As with pointing often accompanied by suggestion or instruction.
7. Teaching Questions: questions and comments made by teacher that give meaning to student's experience. Could include comments made to put student's actions into context, e.g. everyone's gone out, maybe that's why the doors are shut; what is that? what colour are the pencils?
8. Teacher's Move: teacher moves mouse or joystick or uses keyboard.

Students' activity was coded into three categories:

9. Spontaneous 3D Movement: any movement made by student using joystick.
10. Spontaneous Non 3D Movement: any movement made by student using mouse or keyboard.
11. Student's Initiative Completed by Teacher: student starts an action which is completed by teacher.

Behaviour was recorded from the videotape as frequencies, i.e. duration was not taken into account, purely the number of times each category occurred. As sessions differed in length, frequencies of behaviour categories were converted to rates (frequency of category divided by duration of session).

3. RESULTS

Intra-rater reliability was assessed by the same rater repeating the coding of ten randomly chosen sessions after an interval of one week. All eleven categories were significantly correlated (Spearman's rank correlation) between the two sessions but category seven (teaching questions) was significantly different on the t-test so this was omitted from further analysis. As very few pairs completed more than seven sessions, analysis was carried out on these only.

3.1 Teachers' Behaviours

Using regression analysis, significant decreases in rate over repeated sessions was found for all of the teacher's behaviours with the more didactic categories (e.g. instruction and physical guidance) decreasing at a faster rate than suggestion and pointing (see Table 1).

Table 1. *Results of regression analysis for teachers' behaviours*

Behaviour category	R2	probability
non 3D instruction	0.17	p<0.0001
non 3D suggestion	0.15	p<0.00001
3D instruction	0.03	p<0.00001
3D suggestion	0.19	p<0.0001
pointing	0.11	p<0.0001
physical guidance	0.16	p<0.00001
teacher's move	0.13	p<0.00001

3.2 Students' Behaviours

Regression analysis could not be used for the students' results because the data were slightly skewed and so rates for the first and last sessions were compared using a Wilcoxon one tail test. A significant increase in rate was found for spontaneous 3D movement (p<0.03) and spontaneous non 3D movement (p<0.04) but rates remained the same over sessions for Student's Initiative Completed by Teacher.

If a composite score is formed for all teacher categories and then all student categories and these figures are plotted against session, teacher activity can be seen to decrease as student activity increases (Fig 1).

Figure 1. *Graph of medians of behaviour rates against sessions*
Comparison between teachers and children's behaviours

3.3 Knowledge of Makaton

Using a Wilcoxon test, a significant (p<0.03) increase was found in the number of Makaton symbols the students knew at the end of the sessions when compared with their baseline scores.

4. DISCUSSION

These results showed that students made more self-initiated actions as sessions progressed. The only student category which showed no increase was Student's Initiative Completed by Teacher for which there were low rates throughout. This supports the often made claim for virtual environments in education: that they promote self-directed learning (Pantelidis, 1993). The increase in student activity was in 3 dimensional moves as well as non 3 dimensional moves showing that they are taking advantage of the characteristic of virtual environments that distinguishes them from other interactive packages.

Teachers contributed less as sessions progressed as shown by the composite score formed from all the teacher behaviour categories. However this drop was not as great as the increase in rates of student behaviour. This is because some behaviours (both categories of instruction, teacher's move, physical guidance), dropped at a faster rate whereas others (suggestion, pointing) hardly changed. The interpretation of this could be that teachers are not just becoming fatigued but selectively dropping the more didactic and controlling behaviours.

This can be explained with reference to the term "scaffolding" identified by Wood (1976) as one of the functions of tutoring. When a beginner starts to learn a task help (a scaffold) is provided to enable the beginner to make progress by controlling those elements of the task that are initially beyond the beginner's capability. As the beginner becomes more familiar with elements of the task and develops the ability to carry it out independently the tutor intervenes less, i.e. the scaffolding or training support is removed little by little. Another function of the tutor according to Wood (1976) is to maintain the learner's interest and motivation, marking relevant features of the task and interpreting discrepancies between the child's productions and correct solutions and controlling the level of frustration experienced by the learner. This is represented in the present study by the categories of pointing and suggestion which decreased at a slower rate.

To be able to interpret more accurately the role of the teachers' behaviour a sequential analysis would need to be carried out to demonstrate what happened after a teacher made a particular action. This would throw more light on whether teachers reduced their rate of behaviour in response to increasing initiative on the part of the student or whether the increase in student activity occurred after the teachers deliberately reduced their own. Although more informative, this approach is obviously more time consuming. However the method used in the present study is simple enough to be usefully employed to evaluate other educational virtual environments.

5. REFERENCES

J S Bruner (1968) Processes of Cognitive Growth: Infancy, Clark University Press, U.S.A.

J J Cromby, P J Standen and D J Brown (1996) The potentials of virtual environments in the education and training of people with learning disabilities, J. Intell. Disab. Res. (in press).

R Held and A V Hein (1963) Movement-produced stimulation in the development of visually guided behaviour. J. Comp. Physiol. Psychol, 56, pp. 872-876.

V S Pantelidis (1993) Virtual reality in the classroom, Educ. Technol, April, pp. 23-27.

D A Powers and M Darrow (1994) Special education and virtual reality: challenges and possibilities, J. Res. Comp. Educ, 27, pp 111-121.

M Salem-Darrow (1996) Virtual reality's increasing potential for meeting needs of persons with disabilities: what about cognitive impairments? Proc. Third Intl. Conf. Virtual Reality and Persons with Disabilities (H J Murphy, Ed), Northridge CA: California State University Center on Disabilities (in press)

D Sims (1994) Multimedia camp empowers disabled kids, IEEE Computer Graphics and Applications, January, pp. 13-14.

L S Vygotsky (1978) Mind in Society: the Development of Higher Mental Processes Harvard University Press, Cambridge, Mass.

Welsh Health Planning Forum (1992) Protocol for Investment in Health Gain: Mental Handicap, Welsh Office, Cardiff.

D J Wood (1976) cited in: J D Day, L A Cordon and M L Kerwin (1989) Informal instruction and development of cognitive skills: a review and critique of research Cognitive Strategy Research: From Basic to Educational Applications (C B McCormick, G E Miller & M Pressley, Eds), Springer-Verlag, New York, pp. 84-85.

D J Wood (1988) How Children Think and Learn, Blackwells, Oxford, U.K.

ECDVRAT '96

Technology

Day 2: 9 July, 1996, Afternoon

A virtual reality training tool for the arthroscopic treatment of knee disabilities

R J Hollands[1] and E A Trowbridge[2]

[1]Department of Automatic Control & Systems Engineering, University of Sheffield,
Mappin Street, Sheffield, UK

[2]Department of Medical Physics, Royal Hallamshire Hospital, University of Sheffield
Glossop Road, Sheffield, UK

[1]*r.hollands@sheffield.ac.uk*, [2]*e.a.trowbridge@sheffield.ac.uk*

ABSTRACT

Knee injuries are a common form of disability, but many can be treated using surgery. The minimally invasive approach of arthroscopy means faster recover times for the patient than when using open surgery, but the skills required by the surgeon are radically different. Although a number of arthroscopic training techniques are available all have problems either with cost, maintenance or availability. Through the medium of virtual reality, a computer based system can recreate the three dimensional geometry inside the knee, and allow the trainee surgeon to practice on it using replica instruments. In order to provide this facility at a feasible price, this paper describes work under way at Sheffield to develop a PC-based virtual reality arthroscopic training system. The resulting trainer has been tested by surgeons, and despite compromises made to accommodate the PC-platform, has been found to be extremely realistic at replicating some of the standard tasks of arthroscopy.

Keywords: surgery, training, arthroscopy, low-cost, PC

1. INTRODUCTION

Although extremely common, knee injuries can cause a great deal of distress and debilitation. In the case of professional sports people, an uncorrected knee injury can mean loss of livelihood for a time or the end to a career. In the case of ruptured ligaments, leading to knee instability, physiotherapy can often be employed to train the surrounding muscle structures to support the knee. However, for maximum recovery of knee performance, the damaged ligament must be reconstructed. In other cases, it may be necessary to surgically inspect the interior structures of the knee to ascertain the exact nature of the injury, and there will often be no alternative but to correct the fault by surgery.

Examination and repair of knee disabilities can be performed using open surgery. However, there are many benefits to using the minimally invasive approach of arthroscopy. In addition to the benefits to both patient and health care provider of significantly reduced recovery time (Banta, 1993), arthroscopic techniques minimise loss of synovial fluid and allow many of the knee structures to be examined in a relatively natural state. However, as with other 'key-hole' type operations arthroscopy can demand skills in excess of their open-surgery counterparts.

The arthroscope is a small metal tube containing a series of optics, creating a light source at the tip of the tube, and relaying the image from the tip of the tube back along optical fibres to a solid state camera, which then displays the image on a monitor. Although it is possible to get 0° arthroscopes where the viewpoint looks directly in front of the arthroscope, most use a view set at 30° or 70° off the axis. This allows the surgeon to be able to look around the immediate vicinity of the arthroscope tip simply by rotating the arthroscope around its own axis. Unlike open surgery where a large incision is made to provide a working area, the arthroscope is inserted through a small incision, usually just below the patella. Because the large incision in open surgery does not restrict the surgeon, he can manipulate the tools in a simple intuitive manner while directly observing the result. In arthroscopic surgery, however, any extra tools are inserted through a second small incision in a different location on the knee to the arthroscope insertion. The surgeon must navigate the tool blindly through the knee structures to the location of the procedure. The camera must also be navigated to the correct location. The arthroscopic surgeon will rarely be watching either the physical knee or his own hands, but instead must observe the operation on the view relayed to the monitor. Unlike the intuitive manipulation of

 ISBN 0 7049 1140 X

open surgery, the 'key-hole' approach of arthroscopic surgery means that both tools and camera can only be manipulated by pivoting around the entry point while inserting further or withdrawing. The arthroscopic surgeon must therefore decipher a monitor image representing an arbitrary orientation pointing away from the direction of travel, navigate the tip of the rigid camera and tools through the labyrinth of structures in the knee, and perform the operation with two degrees of freedom less than in open surgery. Despite the complexity of the arthroscopic task, the procedure must be carried out with exact precision. It is easy to damage the sensitive components of the knee, and even slight scuffing of the articular surfaces will inevitably lead to arthritis in later life (Bamford, 1993). The learning curve for arthroscopic techniques is long, and to minimise the risk to real patients a number of training techniques are currently used to allow trainee surgeons to enter real surgery higher on the curve.

2. CURRENT TRAINING TECHNIQUES

Current training techniques for minimally invasive surgery encompass a number of different methods. For example triangulation, the ability to get both camera and tool to the same point from different entry points, can be learned by inserting the equipment through the side of a cardboard box. Simple procedures can also be initially practised using the same system, for example the skill of incision is often practised by peeling a grape within the box!

More realistic forms of training are afforded by the use of physical models and cadavers. Physical models recreate the main features of the anatomy of interest using synthetic materials. Simple physical models of the knee may not be much more than the tibia and femur held together by two bits of string representing the cruciate ligaments. Sophisticated physical knee models contain most of the ligaments and tendons available, the meniscii, patella, and surround them with artificial muscle and synthetic skin. Standard inspection and simple surgical procedures are practised on the physical models using a real arthroscope and tools. Although this allows familiarity with the standard operating equipment, it removes the equipment from use for real surgery, and also risks damage to the sensitive arthroscope in the early stages of training. In addition, physical models rarely replicate real anatomy or tissue properties. The models can be fragile in use and any damage to the internal components, whether accidental, or intentional as part of a practised procedure, requires that the model be disassembled, the damaged parts replaced and the model rebuilt. The constant need for repair and maintenance means that the use of physical models can be both expensive and time consuming.

Although cadavers provide a real joint to practice in, the internal structures can be significantly changed by the effects of death and preservation. Additionally, cadavers of healthy individuals are hard to come by, so most cadavers represent only a small percentile of the population. Cadavers are expensive to keep and the strict rules regarding their care makes their use for regular training troublesome.

One of the most successful methods for training surgeons in arthroscopic procedures is the use of live animals, usually pigs. However, various concerns about the welfare of animals in such training means that this technique is outlawed in the UK.

Since virtual reality allows the recreation of the three dimensional geometry of anatomy, and computer simulation techniques can be employed to model the dynamics within the body, it seems reasonable that a VR surgical trainer could provide a viable alternative to some of the existing training techniques.

3. VIRTUAL REALITY SURGICAL TRAINER

3.1 Current Surgical Simulators

Virtual reality trainers for traditional open surgery would rely on the trainee wearing instrumented gloves and a stereoscopic head mounted display. Apart from the problem of encumberment, head mounted display technology is not sufficiently advanced to produce an image of adequate clarity over a wide enough field of view to reproduce the normal working conditions of the surgeon. However, minimally invasive surgery displays the view of the operation directly onto a monitor, and is therefore is ideal for replication using desktop VR which replaces the video monitor with a computer monitor. Unlike open surgery, where the instruments may be manipulated over a substantial working area, minimally invasive surgery uses instruments inserted through a small incision, making computerised measurements and feedback on the instruments easier. Preliminary work has been carried out at a number of sites (Hon 1992; Coleman 1994) reproducing laparoscopic surgery. The virtual environment arises from monitoring the position of the laparoscope and other instruments in an artificial abdomen, displaying the appearance of the internal organs from the viewpoint of the laparoscope, and simulating the interaction of the surgical instruments with the internal organs, including the effect of gravity, blood flow etc. Despite the vast amounts of computer power lavished on these simulations, none can so far manage to include the complex particle flow models of blood release from incisions, and most surgeons find the dynamics of the internal organs vastly unrealistic.

Arthroscopic surgery has all of the problems of laparoscopic surgery but also has one other major complication - in addition to manipulation of the camera and instruments, a key part of arthroscopic surgery is manipulation of the limb itself. This movement is necessary to open up the joint for observation and so the eye/hand co-ordination required for minimally invasive surgery is compounded by additional proprioceptive and visual cues. This aspect does not yet appear to be addressed by other virtual reality arthroscopic simulators being developed (Logan, 1995; Ziegler, 1995) despite their use of high powered graphics workstations.

3.2 Arthroscopic Knee Surgery

The knee is one of the most complex joints in the human body. The two surfaces of the femoral condyles rotate and possibly slide on the surfaces of the tibial plateau, while the patella moves along the femoral groove. The ligaments limit the movement of the bones relative to each other, with four ligaments having specific restraining functions: the anterior cruciate ligament resists anterior subluxation of the tibia; the posterior cruciate ligament resists posterior subluxation; the medial collateral ligament resists abduction; and the lateral collateral ligament resists adduction, see figure 1. In addition to these limitations, the tibial-femoral interface is such that in full extension, the femur effectively locks into place on the tibia, after internal rotation of the tibia relative to the femur.

A standard inspection inserts the arthroscope slightly above the anterior horn of the lateral meniscus and close to the patellar tendon. The camera is passed up the intercondyle notch and examination starts at the apex of the suprapatellar pouch. The surface of the patella can be inspected by rotating the camera and withdrawing it slightly. Next the medial compartment is entered by turning the telescope downwards and flexing the knee. To examine the edge of the medial meniscus a lateral force must be applied to the patient's ankle with the knee in about 30° of flexion. This then also allows the camera to enter the lateral compartment and further inspection of the articular surface of the lateral condyle is achieved by both extending and flexing the knee. The postero-medial compartment can also be examined after flexing the knee to approximately 30°.

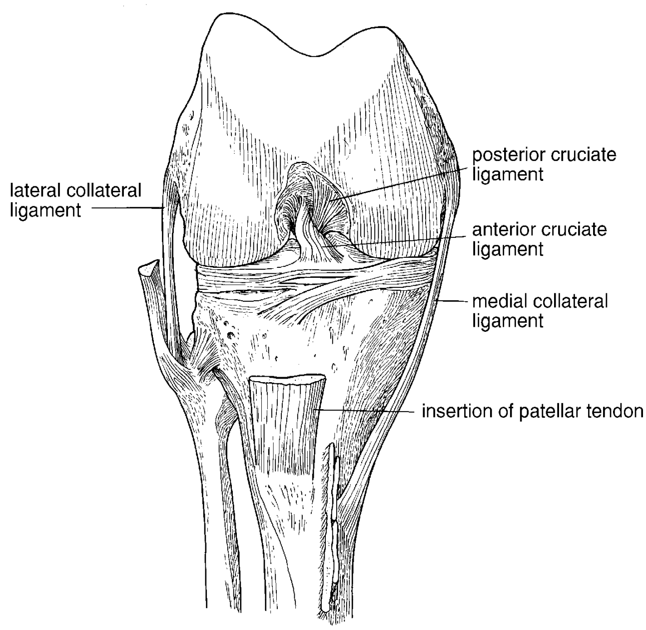


Figure 1. *The ligaments of the knee*

3.3 Arthroscopic Simulator Requirements

Eventually a fully functionally simulator is planned, able to represent many pathologies and patient types, and permitting the practise of diverse surgical procedures. However, initially the simulator must simply allow the trainee surgeon to perform a standard inspection similar to that summarised in section 3.2. Although fairly simple, a standard inspection requires that the trainee be able to recognise the major landmarks in the knee, triangulate both the arthroscope and a probe, and be able to move both around the compartments of the knee without damaging either the arthroscope or the knee itself.

If the resulting trainer is to stand any chance of being adopted as a commercial product, it must be priced at a level which would make it competitive with similar training technologies. The dynamic nature of the virtual environment in the arthroscopic trainer would normally require the use of high powered graphics supercomputers. However, to deliver the system at a competitive price point requires the use of a PC platform. The use of a PC is not as restrictive as it may first appear. The high-end Pentium PCs currently available have computational power at least equal to the lower cost graphics workstations, and there are an increasingly large variety of after-market graphics accelerators becoming available which deliver high-end workstation performance at a fraction of the price. However, careful design is still required to produce a convincing experience with the minimum of hardware. For this reason the approach taken is that the simulation need only *appear* to be correct as opposed to the high level of accuracy found in numerical models, which would be visually redundant.

Regardless of how complex the simulation of the environment may be, the effect would be completely lost if the trainee surgeon had to operate it in a different way to real surgery. Since all arthroscopic knee surgery inevitably requires the manipulation of the leg, a physical leg replica must be part of the system, although this need not be as complex as the physical models mentioned earlier, because all the internal components are modelled with the computer. Similarly, although only a synthetic arthroscope and tools are required in the simulator, they must represent their physical counterparts as much as possible.

Although the use of force-feedback is desirable in some other surgical simulators, the orthopaedic surgeons consulted did not think that it was significant for inclusion in the arthroscopic simulator. The leverage action of the tools around their entry points mean that, on the whole, only large forces would be felt by the surgeon. Given the sensitive nature of the knee's components, any action resulting in a force large enough to be felt would mean that the trainee had already made a mistake and a simple audio or visual cue would be sufficient feedback. However, the general feeling of the environment was thought to be important. Instruments should be weighted similarly to the real items, the artificial leg should be of the same heaviness, and the general damping action of moving the arthroscope and tools in muscle and other surrounding structures should be replicated.

The use of a computer as a mediator for the simulation means that extra facilities are available in excess of what would be with current training systems. Ideally the arthroscopic trainer could record the trainee's progress, provide multi-media tutorials and perhaps even automatic appraisal of the trainee's progress.

The virtual reality arthroscopic trainer currently being developed at Sheffield meets most of the above objectives. The following two sections describe the hardware and software components of the simulator.

4. SYSTEM HARDWARE

Because most of the features of the simulator are obtained using software, the amount of hardware has been reduced to a minimum. On the current system, the host computer is a standard entry-level Pentium equipped with 16Mb RAM and sound card. The graphics Matrox Millennium graphics accelerator card used is extremely low-cost, being intended primarily for computer games and multimedia, but nevertheless highly effective.

To create at least the same sense of realism as current training methods, an artificial leg and replica instruments are used. These allow the trainee surgeons to work into same way with the computer based trainer (figure 3) as they would with current physical models (figure 2).

Figure 2. *Physical model trainer* **Figure 3.** *Virtual reality trainer*

4.1 Artificial Instruments and Tracking

In order to allow the computer to track the position of the arthroscope and tools, these are fitted with receivers for the Polhemus Fasttrak electromagnetic tracking system. Although other tracking technologies are available, none were particularly suited for the task. Linkage based electromechanical trackers do not easily allow the use of replica instruments, and the connecting linkages can be obstructed by other parts of the system, and can also obstruct the surgeon in his task. There are some very good miniature 4 degree of freedom electromechanical trackers which are ideal for positioning over an entry point and tracking any instruments inserted, however the size and cost makes them inappropriate for use on the many possible entry points around the knee. Most other tracking technologies only measure orientation and not position, or require a constant line of site between the transmitter and receiver. The electromagnetic tracking system used does not require a clear line of site and is accurate to 0.03” RMS in measuring position and 0.15° RMS in measuring orientation. Because the instruments themselves are tracked, no obstructive hardware needs to be fitted to the artificial leg and multiple entry points can be easily accommodated. The small size of the receivers, about the size of the die, means that they can be easily mounted to artificial instruments. The disadvantage of using electromagnetic tracking is that the accuracy of the results can be adversely effected by the nearby presence of ferrous metals, or other magnetic fields. Even in ideal conditions, accuracy becomes worse if the receivers are further than 30 inches away from the transmitter. To minimise the effects of any environmental interference the transmitter unit is mounted inside the upper part of the artificial leg, just above the knee. Considerations of symmetry in the tracking mathematics means that at any time a receiver could be in either of two locations, and therefore a hemisphere of operation is declared for the tracker. Although the transmitter must be close to the knee for maximum accuracy, its position has to be carefully fixed to minimise the occurrences of an instrument being used behind it and generating an erroneous result.

The need to minimise ferrous metal in the environment also means that simple plastic replicas of the arthroscope and tools are used. Currently, the tools still need to be artificially weighted to a level similar to the real tools, but the shapes are identical to their real counterparts. This is especially important in the case of the arthroscope, where certain features of the tool, such as the insertion lug for the light source, are used by the surgeon to determine the roll orientation of the end optics. In addition to replacing the redundant real items, the plastic replica of the arthroscope and tools are cheap enough to be disposable in the event of breakage by unskilled handling, unlike their real counterparts.

4.2 Artificial Leg

The artificial leg is different from the physical models currently used in training, in that it has effectively no internal components. A simple leg shell can be used since all of the structures inside the knee are modelled within the computer. However arthroscopic surgery inevitable requires the manipulation of the lower leg to open up the various compartments of the knee for access and inspection. Although a real knee can be made to rotate and translate in a number of planes, the two key movements are flexion (bending forwards and backwards) and abduction (side-to-side). The amount of abduction available is dependant upon the degree of flexion, with no abduction possible at full extension. This movement is modelled in the artificial leg by having small cams on the flexion joint lock the joint in full extension, but permit increased abduction-adduction with increased flexion. To simplify design, the abduction-adduction is around a single axis in the centre of the joint, unlike the real case of rotating around the left or right femural condyl depending on current displacement. Additionally, flexion is physically modelled as a pure rotation, as opposed to the combination of rotation and translation in a real knee, The difference in the feel of the manipulation with this simplified mechanism is not thought to be a significant factor in operational reconstruction. The artificial leg can be easily weighted by using plaster of Paris in the lower, and the upper leg is clamped to a table, in a similar way to the surgeon strapping the real upper leg, or wedging it against a stop. With the current prototype, figure 4, there have been problems with the internal hinging components occasional obstructing the movement of the instruments within the knee cavity. Another prototype is currently under design with a redesigned mechanism to minimise obstruction, and also to include a thick synthetic skin, to provide the damping action of the muscle structures, and provide a more realistic entry point than the simple hole in an extension of the lower leg currently used.

Figure 4. *The artificial leg, showing hinge mechanism and entry point*

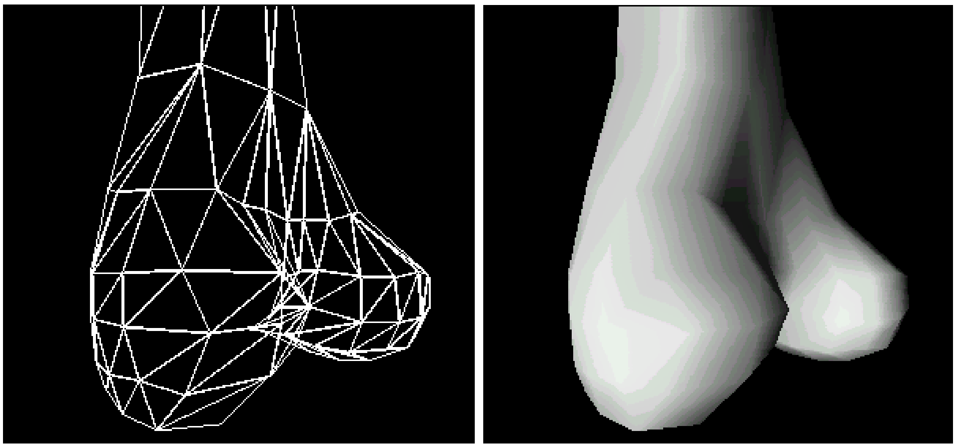
Currently, a third electromagnetic receiver is used on the lower leg to monitor flexion and abduction. However, since the leg can only be moved in two degrees of freedom, each of which could potentially be monitored using simple electromechanical trackers on the hinges, this system is liable to change in the near future. Although this adds another component to the hardware of the system, it also allows the use of a lower cost electromagnetic tracking system for the arthroscope and tool.

5. SYSTEM SOFTWARE

The VR toolkit used as the basis of the simulation is the Visual C++ based WorldToolkit from Sense8, running under Windows NT. The toolkit provides functions for most of the high level tasks of handling the peripheral devices, manipulating objects, detecting collisions etc. in addition to allowing access to lower level facilities of OpenGL.

5.1 Graphical Objects

The geometry of the knee components featured in the trainer were obtained by digitising plastic replicas using the 3Draw from Polhemus. Although this required the writing of digitising software for the product, it was thought the optimum route. Commercial data sets are available but tend to be expensive and overly detailed. Conversion of volumetric data from MRI and CT scans to surface data is also problematic, often produces inexact images, and is surrounded by difficulties caused by medical ethics considerations. Digitisation of the replicas by hand allowed customisation of the digital object, including having the edges of the polygons follow key contours on the real objects, and having increased polygon density around the main area of interest, the articular surfaces. It is essential to only have the minimum of polygons necessary to provide a sufficient representation of the object. Higher polygon counts mean slower rendering rates on the screen, can cause transformation delays when the objects are moved, and cause similar penalties when other polygon specific functions are used, e.g. texture mapping, smooth shading, collision detection etc. Although the model used has only a few hundred polygons in each object (figure 5a), the use of smooth shading makes the object look much more detailed (figure 5b). The low resolution of the data set is only revealed when looking at the object profile, and this is not thought to cause significant problems, although there may be a risk of trainee surgeons learning to use the discontinuities of the profile as landmarks, instead of the structures themselves.



a) Wireframe *b) Smooth shaded*

Figure 5. *Graphical object display*

5.2 Kinematic Simulation

The initial model of knee movement developed was an extension of the four bar link model used by O'Connor and Zavatsky (1993). This only models the two dimensional flexion-extension movement of the knee. The four links in the model are the two cruciate ligaments, assumed to always be under tension, and their two pairs of connection points, see figure 6. The two cruciate ligaments actually provide the limits to the flexion-extension movement of the knee, with the posterior cruciate ligament under maximum tension at full flexion, and the anterior cruciate ligament under maximum tension when the leg is straight. This model has been extended to provide a similar action to O'Connor and Zavatsky's model but with the cruciate ligaments undergoing changes in tension, and to provide the flexion dependant abduction-adduction limiting effects of the collateral ligaments (Hollands, 1995).

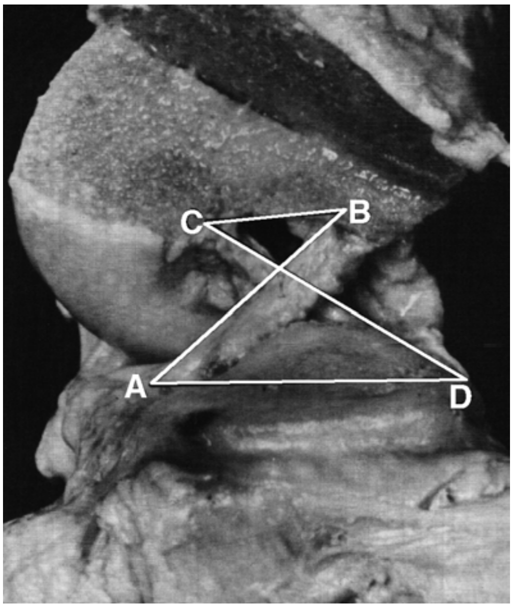


Figure 6. *Four bar link structure of the cruciate ligaments*

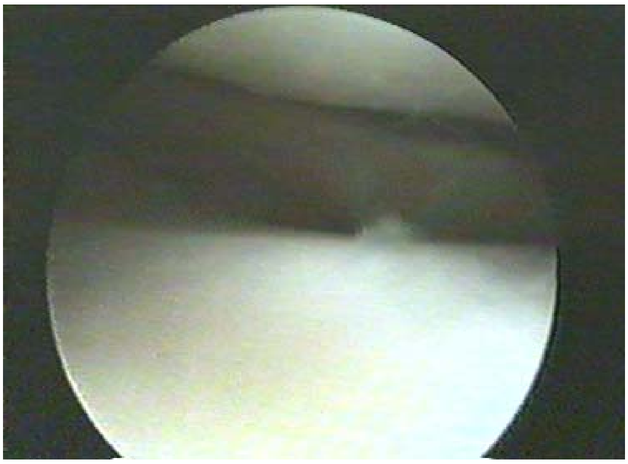
When the training system was being initially developed, the computing platform used then was unable to perform complex transformations at an acceptable rate, so for reasons of necessity the four bar link flexion movement was converted to a simple rotation around a point in the femural condyles. It was found by experimentation that a similar translation and rotation movement to the bar link model could be obtained by rotating around a point offset behind and above the centre of the femoral condyl. In the current model, the offset rotation has been retained, but the limiting action of the ligaments included, similar to those in the extended four bar link model.

Each cruciate and collateral ligament has fixed connection points on the femur and tibia. After any relative movement of the bones, the distance between the two end-points of each ligament can be calculated to determine

 ISBN 0 7049 1140 X

ligament tension and provided limitations of movement, as in the four bar link model. Since the collateral ligaments are not seen in normal arthroscopic surgery, only the cruciate ligaments are represented in the virtual joint. The ligaments must stretch and contract with joint manipulation, however the usual techniques for modelling deformable objects are numerically too intensive to keep a satisfactory update rate on the PC platform. Instead, a generic shape for each ligament is constructed and placed so that one end of the object coincides with its connection point on one of the articular surfaces. The ligament object is then rotated so that it lies along the vector linking the two connection points on each articular surface. The 'ligament' is then scaled by the amount necessary to cause the other free end to 'butt up' to its second connection point on the other bone surface. If the generic shape chosen is similar to a spiral, the scaling factor then causes the spiral to extend and contract in a manner which mimics the real ligament bundle.

Although the appearance of the objects within the display could be further enhanced with the use of more advanced shading techniques and extensive texture mapping, this would create an appearance far clearer than in real arthroscopic surgery, where the view is obscured both by the very narrow focal depth of the camera, and various bits of debris floating around in the saline solution being pumped though the joint. Figure 7 shows a snapshot from a real arthroscopic inspection, and figure 8 shows a comparative view from the virtual arthroscopic trainer. The use of attenuated light sources, scene fogging, visual noise and other techniques are currently under investigation in the Sheffield simulator to try to replicate the 'poor' image quality found in real systems.



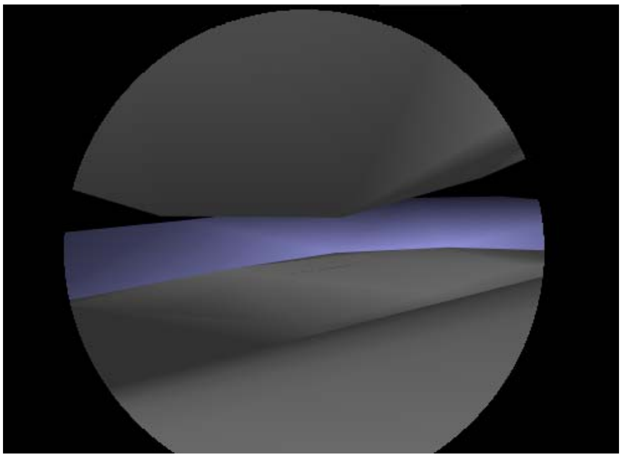


Figure 7. *Real arthroscopic view* **Figure 8.** *View from virtual arthroscopy trainer*

5.3 Under Investigation

In addition to the visual effects mentioned in section 5.2, a number of other features are also currently being experimented with. Although the knee does not contain a great number of deformable objects, some of the key components of the knee have deformable characteristics which need to be modelled even in a simple inspection. Although a number of techniques are available to model deformable objects, all are relatively computer intensive and their inclusion in the PC-based model would slow it down by an unacceptable amount. In the same way as a simple solution was found to create the illusion of deforming the ligaments, investigations are under way to devise simple solutions to recreate the highly constrained deformations found in the other knee components.

Collision detection is an integral part of the trainer. Constant checks must be made to monitor the position of the instruments within the knee and any collision with the internal structures is indicated by aural and visual cues. However, even with the low polygon density of the graphical objects used in the trainer, the time required to detect collision using a polygon intersection technique is far too long. The simple bounding boxes used for collision detection in large scale systems are inappropriate in the knee model because the irregular surfaces of the objects combined with their close proximity would yield constant collision flags. Voxel based collision detection of non-deformable objects have been reported as being very efficient in other surgical simulators (Logan, 1996), and their suitability for inclusion in the PC-based trainer is currently being investigated.

5.4 The Display

Obviously, the key feature of the display is the simulated view through the arthroscope. However, a number of additional displays are also available, see figure 9. A second view onto the knee environment can be requested showing an overview of the entire knee for orientation purposes. Unlike the primary window, whose viewpoint position is determined by the position of the synthetic arthroscope, the viewpoint in the overview window is controlled by standard mouse. There are facilities available for the user to record a session, and then play it back at a later time, and the orientation window then provides valuable information about what was actually happening to the arthroscope to produce the pictures seen in the primary window. In addition to the user being able to play back their own sessions, expert sessions can be examined for comparison.

A third window can show captured images from inside a real knee. A previous project (Hollands, 1995) examined the feasibility of developing a trainer only using a large database of captured images. It was determined that a prohibitive amount of space was required to store all the images necessary for a pseudo-continuous simulator. However, the database developed is used in the current simulator to provide a view representative of the view in a real knee which would be seen at, or near, the position of the synthetic arthroscope. As the synthetic arthroscope is moved around inside the knee, the image database constantly updates the window with the picture taken at the location closest to that of the current viewpoint. The result is a discontinuous, but photorealistic, arthroscopic simulation in synchronisation with the continuous virtual simulation.

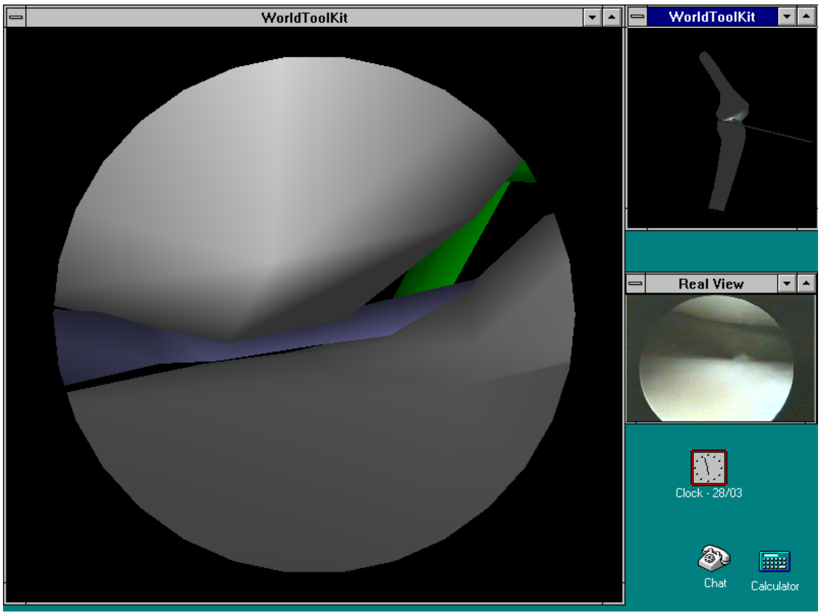


Figure 9. *Virtual arthroscopy trainer display*

6. CONCLUSIONS

Knee injuries are an exceedingly common form of disability, and in those requiring great mobility the consequences can be severe. Many knee disabilities can be fixed by corrective surgery and arthroscopy offers an approach with minimum intrusion and fast recovery times. The skills required for arthroscopic surgery are radically different from those for open surgery and a virtual reality trainer could provide a cheap, effective alternative to current training techniques. Its closest counterpart in traditional training techniques is the physical model which can be expensive to maintain since it must be disassembled and damaged parts replaced after any surgical procedure has been practised. In the virtual counterpart, the joint anatomy can be ‘repaired’ simply by resetting the software. Since the virtual arthroscopic trainer only uses low-cost plastic replicas of the arthroscope and instruments, it has also been identified by training centres as a lower risk method of training than using the physical model. Newly qualified surgeons often damage the expensive real surgical instruments through inexperience when using the physical models.

We believe that the combination of virtual reality and simulation allows the creation of an extremely flexible and useful training tool for surgery. However, we also believe that any such tools must be created in a manner as to be economically feasible, or they will stand no chance of being incorporated into standard training programmes.

The prototype system described here has been used by a number of surgeons from the Royal Hallamshire Hospital and Northern General Hospital in Sheffield and all have been impressed by the sense of realism afforded by the necessity for correct hand-eye co-ordination when manipulating the dummy instruments together with the leg model within the virtual environment. Development of the virtual training system is fully supported by Smith and Nephew Surgical who have expressed an interest in incorporating the final trainer in their arthroscopic courses and training centres throughout the country.

Acknowledgements: The virtual arthroscopic simulator project is supported by Virtual Presence Ltd. and Smith and Nephew Surgical, and funded by a grant from the Engineering and Physical Sciences Research Council.

7. REFERENCES

D J Bamford, S A Paul, J Noble and D R Davies (1993), Avoidable complications of arthroscopic surgery, *Journal Royal College Surgeon Edinburgh*, *37*, pp. 92-95

H D Banta (1993), Minimally invasive surgery - implications for hospitals, health workers, and patients, *British Medical Journal, 307*, pp. 1546

J Coleman (1994), Virtual reality in medicine, *Proc. London Virtual Reality Expo.*, London, pp. 169-173

R J Hollands and E A Trowbridge (1996), A PC-based virtual reality arthroscopic surgical trainer, *Proc. Simulation in Synthetic Environments*, New Orleans, pp. 17-22

R J Hollands, E A Trowbridge, D Bickerstaff, J B Edwards and N Mort (1995), The particular problem of arthroscopic surgical simulation - a preliminary report", In *Computer Graphics : Developments in Virtual Environments* (R A Earnshaw & J A Vince, Eds), Academic Press, London, pp. 475-482

D Hon (1992), Tactile and visual simulation: a realistic endoscopy experience, *Proc. Medicine Meets Virtual Reality*

I Logan (1995), Virtual reality training simulator, *http://www.enc.hull.ac.uk/CS/VEGA/medic/surgery.html*

I Logan, D P M Wills, N J Avis, A M M A Mohsen and K P Sherman (1996), Virtual environment knee arthroscopy training system, *Proc. Simulation in Synthetic Environments*, New Orleans, pp. 11-16

J J O' Connor and A Zavatsky (1993), Anterior cruciate ligament function in the normal knee, In *The Anterior Cruciate Ligament: Current and Future Concepts* (D W Jackson *et al.* Eds), Raven Press Ltd, New York, pp. 39-52

R Zeigler, G Fischer, W Müller and M Göbel (1995), Virtual reality arthroscopic training simulator, *Journal Comput.Biol.Med.,* **25**, *2*, pp. 193-203

Adaptive multimedia interfaces in PolyMestra

Ephraim P. Glinert and G. Bowden Wise

Computer Science Department, Rensselaer Polytechnic Institute,
Troy NY 12180, USA

glinert@cs.rpi.edu, wiseb@cs.rpi.edu

ABSTRACT

An architecture for a new generation of multimedia systems is presented based on the concept of metawidgets, which are collections of alternative representations for information, both within and across sensory modalities, along with user—transparent mechanisms for choosing among them. The proposed architecture allows us to overcome certain drawbacks of today's systems, where the designer typically must assign each component of the display to a specific modality in a fixed and inflexible manner. The design of the PolyMestra environment based on our architecture is next described in detail, with particular emphasis on the layered development approach, core software tools and inter—application communication. Finally, we discuss the current status of the implementation, and outline plans for distribution of the prototype later this year to get user feedback.

Keywords: multimedia systems, multimodal widgets, object oriented frameworks, C++, Standard Template Library (STL)

1. INTRODUCTION

Multimedia systems which include sophisticated audio output capabilities are fast becoming the standard platform for most users, both at work and at home. No longer restricted to a visual medium, these systems allow users to not only see the information presented to them but to hear it as well. Research into so—called virtual reality hints that it may not be long before our repertoire of standard interaction techniques is further augmented to include touch, gestures, voice and 3D sound.

The expanded palette of interaction technologies is attractive, in that they may enable users to communicate with their computers in a more "natural" way. But this additional freedom also presents new challenges to software designers, who must now develop applications which deliver information to end users in the most effective manner possible in a multisensory realm that encompasses text, graphics, speech, nonspeech audio, etc.

Successful concurrent exploitation of several modalities requires careful planning, otherwise some information may not be perceived by the user. Today's multimedia applications typically prescribe the modality in which any given information is presented in a hard coded (predetermined) manner. The drawback of this strategy is that, no matter how much or how well the systems organizer wrestles with the issue of how best to design the output, he/she is fighting a losing battle.

The reason is simple. An inflexible assignment of information to a particular modality by the designer must eventually lead to situations in which the chosen output modality is unacceptable to a given user and/or the circumstances at hand. For example, environmental conditions such as a noisy factory floor can preclude the use of sound. Even disregarding such "external" sources of interference, users often run several applications simultaneously and must absorb information from all of them; one issue to consider in this case is that too much (cumulative) information in a single modality can lead to sensory overload.

Most importantly, from our viewpoint, many users have sensory impairments and cannot interpret information in one or more modalities. Application designers who create systems whose output is rigidly distributed over multiple modalities therefore sharply reduce the potential customer base for their products! A better approach, in our opinion, is for applications to become multimodal at a high (abstract) level so that they are able to display all information by exploiting alternative and/or complementary sensory modalities, as (changes in) the working environment, individual users' needs and preferences, and other factors (both extra— and intra—system) dictate.

This report describes ongoing research, in which we are exploring a new architecture for multimodal systems based upon the preceding ideas. Mestra was a figure in ancient Greek mythology who, in order to escape the slavery into which she had been sold by her father, prayed to Poseidon who loved her and conferred upon her the power of metamorphosis so that whenever she was sold she could escape by changing her form. Our choice of name for our system thus reflects our goal of freeing users of multimedia systems from the confines of inflexible information displays by supporting metamorphosis of information from one modality to another.

In what follows we first present and justify our architecture for the new generation of multimedia systems we envisage. We then discuss in detail the design of the PolyMestra environment for IBM—compatible PCs which is based upon our architecture, and whose implementation is currently in progress. Finally, we outline our plans for distribution of the prototype to get user feedback later in the year.

2. AN ARCHITECTURE FOR MULTIMODAL SYSTEMS DEVELOPMENT

Multimodal applications (i.e., those which must or can present information in two or more modalities) require a very different environment than that now provided by graphical user interfaces (GUIs), because the simultaneous presentation of information in more than one modality requires careful management if all the information is to be understood rather than lost. In this section, we define concepts and present a development methodology for end user applications that are empowered to present information in alternative modalities. However, before presenting our architecture we briefly review pertinent findings from the science of human perception and from cognitive psychology, in order to get a better understanding of how we process information in a multisensory environment.

2.1. Cognitive Aspects of Multimodal Interaction

At any given moment, each of us is performing a variety of tasks. Sometimes, we are able to easily perform two or more tasks at once (e.g., driving a car, listening to music, and drinking a cup of coffee). At other times, we are only able to perform one task at a time because it requires so much concentration (e.g., operating a power saw). How can we explain these apparent limitations and discrepancies in the behavior of the human cognitive system? Psychologists call this mental attention.

The user of a multimodal computing environment is bombarded by a multitude of informational messages, each of which may be presented in a different modality. The user must attend to each message, in order to comprehend it. When there are insufficient attentional resources, information will be lost. If a multimodal interface is to be both usable and efficient, the probability of the user missing a message or not understanding it must be minimized.

To gain a better understanding of how to present information more effectively using several modalities, we examined a number of theories of attention, including: capacity theory, resource theory, confusions theory, and compatibility of proximity. Fracker and Wickens (Fracker and Wickens, 1989) provide an excellent discussion of these ideas; we summarize here just that work which we found most applicable to our research.

Most modern theories of attention are based on capacity theory (Knowles, 1963; Moray, 1967), which contends that information related to simultaneous tasks is processed in parallel until the mental load forces a bottleneck to ensue, after which only one task may be performed at a time. However, the precise nature of this “bottleneck” has been a source of conflicting opinion among psychologists.

A number of theories assume that task performance can be related to the task’s demand for processing capacity, commonly referred to as “resources.” When multiple tasks are performed simultaneously they compete with one another for the limited mental resources available, which may ultimately cause performance to deteriorate. These theories have been collectively called resource theories (Kahneman, 1973; Navon and Gopher, 1979; Wickens, 1984; Wickens, 1987).

Early resource theory (Kahneman, 1973) hypothesized that there is a single, undifferentiated pool of resources available to all tasks and mental activities. The available resources can be used for several tasks, provided the demands imposed by the tasks do not overly deplete the finite pool of resources. Attention is envisaged as a resource that can be distributed among different stages of processing, depending upon task demands.

More recent theories, however, suggest that there is more than one pool of resources. Multiple resource theory (Wickens, 1984; Wickens, 1987; Boles and Law, 1992) shares the philosophy that several tasks can be performed concurrently as long as they do not compete for the same resources. Unlike the earlier models, however, all tasks are not assumed to compete for a single, undifferentiable pool of resources. Rather, the hypothesis now is that there are multiple pools of resources within the human processing system which span three orthogonal dimensions (see Fig. 1):

(1) Stage of information processing (encoding, central processing, or response selection).

(2) Modality of input (visual or aural).

(3) Code of information processing (spatial or visual).

Each dimension can be thought of as a separate pool of resources. For the stage—defined resource pools, information is assumed to be processed by humans in stages, going from encoding to central processing to response selection. Tasks in the different stages use different resources for the most part. This implies that humans can encode or centrally process information for one task at the same time they are responding to another.

Proc. 1st Euro. Conf. Disability, Virtual Reality & Assoc. Tech., Maidenhead, UK, 1996

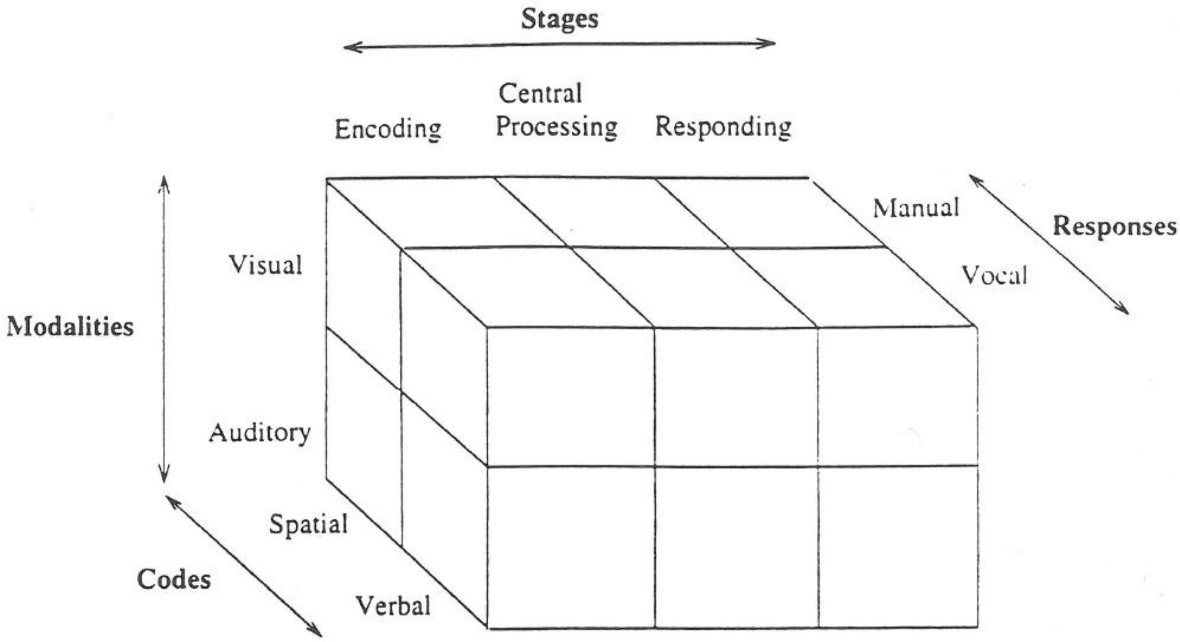


Figure 1: *Dimensions of resources found in the human processing system as suggested by Wickens (Wickens, 1984).*

For the modality—defined resource pools, tasks can be more easily performed when each utilizes a different modality. Thus, humans are able to attend to a visual message and an auditory message at the same time. This is directly relevant to our multimodal architecture, since it suggests that users will be able to comprehend more information when it is divided among several modalities rather than presented in a single modality.

Information can also be presented either visually or spatially. Performance will be better when information for different tasks is presented using different codes rather than all being presented in the same manner. This dimension impacts all stages of processing: encoding (speech vs. graphics); central processing (spatial working memory vs. linguistic working memory); and response (speech output vs. manual output).

2.2. Metawidgets and Representations: A Substrate for Multimodality

Typically, applications which run under today's GUIs are developed using libraries of widgets which represent familiar user interface elements such as windows, icons, and buttons. However, because they are constrained a priori to the visual modality due to their graphical nature, widgets are too low level to be of use directly by a multimodal application, which must communicate with the environment in which it lives at a higher level of abstraction in order to interact with the user in any of several modalities in a manner transparent to the applications programmer.

Metawidgets, proposed by Glinert and Blattner (Blattner, Glinert, Jorge and Ormsby, 1992; Glinert and Blattner, 1993), are multimodal widgets which represent some informational abstraction of the application that is to be presented to the user. Each metawidget contains a repertoire of representations, each of which may be in a different sensory modality (or combination of modalities), as well as methods for selecting among them. By taking into account the user's preferences, along with relevant extra—and intra—system factors, the selection mechanism can determine which representation is optimal for presentation when the metawidget is displayed.

From a functional point of view, metawidgets must be able to select a suitable representation, display it to the user, compute the cognitive resources this representation will consume, and communicate with the underlying core software. Many of these operations, if not all, will be performed the same way for every metawidget, indicating that much of the metawidget code may be reused.

2.3. Modalities and Cognitive Resources

Users often run several applications simultaneously, and must absorb information from all of them. Too much (cumulative) information in a single modality leading to sensory overload is one potential problem to be overcome. Cross—modality conflicts, as when the visual dominance (Wickens, 1992) of the human perceptual system causes an auditory stimulus to be ignored, are another.

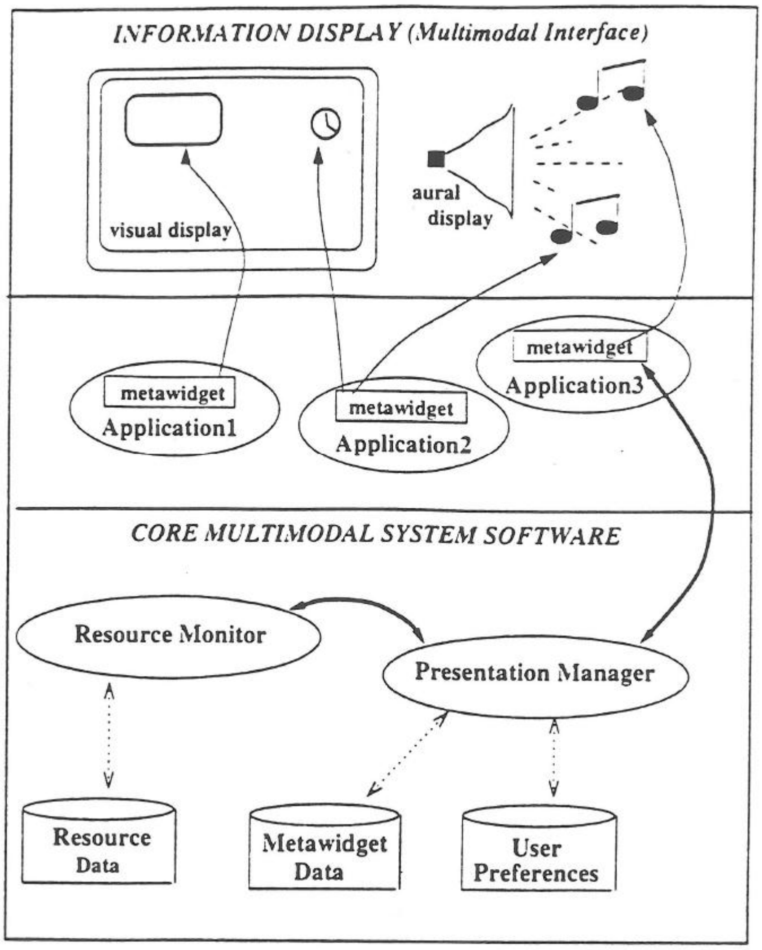


Figure 2: *Run time view of a multimodal environment. Thin lines indicate the representations associated with a metawidget. Thick lines show inter—application communication paths. Dotted lines show internal data paths. Although each metawidget will have a similar communication path to the Presentation Manager, only one is shown in the figure for clarity's sake.*

To eliminate or at least alleviate such phenomena, the demands on the user's cognitive system must be periodically monitored by the computing environment. We define a cognitive resource as some aspect of the user's perceptual or attentional processes that is to be monitored for this purpose.

Each representation of a metawidget resides in one or more modalities. Each modality of interest will have a set of cognitive resources defined for it, along with heuristics for determining the contribution of each such resource to the user's total cognitive load. If a representation resides in more than one modality, it will consume resources from all of them. It is the responsibility of the underlying core software to monitor which representations are currently active, measure their contribution to the overall cognitive load, and take action where appropriate to prevent the user from becoming overloaded.

In our architecture, data on currently active applications, their metawidgets and associated representations are collected and assessed by a pair of core system tools which run in the background concurrently with the user's applications (cf. Fig. 2):

- Resource Monitor. This tool is responsible for determining the contribution to the total cognitive load by each application running in the environment, on the basis of its consumption of cognitive resources relevant to the user.
- Presentation Manager. This tool handles the presentation of information so that the user does not become overloaded. The cognitive load, user preferences, and other system data are all weighed to determine whether some information needs to be transformed into a representation in an alternative modality, and if so which modality is the optimal choice for this user at this time.

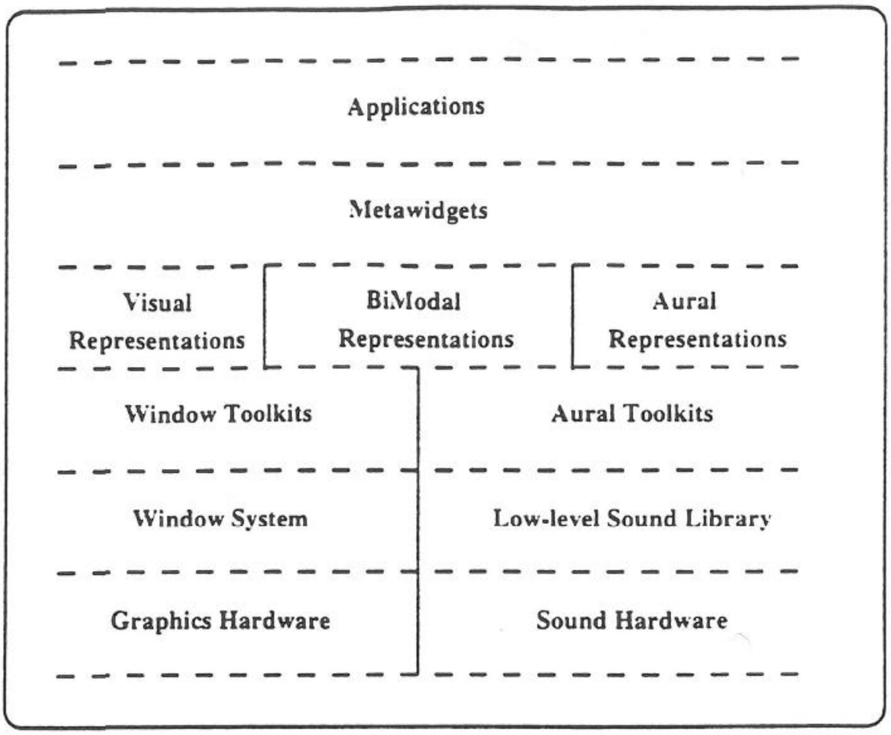


Figure 3: *A layered approach to multimodal systems development.*

2.4. A Layered Development Approach

Metawidgets, with their associated representations, provide natural building blocks for developing multimodal applications. Thus, at the heart of our architecture for multimedia computing environments lies a layered multimodal framework, as shown in Fig. 3. The representations of a metawidget are developed using toolkits containing widgets in a particular modality. Visual representations, for example, can be constructed using a toolkit of widgets for the underlying window system. Similarly, aural representations can be developed from audio toolkits that utilize the underlying sound hardware.

3. POLYMESTRA: A MULTIMODAL ENVIRONMENT FOR MICROSOFT WINDOWS

PolyMestra is a multimodal environment based upon our architecture currently under development for IBM—compatible PCs. PolyMestra provides application designers with a C++ framework for building multimodal applications for the Microsoft Windows 3.x platform. In this section, we discuss the implementation of PolyMestra in detail, with particular focus on some of the challenges encountered and our solutions for overcoming them.

3.1. Metawidgets and Representations

One of the early crucial decisions was how to associate a metawidget with its representations. What is needed is a mechanism that allows the current representation of a metawidget to change, without the metawidget having to be aware of the precise representation chosen (the metawidget should only be concerned with whether the selected representation is in an acceptable format to, and does not overload, the user).

Four basic assumptions drove the metawidget—representation design:

(1) Each metawidget must have one or more representations.

(2) These representations may be in the same or (preferably) different modalities.

(3) The representation to be displayed by a metawidget is selectable at run time, based on user preferences as well as extra—and intra—system information.

(4) The representation of a metawidget may change (metamorphose) while it is on display.

These assumptions imply that the coupling between metawidgets and their representations does not depend on the content of the information being displayed or its modality of presentation. This suggests that the mechanism employed to associate a metawidget with its representation can be common to all metawidgets.

©1996 ECDVRAT and University of Reading, UK ; ISBN 0 7049 1140 X

In C++, such a mechanism is easily implemented by having each metawidget contain a base class pointer to its representation. By deriving all representations from a common base representation class (called BaseRepresentation in PolyMestra), the metawidget can point to an instance of any one of the derived classes. Furthermore, this approach ensures that all representations for a metawidget share the same interface and reuse common functionality.

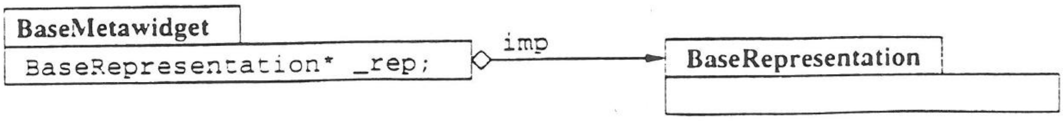


Figure 4: *A metawidget in PolyMestra contains a pointer, or handle, to its representation. The arrowhead line marked 'imp' with a diamond at its tail indicates that metawidgets are implemented by one or more representations.*

Once an optimal representation for a metawidget is chosen, the base class pointer is updated to point to that representation. Whenever a metawidget is told to display/hide itself, to calculate its resources requirements, or to send these requirements to the core software, the operation is forwarded to the active representation (via virtual functions).

Similarly, all metawidgets are derived from a common base class BaseMetawidget, so that every metawidget also shares the same interface and reuses functionality common to all metawidgets. BaseMetawidget contains a data member, namely a pointer of type BaseRepresentation to hold the currently selected representation. In addition to methods for selecting a representation for a metawidget, the operations provided by every representation are also implemented in BaseMetawidget except these methods simply forward the operation to the metawidget's representation through the base class pointer.

Pointers used in this way are commonly called handles (Coplien, 1992; Murray, 1993). Handles provide the architecture that enables metawidgets to utilize multiple representations, to delegate operations to the current representation, and to change the representation at run time. Handles may also be used to hide implementation details, to minimize the impact of changes during development, to determine the type of an object from the context in which it is constructed, and to create objects when an object's size is unknown.

Fig. 4 shows how the representation of a metawidget is implemented as a pointer to a BaseRepresentation. The object oriented class diagrams found in this figure and the next are patterned after Rumbaugh's Object Modeling Technique (OMT) (Rumbaugh, Blaha, Premerlani, Eddy and Lorenson, 1991). Classes are drawn in rectangular boxes; arrowhead lines are used to denote relationships between classes. Simple arrowhead lines indicate inheritance, while arrowhead lines with a diamond at the tail denote aggregation. Pseudocode for class methods is sometimes shown in a dashed box connected by a dashed line to the operation it represents.

3.2. The Representation Repository

Another design issue concerned how to manage the collection of available representations for each metawidget. Not only will each metawidget have more than one representation, but a particular representation may be useful to more than one metawidget! Complicating the matter further was our desire to make it possible for developers of a multimodal application to easily incorporate additional representations into a metawidget's existing palette of alternatives, without breaking existing application code.

Given that the representations for a particular metawidget are cataloged in some way, a mechanism is also needed for a metawidget to obtain a list of these representations at run time so that one of them may be chosen. This same list of representations can also be consulted to select a new representation during metawidget metamorphosis.

In PolyMestra, representations for a particular informational concept share a common base class, from which all representations in the collection are derived. A metawidget is associated with a particular collection of representations by identifying itself with a base representation class. Each metawidget contains a virtual method, BaseRepName(), which returns the name of the metawidget's base representation class.

Fig. 5 demonstrates how a metawidget class for text strings, MTextString, is associated with a collection of text strings derived from a common base class BaseTextString. In this example, four different representations of a text string are derived from BaseTextString: two visual representations VTextLabel and VTextButton, one aural representation ATextAnnounce, and one bimodal representation, BTextView. By specifying its base representation name as BaseTextString, metawidget MTextString may present itself using any of the four representations derived from BaseTextString.

As the reader may already have noticed, we have adopted a naming convention for class names in the PolyMestra framework. The prefix of each class name indicates its category: "M" for metawidget classes;

"Base" for base representation classes; "V" for visual representations; "A" for aural representations; and "B" for bimodal representations. These naming conventions are employed in several of our figures.

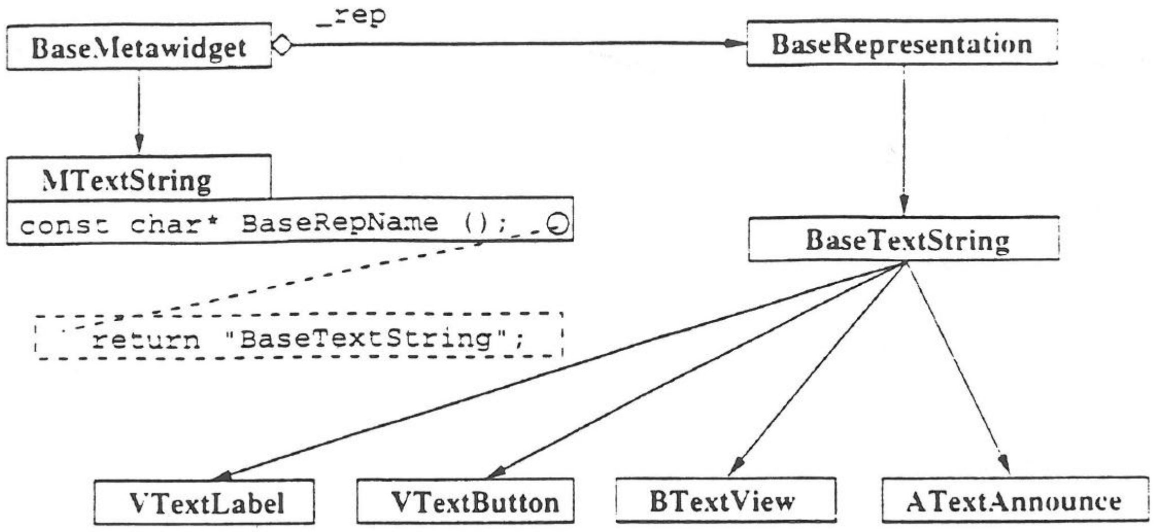


Figure 5: *Representations of a metawidget are associated by the name of the representation's base class.*

Whenever a metawidget selects a representation, its base representation name is used to query the representation repository for a list of possible representations. The list is then examined to select and create a representation for the metawidget. But how does an instance of a representation get created using this list? What kind of object is stored in the representation repository?

What is needed is an object from which other object instances can be created. The list of representations obtained from the repository contains objects from which instances of the particular representation types can be created. Object oriented languages may be classified on how they relate classes and objects. Single hierarchy languages, like Smalltalk and Self, treat classes and objects as essentially the same. This allows some objects to be able to create other object instances. These special "factory" objects are called exemplars (Coplien, 1992).

C++, which treats classes and objects as distinct entities, is a so—called dual hierarchy language. Classes are similar to abstract data types. Classes are fixed at compile time, and objects do not exist until run time. The compile time intensive nature of C++ furthermore dictates that objects may be created only through declarations using built in types or classes, or through the use of the operator new. C++ does not have an exemplar facility analogous to that available in single hierarchy languages.

Classes in C++, however, can be imbued with exemplar facilities so that object instances may be created from another object via a virtual function call. The virtual function simply returns a pointer to the newly created object, usually of type equal to the base class from which the exemplar class has been derived. In PolyMestra, each distinct representation has an exemplar object from which representations of that type may be created. By making exemplar objects static, only one exemplar is created and placed in the repository for each type of representation.

In the example representation hierarchy for BaseTextStrings shown in Fig. 5, there will be an exemplar for each of the four types of possible string representations. An exemplar object for VTextButton creates an instance of VTextButton, but returns a pointer of type BaseTextString to it. In C++, this is possible because a base class pointer may point to any instance of a class that is derived from the base class.

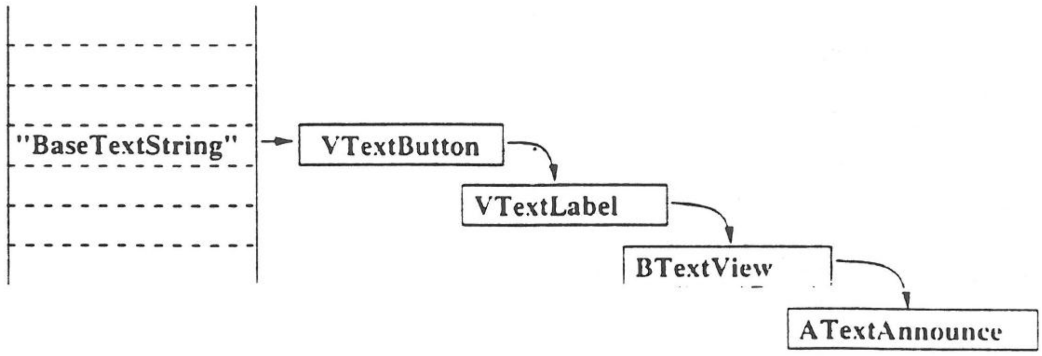


Figure 6: *The repository contains lists of exemplars for each collection of representations, indexed by the base representation name.*

The representation repository contains a list of exemplars indexed by the base representation class name. Fig. 6 depicts how the text string representations are stored in the repository. When a designer creates a new representation, a static exemplar instance of the representation is created which automatically registers itself with the repository. Because the representation exemplars are static, we are guaranteed that the repository entry will be created before any metawidget can use the representation in an application.

3.3. Modalities and Cognitive Resources

A key goal was to exploit the insights into the human cognitive—perceptual system previously discussed in Section to develop mechanisms for measuring the user's cognitive load as information is presented by the multimodal environment. These mechanisms could then be employed by the PolyMestra run time tools to present information in such a way that the user can assimilate it all. Using multiple resource theory and others (Wickens, 1987; Wickens, 1992; Boles and Law, 1992) as a guide, we devised the following assumptions about cognitive resources:

(1) Each modality has a fixed, predetermined set of cognitive resources.

(2) Cognitive resources may be unique to a particular modality or shared across more than one modality.

(3) Representations consume cognitive resources from each modality in which they reside.

Our original design involved the use of separate classes for each modality of interest in the system. Each modality class would then be responsible for measuring its own resources. The drawback of this design was that it required representations to inherit not only from BaseRepresentation but also from a specific modality class. We were concerned that the resulting class hierarchy was too cumbersome, and would cause developers to not want to use our framework.

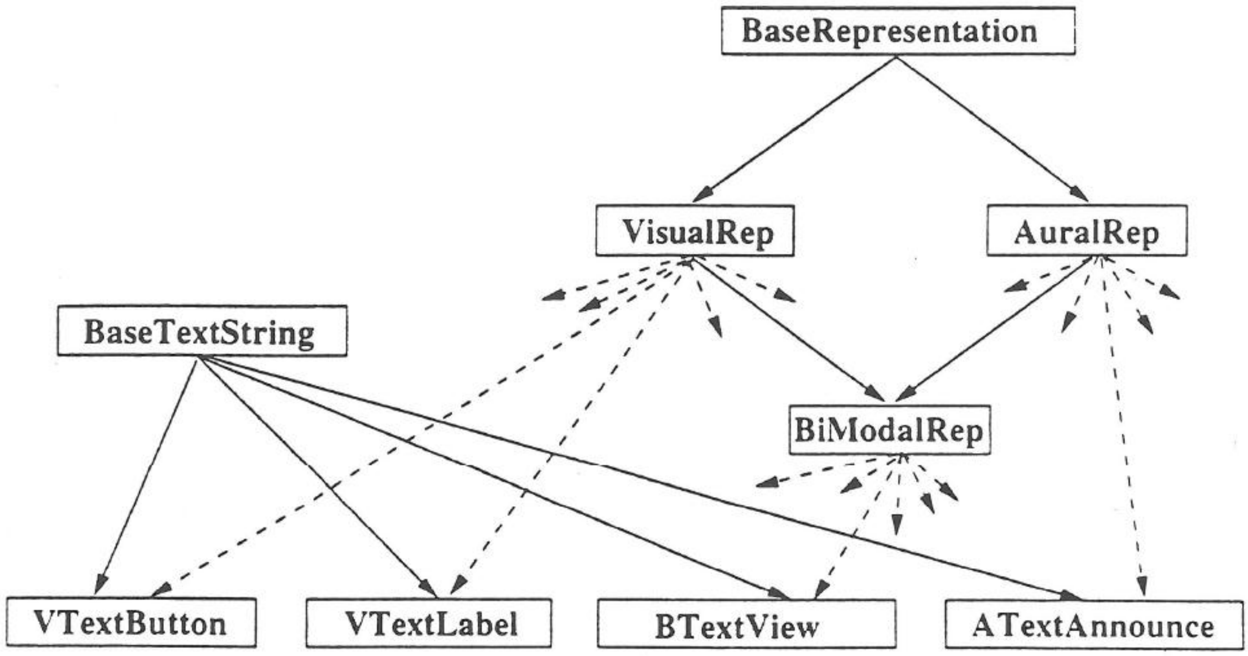


Figure 7: *The hierarchy that results when modality characteristics are absorbed by representation classes. All lines denote inheritance. The dashed lines are used to show inheritance from the representation classes to actual representations in each modality.*

Instead of defining separate classes for the various modalities, we therefore chose to build cognitive resource functionality into the representations. The generic BaseRepresentation class contains virtual functions for calculating cognitive resources, but since no modality has yet been chosen, these methods do nothing. Modality specific representation classes are then derived from BaseRepresentation. Fig. 7 shows the class hierarchy that results when classes VisualRep, AuralRep and BiModalRep are added to the hierarchy.

3.4. Inter—application Communication

In order for the core software tools to monitor the load imposed by client applications, some form of inter—application communication is needed. PolyMestra uses Microsoft Window`s Dynamic Data Exchange Management Library (DDEML) for this purpose.

The Presentation Manager serves as the locus of communication for all metawidgets. When a metawidget is first created, it registers with the Presentation Manager. The Presentation Manager saves the registration information to maintain an internal database of active metawidgets. Later, when a representation must be selected for presentation to the user, the appropriate metawidget queries the Presentation Manager to determine if addition of this representation to the information display would overload the user. If the Presentation Manager decides it would not, the currently selected representation may be presented, otherwise, the metawidget must resubmit its query after selecting an alternative representation from among those available to it.

The Presentation Manager does not maintain information regarding the user's current cognitive load. The Resource Monitor is responsible for collecting the relevant data and computing this value when required, taking into account cross—modality and within—modality effects. By encapsulating knowledge relating to the calculation of cognitive load inside the Resource Monitor, separate from the Presentation Manager, it becomes possible to easily modify/update this part of the system when necessary without otherwise impacting system operation.

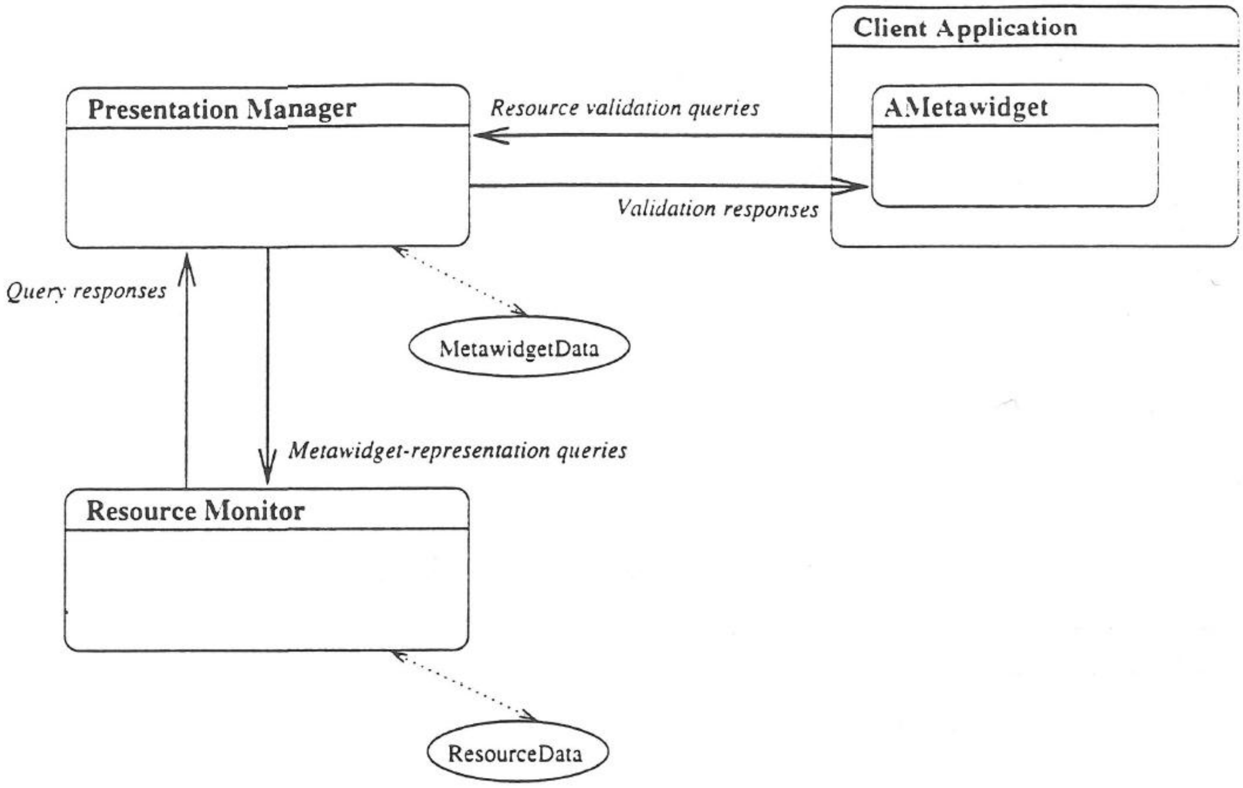


Figure 8: *Inter—application communication in PolyMestra. Solid lines indicate communication paths and dotted lines indicate internal data paths.*

Each query that the Presentation Manager receives from a metawidget includes the anticipated incremental cognitive resource consumption for the representation whose display is proposed. These values are passed by the Presentation Manager (using the DDEML) to the Resource Monitor, which uses them to determine what the new cognitive load on the user would be if this representation were in fact displayed. Fig. 8 shows the PolyMestra communication protocol in detail.

4. CURRENT STATUS

Implementation of PolyMestra is well under way in C++ using the Standard Template Library (STL) extensions (Stepanov and Lee, 1995; Nelson, 1995). Our visual representations are being developed using Borland's Object Windows Library (better known as OWL) and we are developing a toolkit of aural widgets on top of Microsoft Window's Multimedia Control Interface. The system currently consists of over 15,000 lines of code (excluding comment lines). We expect to have the first working prototype, including utilities for web browsing and e—mail, available for distribution to alpha test sites by late summer. We eagerly seek applications programmers and users who are willing to experiment with our system and provide feedback about their experiences, which will then enable us to refine our design and implementation. Interested parties are invited to contact the second author for more information.

Acknowledgement: This research was supported, in part, by the National Science Foundation under contracts IRI--9213823 and CCR--9527151.

5. REFERENCES

M. M. Blattner, E. P. Glinert, J. A. Jorge and G. R. Ormsby (1992). Metawidgets: Towards a theory of multimodal interface design. In Proceedings COMPSAC'92, Chicago, September 22--25, pp. 115--120. IEEE Computer Society Press, Los Alamitos, CA.

D. B. Boles and M. B. Law (1992). Orthogonal lateralized processes have orthogonal-attentional resources, In The Annual Meeting of the Psychonomic Society. St. Louis, November 13.

J. O. Coplien (1992). Advanced C++ Programming Styles and Idioms. Addison Wesley Publishing Co., Reading, MA.

M. L. Fracker and C. D. Wickens (1989). Resources, confusions, and compatibility in dual—axis tracking: Displays, controls, and dynamics. Journal of Experimental Psychology: Human Perception and Performance, 15(1), pp. 80--96.

E. P. Glinert and M. M. Blattner (1993). Programming the multimedia interface. In Proceedings First ACM International Conference on Multimedia (MULTIMEDIA'93), Anaheim, August 2--6, pp. 189--197. ACM Press, New York City, NY.

D. Kahneman (1973). Attention and Effort. Prentice Hall, Englewood Cliffs, NJ.

W. B. Knowles (1963). Operator loading tasks. Human Factors, 5, pp. 151--161.

N. Moray (1967). Where is attention limited? A survey and a model. Acta Psychologica, 27, pp. 84--92.

R. B. Murray (1993). C++ Strategies and Tactics. Addison Wesley Publishing Co., Reading, MA.

D. Navon and D. Gopher (1979). On the economy of the human processing system. Psychological Review, 86, pp. 214--255.

M. Nelson (1995). C++ Programmers Guide to the Standard Template Library. IDG Books, Foster City, CA.

J. Rumbaugh, M. Blaha, W. Premerlani, F. Eddy, and W. Lorenson (1991). Object Oriented Modeling and Design, Prentice Hall, Englewood Cliffs, NJ.

A. Stepanov and M. Lee (1995). The standard template library (ANSI/ISO document). Available electronically from ftp://butler.hpl.hp.com/stl/stl.zip.

C. D. Wickens (1984). Processing resources in attention. In Varieties of Attention (D. R. Davies, Editor), pp. 63--258. Academic Press, New York City, NY.

C. D. Wickens (1987). Attention. In Human Factors Psychology (P. A. Hancock, Editor), pp. 29--80. North Holland, Amsterdam.

C. D. Wickens (1992). Engineering Psychology and Human Performance. Harper Collins, New York City, NY.

Myoelectric signals pattern recognition for intelligent functional operation of upper-limb prosthesis

N Chaiyaratana[1], A M S Zalzala[2] and D Datta[3]

[1,2]Robotics Research Group, Department of Automatic Control and Systems Engineering, University of Sheffield, Mappin Street, Sheffield, S1 3JD, UK

[3]Disablement Services Centre, Northern General Hospital, Herries Road, Sheffield, S5 7AT, UK

[1,2]*rrg@sheffield.ac.uk*

ABSTRACT

This paper presents a comparative study of the classification accuracy of myoelectric signals using multilayer perceptron with back-propagation algorithm and radial-basis functions networks. The myoelectric signals considered are used to classify four upper-limb movements which are elbow bending, elbow extension, wrist pronation and wrist supination. The network structure for multilayer perceptron is a fully connected one, while the structures used in radial-basis functions network are both fully connected and partially connected. Two learning strategies are used for training radial-basis networks, namely supervised selection of centres and fixed centres selected at random. The results suggest that radial-basis function network with fixed centres can generalise better than the others without requiring extra computational effort.

Keywords: multilayer perceptron, myoelectric signal, pattern recognition, radial-basis function network, upper-limb prosthesis

1. INTRODUCTION

Myoelectric signals are the signals which are generated by muscles when they contract. They have been used in various aspects of medical and biomedical engineering, for example, the diagnosis of neuromuscular disease such as polymyositis (Kumaravel and Kavitha, 1994). One of the most common uses of myoelectric signals is for controlling prosthesis manipulators (Scott and Parker, 1988). Myoelectric control prostheses have received widespread use as devices for individuals with amputations or congenitally deficient upper limbs (Scott and Parker, 1988), and many systems are now available commercially to control a single device such as a hand, elbow and wrist. Each myoelectric signal, generated by the muscle in performing different task, has a unique pattern which contains the information about the direction of movement and the speed of action. To be able to control the prosthesis successfully, the microprocessor, which is part of the prosthesis, must be able to classify these patterns accurately; this results in a pattern recognition problem. The myoelectric signal is essentially a one-dimensional pattern, and the methods and algorithms developed for one-dimensional pattern recognition can be applied to this analysis. The information extracted from the myoelectric signal, represented in a feature vector, is chosen to minimise the control error. To achieve this, a feature set which maximally separates the desired output classes must be chosen. The need for fast response of the prosthesis limits the period over which these features can be extracted.

A number of researchers have discussed the possibility of using neural network in solving myoelectric signal pattern recognition problem. Most research had been carried out by using a multilayer perceptron which contains one hidden layer in conjunction with back-propagation algorithm, except the work by Costa and Gander (1993) which had been carried out by using a multilayer perceptron with two hidden layers. Myoelectric signals can be drawn from various locations on subject's body, an application dependent criterion. For example, the signals from flexor digitorium superficialis are used in classification of finger movements (Hiraiwa et al., 1989) or the signals from biceps and triceps branchii are used to determine the arm movements (Hudgins et al., 1993; Ito et al., 1991; Kuruganti et al., 1995). The control signal can be derived from a single myoelectric channel (Costa and Gander, 1993; Hudgins et al., 1993; Karlik et al., 1994; Kelly et al., 1990), or from multichannel such as two channels (Kuruganti et al., 1995; Yeh et al., 1993), four channels (Ito et al., 1991) or five channels (Doerschuk et al., 1983). Using a single channel myoelectric signal would result in less complexity in neural network structure. However, using multichannel signals makes the positions of electrodes become less critical to the experiment and increase the classification accuracy (Kuruganti et al., 1995). Most of the previous research is concerned with the arm movements classification task. Usually, the experiments are carried out in two possible ways: by exposing the subject's arm movements to a weight constraint (Costa and Gander, 1993) or by allowing the subject to perform the movement naturally (Hudgins et al., 1993; Kuruganti et al., 1995).

To control *n* functions in the prosthesis requires *n* unique patterns of muscle contraction. The control schemes have been based almost entirely on the discriminant approach to pattern recognition, in which each pattern is described by a set of features. Features set can be obtained by using various methods. For example, the parameters of some stochastic models such as an autoregressive (AR) model or autoregressive moving average (ARMA) model can be used as features set. A number of research has been done by utilising AR model (Doerschuk et al., 1983; Graupe et al., 1985; Karlik et al., 1994; Kelly et al., 1990; Kiryu et al., 1994; Latwesen and Patterson, 1993; Merletti and Lo Conte, 1995; Zardoshti-Kermani et al., 1995). All of these works are based upon the research by Graupe and Kline (1975) which involves modelling myoelectric signals as ARMA models. The later research has shown that the use of AR model is sufficient for modelling myoelectric signals. Graupe et al. (1985) have proved that a myoelectric signal within a period of 0.2-0.3 second can be model as a 4th order AR model.

Various methods have been used to obtained the parameters of AR model. One of the most common methods which has been used is recursive least squares (RLS) or sequential least squares (SLS) (Graupe et al., 1985; Kiryu et al., 1994; Latwesen and Patterson, 1993; Zardoshti-Kermani et al., 1995). This method has been proved to be very reliable and has capability in dealing with noisy myoelectric signal. The extension of recursive least squares algorithm to accommodate multivariable autoregressive model is proposed by Doerschuk et al. (1983). This extension is done such that all parameters of the AR models from different channels of the signals can be computed simultaneously. Other methods which have been used to obtained the parameters of AR model are the application of discrete Hopfield network (Kelly et al., 1990) and the PARCOR algorithm (Karlik et al., 1994). Unlike recursive least squares algorithm which is based on the principle of minimising the error between the estimate value and the actual value of each signal in each iteration, discrete Hopfield network is used under the principle of multivariable optimisation. One advantage in using discrete Hopfield network is that the convergence rate of computation by discrete Hopfield network is higher than that of the recursive least squares algorithms (Kelly et al., 1990).

Some other characteristics of myoelectric signal can also be used as features for the neural network input. For example, the time domain characteristics of the signal such as mean absolute value, mean absolute value slope, zero crossings, slope sign changes and waveform length have been used by Hudgins et al. (1993) and Kuruganti et al. (1995). Other time domain characteristics such as the turns and mean amplitude and normalised time average have been utilised by Yeh et al. (1993) and Ito et al. (1991), respectively. An advantage which can be gained by using time domain characteristics instead of parameters of stochastic models is the complexity reduction in feature extraction process. One drawback is the increase in input layer of the network. Parameters from fast Fourier transform (FFT) can also be used as the features set (Hiraiwa et al., 1989). Del Boca and Park (1994) have extracted the signal features through Fourier analysis. Unlike any other previous research, the features are then unsupervised clustered by using fuzzy c-mean algorithm before they are presented to the neural network for pattern recognition. This method has been proved to be very efficient for real-time operation.

The assumption that myoelectric signal is a stochastic signal has always been made when the parameters of AR model, time domain characteristics or results from Fourier analysis are used as features. Costa and Gander (1993) have proposed a new assumption that myoelectric signal should be treated as a chaotic signal. This assumption leads to a totally different type of features that can be used as input to the neural network. They use the Poincaré sections of the chaotic myoelectric signal as the features set. They also suggest the use of their result in automated myopathy diagnosis.

This paper is concerned with the use of neural networks in myoelectric control system. The myoelectric data used in the experiments are the data obtained from the journal article by Graupe et al. (1985). These data are the autoregressive (AR) model parameters of the single-site myoelectric signals obtained by measuring the signals from the location between biceps and triceps with sampling period of 2 ms. The location of the electrode is expected to be the place where the maximum cross-talk between signals occurs. Four contraction types are concerned in the experiments. These contraction types are elbow bending (EB) or elbow flexion, elbow extension (EE), wrist pronation (WP) and wrist supination (WS) (Graupe et al., 1985). The AR model parameters are used as feature for the feature sets. These feature sets are divided into two groups, one for training the networks, another for testing. The feature sets are used as input to two types of neural network. They are a multilayer perceptron with back-propagation algorithm network and a radial-basis function network. Comparative study on myoelectric signal classification accuracy performance between these two types of network is the major objective in this paper. This objective rises from the fact that nearly all research in this area had been carried out by using multilayer perceptron which contains one hidden layer in conjunction with back-propagation algorithm. The experiments which had been done can be divided into two sections. They are determination on multilayer perceptron structure and determination on learning strategies and structure topologies in radial-basis function network.

2. DETERMINATION ON MULTILAYER PERCEPTRON STRUCTURE

The cost function of multilayer perceptron is given in Eq. (1),

$$\varepsilon(n) = \frac{1}{2}\sum_{k=1}^{N_o} e_k^2(n) \tag{1}$$

where $\varepsilon(n)$ is the instantaneous cost function at iteration n,
$e_k(n)$ is the error from output node k at iteration n and
N_o is the number of output nodes.

The error from each output node can be defined as follows,

$$e_k(n) = d_k(n) - y_k(n) \tag{2}$$

where $d_k(n)$ is the desired response of output node k at iteration n and
$y_k(n)$ is the output of output node k at iteration n.

Haykin (1994) gives summary on back-propagation algorithm as follows.

1. *Initialisation.* Set all the weights and threshold levels of the network to small random numbers that are uniformly distributed.

2. *Forward Computation.* Let a training example be denoted by [**x**(n), **d**(n)], with the input vector **x**(n) applied to the input layer and the desired response vector **d**(n) presented to the output layer. The net internal activity level $v_j^{(l)}(n)$ for neuron j in layer l is given by

$$v_j^{(l)}(n) = \sum_{i=0}^{p} w_{ji}^{(l)}(n) y_i^{(l-1)}(n) \tag{3}$$

where $y_i^{(l-1)}(n)$ is the signal from neuron i in the previous layer l-1 at iteration n and
$w_{ji}^{(l)}(n)$ is the weight of neuron j in layer l that is connected to neuron i in layer l-1 at iteration n.

For $i = 0$, we have

$$y_0^{(l-1)} = -1 \tag{4}$$

and

$$w_{j0}^{(l)}(n) = \theta_j^{(l)}(n) \tag{5}$$

where $\theta_j^{(l)}(n)$ is the threshold applied to neuron j in layer l.

With the use of a logistic function for the sigmoidal non-linearity, the output of neuron j in layer l is given by

$$y_j^{(l)}(n) = \frac{1}{1+\exp(-v_j^{(l)}(n))}. \tag{6}$$

If neuron j is in the first hidden layer (i.e., $l = 1$), set

$$y_j^{(0)}(n) = x_j(n) \tag{7}$$

where $x_j(n)$ is the jth element of input vector **x**(n). If neuron j is in the output layer (i.e., $l = L$), set

$$y_j^{(L)}(n) = o_j(n). \tag{8}$$

The error can be computed as follows,

$$e_j(n) = d_j(n) - o_j(n) \tag{9}$$

where $d_j(n)$ is the jth element of the desired response vector **d**(n).

4. *Backward Computation.* Compute the local gradients (δ) of the network by progressing backward, layer by layer. For neuron j in output layer L, the local gradient is given by

$$\delta_j^{(L)}(n) = e_j^{(L)}(n) o_j(n) \left[1 - o_j(n)\right]. \tag{10}$$

For neuron j in hidden layer l, the local gradient is given by

$$\delta_j^{(l)}(n) = y_j^{(l)}(n) \left[1 - y^{(l)}(n)\right] \sum_k \delta_k^{(l+1)}(n) w_{kj}^{(l+1)}(n). \tag{11}$$

The weight of the network in layer l can be adjusted according to the generalised delta rule as follows,

$$w_{ji}^{(l)}(n+1) = w_{ji}^{(l)}(n) + \alpha \left[w_{ji}^{(l)}(n) - w_{ji}^{(l)}(n-1)\right] + \eta \delta_j^{(l)}(n) y_i^{(l-1)}(n) \tag{12}$$

where η is the learning rate parameter and
α is the momentum constant.

In this study, there are three parameters which are needed to be considered. They are learning rate parameter, momentum constant and number of hidden nodes. Previous research has shown that only one hidden layer is sufficient for this application. The ranges of value for both learning rate parameter and momentum constant are typically lied between 0 and 1. These two parameters cannot be chosen independently. Three observations on choices of learning rate parameter and momentum constant are given by Haykin (1994) as follows.

- A smaller learning rate parameter leads to slower convergence. The search with smaller learning rate parameter can cover more error surface than the search with larger learning rate parameter.
- For learning rate parameter approaching zero, the use of momentum constant with the value near one will increase the speed of convergence. On the other hand, for learning rate parameter approaching one, the use of momentum constant with the value near zero will ensure the stability of learning.
- The use of large value of learning rate parameter in conjunction with large momentum constant can leads to oscillations in the mean-squared error during learning process and a high value of final mean-squared error (Haykin, 1994).

From these three observations, learning rate parameter are chosen to be 0.1 and momentum constant is chosen to be 0.9. It will lead to a good coverage of error surface and fast convergence. The number of hidden nodes is determined via experiment. The testing range is chosen to be from 10 to 15. Classification results are shown in Fig. 1.

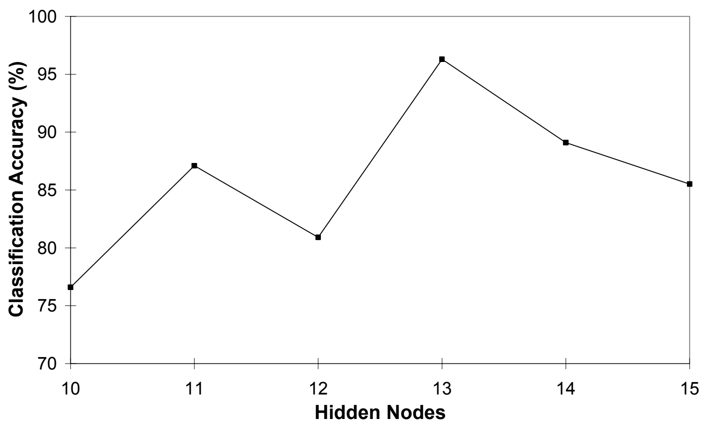


Figure 1. *Classification results of multilayer perceptron with back-propagation algorithm.*

Experiment results show that the optimum number of hidden nodes is 13 with the highest classification accuracy of 96.3 %.

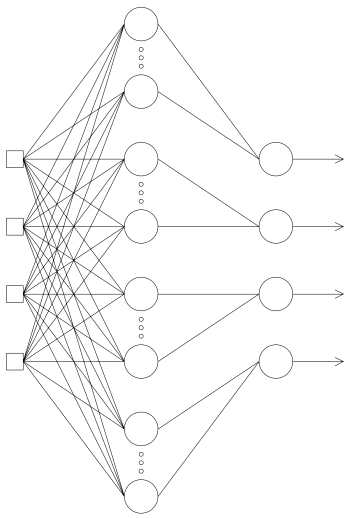
3. DETERMINATION ON LEARNING STRATEGIES AND STRUCTURE TOPOLOGIES IN RADIAL-BASIS FUNCTION NETWORK

Two learning strategies are used in this study. These learning strategies are supervised selection of centres of radial-basis function and fixed centres of radial-basis function selected at random.

3.1 Supervised Selection of Centres

In this learning strategy, all free parameters in the network undergo a supervised learning process. These free parameters are the weights connecting the hidden layer to the output layer, the centres of radial-basis function in the hidden layer and the inverse covariance matrices of the radial-basis function. This is the most generalised learning strategy for radial-basis function network. The algorithm used in the learning process is derived from the method of steepest descent. In this study, all learning rate parameters are set at 0.1. Two topologies of the network are used under this learning strategy. These topologies are partially connected network and fully connected network.

3.1.1 Partially Connected Network. In this topology, hidden nodes and output nodes are partially connected. Since there are four patterns in concerned, the number of output nodes in this case is four. Each hidden node will be connected to only one output node. Output from the first output node will be used to determine whether input pattern belongs to class 1 or not. Other output nodes work in similar fashion. Combination of result from all output nodes will be used to determine the class of input pattern. The topology of the network is shown in Fig. 2.



Input layer Hidden layer Output layer

Figure 2. *Partially connected radial-basis function network.*

This network topology can be viewed as four radial-basis function networks with on output node, sharing the same input working in co-operation. In this study, each part of network that fully connected to a single output node is call sub-network. In this case, there will be four sub-networks.

The cost function for each sub-network is defined as follows,

$$\varepsilon_k(n) = \frac{1}{2} e_k^2(n), \qquad k = 1, 2, 3, 4 \tag{13}$$

where $\varepsilon_k(n)$ is the instantaneous cost function of sub-network k at iteration n and
$e_k(n)$ is the error from output node k at iteration n.

The error from each output node is defined as follows,

$$e_k(n) = d_k(n) - \sum_{i=1}^{M_k} w_{ki}(n) G(\|\mathbf{x}(n) - \mathbf{t}_i(n)\|_{\mathbf{C}_i}), \qquad k = 1, 2, 3, 4 \tag{14}$$

where $d_k(n)$ is the desired output of output node k at iteration n,
$w_{ki}(n)$ is the weight of sub-network k which connects to the ith Green's function in the same sub-network,

$G(.)$ is the Green's function and
M_k is the number of Green's function in sub-network k.

Define Green's function in radial-basis function network as follows,

$$G(\|\mathbf{x}-\mathbf{t}_i\|_{\mathbf{C}_i}) = G((\mathbf{x}-\mathbf{t}_i)^T \boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\mathbf{t}_i)) \tag{15}$$

where $\mathbf{x}$ is the input pattern,
$\mathbf{t}_i$ is the ith centre of radial-basis function network and
$\boldsymbol{\Sigma}_i^{-1}$ is the ith inverse covariance matrix.

In this study, Green's function is chosen to be Gaussian function which can be shown in Eq. (16),

$$G(\|\mathbf{x}-\mathbf{t}_i\|_{\mathbf{C}_i}) = \exp(-(\mathbf{x}-\mathbf{t}_i)^T \boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\mathbf{t}_i))\,. \tag{16}$$

Steepest descent algorithm is used to determine algorithm for adapting the value of weights, centres and inverse covariance matrices. This leads to the use of partial derivatives of the cost function with respect to these free parameters. The adapting formula would be in the form shown in Eq. (17),

$$p(n+1) = p(n) - \eta \frac{\partial \varepsilon(n)}{\partial p(n)} \tag{17}$$

where $p(n)$ is free parameter at iteration n and
η is learning rate parameter.

These partial derivatives are shown in Eqs (18-20),

$$\frac{\partial \varepsilon_k(n)}{\partial w_{ki}(n)} = -e_k(n) G(\|\mathbf{x}(n)-\mathbf{t}_i(n)\|_{\mathbf{C}_i}), \quad k = 1, 2, 3, 4;\ i = 1, 2, 3, ..., M_k \tag{18}$$

$$\frac{\partial \varepsilon_k(n)}{\partial \mathbf{t}_i(n)} = 2 w_{ki}(n) e_k(n) G'(\|\mathbf{x}(n)-\mathbf{t}_i(n)\|_{\mathbf{C}_i}) \boldsymbol{\Sigma}_i^{-1}(n)\left[\mathbf{x}(n)-\mathbf{t}_i(n)\right], \quad k = 1, 2, 3, 4;\ i = 1, 2, 3, ..., M_k \tag{19}$$

$$\frac{\partial \varepsilon_k(n)}{\partial \boldsymbol{\Sigma}_i^{-1}(n)} = -w_{ki}(n) e_k(n) G'(\|\mathbf{x}(n)-\mathbf{t}_i(n)\|_{\mathbf{C}_i}) \mathbf{Q}_i(n), \quad k = 1, 2, 3, 4;\ i = 1, 2, 3, ..., M_k \tag{20}$$

where $G'(.)$ is the first derivative of the Green's function $G(.)$ with respect to its argument and

$$\mathbf{Q}_i(n) = \left[\mathbf{x}(n)-\mathbf{t}_i(n)\right]\left[\mathbf{x}(n)-\mathbf{t}_i(n)\right]^T. \tag{21}$$

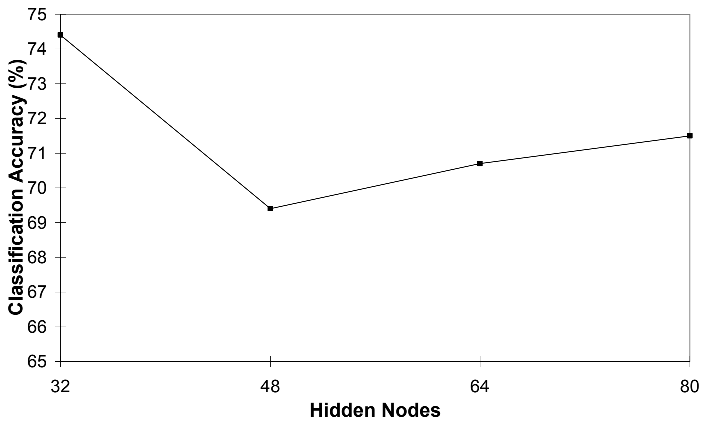


Figure 3. *Classification results of partially connected radial-basis function network with supervised selection of centres (centres in each sub-network are selected from all input classes).*

Two sets of experiments are conducted under this structure type of network. In the first set of experiment, the centres in each sub-network are chosen from all four pattern classes with equal number of centres from each pattern. The number of hidden nodes in each sub-network is chosen to be 8, 12, 16 and 20. In other words, the total number of hidden nodes in the entire network in each case would be 32, 48, 64 and 80, respectively. The classification results are shown in Fig. 3.

Experiment results show that the optimum number of hidden nodes in this case is 32 with the highest classification accuracy of 74.4 %.

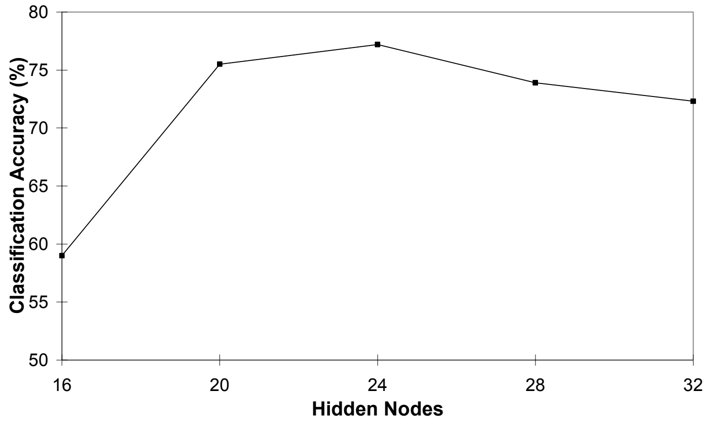


Figure 4. *Classification results of partially connected radial-basis function network with supervised selection of centres. (centres in each sub-network are selected from one input class only).*

In the second set of experiment, the centres in each sub-network are chosen from only one pattern class. The number of hidden nodes in each sub-network is chosen to be in the range of 4 to 8. This is corresponding to the total hidden nodes of 16, 20, 24, 28, and 32. The classification result is shown in Fig. 4. Experiment results show that the optimum number of hidden nodes in this case is 24 with the highest classification accuracy of 77.2 %.

3.1.2 Fully Connected Network. In this topology, hidden layer is fully connected to output layer. The cost function in this case is given in Eq. (22),

$$\varepsilon(n) = \frac{1}{2}\sum_{k=1}^{N_o} e_k^2(n) \tag{22}$$

where $\varepsilon(n)$ is the instantaneous cost function at iteration n,
$e_k(n)$ is the error from output node k at iteration n and
N_o is the number of output nodes.

The error from each output node is defined as follows,

$$e_k(n) = d_k(n) - \sum_{i=1}^{M} w_{ki}(n) G(\|\mathbf{x}(n) - \mathbf{t}_i(n)\|_{\mathbf{C}_i}), \qquad k = 1, 2, 3, 4 \tag{23}$$

where M is the number of Green's function in the network.

Steepest descent algorithm is used to adapt free parameters in the network. Partial derivatives of free parameters are given in Eqs (24-26),

$$\frac{\partial\varepsilon(n)}{\partial w_{ki}(n)} = -e_k(n) G(\|\mathbf{x}(n) - \mathbf{t}_i(n)\|_{\mathbf{C}_i}), \qquad k = 1, 2, 3, 4;\ i = 1, 2, 3, ..., M \tag{24}$$

$$\frac{\partial\varepsilon(n)}{\partial \mathbf{t}_i(n)} = 2\sum_{k=1}^{N_o} w_{ki}(n) e_k(n) G'(\|\mathbf{x}(n) - \mathbf{t}_i(n)\|_{\mathbf{C}_i}) \boldsymbol{\Sigma}_i^{-1}(n)\left[\mathbf{x}(n) - \mathbf{t}_i(n)\right], \qquad i = 1, 2, 3, ..., M \tag{25}$$

$$\frac{\partial\varepsilon(n)}{\partial \boldsymbol{\Sigma}_i^{-1}(n)} = -\sum_{k=1}^{N_o} w_{ki}(n) e_k(n) G'(\|\mathbf{x}(n) - \mathbf{t}_i(n)\|_{\mathbf{C}_i}) \mathbf{Q}_i(n), \qquad i = 1, 2, 3, ..., M. \tag{26}$$

For this structure type of network, all learning rate parameters are set to 0.1 as well. The number of hidden nodes is chosen to be 8, 12, 16 and 20. The centres are selected from all pattern classes with equal number of centres from each pattern. The classification results are shown in Fig. 5.

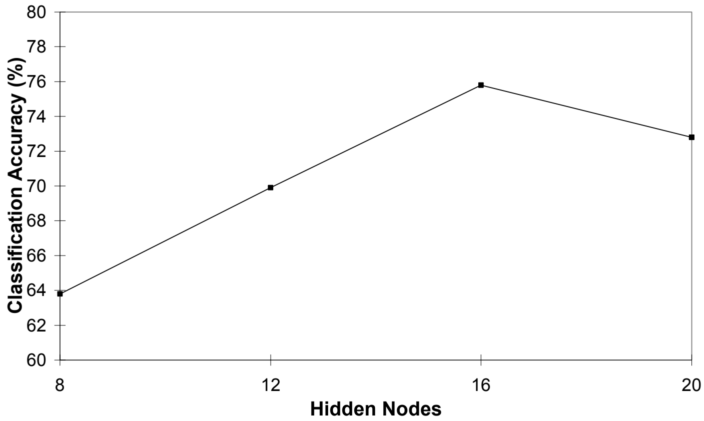


Figure 5. *Classification results of fully connected radial-basis function network with supervised selection of centres.*

From experiment results, the optimum number of hidden nodes is 16 with the highest classification accuracy of 75.8 %.

3.2 Fixed Centres Selected at Random

In this learning strategy, the centres of radial-basis function are fixed. These centres are selected at random from all pattern classes with equal number of centres from each pattern. In this case, the network is fully connected. The Green's function is defined as shown in Eq. (27),

$$G(\|\mathbf{x}-\mathbf{t}_i\|^2) = \exp(-\frac{M}{d^2}\|\mathbf{x}-\mathbf{t}_i\|^2) \tag{27}$$

where M is the number of centres and
d is the maximum distance between the chosen centres.

The only parameter that undergoes supervised learning is the weight in the network. Haykin (1994) suggests that one straightforward procedure for finding the weight of the network is to use pseudoinverse method. The weight solution is given by Eq. (28),

$$\mathbf{w} = \mathbf{G}^{+}\mathbf{d} \tag{28}$$

where $\mathbf{d}$ is the desired response vector in the training set and
$\mathbf{G}^{+}$ is the pseudoinverse matrix of matrix **G**.

The matrix **G** is defined in Eqs (29-30),

$$\mathbf{G}^{+} = \{g_{ni}\} \tag{29}$$

$$g_{ni} = \exp(-\frac{M}{d^2}\|\mathbf{x}(n)-\mathbf{t}_i\|^2), \qquad n = 1, 2, 3,..., N;\ i= 1, 2, 3,..., M \tag{30}$$

where N is the total number of training input pattern.

One way of solving Eq. (28) is to use recursive least square (RLS) algorithm. In this experiment, the number of hidden nodes is chosen to be 8, 12, 16, and 20. The classification results are shown in Fig. 6.

All experiment results in this case are over 99 % classification accuracy. The optimum number of hidden nodes is 8 with classification accuracy of 99.0 %.

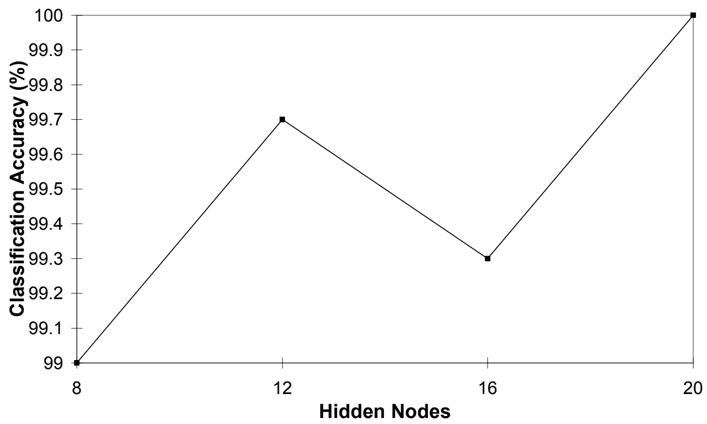


Figure 6. *Classification results of radial-basis function network with fixed centres selected at random.*

4. CONCLUSIONS

Radial-basis function network with fixed centre selected at random has the highest classification accuracy. This is true even with the number of hidden nodes as small as 8. One possible reason that radial-basis function network with supervised selection of centres cannot generalise as well as the one with fixed centres is that the centres of network are clustered together during training process. These centres cannot distribute themselves to cover the necessary pattern space for correct classification. This is a normal effect which is caused by steepest descent algorithm. Multilayer perceptron still be able to generalise with very high classification accuracy. Compare with radial-basis function network with supervised selection of centres, multilayer perceptron still be a better choice in this application. Since radial-basis function network with fixed centres does not require a bigger network structure or computational effort to outmatch multilayer perceptron, it should be considered for the possibility of microprocessor implementation in prosthesis devices.

5. REFERENCES

J D Costa and R E Gander (1993), MES classification using artificial neural networks and chaos theory, *Proc. Intl. Joint Conf. Neural Net.,* **3**, Nagoya, Japan, pp. 2243-2246.

A Del Boca and D C Park (1994), Myoelectric signal recognition using fuzzy clustering and artificial neural networks in real time, *Proc. IEEE Intl. Conf. Neural Net.,* **5**, Orlando, FL, pp. 3098-3103.

P C Doerschuk, D E Gustafson and A S Willsky (1983), Upper extremity limb function discrimination using EMG signal analysis, *IEEE Trans. Biomed. Eng.,* **30**, *1*, pp. 18-29.

D Graupe and K W Cline (1975), Functional separation of EMG signal via ARMA identification methods for prosthesis control purposes, *IEEE Trans. Sys. Man Cyber.,* **5**, *2*, pp. 252-259.

D Graupe, J Salahi and D Zhang (1985), Stochastic analysis of myoelectric temporal signatures for multifunction single-site activation of prostheses and orthoses, *J. Biomed. Eng.,* **7**, *1*, pp. 18-29.

S Haykin (1994), *Neural Networks: A Comprehensive Foundation,* Macmillan, New York.

A Hiraiwa, K Shimohara and Y Tokunaga (1989), EMG pattern analysis and classification by neural network, *Proc. IEEE Intl. Conf. Sys. Man Cyber.,* **3**, Cambridge, MA, pp. 1113-1115.

B Hudgins, P A Parker and R N Scott (1993), A new strategy for multifunction myoelectric control, *IEEE Trans. Biomed. Eng.,* **40**, *1*, pp. 82-94.

K Ito, T Tsuji, A Kato and M Ito (1991), Limb-function discrimination using EMG signals by neural network and application to prosthetic forearm control, *IEEE Intl. Joint Conf. Neural Net.*, Singapore, pp. 1214-1219.

B Karlik, H Pastaci and M Korürek (1994), Myoelectric neural networks signal analysis, *Proc. 7th Mediter. Electrotech. Conf.,* **1**, Antalya, Turkey, pp. 262-264.

M Kelly, P A Parker and R N Scott (1990), The application of neural networks to myoelectric signal analysis: A preliminary study, *IEEE Trans. Biomed. Eng.*, **37**, *3,* pp. 221-227.

T Kiryu, C J De Luca and Y Saitoh (1994), AR modelling of myoelectric interference signals during a ramp contraction, *IEEE Trans. Biomed. Eng.,* **41**, *11*, pp. 1031-1038.

N Kumaravel and V Kavitha (1994), Automatic diagnosis of neuro-muscular diseases using neural network, *Biomed. Sci. Instrum.,* **30***,* pp. 245-250.

U Kuruganti, B Hudgins and R N Scott (1995), Two-channel enhancement of a multifunction control system, *IEEE Trans. Biomed. Eng.,* **42**, *1*, pp. 109-111.

A Latwesen and P E Patterson (1994), Identification of lower arm motions using the EMG signals of shoulder muscles, *Med. Eng. Phy.,* **16**, *2*, pp. 113-121.

R Merletti and L R Lo Conte (1995), Advances in processing of surface myoelectric signals: Part 1, *Med. Biolog. Eng. Comp.,* **33***,* pp. 362-372.

R N Scott and P A Parker (1988), Myoelectric prostheses: State of the art, *J. Med. Eng. Tech.,* **12***,* pp. 143-151.

E C Yeh, W P Chung, R C Chan, and C C Tseng (1993), Development of neural network controller for below-elbow prosthesis using single-chip microcontroller, *Biomed. Eng. Appl. Bas. Comm.,* **5**, *3*, pp. 340-346.

M Zardoshti-Kermani, B C Wheeler, K Badie and R M Hashemi (1995), EMG feature selection for movement control of a cybernetic arm, *Cyber. Sys. Intl. J.,* **26**, *2*, pp. 189-210.

CCD-Camera Based Optical Tracking for Human-Computer Interaction

Franz Madritsch
Inst. f. Computer Graphics, Graz Univ. of Technology
Münzgrabenstr. 11, A-8010 Graz, AUSTRIA
(franz@icg.tu-graz.ac.at)

ABSTRACT

We are investigating the application of CCD-camera based optical-beacon-tracking-systems in 3-d interactive environments. An optical tracking system has been developed which serves as a testbed for tracking algorithms and accuracy investigations. The 3-d interactive environment features tracking of the observer's viewpoints, stereo-scopic visualization and direct 3-d pointing. The focus is set on high accuracy of both, tracking and stereo-scopic visualization. Algorithms which can track optical beacons with sub-pixel resolution at low noise help to reduce hardware expenditure in terms of camera resolution and computing power. Theoretical considerations about resolution are made and practical experience is presented. Furthermore, user studies are performed to test the created interface environment with regard to immersive interfaces and direct interaction in space.

Keywords: optical tracking, interactive visualization in 3-d, input technology, fish tank VR

1. INTRODUCTION

Optical Tracking has already found a variety of applications for specific tracking problems. However, reports about its usage for standard tracking tasks in human computer interfaces are sparse. The main argument against optical methods is poor economy. Viewpoint, head, 3-d cursor tracking, etc. are dominated by commercially available magnetic, acoustic, and mechanical trackers. Deering [Deering, 1992] outlined the demands to create a high fidelity visual interface environment. The requirements concerning tracking accuracy are not met very well with todays magnetic and acoustic tracking devices. Magnetic systems have to deal with high noise and distortion problems especially in the vicinity of CRTs. This is of some importance because of the high retinal resolution achieved by stereo CRT compared to other current VE display technologies [Ware et al., 1993]. Intensive research on optical tracking is done at the University of North Carolina, where several high quality systems have been developed [Ward et al., 1992]. OBT (Optical Beacon Tracking) already has found applications in commercial motion capture systems [AOA, 1994]. Most of them do recording rather than tracking and compute 3-d coordinates off-line. For a comprehensive survey of optical and other tracking systems see also [Meyer et al., 1992].

We shall present a simple optical tracking method which performs exceptionally well in terms of noise, accuracy, and registration. The system is based on optical beacons which are tracked by two fixed color CCD-cameras. Through adapted algorithms, known as centroid calculation in photogrammetry, sub-pixel resolution is achieved [Trinder, 1989, DeMenthon and Fujii, 1994]. These algorithms are computationally inexpensive and lead to a good real-time behavior. Especially for so called "Fish Tank VR" where the user interacts within a limited space in front of the screen, a pair of ceiling mounted cameras can overview the working area perfectly. Positions of interest like shutter-glasses or 3-d cursors can be marked/tracked robustly by e.g. LEDs (light emitting diodes). Through careful determination of the geometric parameters of the cameras and the display screen and their relative alignment to each other, the loop from the eye-points to tracker coordinates and spatial visualization is closed. Accuracies of one part in 3000 are obtained although images with 640 x 480 pixels are used.

2. OPTICAL TRACKING SYSTEM

From the beginning on, we want to avoid specialized hardware and focus our work on the development of fast algorithms. Secondly we want to enable economic software solutions. The system configuration consists of two graphics workstations with live video input, two color-CCD-cameras, and one high resolution computer screen. To avoid high-cost image processing, objects to be tracked are marked with LEDs as optical beacons. Two cameras are fixed at certain positions to overview a defined volume of interest. The whole system rests on the ability to distinguish the beacons from the background and to determine their position within the image. The method of centroid calculation was found to perform this task in a stable and precise manner. The center of a target is defined as the center of gravity of pixels above a certain threshold. Through the right choice of these thresholds in 'RGB' values, red LEDs can be identified robustly in a slightly darkened environment. The centroid calculation is fast and leads to sub-pixel image coordinates of the targets [Madritsch et al., 1996].

Several prototypes were built. Currently our system consists of two CCD-cameras capturing the position of several red LEDs. The position of the beacons in 3-d space is reconstructed out of the two 2-d images provided by the cameras. In principle the position and orientation of the two cameras can be chosen freely; however, to achieve a proper working volume and a precisely measurable geometric configuration, the following setup was found useful.

2.1 Geometric Camera Configuration

The camera coordinate systems are defined so that the y-axes are coincident with the optical axes and the z-axes represent the cameras' view-up vectors. The x-axes complete the right handed Cartesian coordinate systems. The geometry as depicted in Fig. 1 shows the relationship between the two camera coordinate systems, beacons, and the working volume. The motivation for such a set-up is to keep the geometric parameters simple and precisely determinable. To align the xy-planes and the y-axes here, the images of the on-line cameras or even 2-d tracking data can be used for feedback. As a result, the summation of errors is avoided. After that adjustment only three parameters have to be determined by measurement, namely the angle α and the distances from the cameras' first nodal points to their common viewing centers a and b respectively.

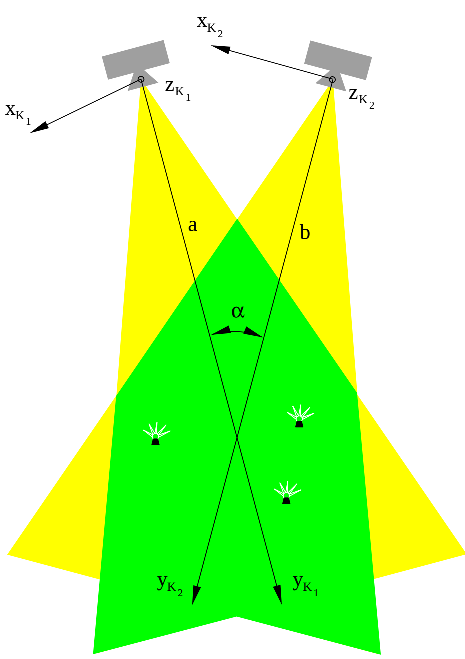


Figure 1: Two cameras with overlapping viewing volume are used to determine the position of several optical beacons

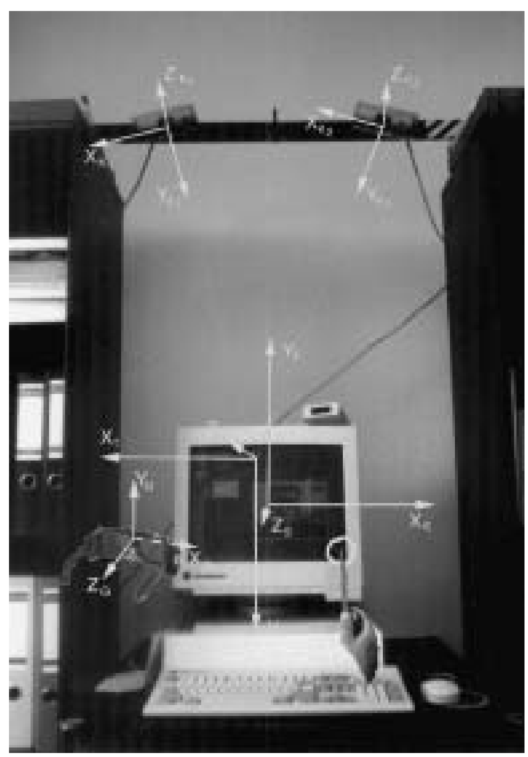


Figure 2: The human-computer interface environment; Two CCD-cameras track the observer's viewpoints and a 3-d pointing device. Visual output via stereo-scopic display.

The position of a beacon in space is determined by inverting the perspective projection equations of each camera [Madritsch and Gervautz, 1996]. Through proper calibration optical distortions of the cameras can be mostly eliminated [Madritsch and Gervautz, 1996, Weng et al., 1992].

2.2 *Tracking Algorithm*

In general optical tracking of objects is a complex problem the pattern recognition and active vision community has been dealing with for years [Blake and Yuille, 1992, Blake et al., 1993, Broeckl-Fox et al., 1993]. Through easily distinguishable and identifiable beacons marking the objects to be tracked, a method called Optical Beacon Tracking is enabled [Madritsch, 1995, Madritsch and Gervautz, 1996]. Algorithms which are fast and robust can be employed to extract the projections of such beacons out of real-time digital images [Madritsch et al., 1996].

The fish-tank VR system which is described here works with red LEDs as beacons. Two LEDs are placed on top of the stereo-glasses, a third serves as a 3-d pointer. The task of the tracking process is to determine the image coordinates of 3 LEDs within the live-video images. A neighborhood search algorithm is employed in a way similar to the method described in [Madritsch et al., 1996]. A quadratic region of fixed side length is searched. As an estimate of the beacon position the center of the search-region is located at the last known beacon position. So far no prediction is used. In the basic step the whole image is searched till all three beacons are found. In subsequent steps corresponding to subsequent frames only the neighborhood search-region is considered first. If the neighborhood search fails for one or more beacons, the remaining image is searched line by line causing a temporal distortion in the flow of the tracking data due to the increased search time. Fig. 3 shows an image of three beacons as seen by one camera.

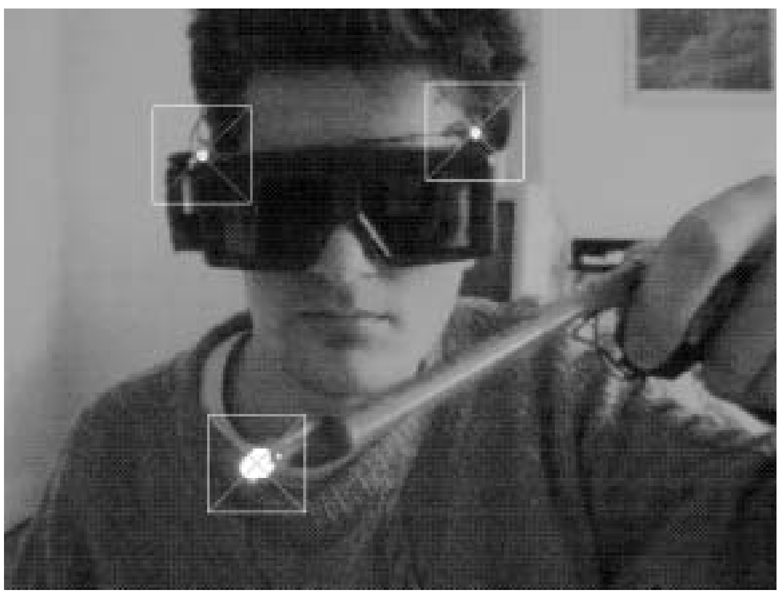


Figure 3: Optical beacons from the viewpoint of the camera. Extracted centroids deliver sub-pixel image coordinates.

The size of the neighborhood search-region influences both the maximum trackable beacon speed and the search time and consequently the update rate. Reflections about the optimal search-region size and alternative algorithms have been made in [Madritsch et al., 1996].

In the fish-tank application there are two kinds of beacons with different temporal behavior; the two LEDs connected to the shutter-glasses worn by the user, and one LED on top of the 3-d pointer. From data recorded during several tests and our practical experience during the evaluation of our prototype we found that the neighborhood assumption is usually correct. However, rapid user motions may cause a temporal loss of the beacons. This case occurs when the 3-d pointer is moved abruptly rather than when the user's head is moved. The spatial impression of the scene is usually maintained steadily.

2.3 *Software system configuration*

Within our environment a usual optical tracking task consists of three processes: Two identical image processing tasks extract the beacons in a neighborhood search and perform the centroid calculation (see Tracking Algorithm). One further process combines the image coordinate pairs to obtain Cartesian 3-d tracker coordinates and delivers position and orientation data to an update and rendering process. In our implementation the image processing runs on two separate workstations. To avoid the transmission of images over the network image processing is done locally on the workstation where the frame-grabbing occurs. All interprocess-communication is currently done by Berkeley datagram sockets. Such data links are fast but not completely safe. Receiving real-time position data without delay is, however, usually as important as its correctness.

2.4 Accuracy

In OBT-systems, as in most other optical systems, the absolute accuracy is determined by the system's "resolution" and the working volume because accuracy can always be scaled by the focal length of the cameras. The term "resolution" must be treated with care in OBT because the sizes of the tracking steps vary. We would like to define resolution as the reciprocal value of the uncertainty multiplied by the range. (and not by the number of distinguishable steps) There can be big discrepancies between these definitions in OBT. The resolution of a OBT-system is given by the resolution of the camera multiplied by a sub-pixel factor. Through centroid calculation the image coordinates of the beacons can be determined with sub-pixel accuracy. The value of the sub-pixel factor is mainly dependent on the relative beacon size within the digital image. In the fish-tank application we worked with a sub-pixel factor of about 4 to 6. Fig. 4 shows how a sub-pixel region is resolved by unweighted discrete centroid calculation for a circular target of 4 pixels in diameter. This pattern is the result of a simulation. A detailed investigation of the discrete centroid calculation together with results from physical measurements can be found in [Madritsch et al., 1996].

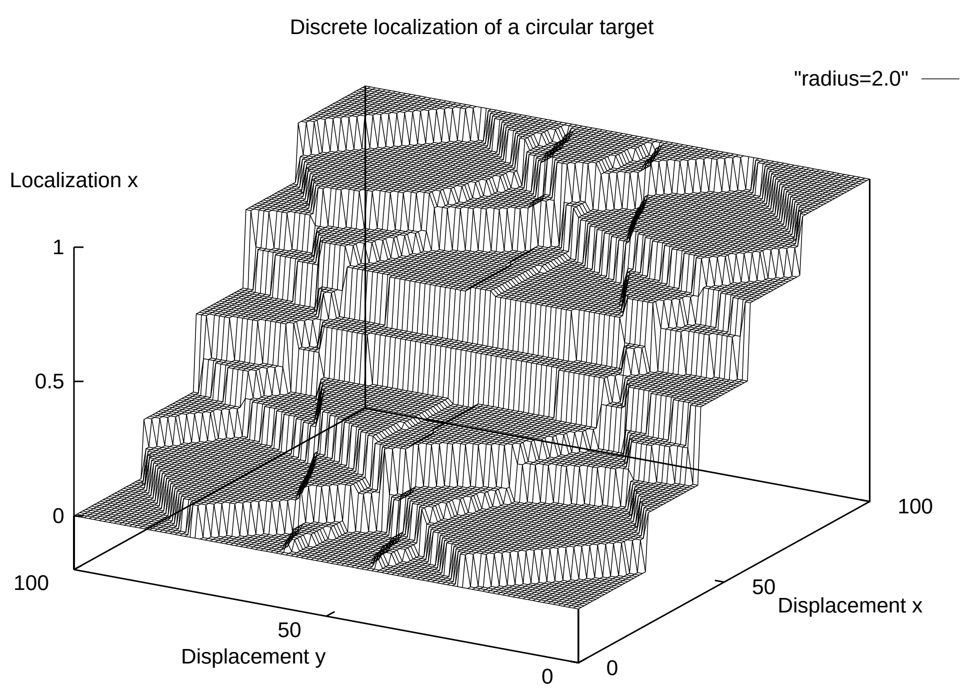


Figure 4: Simulation of target localization by discrete centroid calculation for a relative target diameter of 4 pixels (100 x 100 steps; "Localization x" in grid units)

Using CCD-cameras with a resolution of 640 x 480 pixels a two dimensional tracker resolution of 3200 x 2400 is reached. The resolution in the third coordinate direction is similar but differently scaled depending on the geometric camera arrangement. Our configuration as depicted in Fig. 1 with the parameters $a = 85cm; \quad b = 85cm; \quad \alpha = 36°$ led to tracker accuracy values of 0.17 x 0.18 x 0.6 mm RMS. The absolute registration accuracy between tracker and spatial visualization is ± 1 mm as evaluated by visual registration tests.

3. SPATIAL COMPUTER INTERFACE

Interfaces with 3-dimensional input and output capabilities are becoming more and more common. Perhaps the biggest shift in public consciousness happened when J.D. Foley published their work at NASA [Foley, 1987]. Since the extensive press and media coverage in 1990-1991 the term "Virtual Reality" has become common currency. The historical evolution and state of the art is summarized in [Krueger, 1990, Rheingold, 1991, Ellis, 1991, Kalawsky, 1993]. Stereoscopic visualization and 3-d pointing are small fields within this huge area but nevertheless key technologies in 3-d interfaces. The fish-tank concept used here for spatial visualization is not immersive but offers high retinal resolution to the user and meets the criteria for high definition VR as outlined by [Deering, 1992] quite well. The tracking task done by OBT also fits in, although there is room for improvements in temporal performance as well as in the performance of the graphics workstation. In the following sections a spatial visualization system will be outlined which generates a limited volume where virtual objects can be placed and viewed by one observer wearing tracked stereo-glasses. Furthermore a tracked 3-d pointer enables the observer not only to view the scene passively but also to interact with it in a certain way. Let us first take a look at perceptual issues of spatial visualization.

3.1 Depth Perception and Stereoscopic Viewing

Mostly based on our visual sense the human perceptual system builds a "model in mind" of the 3-dimensional world around us. There are various visual depth cues which allow us to estimate distances and spatial relations with different certainty and reliability. Motion parallax, stereopsis, linear perspective, size constancy, and accommodation are some of the most important visual depth cues [Storey and Craine, 1985]. Most stereoscopic viewing techniques cannot reproduce all of them. In the case of stereoscopic color displays accommodation is always bound to the distance between the eyes of the observer and the screen in order to produce sharp images on the retinae. Stereopsis is achieved by presenting each eye with an image in a perspectively correct way. By using perspectively correct images the "size constancy" depth cue is also reproduced. To calculate correct stereo pairs the position of the observer relatively to the display device must be known. This position must either be assumed to be constant or tracked [Meyer et al., 1992]. In the second case, the rendering routine can use real-time information about the observer's viewpoints to produce current stereo-pairs on-line. This method further enables the depth cue of observer induced motion parallax. Other depth cues such as lighting, shading, atmospheric effects, hidden surface removal, etc. are matters of scene description and graphics performance. In conclusion it can be said that a color stereo screen combined with tracked goggles can principally present the observer with an impression which has almost all visual properties a real object or scene would have. One problem that cannot be solved in this way is the accommodation-convergence breakdown. The eyes converge at the point in space where the virtual object appears while still focused to the screen. This effect troubles the learned coordination between the muscles moving the eyes and the ciliar muscles focusing the lenses. The result of this discrepancy is inconvenience of the viewer when looking at stereo pairs with big vertical parallaxes. Therefore stereo-visualizations should inhabit only a limited volume in front of and behind the screen. According to [Lipton, 1993] both positive and negative parallax should not exceed 1,5°. Other restrictions are finite resolution, imperfect color and brightness reproduction and limited scene detail, as well as temporal performance restrictions. Mathematics of stereo image computations are given in [Deering, 1992, Hodges and McAllister, 1993].

The importance of different depth cues was studied by [Ware et al., 1993]. In their investigation, several subjects were asked to trace 3-dimensional tree structures from certain branches to the corresponding root. The viewing conditions were: monocular, stereo only, head-coupled only, stereo and head-coupled. Although the test persons' subjective opinion preferred head-coupled only over head-coupled and stereo, the timing and error data of the experiments showed that performance increased from stereo to head-coupled only and was best with head-coupled stereo.

Especially in the grasp-range of a user the binocular disparity and vergence (stereopsis) provide an almost absolute cue for depth estimation. Short range depth estimation is used in every-day life for every grasping action and therefore this ability is highly trained. The advantage of stereopsis compared to motion parallax is the instant spatial impression provided without the need of an exploring motion.

3.2 Position of the user's eye-points

For precise registration of virtual and physical objects the position of the first nodal points of the eyes has to be tracked in an ideal case [Deering, 1992]. On the other hand in [Madritsch, 1995] it has been shown that very little information, namely only the direction vector from the display to the user is sufficient to produce eye-point corrected stereoscopic image pairs. In practice it proves to be difficult and technically expensive to track the eye-points (an additional tracker for gaze direction would be necessary [Starks, 1991]). The knowledge of the gaze direction only adds little to the precision of the position of the eye-points since the center of rotation of the eye and its first nodal point lie closely together and the user looks straight at the screen rather than with her/his head turned away from it. For our fish tank application we use the eye-point model described in [Madritsch et al., 1996]. Two LEDs connected to the stereo-goggles are tracked 3-dimensionally. The geometric relationship between the markers and the eye-points is defined in a head aligned coordinate system, see Fig. 5. The vectors from the LEDs to the eye-points are determined from an average head size. Due to the fact that the transformation from tracker to head coordinates cannot be stated completely by two tracked points, the user's head is assumed to be in an upright position. Nodding motions would induce errors depending on the separation between the beacons and the eyes. In practice this error can be neglected.

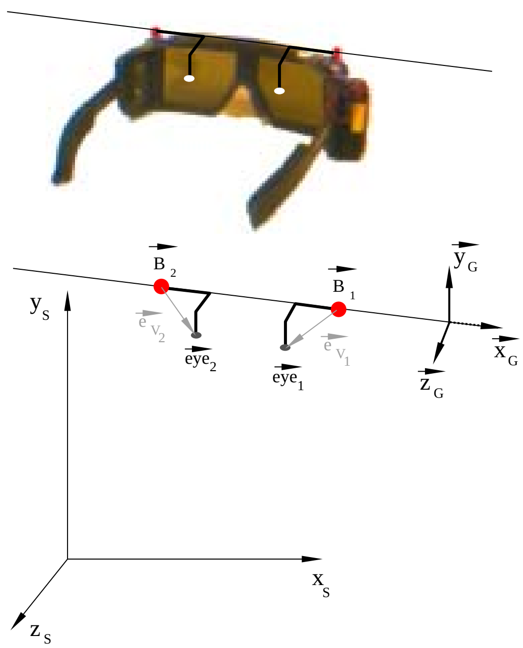


Figure 5: Geometric model of the relation between beacons and eyes

$$\vec{eye_i} = \vec{B_i} + \vec{e_{v_i}} \qquad i = 1, 2 \tag{1}$$

with

$$\vec{e_{v_i}} = M\vec{e_{v_i}} \tag{2}$$

The transformation matrix M is obtained by expressing the goggle coordinate system's basis vectors $\vec{x_G}$, $\vec{y_G}$, and $\vec{z_G}$ in system coordinates

$$R = (\vec{x_G} \quad \vec{y_G} \quad \vec{z_G}) \tag{3}$$

where

$$\vec{x_G} = \frac{\vec{B_1} - \vec{B_2}}{\left|\vec{B_1} - \vec{B_2}\right|}, \tag{4}$$

$$\vec{y_G} = \vec{z_G} \times \vec{x_G}, \tag{5}$$

$$\vec{z_G} = \left(-\frac{x_{G_z}}{\sqrt{x_{G_x}^2 + x_{G_z}^2}} \quad 0 \quad \frac{x_{G_x}}{\sqrt{x_{G_x}^2 + x_{G_z}^2}} \right)^T . \tag{6}$$

Note that $\vec{y_G}$ is kept aligned with the $y_S - axis$ due to the missing rotational information around the $x_S - axis$.

3.3 Registration of virtual and physical objects

To calculate current perspectively correct image pairs the position of the eye-points and visualization coordinates must be expressed in the same coordinate system. We chose a screen oriented coordinate system as the common basis for all coordinates. Such coordinates will be referred to as system coordinates.

A priory the transformation between tracker coordinates and visualization coordinates is unknown. In our case where the tracker supports 3-d pointing it suggests itself to establish a connection between these coordinate systems by selecting certain points on the display as anchors for the transformation. We use the 3 corners P_1, P_2, P_3 of the usable screen area as anchor points, see Fig. 6. Since these anchor points also correspond to the physical viewport limits they can be used to scale the visualization to physical units, provided that the output data of the tracker is calibrated properly.

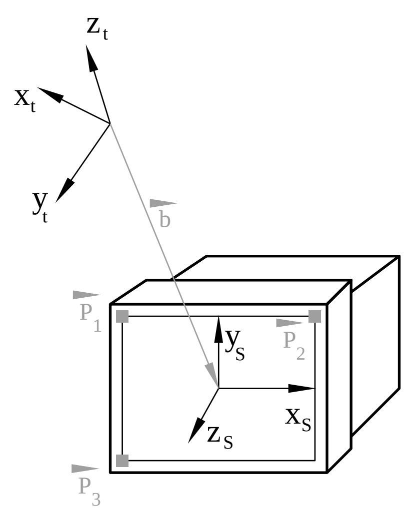


Figure 6: Geometric model of the relation between tracker coordinates and visualization (system) coordinates. For simplicity all coordinates (tracking related and visualization related) are transformed into system coordinates.

Tracker coordinates transform into system coordinates by

$$\vec{P_S} = S^T \vec{P_t} - S^T \vec{b} \tag{7}$$

with

$$S = (\vec{x_s} \quad \vec{y_s} \vec{z_s}), \qquad \vec{b} = \frac{\left(\vec{P_2} + \vec{P_3}\right)}{2}, \tag{8}$$

and

$$\vec{x_S} = \frac{\vec{P_2} - \vec{P_1}}{\left|\vec{P_2} - \vec{P_1}\right|}, \qquad \vec{y_S} = \frac{\vec{P_1} - \vec{P_3}}{\left|\vec{P_1} - \vec{P_3}\right|}, \qquad \vec{z_S} = \vec{x_S} \times \vec{y_S}. \tag{9}$$

This semiautomatic "self-adjustment" is done by the user by pointing at the corners P_1, P_2, P_3. It is only necessary after changes in the arrangement between display device and tracker have occurred.

3.4 Interacting in 3D

Venolia [Venolia, 1993] showed how to build up an interface for direct interaction in 3-d using only a minimum set of interface elements. The system was based on the "roller mouse" (a mouse enhanced with additional wheels to control the third dimension as well). Various 3-d input devices and techniques such as a Bat (flying mouse) [Ware and Jessome, 1988], a data glove [Sturman and Zeltzer, 1994] or a SPACEBALL[1] [Zhai, 1993], just to mention a few, have been investigated. This section will mainly deal with 3-d pointing and some of our experience with this kind of interface.

The system described above establishes a "virtual volume" where computer generated 3-d objects can be placed at will. This volume is a limited space in front of and behind the display screen. The condition for stereoscopic visibility is that from the view of each eye the entire virtual object can be "projected" onto this screen. A three dimensionally tracked stylus allows the user to interact directly within this volume.

3.4.1 3D Pointing We use an LED mounted on a pencil together with the supplying battery as a 3-d pointing device. The position of the LED is tracked and available in system coordinates. A virtual cursor can be placed at the identical position in space. (In this way the registration of physical and virtual coordinates can be checked easily.) In an application where a racket of a squash game was controlled by the 3-d pointer directly the end of the virtual handle was aligned with the pointer. It turned out to be disturbing when the virtual representation of the pointer was obscured by its physical counterpart. If proper registration is given no representation of the physical pointer is needed. But a virtual representation gives confidence to the user that she/he really points to the place in the virtual scene where it appears. This feedback seems to be helpful especially to novice users who are usually surprised that they can really "touch" a virtual object. So we found it useful to simply extend the physical pointer by a virtual top. By further extending the physical pointer the space behind the visualization screen can also be used for interaction, thereby increasing the working volume. It is not completely clear whether a direct one to one correspondence between physical and virtual pointer improves the user's pointing performance. The user might, for example, also control the virtual cursor remotely in a way similar to the situation of controlling a 2-d cursor by a 2-d "mouse". However, direct interaction seems to be more intuitive.

3.4.2 Interactions We implemented several applications utilizing the fish tank VR environment to perform simple interaction experiments. The 3-d squash game mentioned above was one of them. This example where the user can only hit a bouncing ball by a racket controlled by the 3-d pointer showed that a naive user can understand and adapt to such a scenario without any instruction.

As a student project a simple 3-d editor was designed. Models consisting of quaders and ellipsoids were created, stored, and reloaded. Interactions such as moving, scaling, and rotating were supported. Both-handed input was used. One hand controlled the stylus, one hand the "mouse" to select different modes. Manipulations were performed by selecting an object with the pointer and choosing either the move, the scale, or the rotation mode. Moving was done by connecting the center of mass of the object with the pointer and releasing the object at the desired place. Scaling was done by controlling a corner of the object's bounding box while its center stayed fixed. To control the rotation over all 3 axes simultaneously by the pointer is not so straight forward since the "handle on the surface" concept only allowed rotations around 2 axes. In a second approach the mouse which was used as a selection device was replaced by a SPACE MOUSE[2] (6 degree of freedom input device similar to a SPACEBALL™). It improved user performance to select, move, and scale objects directly with the pointer and rotate objects and the scene directly with the SPACE MOUSE ™. We learned that such an interface allows very intuitive handling and manipulation of the objects, requiring little instruction for novice users.

In our opinion both-handed input should be used in such an environment and the user's hands should stay at the input devices. There should be no need of looking away from the center of visual attention while interacting in 3-d. We think that visual discontinuities are even more disturbing in 3-d interfaces than they are in 2-d interfaces. Because of the high number of degrees of freedom it is difficult to position and align objects in space exactly. It is more difficult to align objects in depth [Zhai and Milgram, 1994]. Sometimes it is therefore useful to rotate the whole scene. There is a great need for methods to support the user.

The user tests were performed by observing and interviewing about 50 visitors the system was demonstrated to.

[1]SPACEBALL is a trademark of Spaceball Technologies, Inc.
[2]SPACE MOUSE is a trademark of Deutsche Forschungsanstalt für Luft- und Raumfahrt

4. PERFORMANCE FIGURES

resolution: 3200 x 2400 x 2000 volume elements
accuracy: 0.17 x 0.18 x 0.6 mm RMS
absolute registration: $\pm$ 1 mm
update rate: 30 $\frac{1}{s}$
lag: 40-60 ms
scene complexity: several hundred polygons flat shaded and wire-frame

The absolute registration deviations were determined by visually checking the alignment of a physical 3-d pointer and a virtual cursor within the closed tracking visualization loop of the fish-tank environment.

All results are based on a hardware configuration consisting of two workstations with MIPS™ R4600 100 MHz processors with live video input, two color CCD-cameras (640 x 480 RGB-pixels resampled out of 512 x 492 pixels), one SPACE MOUSE™, and one stereo monitor with LCD shutter-glasses.

5. CONCLUSIONS

The OBT concept offers high resolution and accuracy at acceptable temporal performance. The line of sight restriction of the optical system represents no problem for a fish tank VR configuration. It supports the tracking of the observer's viewpoint and a 3-d stylus robustly. Probably the main advantage of OBT compared to magnetic methods is that there is practically no noise in the tracking data (i.e. the noise is within the error bounds) as long as the neighborhood search works smoothly. Optical beacons are hardly intrusive. No cabling is needed for the goggles and the stylus.

A fish tank VR environment can serve as an intuitive three dimensional interface. We think that today it might be appropriate for naive users rather than for experts because simple tasks can be done very easily. But even in advanced VE systems it might turn out that the more complicated and complex the interface becomes the more of the advantages of a direct interface are lost.

6. ACKNOWLEDGMENTS

Many thanks to Stefan Brantner, Peter Duracher, Harald Ganster, Michael Gervautz, Michael Gruber, Konrad Karner, Franz Leberl, Klaus Wiltschi, and Andrea Zunegg for their support and contribution.

This research project was funded by the ADW under project number 6070/16-II/5/95.

7. REFERENCES

[AOA, 1994] (1994). *Capture Motion, Release Creativity (Trade Literature).* United Technologies Adaptive Optics Associates, 54 CambridgePark Drive, Cambridge, MA 02140-2308.

[Blake et al., 1993] Blake, A., Curwen, R., and Zisserman, A. (1993). A framework for spatio-temporal control in the tracking of visual contours. *Int. Journal of Computer Vision*, 11(2):127–145.

[Blake and Yuille, 1992] Blake, A. and Yuille, A. (1992). *Active Vision.* Massachusetts Institute of Technology Press.

[Broeckl-Fox et al., 1993] Broeckl-Fox, U., Hartensteiner, K., Kettner, L., Nolting, O., G.Selinger, Weiss, B., and Schmitt, A. (1993). Using three dimensional hand-gesture recognition as a new 3d input device. Technical report, Institut fuer Betriebs- und Dialogsysteme; Universitaet Karlsruhe, Am Fasangarten 5; Postfach 6980; Karlsruhe.

[Deering, 1992] Deering, M. (1992). High Resolution Virtual Reality. *Computer Graphics*, 26(2):195–202.

[DeMenthon and Fujii, 1994] DeMenthon, D. F. and Fujii, Y. (1994). Three dimensional pointing device monitored by computer vision. U.S. Patent 5,297,061, (Application 08/063489, May 1993).

[Ellis, 1991] Ellis, S. R., editor (1991). *Pictoral communication in virtual and real environments.* Taylor & Francis.

[Foley, 1987] Foley, J. D. (1987). Interfaces for advanced computing. *Scientific American*, 257(4).

[Hodges and McAllister, 1993] Hodges, L. F. and McAllister, D. F. (1993). *Stereo Computer Graphics and Other True 3D Technologies*, chapter 5, pages 71–89. Princeton University Press, 41 William Street, Princeton, New Jersey 08540.

[Kalawsky, 1993] Kalawsky, R. S. (1993). *The Science of Virtual Reality and Virtual Environments.* Addison-Wesley.

[Krueger, 1990] Krueger, M. W. (1990). *Artificial Reality II.* Addison-Wesley.

[Lipton, 1993] Lipton, J. (1993). *Composition of Electrostereoscopic Displays*, chapter 4, pages 11–25. Princeton University Press, 41 William Street, Princeton, New Jersey 08540.

[Madritsch, 1995] Madritsch, F. J. (1995). Correct spatial visualization using optical tracking. *Virtual Reality: Research, Development & Applications*, 1(2):122–126.

[Madritsch and Gervautz, 1996] Madritsch, F. J. and Gervautz, M. (1996). Ccd-camera based optical beacon tracking for virtual and augmented reality (in print). *COMPUTER GRAPHICS forum*, 15(3).

[Madritsch et al., 1996] Madritsch, F. J., Leberl, F. W., and Gervautz, M. (1996). Camera based beacon tracking: Accuracy and applications (in print). *ACM Transactions on Human-Computer Interaction*, 3(3).

[Meyer et al., 1992] Meyer, K., Applewhite, H. L., and Biocca, F. A. (1992). A Survey of Position Trackers. *PRESENCE*, 1(2):173–200.

[Rheingold, 1991] Rheingold, H. (1991). *Virtual Reality.* Summit Books / Simon & Schuster, New York.

[Starks, 1991] Starks, M. (1991). Stereoscopic video and the quest for virtual reality: an annotated bibliography of selected topics. *SPIE Proceedings, Stereoscopic Displays and Applications II*, 1457:327–342.

[Storey and Craine, 1985] Storey, N. and Craine, J. F. (1985). Interactive Stereoscopic Computer Graphics Display Systems. In Shackel, B., editor, *Human-Computer Interaction - INTERACT '84*, pages 163–168. Elsevier Science Publishers B.V., Amsterdam.

[Sturman and Zeltzer, 1994] Sturman, D. J. and Zeltzer, D. (1994). A survey of glove-based input. *IEEE Computer Graphics & Applications*, 14(1):30–39.

[Trinder, 1989] Trinder, J. C. (1989). Precision of Digital Target Location. *Photogrammetric Engineering and Remote Sensing*, 55(6):883–886.

[Venolia, 1993] Venolia, D. (1993). Facile 3D Direct Manipulation. In *INTERCHI '93*, pages 31–36.

[Ward et al., 1992] Ward, M., Azuma, R., Bennett, R., Gottschalk, S., and Fuchs, H. (1992). A demonstrated optical tracker with scalable work area for head-mounted display systems. In *SYMPOSIUM ON INTERACTIVE 3D GRAPHICS*, Special issue of Computer Graphics, pages 43–52. ACM SIGGRAPH.

[Ware et al., 1993] Ware, C., Arthur, K., and Booth, K. S. (1993). Fish Tank Virtual Reality. In *INTERCHI '93*, pages 37 – 42.

[Ware and Jessome, 1988] Ware, C. and Jessome, D. R. (1988). Using the Bat: a six-dimensional mouse for object placement. *IEEE Computer Graphics & Applications*, 8(6):65–70.

[Weng et al., 1992] Weng, J., Cohen, P., and Herniou, M. (1992). Camera calibration with distortion models and accuracy evaluation. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 14(10):965–980.

[Zhai, 1993] Zhai, S. (1993). Investigation of feel for 6dof inputs: isometric and elastic rate control for manipulation in 3d environments. In *PROCEEDINGS of the HUMAN FACTORS AND ERGONOMICS SOCIETY 37th ANNUAL MEETING.*

[Zhai and Milgram, 1994] Zhai, S. and Milgram, P. (1994). Asymmetrical spatial accuracy in 3d tracking. In *PROCEEDINGS of the HUMAN FACTORS AND ERGONOMICS SOCIETY 38th ANNUAL MEETING.*

Analysis of force-reflecting telerobotic systems for rehabilitation applications

William Harwin[1] and Tariq Rahman[2]

[1]Department of Cybernetics, The University of Reading,
Whiteknights, P.O. Box 225, Reading, ENGLAND

[2]Applied Science and Engineering Laboratories, University of
Delaware and Alfred I duPont Institute, 1600 Rockland Road,
Wilmington, Delaware, USA

[1]*w.s.harwin@reading.ac.uk* [2]*rahman@asel.udel.edu*

ABSTRACT

There is interest in a class of assistive technology devices for people with physical disabilities where a person's existing strength and movement has a direct relation to the force and position of the tool used to manipulate an environment. In this paper we explore the design of a head controlled force-reflecting master-slave telemanipulators for rehabilitation applications. A suitable interface philosophy is to allow the system to function in a way that is conceptually similar to a head-stick or mouth-stick. The result is an intuitive method to operate a rehabilitation robot that is readily learned, and has the ability to provide the person with added strength, range of movement and degrees of freedom. This approach is further expanded for a similar class of assistive devices, power assisted orthoses, that support and move the person's arm in a programmed way. The techniques developed for powered orthoses and telemanipulators can also be apply to haptic displays that allow an individual to feel a virtual environment. The so called two-port model is used to predict the behaviour of telemanipulators, powered orthoses. and haptic interfaces and issues relating to stability are discussed.

Keywords: Haptic, Telerobot, Telemanipulator, Spinal cord injury, Proprioception

1. INTRODUCTION

Two classes of assistive device that require intimate human involvement are telerobotic devices and power assisted orthoses. Telerobotic device are an appropriate technology for individuals with spinal cord injuries where the traumatic spinal damage affects both fine motor skills and sensory channels. When a spinal cord injury level is between C2 and C5 there may be some limited hand function, although typically the individual retains a near to normal range of head movement (Stanger 1994), thus head movements are an obvious candidate for telerobot operation. Individuals with higher-level spinal cord injuries often use head-sticks and mouth-sticks to manipulate their environment, and such devices can be considered as a very simple example of a telemanipulator. Telerobots may also be an appropriate technology for individuals with arthrogryposis where there is a need to map a limited range of input motion to a full range of motion at a remote site. Head-sticks mouth-sticks and telerobots form a class of assistive technology that utilise extended physiological proprioception (EPP) - a concept derived by Simpson (1977). Devices of this nature, which also include prosthetic limbs, allow the user to extend existing proprioceptive skills to the tip of the device and so readily conceptualise the movement of the tip in freespace. In addition the forces that are experienced at the end of the device are relayed back directly to the user. Head-sticks and mouth-sticks have an added advantage of being lightweight and highly rigid, and can therefore convey tactile and kinesthetic information from the environment with high bandwidth. Two limitations of head-sticks and mouth-sticks are that they have a limited workspace, both in positional and orientational degrees-of-freedom, and there is no possibility of increasing the mechanical power a user can transfer from the interface to the environment.

Individuals with muscular dystrophy maintain proprioceptive skills but lose muscle strength throughout the progression of the disease. One possible rehabilitation device that can be prescribed during the stage when an individual has insufficient strength to support his or her arms against gravity, is the balanced forearm orthosis, also known as the ball bearing feeder or the mobile arm support. Because the individual's proprioceptive skills are retained, the balanced forearm orthosis is highly successful as a rehabilitation device. A natural consideration for this class of rehabilitation device is whether the person's strength can be enhanced, or whether greater vertical movement might be possible, or whether an advantage is gained by allowing other mechanisms such as voice commands to assist in the movement of the person's arm. All three cases allow the person's proprioceptive abilities to be exploited to the full (Stroud, 1995; Rahman, 1995).

Proc. 1st Euro. Conf. Disability, Virtual Reality & Assoc. Tech., Maidenhead, UK, 1996

A third instance where proprioception plays an important role in a human machine interface is the haptic displays that are emerging to enhance an operator's interactions in virtual environments. Many immersive virtual reality environments utilise instrumented gloves to allow the individual to interact with virtual objects. Such an interface relies totally on visual feedback to indicate when an object has been grasped, whereas in real environments an individual also uses proprioceptive and tactile cues to determine the location and physical characteristics of an object. A haptic display can be used either in virtual environments or as a telerobotic master where force information is relayed back to the operator.

2. IMPLEMENTATION OF A HEAD OPERATED TELEROBOT

A test-bed head operated telerobotic system has been built at the Applied Science and Engineering Laboratories. It consists of two 6 degree of freedom master and slave robots, each controlled by separate IBM-PC 486DX/66 compatibles and linked with a high-speed parallel data link (figure 1). The master is the PerForce hand-controller (manufactured by Cybernet Systems of Ann Arbor, MI). The PerForce has been mechanically modified so that it is controlled by the user's head movements. PID position controllers are implemented digitally and these can be used under appropriate conditions to apply forces at the user's head. The manipulation tasks in the environment are performed by the Zebra-ZERO (manufactured by IMI, Berkeley, CA) which has a force sensor at its wrist. Figure 2 shows the master and slave of the telerobotic test-bed.

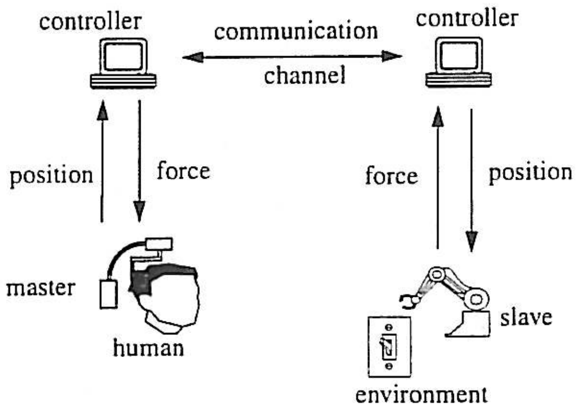


Figure 1. *Configuration of telerobotic system*

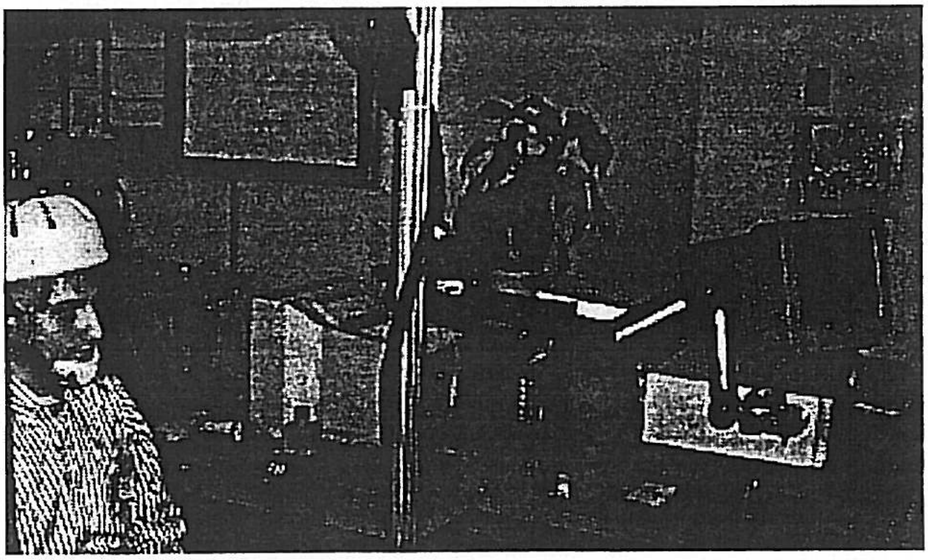


Figure 2. *Master slave telemanipulator setup*

Since an electronic link exists between the two robots, and the force and position data that are collected are available for manipulation by the software, it is possible to evaluate a variety of different control schemes. Additionally, the system is modular allowing for the possibility of different input and output devices, since the kinematics of the master and slave need not be the same.

Proc. 1st Euro. Conf. Disability, Virtual Reality & Assoc. Tech., Maidenhead, UK, 1996

2.1 Head-Stick Control

One natural method of operating the telerobot is to simulate the characteristics of a head-stick or mouth-stick using the telerobot test-bed. An appropriate choice of parameters also allows us to enhance the user's strength and to re-orientate the gripper to the needs of the task.

Figure 3 shows the assignment of coordinate frames on the master and slave robots. The virtual head-stick method of operation involves adding a rigid imaginary link of length l at the origin of the final coordinate frame on the master robot (the centre of the helmet). Using the master robots forward kinematic transform **A** and a transform relating to the positioning of the virtual head-stick **HD** it is possible to determine the target position of the slave robot. This can then be implemented using the slave forward kinematic transform **G**. That is

$$G = CAHD \quad \text{(EQ 1)}$$

When the slave is not in contact with the environment equation 1 determines the trajectory of the slave and the matrix **G** is used to calculate the joint angles of the slave. When the slave is in contact with the environment forces are measured at the wrist sensor and these must be relayed back to the user in an appropriate fashion.

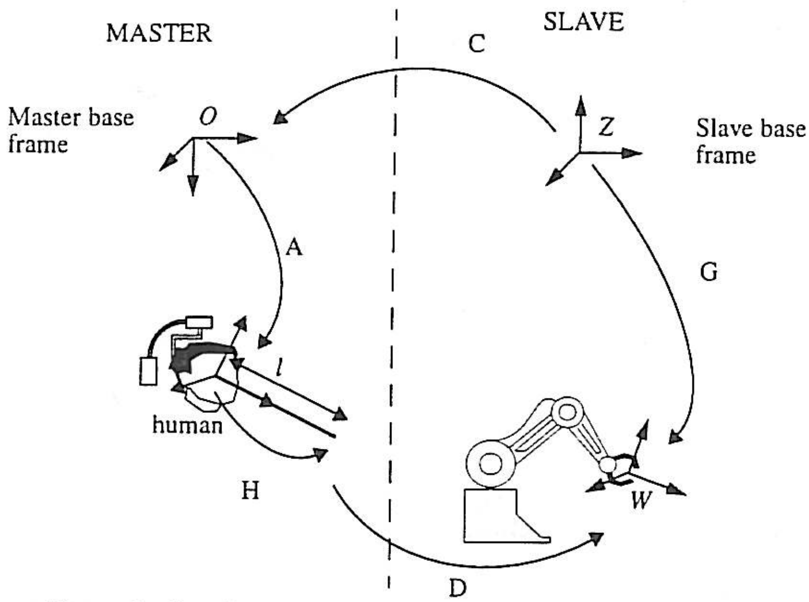


Figure 3. *Coordinate frames used to implement a virtual head-stick*

The sensed forces are measured with respect to the wrist frame and this is first transformed through **G** to estimate forces with respect to the slave base frame and then via the forward master kinematics to determine the orientation of the forces with respect to the tip of the virtual head-stick. A geometrical relationship is then used to determine the torques that must be applied to the master to allow the operator to experience a scaled value of these forces (Salganicoff 1995).

3. IMPLEMENTATION OF A POWER ASSISTED ARM ORTHOSIS

A test-bed power assisted orthosis has also been built at the Applied Science and Engineering Laboratories. It consists of a 6 degree of freedom master (RT200 from Oxford Intelligent Machines Limited) with the end effector replaced by a 6 axis force/torque sensor (Mini from Assurance Technologies Incorporated). A splint assembly is mounted on the force torque sensor and this in turn supports the person's arm (figure 4).

The base level control system first subtracts the weight of the individual's arm from that measured by the sensor, and uses the torque measurements to servo the RT200 robot so that no net torque is experienced by the individual over the combined range of movement of the individual and the robot. Once these servo mechanisms are established additional control algorithms can be instantiated to relate the resultant force that the individual exerts on the force torque sensor to the movement of the robot. Algorithms relating the estimate of the user's residual force to robot position, velocity and acceleration have all been demonstrated (Stroud 1995).

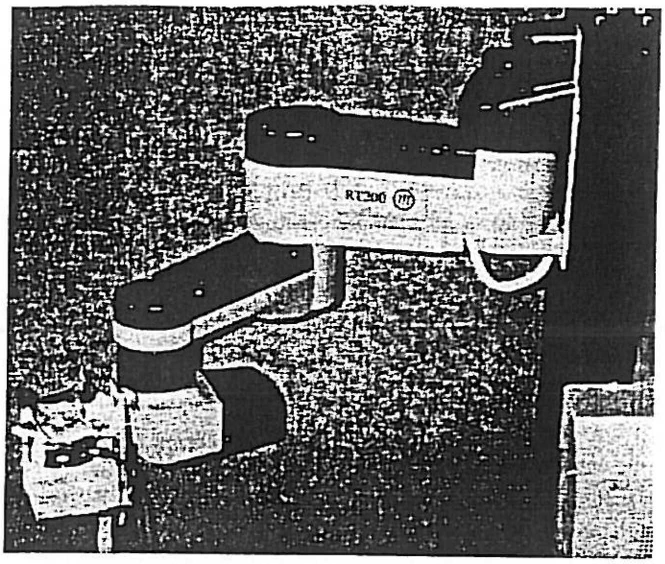


Figure 4. *Power assisted orthosis setup*

4. IMPEDANCE MODELS OF PROPRIOCEPTIVE INTERACTION DEVICES

Hogan (1985) postulated a mechanism for human movement that is now widely accepted, the philosophy of impedance control. In his seminal papers he suggested that human movement is achieved by the central nervous system establishing an end point for the trajectory and an associated movement stiffness, both of which remain constant for an uninterrupted movement and which alone determine the movement trajectory. This concept has gained acceptance in telemanipulator research with Hannaford (1989), Lawrence (1992) and others using it as a criteria for telemanipulator designs. The methods used can give an insight to the operation of telerobots, power assisted orthoses, and haptic interfaces and these methods are developed in the following.

4.1 Two Port telemanipulator model

The Applied Science and Engineering telerobotic test-bed has been designed to study two control architectures, position forward/force reflection, and force forward/position reflection (Rahman 1994). The position forward force reflection is readily implemented on the test-bed since the master contains position encoders and the slave incorporates a 6 axis force sensor. The master has been fitted with a force sensor and work is ongoing to implement the force forward/ position reflection method.

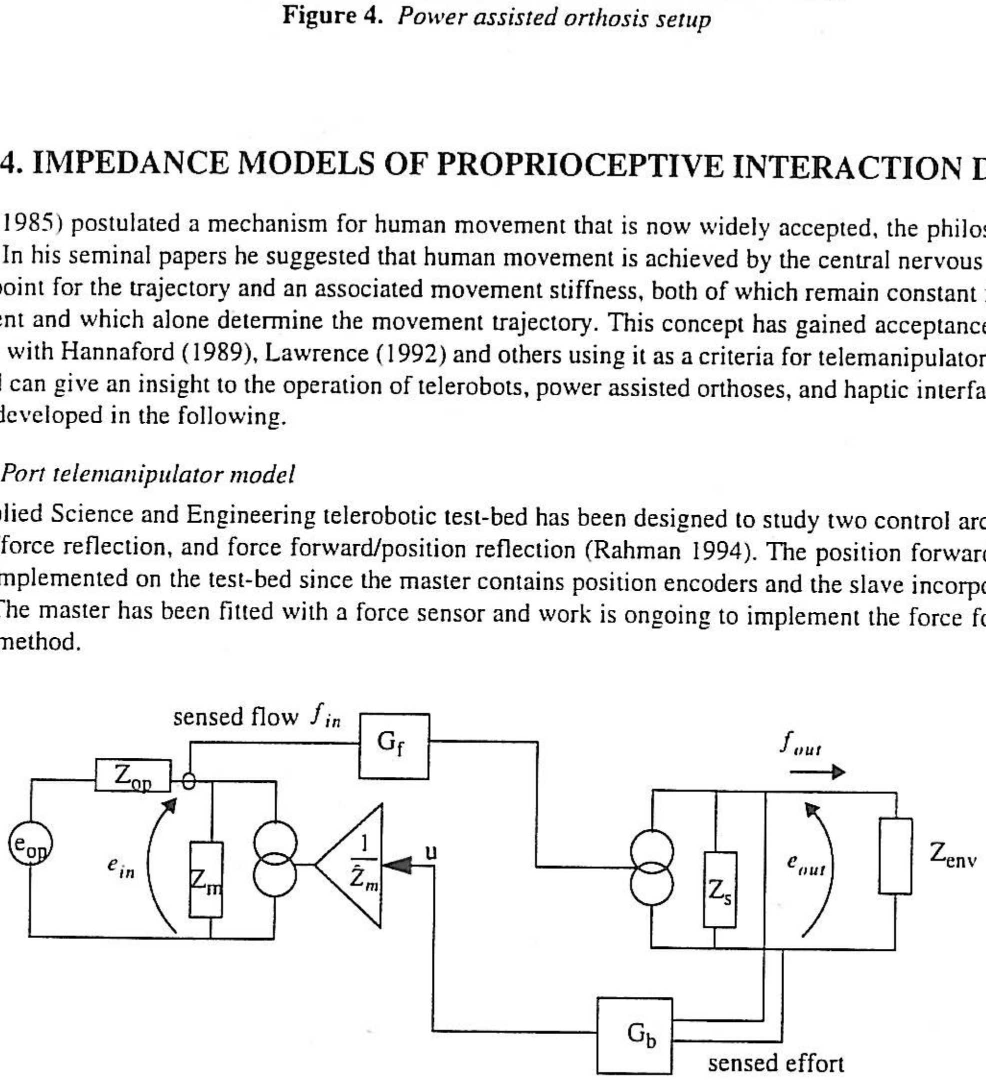


Figure 5. *Two port telemanipulator model*

The Cartesian position forward/force reflection architecture can be modelled in a one-dimensional case using an equivalent network (figure 5) where e is used to represent effort (force or voltage), and f is used to represent flow (velocity or current). For the following discussion force will be used interchangeably with effort, and flow with velocity. Both master and slave use position controllers at the lowest level. Since position relates directly to flow it is appropriate to model the actuators as a dependent flow source. The hybrid parameter matrix (H-matrix) can be used to gain insight into properties of the system and equation 2 shows how effort is related to flow.

Proc. 1st Euro. Conf. Disability, Virtual Reality & Assoc. Tech., Maidenhead, UK, 1996

$$\begin{bmatrix} e_{in} \\ f_{out} \end{bmatrix} = H \begin{bmatrix} f_{in} \\ e_{out} \end{bmatrix} \tag{EQ 2}$$

An ideal telemanipulator system would appear transparent to the operator, although it may magnify the operator's strength and scale the operator's movements. The ideal H-matrix is as follows

$$H_{ideal} = \begin{bmatrix} 0 & \frac{1}{Q} \\ -P & 0 \end{bmatrix} \tag{EQ 3}$$

where Q is the level of force scaling, and P is the level of amplitude scaling. Using the configuration shown in figure 5 it is possible to derive the H-matrix for the telerobot as shown in equation 4.

$$H = \begin{bmatrix} Z_m & G_b(q) \\ -G_f(p) & \frac{1}{Z_s} \end{bmatrix} \tag{EQ 4}$$

The master and slave impedances can be represented as the lumped parameters model $Z=ms+b+k/s$, where m, b and k are the equivalent mass, damping and stiffness of the robot. Using the H parameterisation we can determine the apparent input stiffness of the telerobot as

$$Z_{in} = Z_m + G_f G_b \left(\frac{Z_s Z_{env}}{Z_s + Z_{env}} \right) \tag{EQ 5}$$

Equation 5 can be used to examine the free space and hard contact impedance of the telemanipulator as the task impedance varies from 0 to ∞. For the free space condition the environment we have:

$$\lim_{Z_{env} \to \infty} Z_{in} = Z_m$$

Thus only the impedance of the slave Z_s is felt by the operator in free-space motion, as we would expect, since the force sensor returns no signal. A simple extension to this model would allow scaled values for slave flow to be added to the sensed force thus allowing the operator to experience more of the dynamic character of the slave in free-space movement in terms of a perceived impedance. Although the test-bed architecture allows for this to be implemented the dynamic characteristics of the slave exceed that of the master and human so have little effect on the operator's perception in this instance.

For the hard contact condition,

$$\lim_{Z_{env} \to \infty} Z_{in} = Z_m + G_f G_b Z_s \tag{EQ 6}$$

thus for hard contact the stiffness of the master and slave are additive. In order to feel the environmental impedance, the slave mechanical impedance Z_s must be very high, since it is effectively shunted across the environmental impedance Z_{env}. Secondly, as equation 6 shows, in order for the feel of the task impedance to remain the same, the product $G_f G_b$ must be constant, otherwise impedances presented to the operator will vary. If we increase the forward flow scaling value, we must decrease the reverse force gain proportionally. Obviously, we cannot do this arbitrarily, as the system stability will also be affected by the magnitudes of G_f and G_b.

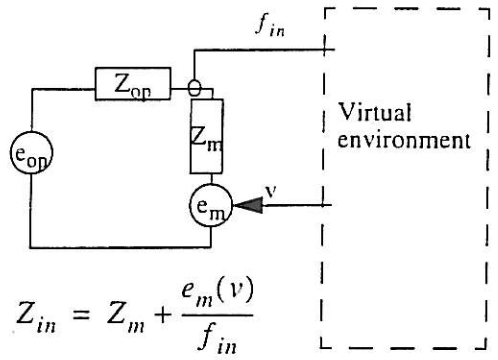
Based on the expected task impedances, we can select the forward position gain G_f and force feedback gain G_b to match comfortable force and displacement levels as measured for the target user population (stanger94). Additionally, closed loop system properties that have user-interface and safety impact, such as bandwidth and stability can be analysed by using the **H** matrix parameters (Hannaford 1989).

4.2 One Port models

The two port model adapts readily to consideration in the design of power assisted orthoses and haptic display designs. A haptic display is equivalent to a situation where the slave and environment are determined solely by the virtual environment. It is easy to simulate infinite stiffness and equation 5 becomes

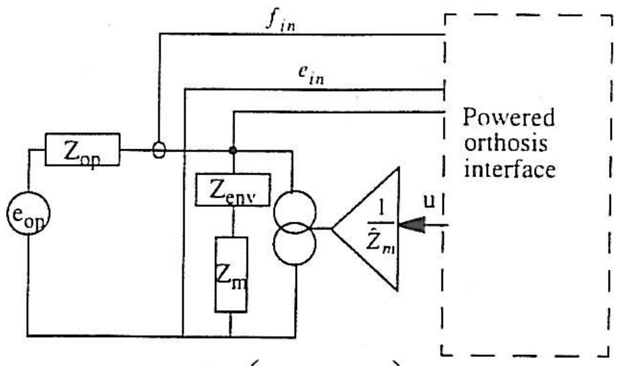
$$Z_{in} = Z_m + G_f G_b Z_{env}$$

where G_f and G_b can be now be chosen to simulate different arbitrary environmental impedances while maintaining system stability. A one port haptic display model is shown in figure 6a. One common lightweight haptic interface is the PHANToM from SensAble Devices. This interface is based on controlling joint torques via motor current hence the effort controlled model is appropriate.



$$Z_{in} = Z_m + \frac{e_m(v)}{f_{in}}$$

Figure 6a



$$Z_{in} = (Z_m + Z_e)\left(1 + \frac{e_m(u)}{\hat{Z}_m f_{in}}\right)$$

Figure 6b

Figure 6. *Controlling impedance from a virtual environment or a powered orthosis*

For the case of the one port model of a power assisted orthosis an additional complication exists in that user may wish to interact with his or her environment through the same port as the arm interface. This can be accomplished by adding the environmental impedance in series with the user impedance and considering the two cases when this tends to zero (no contact) and infinity (contact). This is shown in figure 6b.

5. RESULTS

Good characterisation of the master is essential to the control of the telerobotic system. A system identification of the master robot was carried out to determine the effects of closed loop controller parameters on the impedances along different axes. A PID controller is available on each axis of the robot however only the proportional and derivative terms were used. The individual axes of the master were modelled as a classic closed loop controller and plant. The system identification process then allows the determination of the combined master and controller dynamics and then allows these to be used to establish appropriate values of master impedance.

Step responses were used to command master movement. The resulting movement was measured for different settings of k_p and k_d. The smoothed out data sets were fitted into second-order discrete-time models using MATLAB's system identification package.

The discrete time model was then converted into continuous time format and a closed loop transfer function was generated for each setting of k_p and k_d. Master stiffness was determined in the quasi static case when different vales for k_p were established in the controller and weights used to displace the master from a neutral position. The displacement was measured using the master position encoders.

Results for all axes were determined and it is helpful to consider the characterisation of the linear x axis of the master. Jayachandran (1995) showed that this could be expressed as

$$M = 10^4(0.033k_p - 0.1507k_d + 0.0011)\text{Kg}$$
$$B = 10^4(-1.07k_p - 162.11k_d)\text{ Ns/m}$$
$$K = 10^4(-166.07k_p)\text{ N/m}$$

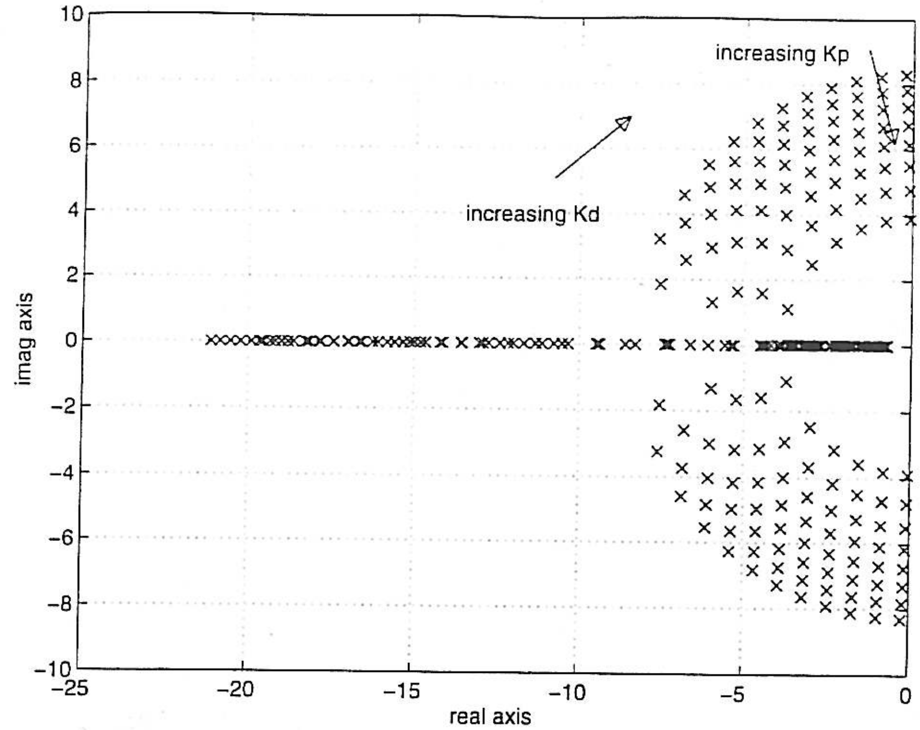


Figure 7. *Stability region for the telerobot test-bed master*

From this it is possible to determine a region of stability for the master in operation and select values for k_p and k_d so as to establish a desired dynamic characteristic for the region of operation. The stability of this mechatronic system can be determined from the defining equations and the movement of the poles in response to settings of k_p and k_d is shown in figure 7.

6. DISCUSSION

It is often not easy to make a direct estimation of the dynamics of a robot, whereas a good model of system dynamics is important in understanding the interaction between the robot and the environment or user. Reasons include the complexity of the mathematics involved, difficulty in determining the position of the necessary mechanical characteristics such as link inertia and actuator characteristics. Also in the test-bed system described, the manufacturers claimed that each axis was controlled via a PID controller, however additional terms were added to improve performance and this was poorly documented. Further it was not possible to access the controller software as the implementation was proprietary. For regions around the operating point it is reasonable to use system identification methods to determine baseline impedance measurements for the purpose of designing the controllers for the master-slave telerobot. This approach also allows us to change the impedance characteristics for the master in a limited fashion by changing the controller parameters.

The results for the PerForce master used in the telerobot test-bed show the range of results for which the master will remain stable and the behaviour of the master in this region (figure 7). Although this guarantees the stability of the master in a particular region of k_p k_d it ignores the effect of the operator modulating his or her impedance and forces (Z_{op} and e_{op} in figures 5 and 6). Experience has shown that this is most likely to occur in the non linear region of the boundary where the environmental stiffness changes radically over a short distance. This is an area of active research in telerobotics, power assisted orthoses and haptic interfaces.

7. CONCLUSION

The telerobotic research test-bed has demonstrated the concept of a virtual head-stick to allow individuals with spinal cord injury to successfully manipulate their environment. The advantages of a telerobotic system of this nature will have over head-sticks and mouth-sticks is the ability to increase the available workspace and degrees of freedom of the individual, as well as increasing power and range-of-movement abilities.

The user's perception of the environment has been shown to be highly dependent on the master and slave impedance characteristics. Ideally, we would like Z_m, the master inertia, to be small, and the stiffness of the slave, Z_s, which is in parallel with Z_{env}, to be as high as possible so that $Z_{in}=Z_{env}$. However making the slave impedance high can lead to high interaction forces. Using the controller it may be possible to change impedance characteristics as a function of the activity thus for free space movement the slave impedance would be reduced so that accidental collisions minimised the forces that resulted. When the operator intends to apply high levels of force to the environment the slave stiffness could be increased to reduce the artifact introduced by the telemanipulator.

It is possible to design telerobotic systems where a mechanical link couples the master and slave robots. This approach will lose the flexibility with regards varying master and slave impedances, however because it is possible to reduce the system complexity this may result in a more effective rehabilitation device.

The discussion of the two port telemanipulator model can be applied readily to the design of power assisted orthosis and haptic display devices in virtual reality. In this instance the operations of the slave are determined by a computer that can either create a virtual environment for the slave or have the environment determined through other commands by the user. For example the user might slide his or her arm along a virtual surface before immobilising it in a convenient region while a particular task is completed.

The two port telemanipulator model allows the operation of a telerobotic system to be modelled easily, and conclusions drawn about the behaviour of a telerobot. Although it is possible to determine the regions of stable operation using this model, it is not sufficiently general to study the conditions where the effort of the operator e_{op} is a function of the overall telerobot dynamics, a situation that may occur both in telemanipulator and virtual environments.

Acknowledgements: The authors gratefully acknowledge support from the National Institute on Disability and Rehabilitation Research, U.S. Department of Education, Grant # H133A20001 and Grant # H113E30013 and Nemours Research Programs. Many individuals have contributed to this work and the authors acknowledge the considerable help of Sean Stroud, Marcos Salganicoff, Daniela Pino, Shoupu Chen, Vijay Jayachandra, Richard Mahoney, Vijay Kumar, and Edwin Hereda. The work was carried out at the Rehabilitation Engineering Research Center for Robotics.

8. REFERENCES

B Hannaford (1989), 'A design framework for teleoperators with kinesthetic feedback' *IEEE Journal of Robotics and Automation*, 5 4 pp 426-434.

N Hogan (1985), 'Impedance Control: An Approach to Manipulation', *Journal of Dynamic Systems, Measurement and Control*, 107 pp 1-24.

V Jayachandran (1995), *A force reflecting assistive telerobot*, Masters Thesis, University of Delaware.

D A Lawrence (1992), 'Designing teleoperator architectures for transparency', *Proc. IEEE International Conference on Robotics and Automation*, pp 1406-1411.

T Rahman, W Harwin, S Chen, R Mahoney, R Glass, and D Pino (1994), 'Rehabilitation robot control with enhanced sensory feedback', In *Fourth International Conference on Rehabilitation Robotics*, Published by the Applied Science and Engineering Laboratories of the Alfred I duPont Institute, Wilmington, Delaware, pp 43-48.

M Salganicoff, V Jayachandran, D Pino, T Rahman, R Mahoney, S Chen, V Kumar, W Harwin, J Gonzalez (1995), 'A virtual head-stick rehabilitation robot system', *Proc. IEEE International Conference on Systems, Man and Cybernetics* Vancouver.

D C Simpson and J G Smith (1977), 'An externally powered controlled complete arm prosthesis' *Journal of Medical Engineering and Technology*, pp 275- 277.

C A Stanger, A C Phalangas, and Michael F Cawley (1994), 'Range of head motion of high cervical spinal cord injured individuals for the design of a testbed robotic system', In *Fourth International Conference on Rehabilitation Robotics, pages* 37-42, Wilmington, DE, June 14-16.

S S Stroud (1995), *A force controlled external powered arm orthosis*, Masters Thesis, University of Delaware

T Rahman, R Ramanathan, R Seliktar and W Harwin (1995), 'A Simple Technique to Passively Gravity-Balance Articulated Mechanisms', *Journal of Mechanical Design* , **117** 4 pp 655-658.

Design of rehabilitation robots using virtual work platforms

J.Ibañez-Guzmán

Ecole Supérieure Atlantique d'Ingénieurs en Génie Electrique
Gavy Oceanis, 44603 Saint Nazaire, FRANCE

javier@esa-igelec.gavysn.univ-nantes.fr

ABSTRACT

The design, implementation and testing of rehabilitation robot like devices is expensive both in time and resources. The use of virtual workspaces based on graphics kinematic-simulation and computer aided design techniques gives qualitatively a design technology in which a full system can be virtually tested and evaluated in advance of actual prototyping. The modelling of the mechanisms plus work environments implies that a virtual test-rig can be built. It allows the preliminary verification of concepts and the generation of useful feedback to designers. The entire cell layout could be changed easily creating complex situations and scenarios.

This paper describes the use of virtual work platforms based on kinematic graphics-robotic simulation for the design of novel rehabilitation devices and other aids, as well as the evaluation and personalisation of existing ones. The objective is to present the techniques available for the design of rehabilitation robots using a software-based approach.

Keywords: kinematic-simulation, design, rehabilitation robots, evaluation

1. INTRODUCTION

The possibility offered by devices that offset the physical-functional limitations of disabled people improves their self-reliance and integration into society in an independent manner. Advances in Information Technology have helped people to overcome the physical limitations affecting their mobility or their ability to hear, see or speak. Much has changed since the first Braille alphabet in the XIXth century, today, it is possible to find equipment for the disabled ranging from robotic devices to virtual reality interfaces (J. Adams, 1994). However, high development costs due to the need of adapting or purposely designing assistive devices to the user needs imply that many of these are still very expensive. Furthermore, much design work and prototyping needs to be done in order to augment the level of functional restoration which would approximate to the capabilities of nondisabled individuals.

The use of rehabilitation robotic enables disabled users to interact with their environments, their physical limitations can thus be extended through the use of manipulator-like mechanisms which respond to their commands. A virtual world enables users to navigate and interact with 3-D, computer generated environments. The immersion capability which provides a synthetic environment could be exploited to design novel assistive devices that coupled to the capability of simulating some of the programmed tasks, constitutes a virtual work platform where novel devices could be designed and evaluated. This paper introduces the use of virtual work platforms that employ graphic kinematics simulation techniques for the design of assistive devices like rehabilitation robots.

The modelling of robots and their work environments means that a virtual test-rig can be built that allows the preliminary verification of concepts and the generation of useful feedback to the designers. In the next section the issues involved in the design of assistive devices are examined, outlining design considerations and objectives. The third section gives a summary presentation of the use of virtual work cells for the design of robotic systems. Finally a description of the manner in which this technique could be used for the design of assistive mechanisms is made. It includes a sample application, the examination of the cell layout of the Master II robotic platform and the simulation of some tasks.

2. DESIGN ISSUES IN REHABILITATION TECHNOLOGY

In this section an overview is given of the design issues. First some design considerations are described and an outline of aims which the designs should pursue given. Finally, the design issues are described focusing on the design of robots like rehabilitation devices.

2.1 Considerations

Advances in new materials and Information Technology provide new possibilities for rehabilitation technology. Systems are being designed to tackle tasks which were impractical a few years ago. These are becoming more sophisticated and user expectations have increased. The design of rehabilitation devices is both complex and laden of contradictory constraints. While commercial pressures would demand systems to be produced massively, each disabled user has individual needs. Differences are not limited to the degrees of disability but also to their cognitive-communication abilities, perception and physical skills. Another challenge is that each product could provide the sole link (lifeblood) of the person with its environment, most users would depend on these devices everyday and for nearly most of the day.

A large number of assistive robotic devices have been developed since the 70s with the aim to provide disabled people with some autonomy. Nevertheless, the commercialisation of this equipment and acceptance by the user is not easy to attain. Often user needs are misunderstood, nor input from medical specialist heard, reasons why many projects are unfinished prototypes (A. Casals, R. Villa and D. Casals, 1993).

The design of rehabilitation equipment is first of all human-centred and relies on the understanding of the users needs as well as on the manner in which his/her disability would evolve in time. Thus a successful design could only be accomplished with continuous input from and interaction with potential users and specialist medical staff.

2.2 Aims

The overall design aims could be summarized as follows (after I. Craig, et al, 1994):

Flexible-Adaptable Systems, as the conditions of the user evolve physically and functionally, rehabilitation systems must allow the development of new functions and different man-machine interfaces without significant redesign of existing systems. To be economically attractive, the rehabilitation devices should be adaptable to widely ranging needs, skills, environments and tasks.

Maximize the Potential, the goal should be to design products which could be used by as many disable people as possible with minimum changes. The incremental cost of many functional redundancies considered at an early stage would be minimal by comparison to one-off development costs (J. Adams, 1994).

Open-Modular Approach, the use of existing components and standards and to allow other developers to contribute with their design. Modularity simplifies the addition of new-function and reduces the cost of constructing individual systems by eliminating unnecessary items.

2.3 Issues

Safety, one of the primary issues, the motion of the device and interaction with the user need to be assured, every possibility of system malfunction has to be explored. Users by their conditions could be very vulnerable, their capacity of reaction in case of malfunction is limited thus safety is a priority.

Effectiveness-Reliability, tailoring of equipment and their commissioning is costly, the proposed systems need to demonstrate their effectiveness against other solutions. Equipment which is to be used daily and that could be the lifeblood of the user requires to be very reliable. The level of confidence of the user must not be eroded.

Man-Machine Interface. CONTROL INTERFACE. The motion command of rehabilitation robot like aids is different from industrial robots, motions are simpler and slower, orders are imprecise and in general the user remains part of the control-loop. The tasks performed serve as interactive aids and for performing personal and vocational activities. In general, keyboards, joysticks, switches or the voice are the most common control interfaces used. Small motions on these devices would enable the user to control motions or actions on the environment.

There are two main constraints physical and cognitive. The physical ones are linked to the upper and lower limbs. These could be due to difficulties found in approaching the interfaces, force control, tremor, amplitude of movement, gestural control, fatigue, etc. Cognition constraints are more problematic and could inhibit any use of these devices (J. Ibañez-Guzmán, and G. Jehenne, 1996). Thus the command of an assistive device presents more difficulties that those encountered in classical tele-operation. Users can generate only very few actions capable of producing a command towards the input device, a universal solution would thus be impossible.

©1996 ECDVRAT and University of Reading, UK ; ISBN 0 7049 1140 X

Feedback To The User. The system must provide a clear indication to the users of what it is doing; the user must be able to perceive how the system acts. The usefulness of an aid is highly linked to its adaptation to the characteristics of the users, not least their sensory means. How successful is the information input to the user depends largely on how well it fits to the properties of the human sense involved (G. Jasson, 1994).

Visibility and the sense of action are very important when using rehabilitation devices, they provide the information by which the user knows what is going on with the system.

Personalisation

All individuals are different, in the case of individuals with disabilities, their physical abilities, perceptual skills and cognition are also different. Therefore, although rehabilitation devices could be designed for groups of individuals, these need to be tailored to the needs of users. For example a wheelchair has to be adapted first to accommodate the user in comfort, then to allow him an ample field of view and easy access to the control interface, next this would be adapted to his physical capabilities. Consequently, rehabilitation equipment should be highly adaptable to user needs.

Training

The use of certain devices require particular skills and hence training. The level of confidence of the user would depend on them. The question is how to coordinate training with the development aspects, knowing that the interface control equipment is to be adapted to user needs concurrently. The manner in which this is approached would influence the success of acceptance of the device.

Environmental Factors

The manner in which the rehabilitation mechanism would be integrated into living/work environment needs to be considered. Questions such as the layout of living spaces, logistic concerns, maintenance and the interaction between the rehabilitation systems and surrounding environment need to be addressed.

In summary, the design-implementation of rehabilitation equipment is multi-disciplinary and very iterative, requiring a high degree of testing/evaluation. That is testing the correctness of the design (safety), reporting the effectiveness of training and support time and costs. The possibilities brought by the use of virtual reality techniques together with simulations tools could be beneficial and permit addressing several of the issues cited previously.

3. DESIGN AND VIRTUAL SIMULATION OF ROBOTIC SYSTEMS

In addition to CAD tools, which allow mechanical components to be drawn on computers screens and then edited and processed, there are simulation tools for refining and competing the design of complex mechanical systems such as robots. Software packages are available for studying assemblies of mechanical components in motion, each component and articulation being described mathematically, with the computer calculating the system response at each operating point.

While the use of available 3-D software enables designers to conceptualize new robots and perceive their operation, the utilization of graphics-based kinematics simulation tools would allow designers to evaluate the operation of the mechanisms and optimize alternative configurations quickly and easily (C. Klein, and A. Maciejwski, 1988). The graphics and immersion aspects of the simulation could be used to represent not only the devices but also their environments. It eliminates the need for using physical models to study the proposed robots and to refine their designs. Data obtained from the simulation test runs, plus estimates of weights and moments of inertia of the robot components enable the dynamics to be taken into account. Therefore designers could calculate the forces/torques required at the robot joints and those acting upon attachment points. The control engineer can design the servos and develop control strategies. Consequently, most of the design and modifications needed can be made prior to starting any potentially expensive development.

The availability of graphics-based kinematic simulation packages together with immersion facilities enables designers to perform feasibility studies on novel robotic systems. Hence, it is possible to anticipate which part of the system could collide, to consider safety issues, to determine motion strategies, calculate tolerances and define cell layouts. Kinematics modelling implies a complete representation of the objects in the work cell including a description of how they change over time. Therefore, a kinematic description of the robot and peripheral equipment, together with their 3-D geometric representations, need to be included in the model. Samples of commercially available 3-D graphic kinematic simulators are ROBCAD from Tecnomatix Technologies, CIMstation from Silma Inc.

The development of kinematic models for use in the simulators implies that the geometric and kinematic details of the entire system must be translated into the simulator language. By building separate models for the robots, peripherals and the work environment and then positioning them so as to emulate the layout of the cell, it is possible to have a virtual test-rig. Next the motions of all the devices need to be specified. A series of tests can be made on these virtual test rigs to assess

 ISBN 0 7049 1140 X

the performance and operational feasibility of robotics devices. Results from the test-runs can be used to refine the robot's specification. The tests enable the limiting load sizes, permitted tolerances, expected cycle times to be determined.

This approach is used to design, evaluate, optimize the configuration, determine motion strategies and study the operability of novel robots prior to any manufacture. In addition, this technique is used for tele-manipulation operations and to assist with the training of their operators.

4. VIRTUAL WORKSPACES FOR THE DESIGN OF ASSISTIVE DEVICES

Professionals working towards the enhancement of the performance and safety of humans in the execution of tasks (human performance engineering) have encountered software which allows them to create human models and to study their behaviour. Software programmes like ADAMS/Android, Mannequin or JACK allow the virtual representation of humanoid figures (based on anthropometric measurements) and to represent their environments, to create virtual workspaces to study humans. Tests are performed to determine: reach, balance, collision detection, fields of view, etc (P. Vasta and G. Kondrsake, 1995).

By contrast the approach presented focuses on the rehabilitation devices having humans as the centre of view. However, several of the performance parameters measured to assist with the design are based on a combination of those used in performance engineering and those of robotic design. Virtual workspaces could be used primarily for design purposes but also to assist with the commissioning and for the evaluation of existing ones.

4.1 Design

As with the design of robots or special machines, their configuration (morphology) is first defined and simulated. For this purpose, different kinematic and-geometric configurations are described in the simulator language. When there is interaction between the device and the user, a geometrical description of the user would be needed. The *effectiveness* of the device could be examined through the simulation of the tasks which it is supposed to perform. This phase is highly iterative and thus changes to the graphical representation and kinematics would be cheaper than prototyping. *Safety* concerns when the device is in motion and the assurance of a minimum space between the device and user can be examined using the software collision detection capabilities.

Once the basic configuration is defined other features of the work environment can be added. Using these virtual test-rigs, it is then possible to define the layout of the work environment, tolerances and motion strategies. Next several device tasks can be programmed and then simulated. This virtual workspace also allows us to concentrate on the user point of view, the *Man-Machine Interaction*. One issue which cannot address conventional design methods fully is visualisation. We must make the user see what he will "see" when operating the device. His motion capabilities and posture would restrict his field of view. Another important issue is the manner in which the control interface would be implemented. Questions such as: Does the new device need a special control interface or could it be controlled using a conventional joystick require an answer. This issue can be addressed by substituting the standard computer interface (mouse) by one used by potential users and generating the view that they would have.

Most tasks could be preprogrammed and then triggered by the users. Changes of motion strategies or the layout of the cell are easier and cheaper in this virtual world. The user can participate in the design.

The plate presented in Figure 1 shows one of the proposed modified cells for the Master II robotic platform for office work and daily living activities (R. Cammoun, J-M. , F. Lauture, B. Lesigne, 1994). The cell is based on the original configuration of the Master II platform and expected automatic tasks. The robot simulation software ROBCAD was used and most of the office work and daily living activities were programmed and simulated. Emphasis was made on the optimisation of the cell space and the visualisation of tasks (O. Flegeau, 1995).

4.2 Integration

In rehabilitation most equipment has to be *tailored*. For example, if one would like to use the MANUS tele-manipulator, a 6 degrees of freedom manipulator designed for daily living activities and that can be attached to a wheelchair (J.-C. Cunin, 1994), it is necessary to define the attachment point and joint limits taking into account the geometry of the wheelchair and the user anatomy. By modelling the manipulator, wheelchair, the body of the user and the defining points towards which the robot end-effector should move, it is possible to use robotic simulation software to determine the optimal location of the arm and define motion strategies.

User *training* and adapting the *control interface* can also be made. Complex situations and scenarios which are difficult or dangerous in the real world could be generated in safety. That is a virtual reality approach serves as a sort of scene authoring to build virtual environments which resemble real situations.

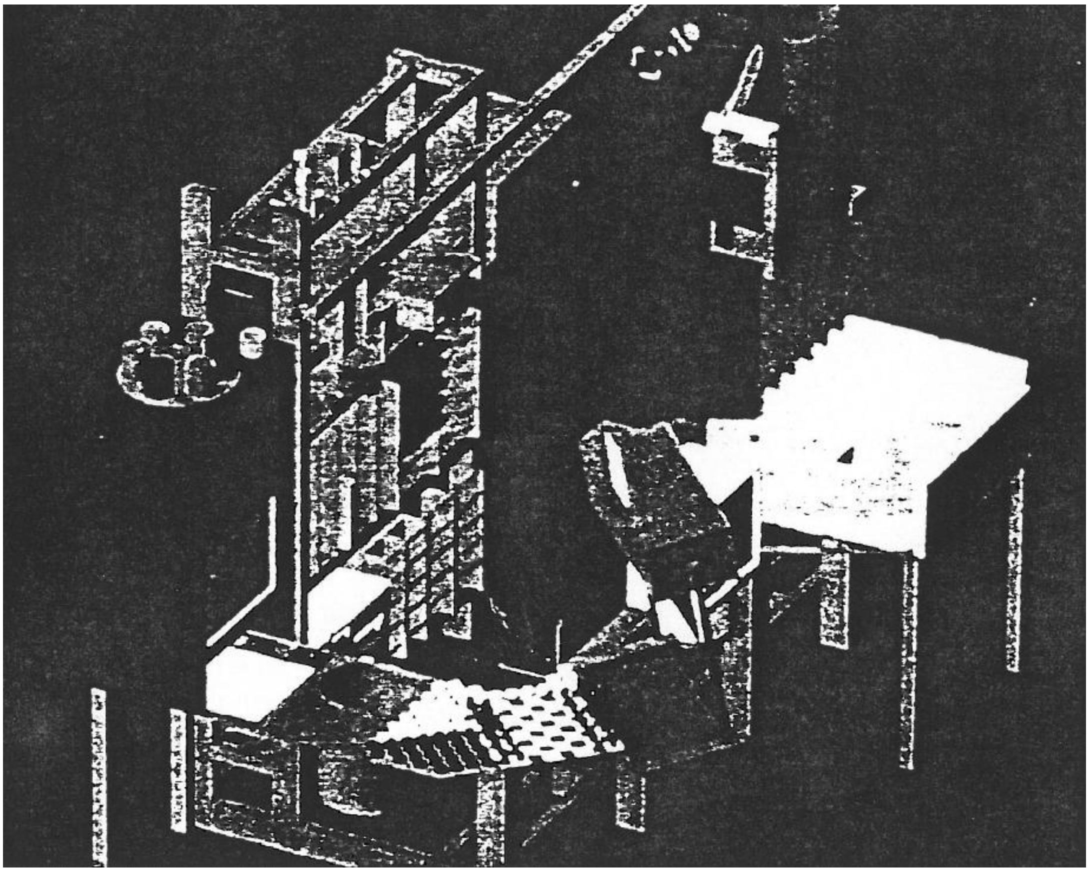


Figure 1. *Proposed cell for the modified Master II platform*

4.3 Evaluation

For a rehabilitation centre testing the correctness of the design, measuring its effectiveness in use, reporting of the effectiveness of training and support time, etc are very important prior to any commitment of usage. The utilization of virtual workspaces offer almost limitless potential for collecting performance data and allowing to perform human-centred test and hence exploring different alternatives in safety.

It should be remarked that the technique described is no substitute for actual prototyping and field trials, however, it reduces the number of interaction and gives the opportunity to examine more alternatives at lower costs.

5. CONCLUSIONS

An approach based on the use of virtual workspaces towards the design of robot-like rehabilitation devices has been described in this paper. Virtual Reality techniques allow users to navigate and interact with 3-D, computer generated environments, these together with graphic kinematic simulators that allows the modelling and simulation of any kinematic chain are exploitable tools for the design of rehabilitation devices. Virtual prototyping of future rehabilitation aids allow us to perform human-centred tests and evaluation of concepts without expensive prototypes. Entire systems can be modelled and exercised on a computer. It is qualitatively a new design technology in which a full system can be virtually tested, operated by the potential users and evaluated. The manner in which safety, effectiveness of design, man-machine interface and personalisation issues linked to the design of rehabilitation robots like systems using a virtual approach has been presented.

The approach can be used also for the evaluation of existing robot like rehabilitation devices. The virtual platforms offer almost a limitless potential for collecting performance data. These platforms can be also used for both selecting the most appropriate device and/or defining the modifications necessary to adapt an existing one to a particular user.

This technique, if compared with actual prototyping does not act as a substitute to the insight which an actual prototype would give. It should be considered as a tool which shortens the number of design changes on the final prototypes and assists with optimisation procedures. It is envisaged that once the connection is established between kinematic graphics simulators and human performance engineering software tools (kinematic models of devices and humanoid models), the possibility of designing rehabilitation devices solely in a virtual manner will be nearer.

6. REFERENCES

J. Adams, (1994), Technology combats disabilities in *IEEE Spectrum*, **31**, 10, pp. 24-26.

R. Cammoun , J-M Détriche, F. Lauture, B. Lesigne, (1994), RAID Une station de travail orientée bureautique pour les personnes handicapées, *Proc. Colloque INT*, Paris.

A. Casals, R. Villa and D. Casals, (1993), A soft assistant arm for tetraplegics, in *Rehabilitation Technology, Studies in Health Technology and Informatics*, (E. Ballabro et all, Eds.), **9**, IOS Press, Amsterdam.

I. Craig, et al, (1994), Tools for Living: Design Principles for Rehabilitation Technology, in *Rehabilitation Technology, Studies in Health Technology and Informatics, op.cit.*

J.-C. Cunin, (1994), Le Projet de télémanipulateur MANUS, *Proc. Colloque INT*, Paris.

O. Flegeau, (1995), Simulation et Evaluation de la Plateforme RAID, rapport interne, ESA IGELEC, Saint-Nazaire.

J. Ibañez-Guzmán, and G. Jehenne, (1996), Development and Field Tests of a Voice Controlled Wheelchair plus its Security Module, *Proc. Vth European Congress of Occupational Therapy*, (to be published), Madrid.

G. Jansson, (1994), In what ways can the psychology of perception contribute to the development of rehabilitation technology?, in *Rehabilitation Technology, Studies in Health Technology and Informatics, op.cit.*

C. Klein, and A. Maciejwski, (1988), Simulators, graphic, in *International Encyclopedia of Robotics and Automation* (R. Dorf, Eds.), **3**, John Wiley & Sons, Inc., New York.

P. Vasta and G. Kondrsake, (1995), Human Performance Engineering: Computer-Based Design and Analysis Tools, in *The Biomedical Engineering Handbook* (J.D. Bronzino, Eds.), CRC Press, London.

ECDVRAT '96

Visual Impairment, Ambisonics, and Mobility

Day 3: 10 July, 1996, Morning

A 3D sound hypermedial system for the blind

Mauricio Lumbreras[1], Mariano Barcia[2], Jaime Sánchez[3]

[1,2]LIFIA - Laboratorio de Investigación y Formación en Informática de Avanzada
Fac. Cs. Exactas - Universidad Nacional de La Plata
Calle 50 y 115 - 1er Piso - (1900) La Plata
ARGENTINA

[3]Department of Computer Science, University of Chile
Casilla 2777, Santiago
CHILE

[1]*mauricio@info.unlp.edu.ar,* [2]*mbarcia@info.unlp.edu.ar,* [3]*jsanchez@dcc.uchile.cl*

ABSTRACT

It has been said that quite often a hypermedial application running over a GUI is somehow inappropriate or unusable. This is the case for end-users with little or no visual capabilities. In this paper we present a conversational metaphor to try to ameliorate this problem. We propose a framework in which the interaction is rendered using 3D sound. Several voices used as media to convey information are placed in the space through this technology. By using a glove the user controls the system manipulating a special version of 3D auditory icons, called *audicons*. Moreover, we introduce a technique called grab-and-drop, equivalent to the visual drag-and-drop. We show that this framework permits the building and adaptation of current hypermedia interfaces in a way that can be used without visual cues.

Keywords: drag-and-drop, 3D sound, aids for the visually-impaired, auditory I/O, virtual reality, metaphor, hypermedia

1. INTRODUCTION

It is widely known that hypermedial applications, in particular those accessed via CD-ROM, are becoming prevalent in current domains such as education and on-line documentation. Unfortunately, the interface metaphors emphasizing graphical images, icons, menus, and the like, do not consider visually-impaired people. Several initiatives enable the visually-impaired to have generic access to computing systems, such as the Mercator Project (Mynatt et al., 1992) and the European GUIB Project (Weber et al., 1993). In this paper we present a rather different and more specific approach by trying to answer questions such as:

- how visually-impaired people can take advantage of 3D sound technology?
- how we can produce suitable metaphors to browse information?, and
- what is the best modality in which to interact with a system without visual cues?.

New virtual reality technologies enable us to produce a spatialized kind of sound over headphones, called 3D sound. This sound processing is achieved by convolutioning the desired sound with certain FIR (Finite Impulse Response) filters, called HRTF's (Head Related Transfer Function), that are calculated with special acoustic recordings taken from real human ears (Wenzel, 1992). As a result, we hear through the use of the headphones a sound in a specific position of the space.

The advantages of these capabilities prompted us to produce some kind of virtual aural environment in order to generate a user interface to be used with information systems for visually impaired people. But nice sounds floating in the space do not provide a usable and friendly information system. How can we produce a useful framework? Several studies talk about the use of 3D sound , but very few deal with models and metaphors to exploit these capabilities.

 ISBN 0 7049 1140 X

2. THE METAPHOR

This study is based on a special version of a well-known hypertext model (Conklin, 1987). The basic architecture consists of a direct graph of nodes and links. Nodes represent certain documents. Links reflect semantic relationships between documents. Usually links can be displayed both explicitly, as a menu of choices on the screen, or implicitly, as an action taken when the user does something to an identifiable target object in the interface. The hypertext model offers many advantages such as the management of linked references, a context in which the content is not bound to a fixed structure, and dynamic interaction.

We adapt the conventional hypermedial model to fit our needs. Each node belongs to some type that reflects the kind of information contained. Each type of node is mapped to a speaker in a determined position in the space. The role of speakers is to engage a conversation, allowing the user to interact with participants by browsing and managing the flow of talking. The most interesting feature is the link selection. If during the talk there is a link to another concept, the speaker in charge of talk, makes a short comment. If the user is interested he or she may indicate the direction of the next desired speaker by using interfaces such as pressing a button in a joystick or grabbing-and-dropping some audicon in a position in the space. The user can do this, because she or he recognizes two cues: the voice, and more critically, the position of the speaker in the space. The voice of each person was sampled from different real subjects, and processed in order to generate 3D voice. Through the assigned type, each speaker plays a particular informational point-of-view and an anthropomorphical characteristic is assigned to the information content (Muller et al, 1992).

Unconsciously, the user is using a hypermedia system, because a hypertext structure was mapped to the conversation. Moreover, each link may reflect several conversational characteristics, such as requesting, acknowledging, counter-offering, etc. As a result, the information presented is fine-grained, nonlinear, and highly interconnected. Additionally the structure of the system and type of information are dependent on the application domain, because different domains may present different classes of information or point-of-view assignments.

3. THE ENVIRONMENT

The environment is controlled by manipulating a special version of 3D auditory icons, called *audicons*. An audicon is an entity that represents functionality. It is rendered with 3D sound without visual cues. Audicons have behavior, a position in space, and a set of sounds that can be played by the user. Thus we extend the Handy Sound environment (Cohen, 1995). Audicons are presented in space within a customizable horizontal plane, and selected by using a grab-and-drop technique. As a result, we present some kind of direct manipulation, a new concept for running a system without visual cues. By taking advantage of 3D sound, the user can select one of several simultaneous audicons. This is also referred to as the "cocktail party effect". In order to avoid annoyance, the user can select either sequential or simultaneous audicon notifications.

In addition, a static surrounding can be simulated by using an auditory version of a room metaphor enabling the modeling of the static environment architecture. Thus the user could move between rooms and a corridor. Rooms are organized along the corridor to provide a type of spatial index. In each room we can have a conversation related to the whole content.

In order to carry out task control , there is a special speaker, the assistant, that consistently remains in a fixed position in relation to the user. The assistant manages tasks such us backtracking and user orientation through context dependent advice. As a result, control task and information content are presented homogeneously. In other words, the user interact with different people.

When desired, the space aural simulation of the environment is reinforced by means of verbal descriptions presented by the assistant. There is evidence that this mode creates isomorphism between the mental model and the simulated space (Denis, 1993).

3.1 The grab-and-drop technique

Usually hypermedia applications are navigated by using the "point and click" technique. Arons (1991) states "...one cannot 'click here' in the audio world to get more information, by the time a selection is made, time has passed, and `here' no longer exists...". Keeping this idea in mind, it seems almost impossible to find a solution.

The well known drag-and-drop technique implies direct manipulation, but it is difficult to apply to a transient medium such as sound. In our approach by using a spatial metaphor, the clickable speaker exists all the time regardless of whether it is speaking or silent. This concept can be of great value when applying the proposed grab-and-drop technique.

3.2 Why Grab-and-Drop?

Usually sound is used to create an awareness of something happening or existing. We extend this notion in order to fully control the system by grabbing-and-dropping audicons, a special type of 3D auditory icons.

Our approach extends the ideas that have appeared in the literature (Gaver, 1986). Actually, in hypermedia applications it is very common to use a pattern of interaction, i.e. to apply an action over a target to backtrack, to stop, to play, and to get the next information chunk. The "drag-and-drop" technique is a powerful tool that matches this pattern of interaction very closely. If we provide the user with a way to pick up a certain entity and associate it with the target, we are offering the same functionality as offered by the drag-and-drop technique.

If the user can control the audicons in a kinesthetic way, by means of a glove, we are taking advantage of one of the major input modalities available to the visually-impaired. Therefore, this design of the user interface is based on the possibility of allowing users to actively interact with the environment instead of wandering around menus and prompts.

3.3 Grabbing audicons vs. dragging icons

Basically, the underlying idea is simple: the user, wearing a VR glove, closes his hand to pick up an audicon, grabs and drops it over the target object. The 3D sound offers the spatial attribute that makes possible the spatial manipulation of the audicon and the glove represents the means of interaction.

3.3.1 Similarities and Differences

In GUIs, the icon representation changes when it is dragged, notifying the user that the drag mode is on. Usually Virtual Reality gloves do not provide kinesthetic feedback. For this reason, the feedback must be rendered in a synesthesic way, replacing the tactile sensation by an adequate sound of something that is closing.

When a standard icon is dragged, its graphical representation is updated in the display to show the new position. When an audicon is grabbed, due to the transient nature of sound, a looping sound characteristic of this audicon is played to replace the update functionality. Here, we have two main ideas. On the one hand, the sound that is played in loop when the audicon is grabbed serves as grabbing feedback. On the other hand, it is also a notifier that the audicon grabbed is the right one. Another point to consider in the graphical domain when the user is dragging an icon that has been erroneously chosen. To cancel this activity, the user drops the icon in a void area of the desktop. Without visual cues, the user can lose confidence since he can not know if he is dropping the audicon in an empty space. To solve this problem, we provide a special audicon called "the trash" that absorbs an audicon dropped near it, and returns the audicon to the rest position, see Fig. 1.

Usually, in a GUI, the user scans the desktop to find the right icon. In the aural environment the user can drop a special audicon, the scanner, over the assistant, a special speaker, in order to get into scan mode. When the user passes the hand near the rest position of some audicon, she or he hears its characteristic sound with an intensity that is proportional to the distance between the hand and the audicon rest position. To cancel this mode, the user again grabs the scanner audicon over the assistant. To recognize or refresh one´s memory of which audicons are in the surrounding environment, the user can grab one of the refreshers over the assistant in order to hear sequentially or simultaneously the different audicons by their characteristic sound.

In a graphical-based hypermedial application the user navigates by clicking buttons. In the acoustic-proposed context, if the user wants to listen to a chunk presented by one of the speakers, he grabs and drops an audicon which is placed in the direction of the desired speaker.

Audicons are manipulated by grabbing and dropping. Each step is rendered with different sounds: 3D auditory icons for either synesthesic feedback or the icon´s own characteristic sound, and a looping *earcon* or musical sound for the grabbing process. Since this interaction belongs conceptually to the same entity, it has a behavior, and is conceptually perceived as a part of a whole, then we call this *audicon*. Thus the term audicon represents both a spatial/functional concept and its behavior.

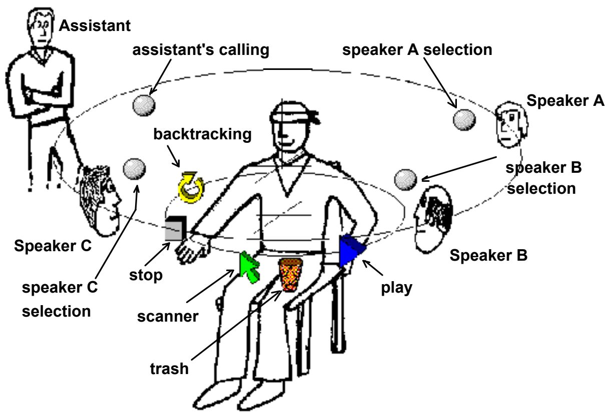


Figure 1. *Artistic version of the audicon´s environment. The audicon´s graphical representation is only illustrative because they are rendered with 3D sound.*

4. IMPLEMENTATION

We built several prototypes testing hardware alternatives. All versions were implemented using a PC clone as the underlying hardware. The first version tested was developed using a Gravis Ultrasound Max sound board as 3D sound generator. This board has a limited capability to render 3D sound, because the sound must be preprocessed and the interpolation technique used limits 3D positioning. The sound is produced by mixing four or six preprocessed sounds that represent above, below, right, left, front and back positions. In the same prototype to manipulate the environment we include a low cost version of a VR glove: the Power Glove. This glove, originally manufactured to be used with Nintendo games, was adapted to be plugged into a PC. The fact that the Power Glove has limited finger flexion detection, allows the accurate detection of two simple gestures like opened and closed hand. A software written in Borland C++ was in charge of querying the glove, playing the 3D sound, and carrying out the interface management and control of the hypermedial engine.

The final prototype was an upgrade of the first versions, because the Ultrasound board was replaced by one Alphatron 3D sound card (Alphatron, 1995). This piece of hardware has the capability to compute 3D sound in real time, by taking advantages of a Motorola DSP 56001. With this configuration the system is capable of rendering 2 or 4 sources simultaneously, depending on the playback frequency (44 or 22 Khz). Head tracking reinforces notably the sensation of immersion and offers a powerful opportunity to eliminate the ambiguity of the sound source position, but we have not included head tracking as yet.

5. DISCUSSION

It is premature to say that this system is fully usable. Initial testing shows that the cross modality could be interesting as a sound-position correlation training application, but long sessions with the system can cause fatigue due to continuous arm movement. Initial results indicate that a joystick or even a keyboard could be suitable devices to interact with the system, but always in the presence of 3D sound.

In addition, it is not currently possible to determine if it is necessary to render 3D sound in real time to build an usable system. This idea is supported by the fact that 3D sound can be computed off-line. Thus we can produce self-contained material packaged in a CD-ROM. With certain trade-offs related to real-time manipulation, the final product will contain a large set of 3D audicons to produce the environment simulation as well as the voice of the different speakers. This option requires basically an inexpensive platform: a personal computer, a CD-ROM player and a sound card. By using a set of HRTF's provided by Dr. Fred Wightman, we are testing the viability of our idea.

When entertainment is concerned, immersion, interactivity, and physical user activity are factors in the acceptance or rejection of a game. The grab-and-drop technique is a good choice for initial production of an unpublished type of entertainment for the visually impaired, taking into account the use of 3D sound technology with glove-mediated interaction.

Another interesting idea is the access to the WWW. It is possible, by using the proposed metaphor, to render synthetic 3D voice in real time, by passing the synthetic voice through a 3D sound processor. Mappings between VRML and an acoustic modality is not an unreachable idea.

At the moment testing is still ongoing. For this reason we have no polished methodologies and well grounded results yet. In the near future we will probably have more data to share and discuss. For now, we believe that the metric organization of the audicons and their modalities suggests a new research dilemma.

Acknowledgment: We would like to thank to Dr. Fred Wightman of the University of Wisconsin for the HRTF set. This work was partially supported by the Chilean Science and Technology Fund(FONDECYT) grant No. 1950584

6. REFERENCES

Alphatron (1995), *Alphatron User's Manual*, Crystal River Engineering Inc., 490 California Ave, Suite 200, Palo Alto, CA 94306

Arons B. (1991), Hyperspeech: Navigating in speech-only hypermedia. *In Proceedings of Hypertext '91*, pp. 133-146.

Conklin J. (1987), Hypertext: An introduction and Survey, *Computer*, September 1987, pp. 17-41.

Cohen M. & Wenzel, E.(1995). The Design of Multidimensional Sound Interfaces. *Unpublished Technical Report 95-1-004*. Human Interface Laboratory, The University of Aizu, Japan.

Denis M. (1993), Visual Images as Models of Described Environments, in *Proceedings of the INSERIM-SETAA conference Non-Visual HCI*, (Paris, March 1993)

Gaver W., (1986) Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction*, **2 (2)**, pp. 167-177.

Lumbreras M., Rossi G. (1995), A Methapor for the Visually Impaired: Browsing Information in a 3D auditory Environment, *in Companion Proceedings of CHI '95*, (Denver CO, May 1995), ACM Press, pp. 261-262.

Muller M., Farrel R. et al (1992), Issues in the Usability of time-varying Multimedia, In *Multimedia Interface Design, ACM Press*, pp. 7-38.

Mynatt E., Edwards W. (1992), The Mercator Environment: A Non Visual Interface to X Windows and Unix Workstation, *GVU Tech Report GIT-GVU-92-05*

Weber G., Kochanek D. et al (1993), Access by blind people to interaction objects in MS Windows, *Proc. ECART 2 European Conference on the Advancement of Rehabilitation Technology* (Stockholm, may 1993), pp. 2.2.

Wenzel E.M. (1992), Localization in Virtual Acoustic Displays, *Presence*, vol. 1 number 1, pp. 80-107.

Animated tactile sensations in sensory substitution systems

D A Eves[1] and M M Novak[2]

School of Applied Physics, Kingston University
Penrhyn Road, Kingston Upon Thames, Surrey, KT1 2EE, United Kingdom

[1]*AP_S447@ crystal.king.ac.uk,* [2]*AP_S412@crystal.king.ac.uk*

ABSTRACT

We have designed and made a computer controlled, electrocutaneously stimulated, tactile display to assist in the interpretation of complex graphical data. This paper covers the initial experiments with the tactile display, complementing the standard VDU. An animated vector was generated and the absolute tactile directional acuity of the observers was measured. Consideration is given as to how the use of animated tactile data would enhance graphical user interfaces for the visually impaired.

Keywords: tactile graphics, directional acuity, blind, multisensory displays

1. INTRODUCTION

This research was conducted to produce a multisensory display system for use in the visualisation of scientific information. The use of visual displays for scientific visualisation is, of course, implicit. Sound display, or sonification, of scientific data has also been researched but is still very much in its infancy. There has been very little research carried out on the use of tactile sensations to complement the visualisation of scientific data, except in the realm of virtual reality where the sense of touch is used primarily for force feedback.

Initially, we considered which aspects of the sense of touch could be of use in the field of scientific visualisation. In the field of scientific visualisation it may also be said that there are many aspects that present problems when using vision only. Thus the definition of the 'problem', as well as the solution, is conveniently loose. This has allowed us to analyze the sense of touch without constraint and then choose an aspect of it that has a potentially high information capacity. It was decided that the best means of maximising the quantity and resolution of data to be displayed tactually would be to have an animated tactile display. The field of scientific visualisation was then explored and a suitable problem chosen. This approach has produced two potential uses for touch in sensory substitution systems, primarily in graphical user interfaces for the blind, and graphical/scientific data displays for the blind.

The tactile display, designed and made by the authors, stimulates the non-dominant hand. In the current configuration, information relating to the region defined by the mouse pointer is displayed on the hand. This information may be the visible data, reemphasised tactually, or it may be another 'dimension' of the information, unseen but tactually correlated with the visual data. The limitations of a purely visual display methodology for scientific visualisation is clear when one compares the multidimensional nature of scientific data with the two dimensions available on a standard VDU. The problems this creates forms the justification for our research and one of these problems is elaborated in the following discussion.

If the screen is initially restricted to black and white only (no shades of grey) then all that can be displayed, without loss of information (except by virtue of pixel resolution) or potentially confusing the observer, is a simple, 1 dimensional, x-y line graph, i.e. 1 dependent variable against 1 independent variable. When the above restriction is removed and colour or grey shading is allowed then a 2 dimensional scalar data set can be displayed. The two independent variables are the horizontal and vertical axes and the dependent variable is mapped onto colour hue or grey shade intensity. Although this method of displaying in three dimensions is loss-less, a degree of error is introduced due to the way the eye and brain interpret visual information (Coren, 1978). There are other methods of displaying three dimensions on a two dimensional screen, each with their own advantages and disadvantages (Walton, 1993). The amount of data displayable at any one time becomes dramatically reduced when vector information needs to be displayed, even when considering only a two dimensional vector on a two dimensional 'field'. Vector data needs encoding; one, often used, method is to use spatially separated arrows. The magnitude is encoded by the colour of the arrow, or size, or length, or doubly encoded using two of these attributes (Collins, 1993). Now, many, and probably a variable number of pixels, are required to display just one point in the data set. The data also requires at least as much interpretation by the observer as in the previous case because

 ISBN 0 7049 1140 X

the user also has to mentally interpolate a vector. A further complexity arises when one, or more, scalar or vector variables are required to be visualised in two or three dimensions. In addition to the above problems, a lot of the data cannot be seen as it is obscured by the data in front of it.

This study explores the animated display of vector information by the sense of touch. A perceived advantage of the resulting multisensory display system can be illustrated as follows. On a surface, it may be useful to display both scalar and vector information. The spatial resolution of the vector data is considerably lower than that of the scalar data. Visually displaying the vector information also masks some of the scalar information. It should be possible to send higher resolution vector information to the tactile display, leaving the VDU clear to display the scalar information.

The structure of this paper is as follows: section 2 covers the choice of stimulation method and the perception of directional information. Section 3 describes the implementation of the computer controlled tactile display. Section 4 discusses the design of the directional acuity experiment and section 5 covers the interpretation of the results. Section 6 explores how these results, along with a previous experiment by the authors, can be used. Some aspects of further work is also outlined.

2. OVERVIEW OF THE TACTILE DISPLAY

2.1 Stimulation Method

There are two main ways currently used to elicit the sensation of touch which are suitable for computer control: electrocutaneous (or electrotactile), and vibrotactile stimulation. Both are widely used in sensory substitution systems for physically handicapped (Kaczmarek, 1991; Szeto & Saunders, 1982). For example, Geake (1994) discussed the use of an array of electrotactile stimulators for giving positional feedback to the user of a motorised upper limb prosthesis. Vibrotactile stimulation is frequently used in displays for the blind.

When compared, each method has advantages and disadvantages depending on the choice of criteria. However, two overriding criteria were the speed at which a system could be implemented and the restriction on cost. This made electrocutaneous stimulation the most attractive alternative for this project.

Electrocutaneous stimulation is the localised stimulation of cutaneous receptors and fibres of the afferent nervous system by means of a small electric charge applied to the skin. The sense of touch can be divided into different modalities i.e. pain, pressure, and temperature. There are several types of cutaneous receptors and nerve endings, including free nerve endings, meissner's, and pacinian corpuscules. Unfortunately, there is no simple relationship between sensation modality and receptor type. For example, in the cornea of the eye there are only free nerve endings, yet it is sensitive to pain, pressure, and temperature (Schiffman, 1995). Electrocutaneous stimulation affects any receptor or fibre which happens to be within range. Thus the quality of the sensation can vary, depending on several parameters, including electrode configuration, skin hydration, and body site. The quality of the sensation can also change with stimulation strength, not surprisingly, if strong enough it will feel painful.

2.2 Perception of Directional Information

From a physiological perspective, tactile directional acuity is governed by two different processes. First, cutaneous stretch receptors are stimulated by the relative frictional motion of an object in contact with the skin. The second is spatial data which changes with time (Norrsell & Olausson, 1994). Our display only exploits the second of these processes.

Gardner & Sklar (1994) have explored directional acuity using a linear vibrotactile array. Such a display does not exploit skin stretch. The object of one of their experiments was to determine the relative importance of density and path length parameters. They concluded that discrimination accuracy depends on the number/density of stimulations, rather than path length. They only studied whether the observer could determine if the stimulus was moving along the finger towards the tip or towards the palm, i.e. a two alternative, forced choice.

There is a sensation known as apparent motion which corresponds fairly closely to visual animation. This sensation appears between two speeds of tactile animation. In an experiment carried out by Sparks (1979), a linear array of electrocutaneous stimulators was used and animated motion was generated by sequentially switching on and off stimulators, very similar to that above. As the speed of this animation was varied, the observers reported two transition points, or thresholds. Below the lower threshold, the sensations were felt as discrete, successive steps. Above this threshold, the animation was felt as a smooth, continuous motion, up to a higher threshold. Beyond this, the sensation was sharp, with no sensation of motion, though the direction could still be determined for a little longer.

By controlling the relative intensities of adjacent stimulators, one can allow the user to mentally reinterpret the two separate simultaneous stimulations as one stimulation located at a point intermediate between the two stimulators. Thus, if one decreases the intensity of one stimulation, whilst simultaneously increasing the stimulation at a different point, one can interpret this as a sensation moving gradually from one point to the other. Alles (1970) explored this 'phantom' sensation using vibrotactile stimulators.

Keyson & Houtsma (1995) studied direction discrimination using a motorised trackerball arrangement. This test makes use of frictional contact with the skin of the fingertip. Tests were made at eight equally spaced directions. At each of these positions a threshold test was carried out by presenting a reference stimulus followed by a variable test stimulus, and asking the observer to tell the difference. Whilst this test examined one's directional acuity at a number of directions, it did not test one's ability to absolutely determine the direction of stimulation. It was considered that a measurement of absolute directional acuity would be desirable.

3. IMPLEMENTATION OF TACTILE DISPLAY

Our tactile display is controlled by a C program running on a '486 based personal computer (PC) under MS-DOS. Our original intention was for the PC to pass brief, high level instructions to a slave microprocessor which would then control the tactile display. This approach was desirable because scientific visualisation is a computationally intensive task and sharing the PC between the animation of the tactile display and graphical rendition would have meant a loss of real time control. However, financial restrictions forced the abandoning of the idea of a slave processor. Experimental work so far has not involved complex visualisation.

The display consists of 9 electrodes in an approximately square, 3x3 grid, thus there is one electrode on each of the proximal, middle, and distal phalanges of the index, middle, and ring fingers of the non-dominant hand. The non- dominant hand is used for ease of implementation and also Wilker et. al. (1991) found no significant difference in observer performance between mounting a vibrotactile display on the non dominant hand, with the dominant hand controlling a mouse, and mounting the vibrotactile display on the mouse, thus displaying to the dominant hand. Each electrode is of concentric design, with a central electrode 2mm in diameter and an outer ground-referenced ring of 9mm diameter. The electrodes are mounted on a clay hand 'jig'. The hand rests on the electrodes. Thus lateral movement of the hand and fingers is restricted, improving placement accuracy and repeatability, whilst allowing easy and quick hand removal, should this prove desirable. A momentary thumb switch, mounted on the jig, has to be pressed and held to allow electrocutaneous stimulation to occur.

One stimulation is, in fact, a brief burst of six 'biphasic' current pulses. The time delay between these pulses is controlled. A biphasic pulse consists of a short positive-going phase, rapidly followed by an equally short negative-going phase of equal amplitude. The phase widths, and inter phase interval are also software controlled. To vary the intensity of the stimulation, the software adjusts the electrode current amplitude, whilst the timing parameters are held constant.

The software to control the device has to give the appearance of doing two things at the same time and in as close an approximation to real time control as possible. The tactile display is initialised and the timing parameters are used to produce a look-up table that is downloaded to the display interface and is stored in RAM there. When a stimulation is initiated by the computer, the RAM addresses are counted through by the hardware and the biphasic pulse burst is generated. Just before the stimulation is required, nine digital to analogue converters (DACs) are given the desired individual amplitudes for the nine electrodes. By having successive pulse bursts and adjusting the relative intensities of the electrodes it is possible to animate the tactile display, so an apparently moving stimulation can be displayed. The vector is encoded as a moving, broad wave. The relative position, at a moment in time, of an electrode to this wavefront is calculated so that arbitrary vector directions can be animated. This is done by having another look-up table, this time residing in the computer. This holds the shape of the wave, which will be used to adjust the amplitudes of the individual electrodes. A wave travels in a direction perpendicular to the line of its wavefront and a wavefront is defined as a contour of equal phase. The perpendicular distance of an electrode to the desired wavefront is calculated and determines an initial offset (phase) into the look-up table for that electrode. The look-up table values are, after various scaling processes, passed to the DACs and the first frame of the animation is initiated. The time between frames is kept constant. This is so that the perceived speed of the complete computer system is not dependent on the speed of the wave stimulation. To control the speed of the wave the size of the increment, added equally to each of the offsets in the look-up table, is controlled. The smaller the increment, the slower the wave.

There are very many parameters which affect the perception of the stimulus and it would be impossible to optimise each of them in any quantitative way. Several pilot studies were conducted to qualitatively optimise these. Consider, for example, the sensation of apparent motion mentioned above. Speeds below the lower threshold cause the wave to be perceived in successive steps, and that was found to be confusing and resulted in difficulties with direction discrimination. Speeds above the upper threshold were felt as a single pulse over the fingers, with no discernable direction.

To maximise the probability of determining the absolute direction of a stimulation wave, the speed should be at the useable minimum. It was considered that the lower threshold of apparent motion could be reduced by using the method of phantom sensation. The required magnitude of the phantom sensation, as it moves from one electrode to the next, affects the shape of the wave. If the sensation is designed to be half way between two electrodes, the magnitudes of the stimulations on each of the electrodes are equal, but each is less than 50% of the stimulation magnitude of a wave, were it located directly under an electrode. Yet another parameter which affects apparent motion is the delay between successive waves. If too short, this can clash with the requirements of wave shape and, therefore, phantom stimulation perception, causing an inability to tell whether the stimulation is travelling forward or backward.

Summation of electrocutaneous stimulation occurs along a finger, but does not occur across fingers. When a wave is travelling perpendicular to the proximal - distal direction all three electrodes on a finger are energised at the same time and the intensity of the sensation is considerably higher than the sum of their individual sensations: a similar effect as the phantom sensation correction required above. To remove this problem another parameter had to be introduced. This attenuated the stimulation by an amount dependant on the direction of the vector wave.

4. EXPERIMENTAL SETUP

To evaluate its potential usefulness, an experiment has been conducted to evaluate users' directional acuity with this device. Quantitative parameters associated with maximising the dynamic range of sensation for an individual electrode were found to be in agreement with work done by Kaczmarek (1990) to optimise them. Qualitative parameters determined in our pilot studies are:

Wave shape: Shifted Cosine

Duty cycle: 0.3

Wave speed: 80mm/sec

Direction-dependent attenuation factor: 0.4

The pilot studies also showed that the length of time for an observer to make a judgement of the direction of stimulation was longer than expected. Psychological loss of attention and irritation are major problems in tedious psychophysical experiments. It was therefore decided that the experiment would be broken down into 5 tests, one per day. Each test consisted of 90 trials. A trial was one judgement of the direction of the stimulation. It was also decided that the experiment would be carried out over only 180 degrees to reduce the volume of data required. The experiment was designed so that it could be considered reasonable to extrapolate the results over 360 degrees.

The decision of the observer was entered graphically on the computer screen. A circle was displayed with an outer arc

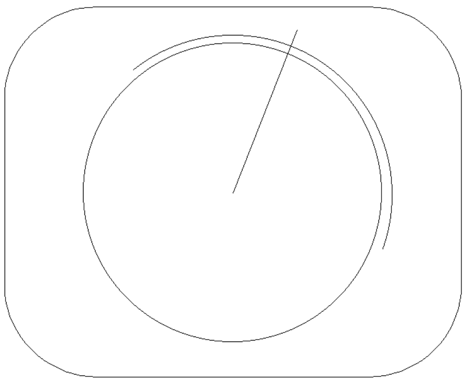


Figure 1. *Experiment input screen showing region of directions and radial arm.*

showing the region of directions to be displayed. This region was chosen arbitrarily from 140 degrees, moving clockwise to 310 degrees. For ease of data analysis, 12 o'clock position was identified with 180 degrees. The direction input was displayed with a radial arm (Figure 1). The direction of the arm was controlled with the mouse, clicking the left button to register the decision. The tactile trial vector was displayed until a decision was made.

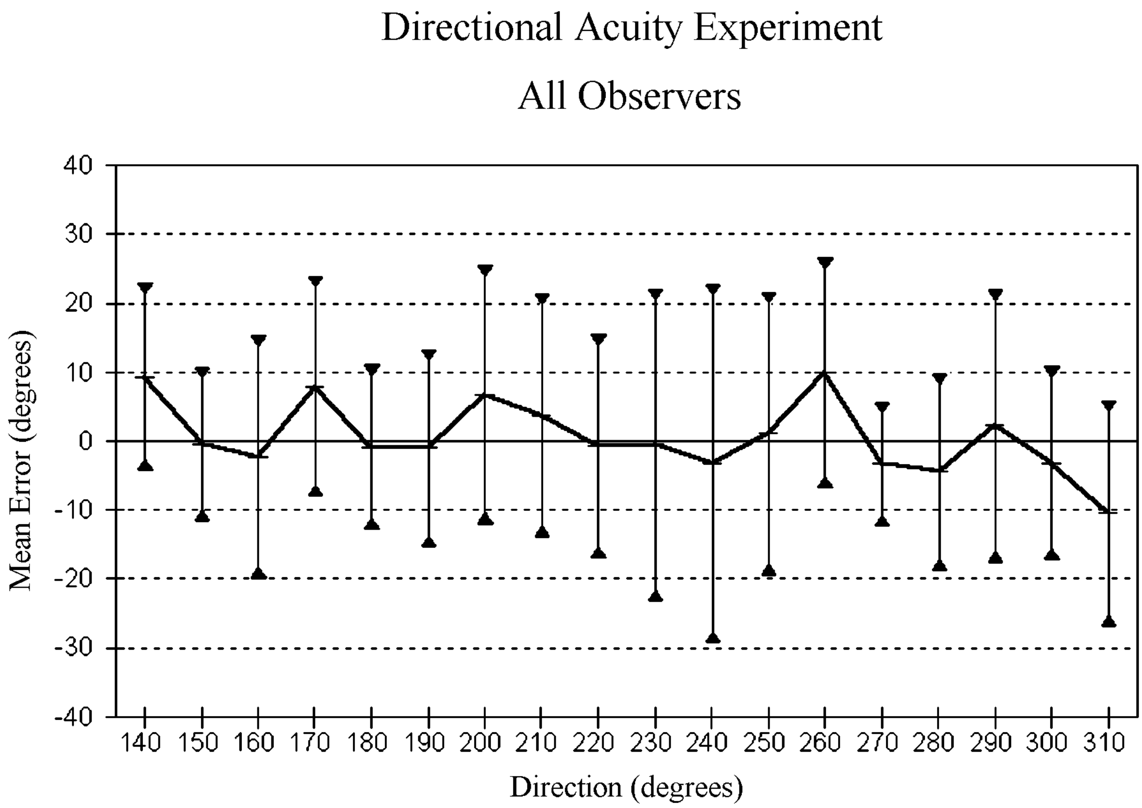


Figure 2. *Mean error for all observers combined (1350 trials). Error bars +/- 1 S.D.*

Unknown to all but one of the observers were the details of the experiment. The region of directions was broken down into 18 discrete directions at 10 degree increments. The trials were block-randomizes and each direction was presented 5 times. The observers were not even told that there were 90 trials to a test.

Feedback was, however, given after each trial. A radial arm pointing in the correct direction was briefly displayed together with the observers judgement so that any error could be visually assessed. A percentage score was also displayed, zero percent denoting an error of 180 degrees. When asked, none of the observers 'guessed' any of the details of the experiment.

The nine electrodes need to be individually calibrated for both zero (just on the threshold of sensation) and span (comfortably below the threshold of pain). The spans on each electrode needed to be adjusted so that their sensations were all about the same. In the calibration procedure each electrode was energised sequentially. Calibration was found to be difficult and was iterated with a 'learn' facility. Before each test the calibration was checked and adjusted as necessary, and learning allowed. The learn facility was visually the same as the test except that the observer controlled the direction of the stimulation via the mouse controlled radial arm.

5. DISCUSSION OF RESULTS

Three observers, including the principal author were used in this experiment. There was no significant difference in their performances, despite one of them knowing the details of the experiment. Within each observer's set of 5 tests there were no significant differences between individual tests. This shows that there was no learning (or otherwise) effect. These two results allowed all the data to be combined, making a total of 1350 trials, 75 at each direction. The mean error between the trial direction and the observers judgement was determined, as was the standard deviation. Figure 2 shows the mean error at each trial direction, the error bars are +/- 1 standard deviation. The mean decision time was approximately 9 seconds, or 12 complete wave cycles. This is a long time but is probably due to the experimental methodology. It is considered that in a 'real' situation judgements would be differential and the interactive nature of the interface would allow much faster assessment of direction.

Two observations can be made about the results from this graph. First, the results at the ends of the range may be affected by 'end effects'. There are fewer 'choices' of direction for a stimulation which is considered to be at, or near, an end. Not only that, but the choices are lop-sided. This manifests itself as a positive mean error at the 140 degree end and negative at the 310 degree end. Whilst this characteristic appeared on each observer's results, it was not found to be significant.

The second observation is that there appears to be no trend, or function, to the results. In vision, there is an oblique effect, whereby one's angular acuity is much better at the horizontal and vertical orientations. If there were a similar tactile effect, the mean error and standard deviation would be much smaller around 180 and 270 degrees. The absence of an oblique effect for this frictionless method of stimulation is in agreement with others, for example Norrsell & Olausson (1974). A tactile oblique effect has been noted when skin stretch has been utilised (Keyson & Houtsma, 1995).

Although many trials have been taken, the mean errors are still quite large. This is considered to be due to variations in calibration between tests. An unnoticed error in adjustment could cause the perceived direction to be consistently different from the true direction. The mean of the standard deviations is 15 degrees. When the dynamic range of the stimulation is taken into account, this could be taken as a measure of absolute directional acuity.

6. CONCLUSION

When evaluating the usefulness of these results, one has to bear in mind that this is an 'absolute' test of directional acuity. In other words, each stimulation is psychologically assessed in isolation. It would be desirable to carry out this experiment's visual analog, however no means could be simply devised to do this without introducing cues which would confound the experimental procedure. An earlier, unpublished experiment by the authors considered near-absolute grey scale intensity on a computer VDU, and will serve as a comparison.

In this grey scale intensity experiment a 256 shade grey scale was displayed on one side of the screen (area 100mm high, 50mm wide) and on the other side of the screen was a test shade (area 100mm high, 50mm wide). A thin horizontal line, under mouse control, was used by the observer to match the test shade to the correct point on the grey scale.

In both the tactile experiment and the visual grey scale experiment there were no trends in the standard deviations. There was also a simple linear relationship between the dependant and independent variables so a simple comparison between the two can be considered. Dividing the dynamic range of the sensory modality by the mean of the standard deviations, gives a dimensionless measure of acuity. In the visual test, the dynamic range was from zero to 255, black to white, and the mean of the standard deviations was 12. This gives a value of 21. Taking the mean of the standard deviations from the tactile directional acuity experiment above and a dynamic range of 360 degrees, the measure of acuity is 24. It should be noted that the acuity of an observer greatly depends on the design of the experiment. Generally, experiments which study relative, or differential, acuity produce 'better' results.

So far, only the direction of the stimulation has been considered. Future experiments should address the interaction between the direction and magnitude of the vector. The magnitude of the vector could be encoded in two ways. Either the overall magnitude of the sensation could be modulated in proportion with the magnitude of the vector, or the speed of the animated wave could be varied, or both.

After magnitude effects have been studied, further work proposed will concentrate on cognitive evaluation of the use of the display, rather than the psychophysical experiment presented above. These experiments will evaluate how well the user understands graphical information displayed tactually and not how well the tactile display performs. Though the exact format has yet to be decided, there are two main areas of interest: tactile only and visual/tactile displays.

'Tactile only' experiments would be especially relevant to visually impaired computer users. Many aspects of the graphical user interface could be tactually displayed and vector information used to guide the user to places requiring attention. Experiments could be conducted to determine the ability to locate 'targets' using tactile guidance. Access to graphical data, such as numerical diagrams, photo-realistic images, etc. could be enhanced due to the greater information capacity of an animated tactile display. To explore this, an experiment using a two dimensional vector field could be carried out. The user is given the task of extracting information from the display. This could be done by the user drawing streamlines on the visual display from information portrayed on the tactile display, identifying sinks and sources, vortices, etc. 'Visual/tactile' experiments would explore the ability to assimilate tactile and visual information at the same time.

One of our previous experiments with this tactile display (Eves & Novak[96]) discovered that there is a strong ability to learn animated tactile symbols, or icons. These could be used in an interactive environment in a similar manner to audible icons. They would have a great advantage over audible icons as they would be silent and so not disturb other people. This would also allow much more liberal use. It is considered that the number of animated tactile icons available on the current 9 electrode display is considerably greater than the number humanly discernable. The current display is theoretically capable of displaying 511 static icons, even without stimulation magnitude encoding. It would be useful to determine the rules and size of a dynamic tactile symbolic vocabulary. Given the great number of animated tactile symbols theoretically available, it should be possible to choose ones that are intuitive, or at least, easily learnt.

Combining directional tactile display and animated tactile symbols may provide better, more complete information display for visually impaired computer users. To give visually impaired people access to graphical information using computer technology requires the translation of graphical information into a form that can be input and processed by acoustical or tactile display devices. These devices have to have as large a repertoire as possible to give the translator of graphical information a large choice of display techniques, thereby helping to maximise the accuracy of rendition.

Acknowledgements: This research is supported by the E.P.S.R.C. through a studentship to David Eves.

7. REFERENCES

D S Alles (1970), Information Transmission by Phantom Sensations, *IEEE Transactions on Man-Machine Systems*, **11**, *1*, pp. 85-91

B M Collins (1993), Data Visualisation, In *Animation and Scientific Visualisation* (R A Earnshaw & D Watson, Eds), Academic Press, London, pp. 23-24.

S Coren (1978), *Sensation and Perception*, Academic Press, London, pp. 150-155.

D A Eves & M M Novak (1996), The Utilisation of Tactile Sensations in the Analysis of Complex Systems, *Proceedings, Eurographics UK Chapter 14th Annual Conference*, London, **2**, pp. 187-194.

E P Gardner & B F Sklar (1994), Discrimination of the Direction of Motion on the Human Hand - A Psychophysical Study of Stimulation Parameters, *Journal of Neurophysiology*, **71,** *6*, pp. 2414-2429.

T Geake (1994), *An Advanced feedback System for Myoelectrically Controlled Prosthesis, MPhil Thesis*, Kingston University, London.

K A Kaczmarek (1990), Maximal Dynamic Range of Electrocutaneous Stimulus Waveforms, *IEEE Transactions on Biomedical Engineering*, **37**, *7*, pp. 701-715.

K A Kaczmarek (1991), Electrocutaneous and Vibrotactile Displays for Sensory Substitution Systems, *IEEE Transactions on Biomedical Engineering*, **38**, *1*, pp. 1-16.

D V Keyson and A J M Houtsma (1995), Directional Sensitivity to a Tactile Point Stimulus Moving Across the Fingerpad, *Perception and Psychophysics*, **57**, *5*, pp. 738-744

U Norrsell and H Olausson (1994), Spatial Cues Serving the Tactile Directional Sensibility of the Human Forearm, *Journal of Physiology*, **478,** *3* pp. 533-540.

H R Schiffman (1995), The Skin, Body, and Chemical Senses, In *Sensation and Perception* (R L Gregory & A M Colman, Eds.), Longman, London, pp. 71-73.

D W Sparks (1979), The Identification of the Direction of Electrocutaneous Stimulation along Multistimulator Arrays, *Perception and Psychophysics*, **25**, *2*, pp. 80-87

A Y J Szeto and F A Saunders (1982), Electrocutaneous Stimulation for Sensory Communication in Rehabilitation Engineering, *IEEE Transactions on Biomedical Engineering*, **29**, *4*, pp. 300-308.

J Walton (1993), Get the Picture - New Directions in Data Visualisation, In *Animation and Scientific Visualisation* (R A Earnshaw & D Watson, Eds), Academic Press, London, pp. 30-32.

S F Wilker, G Vanderheiden, S Lee, & S Arndt (1991), Development of Tactile Mice for Blind Access to Computers: Importance of Stimulation Locus, Object Size, and Vibrotactile Display Resolution, *Proceedings, Human Factors Society 35th Annual Conference*, pp. 708-712.

The generation of virtual acoustic environments for blind people

D A Keating

Department of Cybernetics, University of Reading,
Whiteknights, PO Box 225, Reading, ENGLAND

cybdak@cyber.reading.ac.uk

ABSTRACT

VR systems, like the cinema, tend to concentrate on the visual image, leaving the audible image as garnishing. If blind people are to make any use of virtual environments then the audible image needs to be greatly improved. An ideal system would encode the amplitude and direction of sounds in a way that was independent of the means used to portray the image to the user. The Ambisonic B-format describes four signals: a reference, and three vectors. These are simple to generate and manipulate and may easily be converted into loudspeaker feeds for one, two or three dimensional arrays of speakers.

The generation of binaural signals for headset based systems is made difficult by the complexity of the head related transfer function(HRTF). Steering the virtual image to take account of head movements is computationally expensive and requires detailed knowledge of the users HRTF if the image is to be at all realistic. An alternative is to make use of virtual loudspeakers around which the image is moved but which have a fixed location relative to the users head. If the speaker feeds are derived from a B-format signal this may easily be panned and tilted by manipulating the relative amplitudes of the three vectors according to simple mathematical rules. The HRTF of the user can be measured at the positions of the virtual speakers and the resultant relationship between the combination of all these signals and the two headphone signals need only be computed once.

This paper gives an introduction to the Ambisonic system and shows how B-format signals can be used to encode a virtual acoustic environment. It concludes by describing how the B-format signals may be manipulated and used to generate almost any number speaker feeds or headphone feeds.

Keywords: surround sound, ambisonics, acoustic modelling

1. INTRODUCTION

1.1 What is a virtual acoustic environment?

An acoustic environment (virtual or real) consists of a number of primary sound sources in 3-D space and a number of reflecting surfaces. Each of these surfaces will have characteristics depending on its sonic properties. For example hard solid surfaces will reflect all frequencies well but the addition of a soft covering will increase the absorption at high frequencies and hence lower the reflectance. Suspended panels absorb particular bands of frequencies; glass for example can be a good absorber at low frequencies but a good reflector at high frequencies. The sense of space perceived by an observer is primarily governed by the timing and direction of reflections from the surfaces. A sense of depth in stereo recordings of real sounds is usually due to the timing of early reflections from near surfaces such as the walls or floor of the room in which the recording was made.

1.2 Interacting with an acoustic environment

In traditional audio entertainment, or indeed cinema, the experience is totally passive. In a virtual acoustic environment the user is able to move around the environment thus increasing their involvement and allowing them to build up a better mental image of the acoustic space. To further the ability of the user to build up this image active involvement is required. If the user makes a sound then they should hear the response of the environment to this sound. This means not only that objects in the environment should react to sound (for example birds could stop chirping when disturbed by sound) but that the sound should be treated as any other primary source and appear to be reflected by the many surfaces in the environment. This would allow the ‘acoustic’ of otherwise quiet places to be determined by the user.

2. REPRESENTING DIRECTION

As far as the user of a virtual environment is concerned, the environment consists of a number of sounds each of which comes from a particular direction. The method by which these sounds are finally to be portrayed to the user should not concern us at this moment. The biggest failing in surround sound systems for audio entertainment is that the designers failed to realise that the need was to encode the directions of the sounds, not the method by which the image was finally to be created. The ill fated Quadraphonic system of the 1970s set domestic surround sound back by about 15 years and Dolby Surround threatens to do the same again at the present time. The ambisonic system of representing direction (in 3-D) in audio signals was conceived at the same time as Quadraphonic and thus despite being based on sound psycho-acoustic and engineering principles (Gerzon, 1974) was tarred with the Quadraphonic brush.

2.1 Ambisonic B-Format

The ambisonic B-format representation of sounds including direction consists of four signals. These signals in turn represent the zero and first order spherical harmonics in three dimensions (Gaskell, 1979; Gibson et al, 1972). The zero order harmonic (denoted W) is simply a sphere which has an amplitude which is independent of direction. The three first order harmonics consist of dipoles in three mutually orthogonal directions. These are (by convention) Front-Back (denoted X), Left-Right (denoted Y) and Up-Down (denoted Z). The B-format representation may be expanded to contain the second order harmonics which are quadrapoles. This gives the ability to improve the discrimination of direction (Gerzon, 1976; Bamford and Vanderkooy, 1995) but requires a total of nine signals to represent all three dimensions or five to represent the horizontal plane only. The advantage of increased discrimination is at present outweighed by the disadvantage of having to produce five or nine channels of audio output. The basic B-format representation which consists of a reference and three vectors is mathematically simple to produce and (with the exception of binaural) it is also simple to transcode into other formats.

2.2 Ambisonic UHJ

The ambisonic UHJ format is a hierarchical format which allows surround sound to be encoded in one to four channels. One channel gives a mono (directionless) signal, two give horizontal surround of low directivity but which is stereo compatible, three give horizontal surround of higher directivity and four give full periphonics (with height). Two channel recordings (strictly BHJ but usually referred to simply as UHJ) are available in the domestic audio formats. The two channel format uses the relative amplitudes of the two signals to encode left-right information and the relative phase to encode front-back information. UHJ recordings are usually produced from B-format masters, a soundfield microphone or similar pseudo-coincident microphone technique being used to provide the B-format signals directly for recording of the master.

2.3 Binaural

The binaural representation of direction attempts to mimic that which is available to humans; that is the two signals which are received by the ears. A major problem is that these signals are unique to each of us and so there is no standard definition of them. There are, however, a number of commonly used pseudo-standards; the BBC use a perspex disc of a certain diameter and define the signals as those which would be detected by microphones at a certain distance either side of the disc. The outputs from proprietary dummy heads such as those produced by Bruel and Kjaer or Senheiser might also be considered as standards. The mathematical relationship between the direction of the sound source and the two binaural signals is extremely complex even when generalisations and simplifications have been made. Thus not only is it difficult to produce these representations but if they are not to be used directly to provide the feed signals of headphones it is also difficult to derive any other signals from them.

2.4 Dolby Surround

The Dolby Surround (or Dolby Stereo) formats for cinema or home video surround sound are perhaps the most common surround sound formats in use today. Unfortunately there is no full definition of direction associated with these formats. Four signals are defined: Centre (for the dialogue channel), Right and Left (primarily for the stereo music soundtrack) and Surround (for all sound effects). Two channels are produced from these four by means of a simple phase-amplitude matrix. The primary use of the Dolby formats is to provide a surround effect to augment a dominant visual image and not to accurately reproduce an acoustic environment. Despite the lack of full definition there are proprietary processors available which will produce Dolby signals from mono sources and directional information. In order to reproduce the direction intended, the loudspeaker layout must be identical to that used by the manufacturer who produced the processor. This is not usually defined in detail, and even if it were the basic Dolby format is incapable of positioning signals from the rear with any kind of real accuracy (all the rear speakers carry the same signal). The more recent 5.1 channel (digital) format does allow positioning at the rear but is still not fully defined and suffers from the same problem as the earlier formats in that the speaker positions are fixed by the format.

2.5 Dolby Pro-Logic

Dolby Pro-Logic does not refer to a format in its own right but refers to the type of decoder used. A pro-logic decoder takes the normal two Dolby Surround signals and steers the four output signals depending on the dominant direction of the source. This gives better directivity than the simple decoder when there is a dominant direction but can offer no advantage if many directions are present at once. An additional limitation of the Dolby formats is that they have no provision for height information.

3. MANIPULATING THE DIRECTION

Helmet based VR and telepresence systems (Griffin and Keating, 1992) require acoustic and visual images which track head movements. Such helmet based systems traditionally use headphones for the audio reproduction. VR head tracking systems are available commercially which give information on head position and orientation. If the head orientation signals are used to steer the surround audio then the imagery obtained from the headphones can be good. Any front-back ambiguity may be resolved by small head rotations although the effect of the headphones on the conch resonance (Moore, 1989) needs to be compensated for as this is a potential source of lack of realism. Even if the system is not helmet based then the surround signal may need to be steered as the user is likely to be able to rotate in the virtual environment albeit by joystick or similar control.

3.1 Steering Two-channel Sources

If we are considering a generalised surround signal steering block then the major problem is the large number of different formats which must be catered for. In addition to this, each of the formats has its own set of problems. We cannot, for example, rotate a stereo signal and produce a stereo format output as stereo is not a surround format. In Dolby surround, attempts to mix between speaker feeds (L, R, C & S) to achieve rotation yields strange results as the front and surround signals are simple combinations of left and right. Rotating the sound image by 180 degrees would in any case be impossible as there is no side information at the rear due to the speaker layout used.

Steering UHJ in 2-channel format is difficult as the direction information is encoded using amplitude and phase difference, thus a rotation is not a simple function of the two signals. The same is true to an even greater extent with binaural sources. Not only must the amplitudes and phases be manipulated in a complex fashion but the generalised head related transfer function (HRTF) must be known (Wenzel 1992). This makes steering these formats directly very difficult.

3.2 Steering Multi-channel Sources

Multi-channel formats fair little better if they are defined in terms of speaker feeds. The one format which is a notable exception to this is ambisonic B-format. The components of the B-format can be defined as follows: W (the reference signal) has a gain of 0.7071 regardless of direction, the vector Y (left-right) has a gain of sin(A) * sin(B), X (front-back) has a gain of cos(A) * cos(B) and Z (up-down) has a gain of sin(B). Where A is the angle of the source measured counter clockwise from the front and B is the angle of elevation.

Steering these signals is simplicity itself (Gerzon, 1975; Malham and Clarke, 1992). To rotate a signal C degrees counter clockwise the following equations are used:

$$newW := W \quad (1)$$

$$newX := X.\cos(C) - Y.\sin(C) \quad (2)$$

$$newY := Y.\cos(C) + X.\sin(C) \quad (3)$$

$$newZ := Z \quad (4)$$

In other words W and Z are unchanged and the new X and Y terms are combinations of the old terms modified by gains in the range -1 to +1.

Tumble (rotation about the Y-axis) by an angle D may be achieved as follows:

$$newW := W \quad (5)$$

$$newX := X.\cos(D) - Z.\sin(D) \quad (6)$$

$$newY := Y \quad (7)$$

$$newZ := Z.\cos(D) + X.\sin(D) \quad (8)$$

Tilt (rotation about the X-axis) by an angle E may also be added if required as follows:

$$newW := W \tag{9}$$

$$newX := X \tag{10}$$

$$newY := Y.\cos(E) - Z.\sin(E) \tag{11}$$

$$newZ := Z.\cos(E) + Y.\sin(E) \tag{12}$$

Each of the new vectors is thus a sum of each of the old vectors multiplied by a constant, where the constants depend on the manipulation required. The reference remains unchanged throughout. It makes a great deal of sense to produce our virtual environment using the B-format definition of direction as manipulating the signals to allow for movements of the user is so easy. If, however, we wish to incorporate recorded sounds which are in another format into our virtual environment they must first be converted into B-format.

4. CONVERTING SURROUND SOURCES TO B-FORMAT

Fortunately there already exist methods of producing pseudo B-format signals from stereo and UHJ sources. These are not 'pure' B-format as the directional information is contained in both amplitude and phase. The steering algorithms quoted earlier, however, are unaffected by this and give the same results as they would with 'pure' B-format.

4.1 Converting Dolby Surround to B-format

It remains necessary to develop methods of converting Dolby surround, binaural and the multi-channel formats into B-format. Although it would be possible to decode the 2-channel Dolby surround signals directly the resultant surround channel (L-R) requires decoding by a modified B-type noise reduction decoder. It is therefore better to take the centre (C) and surround (S) channels from an existing Dolby-surround decoder with zero delay on the surround channel..

A transcoder developed by the author (Keating and Griffin, 1993) which was found to give acceptable results used the following equations:

$$W = 0.5\,(C - jS) \tag{13}$$

$$X = 0.7\,(C + jS) \tag{14}$$

$$Y = S \tag{15}$$

These are similar to the 'super stereo' or 'enhanced' equations usually used in ambisonic UHJ decoders but with increased difference signal. It is not suggested that these are optimal but they give reasonable results all the same.

4.2 Converting Dolby Pro-logic to B-format

Using a Dolby Pro-Logic decoder presents some problems as the outputs may or may not be 'steered' electronically by the decoder depending on whether there is any dominant direction present. A converter must therefore give acceptable results when presented with simple surround sound signals or (in the extreme case) only one active output. If for example there is only dialogue present the pro-logic decoder will reduce the left, right and surround output signal amplitudes considerably, leaving only the centre channel active.

A set of transcoder equations which give the same results as those given earlier for surround signals but also copes with Pro-Logic signals is as follows :

$$W = 0.25C + 0.177(L + R) - 0.5jS \tag{16}$$

$$X = 0.467C + 0.165(L + R) + 0.5jS \tag{17}$$

$$Y = 0.7(L - R) \tag{18}$$

4.3 Converting UHJ to B-format

Standard ambisonic decoder equations may be used to produce B-format from UHJ. These are as follows:

$$W = 0.667\Sigma + 0.11j\Delta \tag{19}$$

$$X = 0.556\Sigma - 1.099j\Delta \tag{20}$$

$$Y = \Delta + 0.513j\Sigma \tag{21}$$

where:

$$\Sigma = (L + R)/2 \tag{22}$$

$$\Delta = (L - R)/2 \tag{23}$$

4.4 Converting Stereo to B-format

Stereo signals may be considered as two mono signals and placed anywhere but the following equations allow the stereo image to be positioned to the front with a width (α) of up to 180 degrees:

$$W = 0.65\Sigma - 0.27j\Delta(\alpha/180) \tag{24}$$

$$X = 0.98\Sigma + 0.4j\Delta(\alpha/180) \tag{25}$$

$$Y = 0.75\Delta(\alpha/180) \tag{26}$$

4.5 Converting Binaural to B-format

The problem format is binaural because there is no standard definition of the Left and Right signals. These will in general be similar but will depend on the exact recording/synthesising technique employed. Although side information is at high frequencies carried in the relative amplitudes of the two signals this reverts to time difference (or phase) at frequencies below about 700 Hz. In addition front-back information is carried in subtle clues caused by the head and ear shapes. It is thus very difficult to decode or transcode the directional information present in binaural signals and for this reason this has not yet been considered by the author.

5. CONVERTING B-FORMAT SIGNALS TO SURROUND

In order to give versatility to a virtual acoustic environment generating system it is useful to be able to provide outputs in any of the common surround-sound formats

5.1 Converting B-format to Dolby Surround

The author has used a standard UHJ encoder driving a two-channel Dolby surround decoder with some success. The following transcoder equations for four channel discrete or two channel Dolby surround should be a good starting place for readers wishing to experiment:

$$C = W + 0.71X \tag{27}$$

$$L = W + 0.71Y \tag{28}$$

$$R = W - 0.71Y \tag{29}$$

$$S = jW - 0.71jX \tag{30}$$

or:

$$L = (0.86 + 0.35j)W + (0.25 + 0.25j)X + 0.35Y \tag{31}$$

$$R = (0.86 - 0.35j)W + (0.25 - 0.25j)X - 0.35Y \tag{32}$$

5.2 Converting B-format to UHJ

The standard UHJ encoder equations are:

$$L = (0.4699 - 0.17j)W + (0.0928 + 0.255j)X + 0.3278Y \quad (33)$$

$$R = (0.4699 + 0.17j)W + (0.0928 - 0.255j)X - 0.3278Y \quad (34)$$

5.3 Converting B-format to Stereo

The UHJ encoder equations above may be used to provide a stereo compatible output but the following simpler equations may be used:

$$L = W + 0.35X + 0.61Y \quad (35)$$

$$R = W + 0.35X - 0.61Y \quad (36)$$

These have the advantage that ninety degree phase shifters are not required but suffer from the disadvantage that sounds from directly behind the user are quieter than if the UHJ encoder is used. As stereo is not a surround format the rearward sounds will appear to come from the front

5.4 Converting B-format to Binaural

The translation of B-format signals to binaural could potentially be as difficult as binaural to B-format but fortunately we can cheat. It would be very difficult to derive a generalised transfer function which translated B-format signals directly to left and right ear signals. The complete head related transfer function (HRTF) (Wenzel, 1992) would have to be known and filters derived from this to modify the left and right ear signal amplitudes and phases. Determining the HRTF over a full sphere experimentally would require a great deal of time in an anechoic chamber; whilst implementing the translation functions thus derived would require a powerful DSP system or massive amounts of analog electronics.

The HRTF may of course be computed (Rasmussen and Juhl, 1993) rather than measured but once again the implementation of such a complex transfer function is difficult. The solution is to only measure or compute the function at a few fixed angles. An ambisonic decoder driving four speakers placed around the listener is capable of giving a good illusion of direction in the horizontal plane. It can be argued therefore that if the HRTF is known for these four positions then that is all the information that is required. (In fact symmetry reduces this to two). Calculating the cumulative effect of these speakers at the ears results in right and left ear signals in terms of the four speaker feeds. Each speaker feed is defined in terms of the B-format signals W, X and Y (for horizontal only) and so an overall definition of the right and left ear signals in terms of the B-format signals may be derived very simply. This then gives us the translation of B-format to binaural. The number of virtual speakers is of course arbitrary and a greater number will give better results but will require the HRTF to be known at a greater number of source directions. This method would of course also work for periphonics if a three dimensional array of at least eight loudspeakers was considered. Increasing the order of the spherical harmonics from first to second would give improved directivity but would require a minimum of six virtual speakers in the horizontal plane and twelve if all three dimensions were required.

6. CHOICE OF ACOUSTIC OUTPUT METHOD

The first choice that must be made is whether to use headphones or to use an array of loudspeakers.

6.1 The advantages and disadvantages of using headphones

Headphones have many advantages over loudspeaker arrays. They may already be part of the VR system available to the system developer, and if not they are inexpensive to add. In use they are compact and most importantly they limit the sound they produce to the ears of the user. This becomes of great significance if the user is to interact with their environment by means of a microphone. They do, however, require a head tracking system if head movements are to give the normal additional directional information used to resolve ambiguities. They also require drive signals which are complex to calculate if realism is to be achieved.

6.2 The advantages and disadvantages of using loudspeakers

Arrays of loudspeakers may be driven from B-format signals via a simple decoder (Gerzon, 1980). Many different layouts may be used with different numbers of loudspeakers. At least four loudspeakers are required for horizontal surround sound and at least six for full periphonics (although eight are preferable as the birectangular layout allows compatibility with horizontal surround and stereo). The loudspeakers should be identical but need not be full-range units as a sub-woofer may be driven from the W output of the B-format signals. The cost of such a set-up is still high as four or more 2-channel (stereo) amplifiers are also required to drive these speakers. The set-up is not easily portable and so is best set-up in its own room. This also avoids the sound from annoying other people in the vicinity of the VR set-up. The main advantage of loudspeaker arrays is that errors in decoding degrade the image gradually, whereas those in the binaural (headphone) system can cause massive perceived errors in direction or highly ambiguous results.

6.3 Experimental Results

Initial experiments by the author used an ambisonic decoder driving four speakers placed around a dummy head, which in turn was driving a pair of headphones. The speakers and head were placed in a dead room to prevent room acoustics from having an effect. In this simple experiment the B-format signals were derived from a two channel UHJ source and were not steered.

The image from the loudspeakers was realistic and open. A 'walk-around' test showed that the directions of the sounds were being reproduced correctly. There was some lack of stability of the image if the listeners head was rotated but this should be less of a problem with a true B-format source rather than one decoded from 2-channel UHJ.

The headphone signal was also perceived to be open, with little of the 'in the head' localisation common with headphones. The 'walk-around' test however showed the classic front-back ambiguity to be present, with different subjects perceiving different directions. Is it thought that the addition of a head tracker steering the signals will help considerably. The decoder is now being implemented using DSP to replace the 'acoustic' decoder used in the initial experiment. This will eliminate any unwanted effects due to low frequency room resonances or imperfections in the loudspeakers or dummy head, and should allow the measured HRTF from the user to be used rather than that of the dummy head.

7. CONCLUSIONS

In this paper the problems of providing VR users, particularly those without the benefit of sight, with a satisfactory image of a virtual acoustic environment have been considered. The ambisonic B-format representation of direction of acoustic sources has been suggested as a sensible one to use in a VR system. The methods by which other surround formats may be converted to B-format have been given, as have methods of converting B-format into other formats. A solution to the particular problem of producing binaural signals has been suggested and the early experimental results show this method to hold promise.

8. REFERENCES

J S Bamford and J Vanderkooy (1995), Ambisonic Sound for Us, *Proc. 99th AES Convention*, Pre-print No 4138, October 1995.

P S Gaskell (1979), Spherical Harmonic Analysis and Some Applications to Surround Sound, *BBC Research Dept. Report*, BBC RD 1979/25.

M A Gerzon (1974), Surround-Sound Psychoacoustics, *Wireless World*, **80**, April, pp. 36-39.

M A Gerzon (1975), Panpot and Soundfield Controls, *NRDC Ambisonic Technology Report No 3*, August.

M A Gerzon (1976), Maximum Directivity Factor of n-th order Transducers, *J. Acoust. Soc. Am*, **60**, July, pp. 278-280.

M A Gerzon (1980), Practical Periphony: The Reproduction of Full-Sphere Sound, *Proc. 65th AES Convention*, London, Feb. 1980.

J J Gibson, R M Christensen and A L R Limber (1972), Compatible FM Broadcasting of Panoramic Sound, *J. Audio Eng. Soc*., **20**, Dec., pp. 816-822.

M Griffin and D Keating (1992), Factors Affecting Sound Synthesis and Presentation Within Virtual Reality Systems, *Proc. I.O.A*., **14**, Part 5, pp. 273-280.

D A Keating and M P Griffin (1993), The Production of Steerable Binaural Information From Two-Channel Surround Sources, *Proc. I.O.A*, **15**, Part 7, pp. 173-185.

D G Malham and J Clarke (1992), Control Software for a Programmable Soundfield Controller, *Proc. I.O.A*., **14**, Part 5, pp. 265-272.

B C J Moore (1989), *An Introduction to the Psychology of Hearing*, 3rd Edition, Academic Press, London, 1989.

K B Rasmussen and P M Juhl(1993), The Effect of Head Shape on Spectral Stereo Theory, *J. Audio Eng. Soc*., **41**, No 3, March.

E M Wenzel (1992), Localisation in Virtual Acoustic Displays, *Presence*, **1**, No 1.

Proc. 1st Euro. Conf. Disability, Virtual Reality & Assoc. Tech., Maidenhead, UK, 1996
 ISBN 0 7049 1140 X

Interfaces for multi-sensor systems for navigation for the blind

Penny Probert, David Lee and Gordon Kao

Robotics Research Group, Department of Engineering Science, University of Oxford,
19 Parks Road, Oxford, OX1 3PJ, ENGLAND
Tel: +44-1865-273926 Fax: +44-1865-273908

pjp@robots.ox.ac.uk, david@robots.ox.ac.uk

ABSTRACT

This paper describes work towards a multi-sensor system to assist blind people to move around an urban or indoor environment. The communication of navigation information is constrained by both the type of information imparted and, in many ways more crucially, by the type of information which real sensors can extract. In this paper we describe the use of ultrasound to provide two types of information: first the low level directional information which may be provided currently by a guide dog and then information on features to allow the blind person to locate and direct themselves within the environment. A directional system with a simple vibrator interface is described and its performance assessed by a group of users. Finally we discuss the possibility of feature extraction for navigation by recognition and give some results using frequency modulated sonar.

Keywords: ultrasound, mobility aids, blindness, human interface

1. BACKGROUND AND STRUCTURE OF PAPER

Vision is by far the most important sense for most people, and trends in virtual reality and computer interfaces reflect this emphasis. Sighted people appreciate detailed images and are able to abstract relevant data from images with ease, and the challenge in virtual reality is not in abstracting the data but providing enough detail at sufficient speed. In contrast, the difficulty in presenting information to blind people is reducing information: non-visual channels are less direct and the available sensory channels have much lower bandwidth.

In this paper we address the problem of abstracting and presenting information for navigation. A number of sensors to assist mobility have been on the market for some time, normally using ultrasound as the sensing modality. However they have not met with widespread success. We see two reasons for this. One is that the information is often presented as a continuous, or near continuous, tone, usually through headphones, whereas blind people need their ears to pick up other signals from the environment. The work on the auditory representation of the environment is encouraging as it shows the possibilities of using sound as a rich information system, but does not overcome this problem [Kay95]. Another reason for failure is the restriction on the information which can be gathered. In the ideal world people would like a system which can pick out and convey information on landmarks such as bus stops, on pavement hazards such as parked cars and holes, and on general environmental features such as steps, and to pass on such information in a form so that the user can build up an internal map of his immediate location and probably relate to previous information. Unfortunately this is a very difficult goal to achieve, especially if significant data abstraction has to occur computationally to leave the audio channels free.

Another method of presenting information to the blind, which is more directly analagous to vision in providing spatial awareness, is through a tactile interface. There has been considerable research over the last two to three decades on the provision of an array which might act as a form of vision for blind people and some work suggests that certain functions, for example to do with grasping objects, could be stimulated by this method [Bach-y-Rita95]. However it seems clear that a fully visual sense cannot be conveyed, especially to older people, partly because of the quality of tactile array which can be made and partly because of the problems of interpretation by the brain. Work on the provision of tactile maps has similar problems because of technological problems in making tactile arrays: precision, dynamic range and cost are all major limitations [Frick96]. In addition the tactile sense gets tired easily, and for a large category of the blind, those whose blindness is associated with diabetes, it is often degraded.

There must therefore be a balance between the extent to which information should be abstracted before presentation to the user, which depends on the type of navigational information needed and on the sensing modalities available. Computer vision can provide huge amounts of data but abstraction is difficult and the complexity is unnecessary for simple navigational commands. In this paper we extract two levels of navigational information, and discuss how to present information for each. We describe the progress so far using ultrasound sensing, a compact and low cost technology.

2. INFORMATION PROVISION FOR NAVIGATION

We can view navigation as one of two tasks: a low level task in which someone simply wants directions to follow a safe route along a known path (which we call micro-navigation) , and a more information based approach in which the user wants to move him or herself through the environment by recognising and navigating from landmarks. Current mobility aids fulfil mainly the first of these. Guide dogs direct a person along a path through pressure on a handle and convey no higher level information on context. The long cane can provide more information since as well as warning of hazards it can be used to probe the immediate environment, both through touch and through the echo from a tapping sound. However the information it gathers is restricted and neither aid provides information on anything but the very close environment.

The work we describe in the next sections consists of two systems, one for micro navigation and the other for feature based guidance. For the former we draw from work on robot guidance using time of flight sonar, which has been shown to work reliably in a wide range of environments [Borenstein88]. These sonar devices simply return the range of the nearest obstacle in the path of an ultrasound beam. Ultrasound sensors are cheap, have a wide field of view and reasonably modest power requirements. Feature based guidance is more of a problem. There is no well developed system to draw on here; all available techniques either require very well structured environments or are very restrictive. A different type of sonar sensor, which has been used for blind people through an audio interface (various configurations have been marketed such as the Sonicguide [see RNIB94]), offers an interesting solution, and keeps the cost of the system low. This is the basis of the system we describe in section 4, but, in contrast to the sensor as sold, we reduce the quantity of information presented audibly by including some automatic data abstraction. Ultimately other sensors, such as vision, may be used as well to extend the range of features which can be observed. Computer vision is sometimes seen as the panacea for all these problems but even with full stereo vision it is hard to identify many features in an uncontrolled environment because of the problems of changes in lighting, poor calibration from camera movement and occlusion. The configuration of ultrasound we use has good immunity to noise and, because of the wide beam, can generally handle partial occlusion.

We make no attempt to recognise detailed man made landmarks such as bus stops or telephone kiosks. As well as the fact that there is a huge variety of these and it is hard to build a generic model, a much more sensible approach to providing location is through a talking beacon identifying the landmark (see, for example, [OPEN95]).

For micro-navigation, the user requires a stream of directional information to guide him along a path, and around obstructions. One simple way of doing this, which has proved successful for guide dogs is using a handle. However, as we had expected, and following other studies [Petrie95] potential users favour hands free operation so although the handle seems to have advantages in terms of reliability and ease of use this option was rejected. The channels remaining are tactile or auditory. As we discussed in the last section, each of these presents problems, and detailed information cannot be communicated through a tactile interface. Therefore we felt that there was no point in trying to provide a detailed local image, instead experimenting with a simple reactive interface, in which we simply provide instructions such as: "turn right, turn left, turn around". We describe a system which communicates this information in the next section. The first implementation was simply through an audio interface, and, as predicted, it proved irritating in use. However for so few commands it is easy to implement a tactile interface with vibrators, which are similar to the buzzers in pagers. Experiments showed that for the cheaper vibrators the human skin can detect very little in the way of frequency or amplitude of vibration, so a simple on-off interface was used. In the next section we present results from users of a sonar belt using this simple interface, which has proved very successful.

The higher level of navigation is through feature recognition, maybe to assist in localisation in a known street or office environment. Information is likely to be required less frequently, possibly only on demand and so the auditory interface is acceptable and versatile. At present we are investigating the level at which to process information in this area: whether to attempt full pattern recognition computationally and communicate in words, or to present the information as a sound following a minimum of processing pattern (an option which may offer better feature recognition but at the expense of requiring more of the audio channel). We discuss these issues further in section 4.

3. MICRO NAVIGATION: FOLLOWING A ROUTE SAFELY

The micro navigation task is implemented using time of flight sonar sensors. These emit a brief burst of ultrasound at about 50 kHz and then measure the time until an echo is detected. Multiplying this time by the speed of sound gives the round-trip distance to the reflecting object, which in turn gives its range. The key properties of this sensor are:

- Range accuracy of about 1 cm when the object lies in the centre of the sensor's beam. Overestimates of up to 4 cm can occur when the object is to the side of the beam.
- Low angular precision. The effective width of the ultrasonic beam is often quoted as 30°, although certain types of object can be only be detected through angles as small as 18°.
- Only the first reflecting object within the beam is detected.

A 'sonar belt' was constructed with three time-of-flight transducers mounted on it, directed horizontally. The angle between adjacent transducers was set at 15°. This was found to be an effective compromise between the need to ensure that all obstacles will be detected and the desire to cover a wide area in front of the user. With this separation a region about 60 cm wide was covered 2 m in front of the user.

The three transducers fire in sequence. To avoid crosstalk, enough time was left for an echo to have returned from an object at the maximum range of interest before the next sensor was fired. The shorter the imposed maximum range, the more frequently the sensors can be fired. A maximum range of 3 m was selected (considerably less than the theoretical maximum of about 10 m), making it possible to collect range data at approximately 50 Hz.

Two different versions of the sonar belt were constructed and tested. The first used a set of audible tones and the second used vibrating motors. These two versions are described in the following sections.

3.1 Audible Output

Section 1 described the possibility of analysing the sensor results before passing them to the user. As a first example of this approach, this version of the sonar belt suggested a safe direction of travel to the user, instead of simply reporting the approximate location of obstacles. A threshold value was selected (2 m in the experiments reported here) and the user was given turn instructions depending upon which, if any, of the range readings is less than the threshold. A simple rule base translated the sonar readings into commands.

The turn actions were given to the user through an earpiece. Three different tone frequencies were used for 'turn left', 'turn right' and 'stop'. The system was silent when the user could move straight ahead.

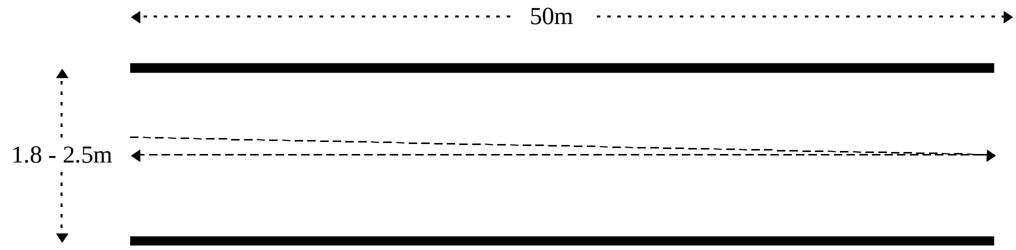


Figure 3-1. *Test run for the audio sonar belt*

To test the system, the user was required to walk blindfold along a test path which was bounded on both sides by walls, turn around at the end (when directed to do so by an assistant) and walk back to the starting point. The user's objective was to avoid collisions both with the walls and with other pedestrians using the path. The dimensions of the test path were as shown in Figure 3-1. To give a 'base case' for comparison, the same test was performed by a trained volunteer using a long cane.

Table 3-1 show the experimental results. A user of the belt was able to complete the test route at an average speed of 0.77 m/s, demonstrating the potential value of time-of-flight sensors in obstacle avoidance and path following. This speed is about 82% of the average speed that was achieved by the long cane user.

Most of the small number of collisions and sudden stops in the test results were caused by interactions with other pedestrians using the path. The narrowness of the path made it difficult to negotiate a clear path past the pedestrian. No meaningful comparison can be made with the number of collisions made by the long cane user because the collisions of the cane are the sensing mechanism itself.

Table 3-1. *Test results from audio sonar belt*

Test Run	Time (sec)	Speed (m/sec)	Collisions or Sudden Stops
1	145	0.69	1
2	130	0.77	2
3	141	0.71	0
4	135	0.74	0
5	154	0.65	0
6	111	0.90	0
7	132	0.76	1
8	129	0.78	0
9	110	0.91	0
10	119	0.84	1
Average	**130.6**	**0.77**	**0.4**

The following observations were made during these experiments:

- Listening to the environment was very helpful. It was, for example, easy to tell when the end of the walled pathway had been reached because the echoes from footsteps became much quieter. The presence of the earpiece and the turn signals could mask some of these audible clues.
- A fraction of a second was needed to convert the tone signal into a direction of travel.
- In some circumstances, the sensors are used for active examination of the environment. The user could, for example, make small body movements to determine the extent of an obstacle. Immediate sensor data, instead of turn signals, is more useful in this situation. The use of the rule base to convert the sonar signals into different turn actions appears unnecessary.
- The same range threshold is not appropriate in all circumstances. It would be helpful for the user to be able to select different thresholds in different environments.

With these points in mind, the tactile version of the belt was created and submitted to tests by blind and visually-impaired users. This device is described in the following section.

3.2 Tactile Output

This version of the sonar belt differs from the previous version in the following ways:

- Each transducer has an associated vibrating motor which is mounted on the inside of the belt, close to the transducer.
- Each motor vibrates constantly when the range measurement of the corresponding sensor is less than the threshold. No pre-processing of the range data is performed.
- The user can adjust the range threshold as necessary.

This system has recently been tested by 16 blind and visually-impaired people at the training centre of the Irish Guide Dogs Association in Cork. Of these people, thirteen used a guide dog and three a long cane, and fourteen made extensive use of these. They had the following profile:

1. Age

under 25	24-45	45-65	65+
0	9	6	1

2. Visual loss

Total	Perception of light only	Poor partial sight	Useful residual vision
10	2	1	3

Proc. 1st Euro. Conf. Disability, Virtual Reality & Assoc. Tech., Maidenhead, UK, 1996
©1996 ECDVRAT and University of Reading, UK ; ISBN 0 7049 1140 X

3. Home environment

City	Urban	Rural
5	9	2

They were asked to negotiate the obstacle course used for training guide dogs and their handlers. The obstacle is delimited by a chain-link fence (which was found to be easily detectable by the sonar transducers, but is a problem for the long cane user as there is no ground plane obstruction), and consists of items such as a dustbin, a typical work in progress sign of the type on pavements, a bicycle and a horizontal barrier. Figure 3-2 shows the belt being tested by two mobility instructors at the training centre.

Figure 3-2. *Testing the tactile sonar belt*

The group of blind people underwent a day's training by the mobility instructors and were then asked their opinion of the belt. The functionality for path guidance was rated highly, with the vibrator operation deemed intuitive and well liked. The only negative comments were to do with the mechanical configuration, which 'did not allow for a sufficient range of stomach shapes!' and there were some problems owing to sonar 'blind spots'. These are easy to adjust and our next version of the belt should have all round coverage. However on the cautious side it must be borne in mind that this group were already highly mobile so the results cannot be taken as indicative of success over the whole range of the visually disabled. Indeed one problem with a tactile interface is the loss of touch sensation which accompanies diabetes, often a cause of blindness amongst the elderly, and for this group an audible or handle type of interface might be more appropriate.

4. COMMUNICATING FEATURES FOR NAVIGATION

For micro-navigation a simple interface seems adequate. The other demand for navigation is for the provision of information on features. The time of flight sonar sensors we have discussed throw away far too much information to be useful in feature recognition. In this section we describe the frequency modulated ultrasound sensor, which is proving useful for extracting certain types of feature information. We describe the sensor briefly, the information which it provides and then possible interfaces to the user.

4.1 Principle of Operation

The sensor consists of a separate transmitter and receiver. The frequency of the transmitter is modulated in a known pattern, and the frequency of the signal being transmitted is compared to that of the signal at the receiver, which was transmitted some time ago and has now returned after reflection. For the linear modulation which is shown in Figure 4.1, the difference frequency is proportional to the time of flight and therefore to range. This difference frequency is extracted automatically at the receiver.

4.2 Direct presentation of the signal

The difference signal is in the audible range for distances of up to a few metres. Sensors such as the Sonicguide, which work on the same principle, transmit the difference signal as a continuous tone to the user through an earpiece, thus presenting an auditory map of the environment. Different ranges appear as different pitches, and the loudness of the sound indicates how large a reflection occurred at that range. The user is provided with a rich source of information, with the possibility of distinguishing between single and multiple objects and of learning the 'sound' of particular feature shapes. However the main problem with the direct presentation of the beat signal is that it monopolises the user's sense of hearing, a vital source of information for a blind person. Interpretation of the signal also requires a lengthy training period before the signal can be well understood.

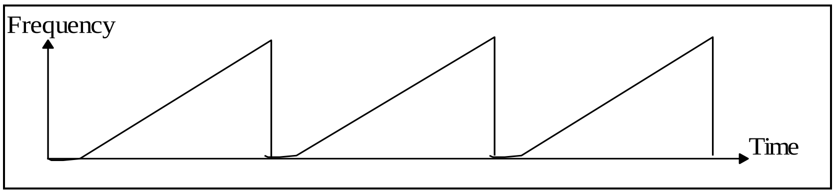


Figure 4.1: *the variation of frequency with time in the transmitter. Three cycles are shown. The frequency varies between 45kHz and 90kHz over a time of 160msec, and there is a blanking time of 24msec between the cycles. The blanking period is made long enough to avoid ambiguity between cycles and so determines the maximum range which can be measured.*

Nevertheless in spite of the disadvantages a small group of blind people use the Sonicguide, which consists of a central transmitter and two receivers built into spectacles, to great effect. Dedicated users report extraordinary perceptive powers with this technology: for example a highly mobile user reported recently to one of the authors that he could detect the difference between conifers and deciduous trees; others have reported being able to detect whether roses were in flower or not. However for most people the device remain inaccessible at least partly because of the interface.

4.3 Extraction of range information

The problem then is in how to extract and present information without overloading the auditory channels. A simple way to do this might be simply to provide information on demand; however for a direct interface it is likely that quite a long burst of the signal might be required. An additional problem, even for experienced and dedicated users, is that small features can get lost in the signal; for example the user is unlikely to pick up a pole in front of a large wall, since the reflection from the pole is swamped.

The range information in the signal at the receiver is in its frequency content, and the relationship between a signal in time and in frequency is given by the Fourier transform. From the Fourier transform it is easy to derive the power spectral density, which shows the amount of energy recovered at each range. For a single well defined surface, for example, the range image (which is equivalent to the power spectral density) shows a clear peak at a single range (Figure 4-2)

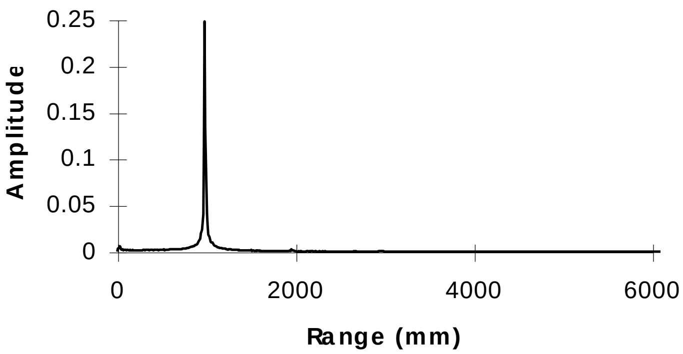
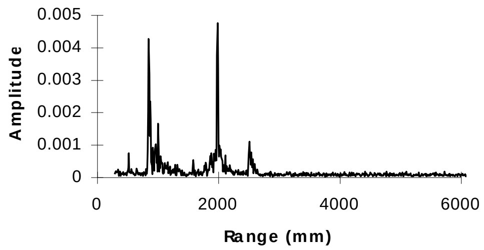


Figure 4-2. *A range image of the sensor pointed towards a wall*

When the sensor is pointing at a surface at normal incidence, the texture of the surface is indicated by the width of the peak. The height of the peak depends on the amount of energy reflected back to the receiver. More complex features are shown below. Note the multiple reflections from the car (from the front of the bonnet, from the windscreen/bonnet join and, the smaller furthest peak, probably from the top of the windscreen). The steps show a characteristic periodic structure.



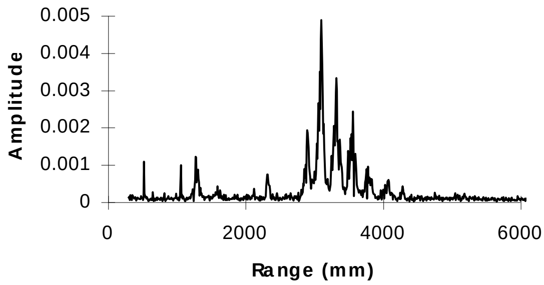


Figure 4-3. *Range maps from (a) the front of a car and (b) a flight of stone steps*

The visual shape of these gives us some ideas for how landmarks may be detected, and the information represented. For example it is easy to pick out some simple features from these range maps: periodicity, the positioning of multiple objects and an indication of texture. Unfortunately, of course, the visual interface is not accessible to the blind (and if it were the sensor would not be needed anyway, except in the dark!) so another method of presentation must be determined. A tactile array which reproduces the pattern is one possibility, but as we said earlier there remain significant technical difficulties in producing arrays with sufficient dynamic range to handle the problem of a high reflectivity background (such as the pole/wall example).

We are looking at two types of auditory interface instead, balancing the possibilities of automatic feature extraction against information bandwidth. One takes information from the range map and uses pitch, volume and rhythm to code distance and peak amplitude. Since most people are more sensitive to differences in pitch than volume it may overcome the background problem. The second attempts to provide direct voice output describing the feature and information on its position. It depends on the success of an automatic feature extraction system to recognise certain categories of object. In the final section we look briefly at the possibilities in this area.

4.4 Extraction of feature description

Automatic feature extraction is not straightforward. As a start in this area we have developed algorithms to detect periodicity (for example steps) and texture. They use the statistical properties of the range map, and the Fourier transform to examine periodicity. In more detail there are three stages:

4.4.1. Determining the area of interest. Experiments on the range maps show that spurious peaks are introduced at close range (less than about 0.5m) from imperfections in the device and in the finite window size of the Fourier transform and so these are not included in the processing. To extract interesting features from the rest of the map a sliding threshold is used, which is based on the intensity integrated over the rest of the range map. Then only those parts of the image above the threshold are retained.

4.4.2. Distinguishing smooth surfaces from groups of objects and textured surfaces. The shape can most easily be characterised by the statistics of the first and second moments of the peaks in the area of interest: i.e. the mean and variance. Using the raw power spectrum however was found to be too sensitive to noise. A better method was to determine the autocorrelation function and then apply tests to examine its width. Figure 4-4 shows the results from several features, which can then be distinguished by a simple classifier.

Taking the autocorrelation function (ACF) provides a significant reduction in the effects of noise. However the method was found most robust if the power spectrum was reduced to a binary image before the autocorrelation (it also, of course, speeds up the algorithm). This is presumably because the thresholding (which was based on the mean intensity over the whole area of interest) acts as a crude low pass filter.

The rate of decay of the ACF is characteristic of the surface, with smooth single surfaces showing a fast rate of decay and periodic surfaces showing clear peaks in the function. Although at first sight the some raw power spectrum data appear as if there may be periodicity, it is clear from the autocorrelation function that such periodicity, if any, is very weak. The Fourier transform of the ACF can be used to picks out any periodicity (Figure 4-5); for non-periodic signals no energy lies above the adaptive threshold.

Because of the long wavelength of ultrasound (a few mm), the amount of energy reflected back to the receiver depends strongly on angle. When there is less energy, the peak spreads even for smooth surfaces because of the effects of noise. However work based on modelling following Kuc [Kuc92] suggest that this can be exploited to determine the angle of the reflecting surface; for example the angle of a wall. This again could be a useful indicator for navigation.

An automatic system for feature recognition based on these ideas, which communicates through a voice box, has been developed. The mean and variance of the ACFs is used to distinguish between smooth surfaces and more complex features. Early trials are promising but low signal levels (caused, for example, because a surface is viewed obliquely) cause errors. Thresholding the peaks on the Fourier transform of the ACF proved a good method of distinguishing periodicity. The user interface provides information on the class of object and its range using words. It would be nice to be able to project position through stereo sound so the user could just hear something shouting out its name, but although it is easy to simulate orientation the problems in projecting range realistically remain unsolved [Loomis95].

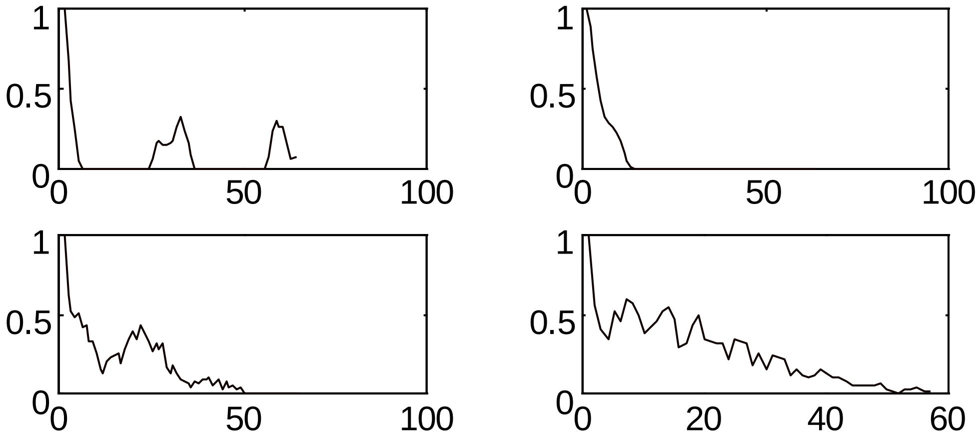


Figure 4-4. *The autocorrelation functions for steps (top right), branches (bottom right), a smooth wall (top left) and a cluster of rubbish bags (bottom right)*

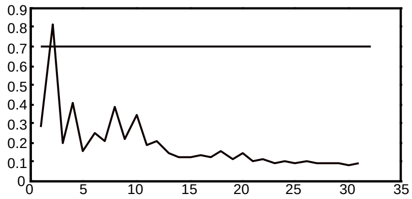


Figure 4-5. *Picking out periodicity: the FFT of the autocorrelation functions for the steps. Note the line marking the threshold, which depends on the average intensity in the image.*

5. CONCLUSIONS

In this paper we have described a two component system for navigation, which imparts directional information for local guidance through vibrators and higher level information on features through an audio interface. One of the ongoing themes in the project is how far information should be abstracted before it is presented to the user. People are far better than computers at recognising patterns, but the presentation of large amounts of raw data can overload the sensing systems available to the blind. The use of vibrators for the low level directional information has proved popular in user trials, but there remain problems of robustness, ease of use and cost if this method is ever going to have widespread acceptance. The early work on the frequency modulated sonar for providing feature information is encouraging but it is too early still to know whether a sufficient range of features can be included, either with the sonar alone or in conjunction with another sensor such as vision.

Acknowledgements: This work is funded by the European Union TIDE programme. One of the authors is supported by an EPSRC grant. We thank Simon Higgs of the Irish Guide Dogs Association for running the trials there and David Witt of Oxford University Department of Engineering Science for the frequency modulated sonar device.

6. REFERENCES

P. Bach-y-Rita and B. Hughes: “ Tactile Vision Substitution: Some Instrumentation and Perceptual Considerations”, in Electronic Spatial Sensing for the Blind, ed. D.H. Warren and E.R. Strelow, NATO ASI Series, Martinus Nijhoff 1985, pp125-140

J. Borenstein and Y. Koren, "Obstacle Avoidance with Ultrasonic Sensors", IEEE Trans. on Robotics and Automation, 42, 1988, pp213-218

O. Bozma and R. Kuc: “Characterising the Environment using Echo Energy, Duration and Range: the ENDURA Method”, IEEE/RSJ International Conference on Intelligent Robots and Systems, 1992, pp813-820

J. Fricke “Tactile Interfaces to Graphical Information”, Rank Prize Fund Symposium on Technology to Assist the Blind and Visually Impaired, Grasmere, UK, 1996

L. Kay: “Sensory Aids to Spatial Perception for Blind Persons: their Design and Evaluation” in Electronic Spatial Sensing for the Blind, ed. D.H. Warren and E.R. Strelow, NATO ASI Series, Martinus Nijhoff 1985, pp125-140

J. Loomis: “Personal Guidance System for Blind Persons”, Conference on Orientation and Navigation Systems for Blind Persons, University of Hertfordshire, 1995

OPEN project: “Audio Interface”, CC-TIDE Project TP 1182, 1995, issued by Royal National Instititute for the Blind, UK

H. Petrie: “User requirements for a GPS Travel Aid for Blind People”, Conference on Orientation and Navigation Systems for Blind Persons, University of Hertfordshire, 1995

RNIB (Royal National Institute for the Blind): “Electronic Mobility Devices or Obstacle Detectors for Blind and Partially Sighted People”, 1994

©1996 ECDVRAT and University of Reading, UK ; ISBN 0 7049 1140 X

ECDVRAT '96

Rehabilitation II

Day 3: 10 July, 1996, Afternoon

An overview of rehabilitation engineering research at the Applied Science and Engineering Laboratories.

William Harwin[1] and Richard Foulds[2]

[1]Department of Cybernetics, The University of Reading,
Whiteknights, P.O. Box 225, Reading, ENGLAND

[2]Applied Science and Engineering Laboratories, University of
Delaware and Alfred I duPont Institute, 1600 Rockland Road,
Wilmington, Delaware, USA

[1]*w.s.harwin@reading.ac.uk* [2]*foulds@asel.udel.edu*

ABSTRACT

The Applied Science and Engineering Laboratories of the University of Delaware and the Alfred I. duPont Institute have an ongoing research program on human machine interactions with a special emphasis on applications for people with disabilities. The Laboratories house two Rehabilitation Engineering Research Centers, the first on Augmentative and Alternative Communication Systems and the second on Applications of Robotics to Enhance the Functioning of Individuals with Disabilities. This paper reviews several projects within these two Centers as well as a project within a programme on Science, Engineering and Math Education for Individuals with Disabilities.

Consumer involvement has always been central to the research philosophy of the Laboratories and is achieved at a variety of levels and this paper includes a discussion of how effective user involvement is achieved.

Keywords: Consumer involvement, Haptic display, Virtual laboratory, Assistive technology, Rehabilitation robotics, Augmentative and alternative communication

1. INTRODUCTION

1.1 The Research Environment

The Applied Science and Engineering Laboratories are situated in the grounds of the Alfred I duPont Institute, a children's hospital in Wilmington, Delaware. This hospital provides a full range of clinical services to children under the age of 18, and along with rehabilitation research also conducts an extensive programme of clinical and medical research, and the Laboratories have benefited from the ready availability of medical expertise.

The Laboratories are jointly operated by the Nemours foundation, who also fund the main hospital, and the University of Delaware. The Applied Science and Engineering Laboratories maintain a satellite laboratory and office space at the University's main campus sited about 15 miles away. The Laboratories have also established and maintained links with the University of Pennsylvania, and Drexel University both situated in Philadelphia, as well as local rehabilitation organisations.

Funding for research comes primarily from grants, and a major portion of these comes from the Federal Government. Two of the 17 Rehabilitation Engineering Research Centers (RERC's) funded by the National Institute for Disabilities and Rehabilitation Research (NIDRR - part of the US Department of Education), are managed by the Applied Science and Engineering Laboratories. These are the RERC on Augmentative and Alternative Communication Systems and the RERC is on Applications of Robotics to Enhance the Functioning of Individuals with Disabilities. The Laboratories also manages a NIDRR grant on the provision of assistive technology, a National Science Foundation project on Science, Engineering and Math Education for Individuals with Disabilities, as well as individually funded research projects

In addition to work outlined in this paper there is ongoing research on access to large data-base, speech processing, natural language processing, telerobots for individuals with disabilities, power assisted orthotics, applications of robots in the school and workplace, design and manufacture of rehabilitation products and enhancing provision of assistive technology services.

The Applied Science and Engineering Laboratories focus on improving the access of individuals with disabilities to enabling technologies. This philosophy has resulted in the Human Machine interface being a primary focus of research, but the Laboratories has also sought to ensure that research is appropriate and in context so an important part of the work is the provision of novel technology related services to individuals with disabilities and related support organisations.

2. RESEARCH UTILISING VIRTUAL REALITY TECHNIQUES

2.1 Computer Recognition of Gestures in American Sign Language

Research on recognising American Sign Language (ASL) grew out of work done in the Applied Science and Engineering Laboratories on transmitting gestures across a telephone channel. Although text based telephone equipment such as TDD and the associated relay service exists, written English is not necessarily a fluent language for an individual who has ASL as their first language. Thus, there is interest in transmitting native ASL from one signer to another via telecommunication services. The problem of this approach is that the bandwidth of a telephone does not allow transmission of images, although Galuska and Foulds (1990) demonstrated that it is adequate for the transmission of sign.

As a consequence of linguistic and virtual reality research, tools and techniques are now sufficiently established to allow a very compact representation of ASL that could be transmitted across a telecommunications network. Since this requires acquisition and representation of ASL by a computer there is also the potential to do language translation to and from ASL. Success in language translation, or at least a subset of this, would not only improve access to telecommunication services, but would also substantially improve a deaf individual's access to other services in the speaking world, in particular the possibility of direct communications with non ASL speaking individuals.

Figure 1: *Animation output from ASL translator*

The gesture research group at ASEL have collated and demonstrated components of a machine to recognise and interpret ASL. A pair of "CyberGloves", from Virtual Technology along with a "flock of Birds" with 3 sensors from Ascension Technology are the primary input mechanism. Simple record and playback has shown that a majority of ASL information is still available by recording hand shapes, and hand and head positions via these transducers.

A neural network recogniser has been demonstrated to recognise the hand shapes of the manual alphabet and work is in progress on translating English via an intermediate form of notation derived from Stokoe's linguistic analysis of ASL, to an animated model of a signer (figure 1).

2.2 Enhanced Sensory Feedback

This is described in an accompanying paper (Harwin and Rahman, 1996), however work in this area was the catalyst for many of the associated projects in investigating haptic systems in telemanipulation and virtual reality.

2.3 A Virtual Interaction Interface

Many individuals need to either do routine movements for the purpose of physiotherapy, or would benefit with a greater understanding of the physical world if they could have meaningful interaction with it. Thus an individual with cerebral palsy may not understand the flight of a ball because he or she may never have had the opportunity to throw one. For such individuals there may be benefit in exploring virtual environments that are tailored to the persons abilities and learning needs.

The virtual interaction interface has been designed primarily using off the shelf components, so that such a device would be affordable by schools, and hospital therapy facilities. It allows "low-immersion" virtual reality that provides many of the benefits of more expensive systems, but without costly and encumbering equipment.

A blue background is provided behind the user and a video camera is positioned to capture any movements. Using a chromatic threshold the background can be removed from the image and replaced with suitable graphics. This combined video and animation image is then displayed to the user who gets immediate feedback on his or her movements and how

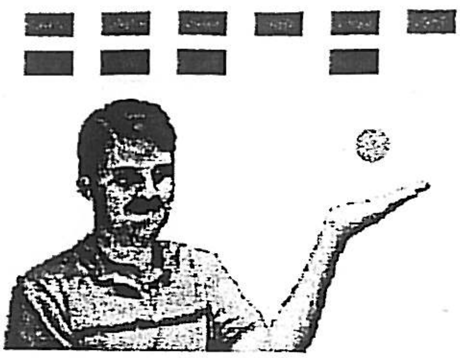


Figure 2: *Composite image from virtual interaction project*

these affect the virtual world.

Algorithms to monitor movements of the silhouette boundary and record when this intercepts the trajectory of any virtual objects are in place and these can be used to program a variety of physical laws such as objects bouncing, rolling or sticking to the 'virtual' individual. A break-out type game has been implemented to demonstrate the potential of this approach and explore the needs of the interface.

2.4 Supporting Large Language Spaces in Augmentative and Alternative Communication

People who use augmentative and alternative communication (AAC) devices need to access a large database of linguistic units as efficiently as possible. The problem is compounded by the fact that most AAC users typically have poor physical skills and that their language is often organised to "fit" into the display or keyboard of the AAC device. The alternative is for the individual to use and remember short forms of the vocabulary, such as words without vowels, or a sequence of pictograms. This approach requires the user to have a good memory for the location of language data and does not give intermediate feedback so often a language element will be chosen incorrectly.

Technology in common use for virtual reality research could support an expanded language space and thus provide a more natural communication for individuals with severe speech and motor impairments. The key to this research is to establish efficient mechanisms for this to happen.

Human Computer interface research is currently moving away from symbolic representations and towards spatial metaphors, the Apple desktop being a typical example. The implications of this trend may have important consequences for AAC research. A central tenet is that spatial metaphors of information can be more efficiently searched than hypertextual representations. Further people with speech and motor impairments can access this information efficiently and the input device can range in complexity from a single switch to a trackballs or a pointing devices.

This laboratory is using a high performance graphics computer to simulate and explore the consequences of spatially based AAC devices. This allows novel approaches to be explored that do not become limited by the computational abilities of the platform. This is also reasonable since it is likely that the speed and storage capacities of computers will continue to increase to allow low cost implementation of facilities such as realistic animation or high speed scrolling. Once promising techniques are established they can be transferred to more practical AAC systems and work is ongoing on a transferring one conceptual AAC device to a practical system based on a low cost head mounted display - the private eye from Reflection Technologies (Demasco et al, 1994).

2.5 Multimodal User Supervised Interface and Intelligent Control (Musiic)

Rehabilitation robots designed to assist an individual with a motor impairment to manipulate their environment have typically depended on one of two interface strategies. The first, command oriented interfaces, are simple for the individual to access but require a high level of structure to the environment and can only perform a small number of tasks (Harwin, 1995). In contrast directly controlled robots have a higher level of flexibility but require a high level of manual dexterity on the part of the operator.

The Applied Science and Engineering Laboratories are developing an interface that attempts to utilise a command level of interface but maintains the flexibility of a direct interface. Central to the interface philosophy is that information that is acquired by the robot should be fed back to the user so that instructions can be given based on a knowledge of what the robot can achieve.

Figure 3. *Real and virtual environments in the Musiic project.*

This project utilises a force sensing robot and optical sensors able to measure surface detail to the millimetre level. The interface uses a speech recogniser and a head mounted pointing device to allow the individual to locate objects to the robot. The system maintains a database and if a match is possible will manipulate the object as commanded. If a match is not possible a dialogue ensues to allow the object to be correlated with information in the database or a new instance of an object to be recorded (figure 3).

This system has been demonstrated in a preliminary form that allows a user to point to an arbitrary object, and move it to an arbitrary location with the command "put that there" (Kazi et al, 1995).

2.6 Haptic Display Systems For Individuals with a Visual Impairment

Teaching methods, especially in science, engineering and maths, tend to discriminate against individuals with disabilities, especially if the individual is blind or visually impaired. This is because of the vast amount of information that must be assimilated visually, including graphs of data, pictures of a physical phenomena, formulae, instrument displays etc. As a result very few individuals with disabilities acquire the scientific skills that allow them to enter employment in a technically skilled work force.

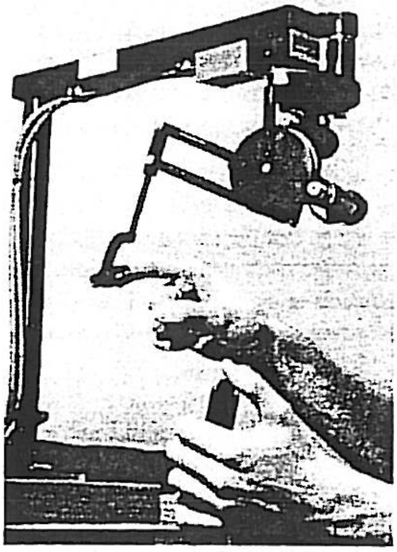


Figure 4. *Haptic representation of scientific data.*

Technology based solutions may have much to offer in making scientific information more accessible to visually impaired individuals. Central to this research is the study of haptic systems that appeal to the sense of feel and touch. This, in combination with auxiliary outputs based on braille displays or synthesised speech, are hypothesised to form a cogent interface for scientific data visualisation. The haptic display system is based on a PHANToM 3 axis interface from SensAble Devices (figure 4). Using this interface Fritz and Barner (1996) have demonstrated a haptic representation of simple xy graphs and work is also ongoing on tactile representation of pictures.

3. CONSUMER INPUT IN RESEARCH AND DESIGN

Because of the low incidence of technical skills in individuals with disabilities, a greater effort is necessary to ensure that research and development is achieving the goal of enhancing the abilities of individuals with disabilities. The Applied Science and Engineering Laboratories has a multi-levelled approach to ensure that research goals are in keeping with the needs of the individual.

The first of these is to ensure that there is a high representation of individuals with disabilities on the staff, and the Laboratories have been pro-active in recruiting professional and motivated individuals with disabilities to its staff. The Laboratories also seek to ensure that the needs of the individual are met with appropriate work-site modifications and with an employment practice that accommodates the needs of the individual.

The Laboratories also run a variety of focus groups related to particular projects where individuals with disabilities relating to the research meet with the researchers and discuss the progress of the work. The frequency of these meetings depends on the research goals, but several projects aim to meet with the consumer advisory group on a fortnightly basis.

The laboratories have also made significant effort to drawn local disability support groups into the research agenda and to liaise with organisations and individuals within the disability community. This approach has been successful in establishing projects that are targeted primarily at enhancing the lives of individuals with disabilities in the local community.

Several research projects need to make measurements on human subject and after review by the hospital ethics committee, this is conducted at the convenience of the individual. Thus, where possible, measuring equipment is designed to be portable so that it can be taken to the individuals home or to a more convenient location.

Formal evaluations of the research goals are made by an advisory committee that includes people with disabilities and part of this process is both open, and closed discussion about the laboratories research goals. Finally the laboratory encourages visits from individuals with an interest in disabilities and technology. Visitors include individuals with disabilities, their family and careworks, clinical professionals and academics. The laboratory tour is customised so that projects relating to the visitors disability are discussed and informal comments allow the research staff to gain a deeper insight to the nature of the disability.

All these methods of consumer interaction have value for the research projects, however it is difficult to quantify the nature and value of these interactions. There is considerable value in engaging a motivated workforce with direct experience of the disability under study however there are very few individuals who meet this criteria. Frequent meetings of a paid review group are also highly effective in gathering data, especially if there is a clear product design in view. Such a group is likely to need to meet less often where the project has more research oriented goals and it is correspondingly harder to keep the group dynamics active and to gather the necessary information on disability concerns. There may also be problems in recruiting and keeping a suitable group of individuals as motivated individuals usually have a large number of other commitments.

4. CONCLUSION

The Applied Science and Engineering Laboratories has an ongoing commitment to supporting appropriate research in enabling technologies for individuals with disabilities. The research environment is highly diverse but with a central focus on the technology needs of individuals with disabilities. Many of the research projects use virtual reality technologies or have developed tools and techniques that are applicable to virtual reality research. The Laboratories involves individuals with disabilities in a wide spectrum of areas to ensure that research maintains focus and the researcher staff gains a thorough understanding of the nature of an individuals disabilities and how that relate to the research.

Acknowledgements: All the research outlined in this paper has benefited from the generous supported of the Nemours Foundation. In addition material relating to the two RERC's is supported by grants from the National Institute on Disability and Rehabilitation Research. Work on Science Engineering and Math education is supported by the National Science Foundation. The authors also acknowledge the helpful cooperation of all the staff at the applied science and engineering laboratories, in particular Patrick Demasco, Kenneth Barner, Garland Stern, and Art Joyce.

Most of the projects discussed in this paper are ongoing and further information may be available via the applied science and engineering laboratories web site http://www.asel.udel.edu/

5. REFERENCES

P Demasco, A F Newell, and J L Arnott (1994), The application of spatialization and spatial metaphor to augmentative and alternative communication, In *Proceedings of ASSETS '94*, Marina del Rey, CA.

J P Fritz and K E Barner (1996). Design of a Haptic Graphing System. In *Proc. of the RESNA 1996 annual conference* Salt Lake City, Utah.

W S Harwin and T Rahman (1996). Analysis of force-reflecting telerobotic systems for rehabilitation applications. In *First European Conference on Disabilities, Virtual Reality and Associated Technologies.*

W S Harwin and T Rahman and R Foulds (1995). A Review of Design Issues in Rehabilitation Robotics with Reference to North American Research. In *IEEE Journal on Rehabilitation Engineering*, **3**.1, pp 3-13.

Z Kazi, and S Chen, and M Beitler, and M Salganicoff and D Chester, and R Foulds (1995). Multimodal User Supervised Interface and Intelligent Control for a Rehabilitation Robot. In *Proc of IJCAI-95 Workshop on Developing AI Applications for the Disabled*. Montreal. Canada.

S Galuska and R Foulds (1990) A real-time visual telephone for the deaf. In *Proceedings of 13th Annual Conference of RESNA*, Washington, DC

Virtual reality technology in the assessment and rehabilitation of unilateral visual neglect

Simon K. Rushton[1], Kate L. Coles[2] and John P. Wann[3]

[1,2]Department of Psychology, University of Edinburgh,
7, George Square, Edinburgh, SCOTLAND

[3]Department of Psychology, University of Reading,
Earley Gate, Whiteknights, Reading, ENGLAND

[1]*Simon.Rushton@ed.ac.uk,* [2]*K.L.Coles@sms.ed.ac.uk,* [3]*J.P.Wann@rdg.ac.uk*

ABSTRACT

Unilateral visual neglect affects a very large proportion of patients immediately after stroke. The presence of neglect has been found to be the major determinant of recovery of everyday function (Denes, Semenza, Stoppa & Lis, 1982). Interventions for the rehabilitation of neglect are not very effective. Picking up from the suggestions of Robertson, Halligan & Marshall (1993) potential ways forward employing VR technology are discussed.

Keywords: virtual reality, unilateral visual neglect, assessment, rehabilitation

1. INTRODUCTION

1.1 Unilateral Visual Neglect

Immediately following an acute hemispheric stroke over 60% of patients exhibit unilateral visual neglect (Stone, Halligan & Greenwood, 1993) - that is the disregard of objects in the space contra-lateral (opposite) to the major cerebral damage. Classic, textbook signs of neglect are eating food on just one side of the plate and then believing it to be empty, combing the hair on just one side of the face. Clinical measures such as line bisection (the centre of the line is marked by the patient and found by the examiner to be far off to the ipsi-lateral side) and object cancellation (objects on the side contral-lateral are omitted when instructions are given to cross-out all the objects present, e.g. all the ‘o’s on a sheet of paper containing ‘o’s and ‘x’s) demonstrate equivalent behaviour. Figure 1. captures the mis-perception of space of a patient with neglect. Traditional accounts would have it that the left side of the bar is ‘neglected’ or not seen, thus the patient is reaching for the middle of the bar that they perceive. An alternative theory (Milner, Harvey, Roberts & Forster, 1993) suggests that the world becomes ‘compressed’ so the true ends of the bar are perceived but the left half appears shorter than the right.

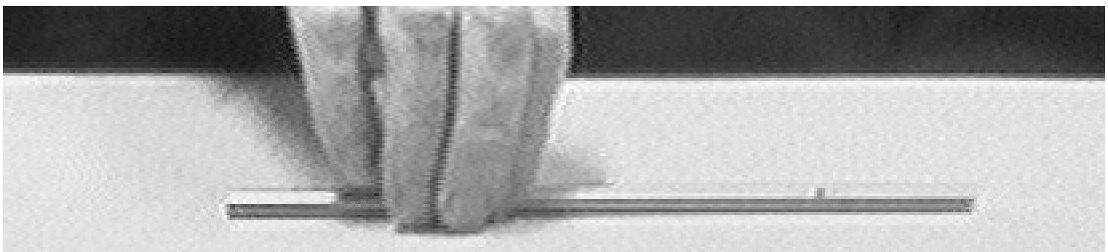


Figure 1. *A patient with neglect attempting to pick up the bar between finger and thumb from the “centre”. (see Rushton, Johnson & Wann, in press)*

Unilateral visual neglect is an attentional or representational deficit, it is not a visual field deficit. Although often accompanied by hemianopia (the loss of half the visual field), unilateral visual neglect dissociates from it. Neglect is also found of auditory and tactile space. Lastly, neglect of ‘imaginary’ space has also found in patients’ incomplete (omission of contra-lateral objects) recall from memory of familiar spaces such as city squares viewed from a certain vantage points (e.g. Bisach, Capitani, Luzzatti & Perani, 1981).

The presence of neglect has been identified as the major determinant of recovery of everyday functions following a stroke (Denes, Semenza, Stoppa & Lis, 1982). Consequently, successful interventions to aid recovery from neglect are vitally important.

1.2 Rehabilitation

Unfortunately, in contrast to the burgeoning number of published case studies and theoretical models, documented work on successful interventions for the rehabilitation of neglect is conspicuously lacking.

In a recent review Robertson, Halligan & Marshall (1993) proposed five possible future ways forward for rehabilitation of neglect: eye-patching, Fresnel prisms, vestibular stimulation, dynamic stimulation and optokinetic stimulation. ‘Forced usage’ and general attention retraining can also be added to this list.

1.2.1 Eye-patching. Eye-patching is founded on differential projection of each eye to the contra and ipsi-lateral hemispheres. By patching the eye that has the stronger connections to the ‘dominant’ intact hemisphere, the hemisphere’s ‘over-influence’ should be reduced so attenuating the attentional bias. Attention should then be more evenly distributed over the left and right hemi-fields.

1.2.2 Fresnel prisms. Fresnel prisms are similar to wedge prisms but are lightweight and thin. Prisms produce a rotation of the visual world. The benefit of doing this is that patients with unilateral visual neglect are not aware that they are neglecting half of the world. By wearing prisms parts of the world are brought into view that were previously unsighted and so the patient’s awareness of the contra-lateral portion of space may be increased.

1.2.3 ‘Forced usage’. An idea that converges with the eye-patching and fresnel approaches is ‘forced-usage’. Patients with hemiparesis, following stroke, (a left or right lateralised loss of control of limbs) show significant improvement can be produced in the effected limbs by the restriction of the good limb and the ‘forced use’ of the impaired limb (Taub, Miller, Novak & Cook, 1993). The explanation for improvement following such an intervention is that of overcoming ‘learned nonuse’ (Taub, 1980). Following major neurological damage a shock-like response occurs in the brain and precludes the use of the limb. After a period of time the effect of the shock diss greater detaily possible to express some movement with the impaired limb. However by this point, a habit of nonuse of the limb has been learnt. ‘Forced usage’ techniques attempt to overcome this learned non-use and have shown impressive results in the studies reported (see Taub, in press).

In the case of unilateral visual neglect, it could be hypothesised that a similar pattern of events is taking place. By restricting sight of certain portions of the visual world and ‘forcing’ the interaction with other parts it may be possible to prevent or overcome learned nonuse, or as it might be more appropriately termed ‘learned hemi-inattention’.

1.2.4 Vestibular stimulation. Rubens (1985) showed that vestibular or caloric stimulation by the injection of cold or warm water to the contra or ipsi-lateral ears reduces visual neglect in many patients. The effect is transient but may be valuable in promoting the patients’ awareness of the neglected contra-lateral portion of space.

1.2.5 Dynamic and optokinetic stimulation. Mattingley, Bradshaw & Bradshaw (1994) presented patients with a line bisection task on a computer. The background comprised of small dots above and below the line that were either static, leftward moving or rightward moving (and a control, neutral background). They found that patients were affected by global motion in the background; a leftward moving background showing a significant shift in bisection error. Butters, Kirsch & Reeves (1990) also presented patients with a line bisection task on a computer screen with either static or moving stimuli on the left side. They found that such stimuli had an influence provided that they were situated in the frame of the task, not just within the neglected side of space. Pizzamiglio, Frasca, Guariglia, Incoccia & Antonucci (1990) looked at the effects of optokinetic stimulation. Patients were asked to bisect a luminous strip in half. The strip was surrounded by luminous dots that were either fixed, left or right moving. The dots produced a slow nystagmus in the direction of movement. It was found that neglect was reduced with a leftward moving background.

There are three simple explanations for this interaction between neglect and visual motion: dragging the eye (optokinetic nystagmus) or attention in the direction of dot motion; the invocation of different perceptual-attentional systems because of the addition of a movement attribute to the test element (relative motion against the background flow); reduction in attentional bias between the neglected and non-neglected fields due to visual motion acting as common stimulus to attention across the whole test display.

1.2.6 General attention retraining. Robertson's (e.g. Robertson & Frasca, 1992) account of unilateral visual neglect suggests that deficits in general attention underlie a major part of unilateral visual neglect. Consequently, an intervention that improves general attention should also reduce neglect. Robertson, Tegner, Tham, Lo & Nimmo-Smith(1995) demonstrated a significant reduction in neglect following training for sustained attention.

2. VIRTUAL REALITY & NEGLECT - POTENTIAL APPLICATIONS

2.1 Assessment

2.1.1 Internal and external validity. A common problem that plagues any assessment is the trade-off between internal and external validity, laboratory or clinical tests provide the former, whereas observation or field study provide the latter (Barker, Pistrang & Elliot, 1994). This is well demonstrated in neglect testing where performance on clinical test batteries may dissociate from every day function.

Virtual Reality Environments offer a testing environment that can be as controlled as the standard laboratory and providing a similar quantity of high quality data. Additionally, however, Virtual Reality based tests can aim for a degree of 'ecological validity' (test that are similar to tasks attempted in real, everyday life) that has never been possible with standard measures.

2.1.2 Depth and motion. Two specific qualities that objects may have in a Virtual Environment that are very difficult to manipulate in standard tests are a 3D position and a movement vector. Objects in the real world are characterised by having both of these attributes. On paper and pencil tasks it is not normally possible to see how neglect interacts with these attributes. Given the dissociation demonstrated between neglect in near and far space (Halligan & Marshall, 1991) and the seemingly privileged status of depth (Nakayama & Silverman, 1986) and motion (Mcleod, Driver & Crisp, 1988) in visual attention this would seem to offer a valuable advancement.

2.2 Rehabilitation

We will consider in turn the potential intervention strategies outlined in the previous section (1.2) and examine how Virtual Environments might offer some advantage.

2.2.1 Eye-patching. Neglect appears to occur across different frames of reference: retinotopic (position on back of eyeball), spatiotopic (position in the world), body-centred (relative to the body's midline) and object centred (relative to the axis of an object of attention). Through use of an head-mounted display (HMD), position tracker, eye-tracker and access to object geometry it is possible to selectively occlude portions of space specified in any one of the four frame relative coordinate system. Additionally, occlusion can be transient with periods of occlusion as short as one frame, a fiftieth or sixtieth of a second. With the use of a see-through HMD it is also possible to occlude parts of the natural world.

2.2.2 Fresnel prisms. In an HMD the rotation or displacement of the visual world relative to the patients' proprioceptive space is a trivial matter and just requires the addition of a constant to the position or orientation of the patients head. Additionally, a virtual body could be rendered in a rotated position relative to the head. An immersive VE offers an enviably flexible rig for exploring the relationship of various frames of reference in unilateral visual neglect.

2.2.3 'Forced usage'. Through the occlusion of the 'good' or ipsi-lateral hemifield it should be possible to force attention to be directed into the 'bad' or contra-lateral region of space for lack of anything to attend to elsewhere. If there is a spatiotopic or body-centric organisation in any areas of cortex this should lead to the use and stimulation of areas that are not otherwise activated due to learned nonuse.

Through the non-tracking of the ipsi-lateral limb it should be possible to further bias attention to the contra-lateral hemifield through the forcing of use of the contra-lateral limb for interaction in the VE (simple smoothing and mapping functions should be able to compensate for impaired motor function in the contra-lateral limb provided impairment is not too great) due to the simple fact that the limb is centred in the contra-lateral field. Additionally, the use of that limb should bring about transient reductions in the bias between hemi-field bias (Robertson & North, 1992).

2.2.4 Vestibular stimulation. Although it is not possible to 'virtually' inject water into a patient's ear it is possible to effect other stimulation similar to that produced by a response from the vestibular system. Wertheim (1994) suggests that the vestibular signal, an eye-position signal , and retinal image flow feed into a single 'reference signal' which indicates self-motion. If this is so then it should be possible to substitute optic accelerating locomotor flow patterns for mechanical stimulation of the vestibular system.

 ISBN 0 7049 1140 X

It is also possible to further pull apart vestibular and visual frames of reference by the rotation of visual world relative to the vestibular frame of reference by rendering a view that is titled relative to vestibular or gravitational coordinates.

2.2.5 Dynamic and optokinetic stimulation. If the superimposition of motion reduces neglect then it suggests two potential uses of VR technology in the treatment of neglect: a prosthetic, a see through HMD that allows the superimposition of motion in depth over the natural visual world, bringing to attention objects in the neglected field; a rehabilitation training task, that uses motion stimuli to manipulate attention and conscious awareness of objects in the neglected field.

2.2.6 General attention retraining. The companion paper by Johnson, Shaw, Rushton & Smyth (this volume) describes an intervention that should produce improved general attention.

3. SOME PROGRESS

The simple super-imposition of motion over the visual field has been found to not be universally effective for patients with unilateral visual neglect. Whether it transpires that it is effective for specific patients has yet to be determined.

As reported in Johnson et al (this volume) the attention retraining underway and ready for a neglect patient.

The authors are not aware of any of the other approaches outlined being attempted in a VE setting as yet. However, work on the interventions using traditional materials and technologies is progressing.

4. CONCLUSIONS

Virtual Reality provides for modifications of the sensory environment with a degree of ease and precision not possible in the natural world. Virtual Reality technology seems ideally suited for attempting the interventions outlined in this paper and also offers the opportunity to combine one or many.

The progress to date is unfortunately related to the number of clinicians and researchers currently employing VR technology for the purpose of the rehabilitation of neglect. Despite this lack of visible success, for both the assessment of neglect and the development of rehabilitation interventions VR offers great promise.

Acknowledgements: The work reported in this paper was supported by MRC Grant 9520612 awarded to the first and third author and EC Telematics grant HC1053HC - VREPAR.

5. REFERENCES

Barker, C., Pistrang, N. & Elliot, R. (1994). *Research Methods in Clinical and Counselling Psychology*, Wiley, Chichester.

Bisach, E., Capitani, E., Luzzatti, C. & Perani, D. (1981). Brain and conscious representation of outside reality. *Neuropsychologia*, **19**, pp. 543-551.

Butters, C.M., Kirsch, N.L., & Reeves, G. (1990) The effect of lateralised stimuli on unilateral spatial neglect following right hemisphere lesions. *Restorative Neurology and Neurosciene ,* **2,** pp. 39-46.

Denes,G., Semenza, C., Stoppa, E., & Lis, A. (1982) Unilateral spatial neglect and recovery from hemiplegia: A follow-up study. *Brain,* **105***,* pp. 543-552

Halligan, P.W. & Marshall, J.C. (1991). Left neglect for near but not far space in man. *Nature*, **350**, pp. 498-500.

Stone, S.P., Halligan, P.W., Greenwood, R.J. (1993). The incidence of Neglect Phenomena and Realted Disorders in Patients with an Acute Right or Left Hemisphere Stroke. *Age and Aging*, **22**, pp. 46-52.

Kinsella, G., Packer, S., Ng, K., Olver, J. &Stark, R. (1995). Continuing issues in the Assessment of Neglect. *Neuropsychological Rehabilitation*, **5**, pp. 239-258.

Mattingley, J.B., Bradshaw, J.L., & Bradshaw, J.A. (1994) Horizontal visual motion modulates focal, attention in left unilateral spatial neglect. *Journal of Neurology, Neurosurgery and Psychiatry,* **57,** pp. 1228-1235.

McLeod, P., Driver, J. & Crisp, J. (1988). Visual search for a conjunction of movement and form is parallel. *Nature*, **332**, 154-155.

Milner, A.D., Harvey, M., Roberts, R.C. & Forster, S.V. (1993). Line bisection errors in visual neglect; misguided action or size distortion? *Neuropsychologia*, **31**, pp. 39-49.

Nakayama, K. & Silverman, G.H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, **320**, 264-265.

Pizzamiglio, L., Frasca, R., Guariglia, C., Incoccia, C., & Antonucci, G. (1990) Effect of optokinetic stimulation in patients with visual neglect. *Cortex,* **26**, pp. 535-540.

Robertson, I.H. & Frasca, R. (1992). Attentional load and visual neglect. *International Journal of Neuroscience*, **62**, pp. 45-56.

Robertson, I.H., Halligan, P.W., Marshall, J.C. (1993) Prospects for the rehabilitation of unilateral neglect. In *Unilateral Neglect: Clinical and Experimental Studies* (1993) I.H Robertson & J.C Marshall. LEA: U.K.

Rubens, A.B. (1985). Caloric stimulation and unilateral visual neglect. *Neurology*, **35**, pp. 1019-1024.

Rushton, S.K., Johnson, D.A & Wann, J.P. (in press). Ecological Correlates of line bisection errors in unilateral visual neglect and their potential application. *European Journal of Neurology*.

Taub, E. (1980). Somatosensory deaffertation research with monkeys: implications for rehabilitation medicine, In *Behavioral psychology in rehabilitation medicine: clinical applications*, (Ince, L.P., Ed), Williams & Wilkins, New York, pp. 371-401.

Wertheim, A.H. (1994). Motion perception during self-motion: The direct versus inferential controversy revisited. *Behavioral and Brain Sciences*, **17**, pp. 293-355.

Virtual reality environments for rehabilitation of perceptual-motor disorders following stroke

John P. Wann

Department of Psychology
University of Reading
3 Earley Gate, Reading RG6 6AL
UK

j.p.wann@reading.ac.uk

ABSTRACT

The incidence of perceptual-motor disorders arising from stroke is steadily increasing in the population of Europe and the USA. This paper outlines the potential role that virtual environments (VE) may play in designing remedial programmes for rehabilitation following stroke. Key principles for the structure of guided learning are identified, but emphasis is also placed on the need to identify when, and how, VE technology can introduce added *value* to the therapy situation.

Keywords: virtual reality, stroke, hemiparesis, rehabilitation

1. INTRODUCTION

To some degree Virtual Reality (VR) has been a technology in search of an application. Central components of the genesis of VR such as the head-mounted displays and dataglove were the result of scientists such as Sutherland (1968) exploring what was technically feasible, rather than from a specific application where computer users needed stereoscopic depth or complex manual interaction. Recently, much has been made of the potential for the use of VR in medical contexts. In general the attention and speculation has been concerned with either remote or augmented surgery, and surgical training. Because many surgical procedures are still critically dependent upon the co-ordination between eye and hand, this is an ideal application for a technology that seeks to provide a reconstruction of visual and haptic arrays and allow direct manipulation. Reciprocally, the precision required in surgical interaction also sets it as a critical test of the extent to which VR systems can adequately simulate natural interactions. Another potential use of VR in the medical industry that has received less attention is rehabilitation and treatment. Although surgery and therapy are both seen as worthy candidates for high-cost computer simulation, there is an important distinction to be made about the information that it is necessary to provide through such simulations. As discussed above, a central component of tele-operation or surgical training is transmitting information to the user that closely approximates the information they would use in a natural setting; so the stereoscopic depth of the patients organs should be appropriately scaled and the users head and hand movements should result in appropriate spatio-temporal changes in the display. The goal of therapy, however, need not be the recreation of "reality". If the goal of VR is to enhance the therapeutic context there is little point in recreating the natural environment that the patient has access to. It may be effective to recreate aspects of the natural environment that a patient has lost access to (e.g. simulated motion) or to transpose a particularly salient source of information to a context in which it would not normally arise (see section on hemiparesis). In this respect that task is not the recreation of "reality" but the recreation of a therapeutic environment, hence we will use the term virtual environment (VE) in the discussion that follows.

Parkinson's Disease has provided an initial focus for the use of VE augmentation (Prothero, 1993). The ability of akinesic patients breakout of a frozen posture and locomote when presented with stepping stone targets is well documented. Prothero (1993) presented Parkinson's patients with virtual stepping stones via a head-up display to enable patients to maintain gait. It was also reported that one patient was then able to spontaneously locomote even without virtual targets. The latter finding is promising but not unusual. There are a number of strategies that arise from the anecdotal reports that Parkinson's patients can use to initiate locomotion (e.g. music and dancing), but they are also invariably transient in their effectiveness. It would be overly ambitious to suggest that VE might have lasting therapeutic value for patients with progressive biochemical disorders such as Parkinson's disease. It may be able to

Proc. 1st Euro. Conf. Disability, Virtual Reality & Assoc. Tech., Maidenhead, UK, 1996

 ISBN 0 7049 1140 X

ameliorate specific deficits or stimulate the discovery of coping strategies, but ultimately it is unlikely to stimulate the natural production of dopamine. There are therapeutic areas, however, where *information* may hold the key to rapid improvement. This paper suggests two examples where the common thread is the need to direct attention. In the case of an attention deficit such as unilateral neglect there is a specific difficulty in focusing attention, where as in hemiparesis it is suggested that a sharpening of the focus may accelerate the re-acquisition of neuromuscular control. In all cases VE may provide a spotlight for the direction of attention.

2. UNILATERAL NEGLECT

A significant percentage of individuals who suffer a stroke, subsequently present with unilateral neglect. This is manifest in an apparent disregard of visual space contra-lateral to the lesion. In many cases reduction of neglect is evident during the first 12 weeks post-CVA. In a small number of cases, however, neglect is experienced as a longer term problem. Classic symptoms of neglect include shaving only one side of the face and eating the food on just one side of the plate. Clinical indicators that have been used are errors in bisecting lines, letter cancellation and figure copying, where an ipsi-lateral bias is generally evident. Neglect may be observed in the absence of visual field deficits and may be manifest in the extinction phenomena in which a visual feature on the contra-lesional side that was previously visible "disappears" when a similar object is presented in the ipsi-lateral visual field. A complementary paper in these proceedings (Rushton, Coles & Wann, 1996) provides a detailed background to existing therapy and the potential of virtual environments. The essence, however, is to capitalise upon the potential of VEs to provide attentional cueing through stimuli that may vary in colour, size, motion and depth. The plasticity of the virtual visual environment allows the tailoring of stimuli to provide greater precision in assessment procedures and a greater variety of therapeutic tasks.

3. HEMIPARESIS

Although unilateral neglect is not evident in all patients following stroke the majority will experience some degree of lateral paralysis, manifest in the loss of muscular control and the impairment of muscular sensation for the contra-lesional side. Hemiparesis will often affect the facial musculature, lower limb and upper-limb, particularly the distal extremities such as the fingers. Recovery can occur quite rapidly over the first 3 months (Fugl-Meyer, 1975; Wing et al, 1990), but thereafter may follow a negatively accelerating curve reaching a plateau below full function. Typically, there may be considerable recovery for proximal musculature such as postural control and relatively poor control attained for hand and finger control or for control of the contra-lesional foot. A primary goal of therapy is to enhance acquisition, particularly during the latter stages of recovery and avoid the early plateauing of active function. The most effective means of achieving such enhancement is open to debate.

Most conventional physical therapies that are used in the treatment of motor pathologies have evolved over considerable periods of innovation and refinement. The effects of conventional therapy for a number of motor disorders, however, are unclear. Despite the convictions of therapists, formal assessments of the efficacy of therapy for disorders such as congenital cerebral palsy have been inconclusive, (Tirosh and Rabino, 1989; Turnbull, 1993). It is not suggested that therapy is futile, quite the contrary, but the goals need to be continually reviewed in the light of current research on motor control and rehabilitation. It is also the case that therapists are severely restricted in the information or stimulation that can be provided to the learner (patient). This position paper considers what a current research in motor control suggests should be the primary goals of motor therapy and proposes a project for the development of tools that would enable a major step in the technological support for therapy

3.1 A Movement Perspective for Therapy

The bulk of research into human motor control has followed the lead of early researchers such as Woodworth (1899) and Lashley (1917) and focused on the efferent and afferent contributions in muscular control. The primary focus has been on the individual's musculo-skeletal system, the patterns of movement produced and neural control of such movements. Furthermore, a common theme across such an approach has been the requirement of some neural representation of the movement to be performed, either in the form of an error comparator for sensory feedback, a program of commands, or schema for generation of a generalizable response and expected sensory consequences. Although there has been substantial disagreement about the precise form that such a representation might take, the process of skill learning from a movement perspective requires that the individual learns about the capabilities and constraints of their motor system. There are hypotheses which avoid the inception of an explicit program, but these do not negate the need to "learn about one's system". Mass-spring models of aimed limb movements (Feldman, 1966, Bizzi 1980) suggested that a limb positioning could be achieved purely by specifying movement endpoints and that movement kinematics can be realised by exploiting the spring-like properties of skeletal muscle. But even this very simple spring-to-end-point model of limb movement requires some implicit knowledge and effective tuning of limb impedance to control trajectory speed and stability . It must be stressed that the acquisition of such knowledge may not be explicit, an individual will seldom "know" the precise configuration of their upper-limb musculature. But our everyday actions where we deftly reach and pick up fragile objects indicates that the control system has implicit knowledge of the dynamic characteristics of the upper-limb. Conversely, the attempts of an individual with spastic quadriplegia to complete the similar tasks highlights that there are problems in the interface between desired action and resultant movement. Computational algorithms for implicit learning of motor characteristics have been made presented by Jordon (1990) who proposed the learning of forward and inverse models of a hypothetical motor system. In simple terms the naive actor uses early explorative behaviour to build a model of the effects that a given set of commands will have in terms of the limb and environment (*forward model*). This is then used to form the *inverse model* of what commands are required to yield a specific desired effect. What can be gleaned from this perspective that would provide some principles for therapy with severe movement pathologies? In the case of hypertonicity or hypotonicity, it is obvious that the dynamic properties (stiffness, muscular recruitment, response characteristics) of the affected limbs are profoundly altered (Brown et al, 1987). Where changes occur following misadventure then it is clear that therapy must allow the individual to progress through a re-education or re-exploration of their motor system.

The focus of the movement perspective, described above, has been criticised as being overly prescriptive. An *action* perspective stresses that an animal's capabilities can only be understood in relation to the environment in which it acts. One should not focus purely on the movements of the body, but rather on how the union between the animal and its environmental *niche* unfolds. A predominant theme in the action perspective is the identification of the *affordances* (potential action) that the environment offers the animal (Gibson, 1979) and how these may be capitalised upon to yield skilled behaviour. The problem with translating this perspective to therapy is that the research effort has been predominantly elitist. Researchers have focused on highly skilled performers such as gannets (Lee & Reddish, 1981), long-jumpers (Lee et al 1982) and jugglers (Beek, 1989). Very little research has focused upon animals/humans who fail to mesh with their environmental niche. The main contribution from the action perspective would seem an emphasis on the role of functional, exploratory interaction between the actor and the environment.

An illustrative example of this point can be found in Conductive Education which stresses goal orientation for the child and that "dysfunction is not a property of the child, but the product of the interaction between himself, or the way he is, and his environment" (Hari & Tillemans, 1984,; p. 27). Hence the focus is on ortho-function and the child actively solving tasks in a way that is most appropriate for him/her. This contrasts sharply with traditional approaches such as Bobath (1973) who proposed the principles of therapy to be the suppression of abnormal reflex activity and facilitation of normal righting and that the CNS could be substantially changed by taking the child passively through postural patterns that conflicted with the abnormal reflex response.

3.2 Principles for action based therapy

1. Provide a therapeutic environment that highlights the consequences a patients actions to allow them to identify the (changed) characteristics of their system (*Movement Principle: forward modelling*)
2. Provide a setting where they can try and test adaptive movement strategies (*Movement Principle: inverse modelling*).
3. Present the patients with functionally related tasks that have clear goals rather than prescribe movement sequences or patterns of contraction *(Action principle 1).*
4. Allow the patient to explore the environment and arrive at solution to his/her specific movement problems, rather than guide them through constrained routines *(Action principle 2)*..

3.3 The virtual learning environment: Identifying "added value"

When considering how a VE can enhance the therapeutic setting one might proposed that it can be used to provide the patient with a virtual body or limb. This avenue does have a number of limitations that require consideration. A patient with hemiparesis has lost the ability to produce sustained contractions of the limb musculature. Over time there is often some recovery of this ability, particularly for the proximal musculature. The re-acquisition of skill may in part be due to simple post-trauma effects, but is also likely to require some degree of neural re-organisation and recruitment in the light of impaired transmission through the cortico-spinal tract. What role would a virtual limb have in guiding the re-acquisition of skill? If the patient is unable to move their natural limb is their any value in providing them with an image of a virtual limb? There would appear to be some potential for this paradigm, but there are constraints to be considered. A virtual limb may be presented as a model for the patients action: Using a see-though head-mounted display, it is viable to present the patient with a virtual limb superimposed in a location close to their own limb. The virtual limb may be moved computationally as a tracking task for the patient to mimic the range and speed of motion. Despite the attraction of the technology, there is no formal basis to presume that this paradigm would be any more effective than presenting the patient with a series of real tasks, such as pushing blocks and wiping tables, that are common in conventional occupational therapy. To justify the use of VE technology there is a need to identify *added value*, where the therapeutic paradigm provides a stimulus that is not available in the natural therapy setting. We may justify the virtual-arm paradigm by making recourse to the early stages of post CVA recovery where the patient may have very little overt movement of a limb. The patient in the early stages of recovery may be able to initiate EMG activity, but the motor unit recruitment is insufficient to overcome the inertia of the limb and produce a change of limb position. A virtual limb, however, may then be a useful and stimulating way of encouraging specific patterns of recruitment. Biceps and triceps EMG can be taken as the peripheral inputs for motion of the virtual limb, such that a rise in triceps EMG, in the absence of biceps EMG produces extension of the virtual limb around the elbow, whereas a reciprocal pattern produces limb flexion. Co-activation of both EMG inputs produces slowing of the limb and eventually virtual-rigidity. Hence the patient who is unable to produce overt limb movement can be presented with a task in which their goal is to re-learn simple patterns of motor unit recruitment, building towards the tri-phasic EMG pattern that is the hallmark of skilled upper-limb movements. This latter example exploits *Principle 1:&2* to provide the patient with *information* that may not otherwise be available to them (e.g. how they are actually contracting their musculature) that may allow them to identify suitable strategies.

This therapeutic approach is somewhat similar to conventional biofeedback therapy. In biofeedback a patient would typically be provided with a specific movement goal, such as the regulation of activity in a specific muscle group. Augmented sensory cues would be provided in the form of vision (of an oscilloscope trace) or sound that reflected the electrical activity monitored from the muscle group. The subject's goal would be to find a control strategy that produced the correct augmented sensory feedback. This could be considered as a simplified case of forward-inverse system modelling, or could equally be considered as little more than operant conditioning. The issue is what can the individual "learn" that is transferable to functional activities ? A number of studies have demonstrated short term gains with biofeedback training of discrete muscle groups (Mroczek et al, 1978; Middaugh & Miller, 1980; Nemec & Cohen, 1984; Mulder et al, 1986), hence augmented feedback can clearly enable subjects with motor pathologies to gain greater control over skeletal muscular responses. There is little evidence, however, that such gains are long-term or that this control can be transferred outside the training context. The stumbling block of conventional biofeedback is that the regulation of contraction in a single muscle group, or even a collection of muscles is far removed from functional goals. The use of VE systems in a traditional biofeedback mode (e.g. Warner & Jacobson, 1992) may provide a useful stimulus to early exploration or provide the patient with a useful computer interface, but falls well short of functional rehabilitation (Wann & Turnbull, 1993).

3.4 Fashioning the learning environment: Providing action feedback

Rather than providing feedback about the patient's biological or muscular-function it is proposed that greater gains may be possible if the environment enhances and highlights the consequences of the patient's action, rather than reflecting aspects of their motor control. In pathology naturally occurring cues, such as proprioception, may often be unreliable or have lost their salience. What is required is augmented feedback that can be fashioned to match the individuals capabilities and prompt the individual towards finding unique solutions to movement problems. Physical laws impose constraints on the manipulation of naturalistic settings. It is difficult to slow down a falling ball to allow a child more time to organise his/her movements, or to stop water from spilling when a cup is moved jerkily towards the mouth. VE technology, however, can provide an interactive environment tailored to the needs and abilities of the patient. Restricted powers of movement and manipulation need not set limits upon the quality of interaction. The mapping between limb movement and movement in the VE is arbitrary, the input from a limb can lose its tremor, through low-pass filtering, or a single finger movement can control virtual locomotion. This flexibility allows the augmentation or amplification of sensory feedback to a perceptual system that may have lost some degree of its natural sensitivity. As the content of a VE can be directly specified, the patient may be given progressively more complex and rich environments. Distracting objects can be banished and the laws of physics bent or relaxed. Simple training tasks can be set in a meaningful context thus developing skills that are immediately useful to the patient, thereby enhancing the patient's sense of accomplishment and engendering future interaction. Without moving from the treatment room the patient can experience information rich environments without the inconvenience often associated with complex natural settings.

4. A POSITION STATEMENT

Virtual environments provide the potential to computationally specify the visual and auditory environment that is presented to an observer. This opens up interesting new avenues for exploring human skill acquisition (Wann & Rushton, 1995a,b; Wann et al 1995), but also presents possibilities for guided learning in a remedial setting (Wann & Turnbull, 1993; Kuhle & Dohle, 1995). This paper identified 4 principles for the design of therapeutic environments that arise from existing work on perceptual-motor learning. There is a clear potential for virtual environments in meeting those goals in a manner that is not possible within the constraints of a natural therapy setting. The power of the VE setting is the ability to selectively tune the environment to provide specific patient goals and the ability to provide real-time feedback with tuneable input-output gains. The principle benefits of the VE environment in a therapy setting are:

i) A 3-D virtual environment can "immerse" the patient to the degree that they will demonstrate appropriate limb & postural corrections in response to virtual environmental (VE) perturbations.

ii) The illusion of (i) coupled with the patients ability to directly manipulate the VE display will produce a considerably stronger learning environment than conventional therapy approaches

iii) The ability of the patient to explore, interact and make errors in the VE environment will provide a facility for motor re-learning unparalleled outside of a VE setting.

iv) The novelty and intrinsic appeal of such an interaction will also provide a powerful motivation factor for rehabilitative exercise.

It is important, however, that the potential of VR is not squandered by "throwing" VE technology at the problem without a formal appraisal of where and how it may provide *added value*. The rehabilitation communities have traditionally welcomed interest from technology providers and therapists are often interested in the potential role that computing or sensor technology may have remedial programs. That goodwill may be lost if virtual reality applications are not clearly focused to capitalise upon the existing knowledge base in motor control and rehabilitation.

Acknowledgements: Work on the design of virtual environments for stroke disorders is being supported by EC Telematics grant HC1053HC - VREPAR.

5. REFERENCES

Beek PJ (1989) *Juggling Dynamics*. PhD Thesis, Free Univ, Amsterdam: Free University Press.

Bizzi E (1980) Central and peripheral mechanisms in motor control. In Stelmach GE & Requin J (Eds.) *Tutorials in Motor Behavior*. Amsterdam: North Holland

Bobath K (1980) *A Neurophysiological Basis for the Treatment of Cerebral Palsy*. Spastics International Medical, London: Heineman

Brown JK, van Rensburg F, Walsh G, Lakie M & Wright GW (1987) A Neurological study of hand function of hemiplegic children. *Developmental Medicine & Child Neurology* 29, 287-304.

Feldman AG (1966) Functional tuning of the nervous system with control of movement or maintenance of a steady posture. II Controllable parameters of muscle. *Biophysics* 11, 565-578.

Fugl-Meyer AR, Jaasko L., Leyman I., Olsson S. & Steglind S. (1975) The post-stroke hemiplegic patient: I, A method for evaluation of physical performance. *Scand. J. of Rehabilitation Medicine* 7, 13-31.

Gibson JJ (1979) *The Ecological Approach to Visual Perception.* Boston: Houghton-Mifflin.

Hari M & Tillemans T (1984) Conductive Education. In Scrutton D (Ed) *Management of Motor Disorders of Children with Cerebral Palsy*. Spastics International Medical, Oxford: Blackwells.

Jordon MI (1990) Motor learning and the degrees of freedom problem. In Jeannerod M (Ed.) *Attention & Performance* XIII, pp 796-836. Hillsdale NJ : Lawrence Erlbaum.

Kuhlen, T & Dohle, C. (1992) Virtual reality for physically disabled people. *Comput. Biol. Med.* 25, 205-211.

Lashley KS (1917) The accuracy of movement in the absence of excitation from the moving organ. *Am. J. of Physiology* 43, 169-194.

Lee DN, Lishman JR & Thompson J (1982) Regulation of gait in long-jumping. *Journal of Experimental Psychology: Human Perception & Performance* 8, 448-459.

 ISBN 0 7049 1140 X

Lee DN & Reddish PE (1981) Plummeting gannets: a paradigm of ecological optics. *Nature* 293, 293-294.

Middaugh SJ & Miller MC (1980) Electromyographic feedback: effect on voluntary muscle contractions in paretic subjects. *Arch Phys Med Rehabil* 61, 24-29

Mroczek N, Halpern D & McHugh R (1978) Electromyographic feedback and physical therapy for neuromuscular retraining in hemiplegia. *Arch Phys Med Rehabil* 59, 258-267.

Mulder T, Hulstijn W & van der Meer JPT (1986) EMG feedback and the restoration of motor control. *American J ournal of Physical Medicine* 65, 173-188.

Nemec RE & Cohen K (1984) EMG biofeedback in the modification of hypertonia in spastic dysarthria. Arch Phys Med Rehabil 65, 103-104.

Rushton S., Coles K.L. and Wann, J. (1996). Virtual Reality technology in the assessment and rehabilitation of unilateral visual neglects. *Proc. ECDVRAT.* Reading UK, July 1996.

Seeger BR & Caudrey DJ (1983) Biofeedback therapy to achieve symmetrical gait in children with hemiplegic cerebral palsy: Long-term efficacy. *Arch Phys Med Rehabil* 64, 160-162

Sutherland I.E. (1968) A head-mounted three dimensional display. Fall Joint Computer Conference, 33, 757-764

Tirsoh E & Rabino S (1989) Physiotherapy for children with cerebral palsy. *American Journal of Diseases of Children* 143, 552-555.

Turnbull JD (1993) Early invention for children with or at risk of cerebral palsy. *Am. J. of Diseases of Children.*

Wann, J.P. & Rushton S.K. (1995a). The use of virtual environments in perception-action research: Grasping the impossible and controlling the improbable. in D. Glencross & J. Piek (Eds) *Motor Control and Sensory-Motor Integration.* Amsterdam: North-Holland.

Wann, J.P. & Rushton S.K.(1995b) Grasping the impossible: Stereoscopically presented virtual balls In B. Bardy, R. Bootsma & Y. Guiard. (Eds) *Proc. of the 8th International Conference on Event Perception and Action.* 207-210

Wann, J.P. & Rushton S.K. & Lee D.N. (1995) Can you control where you are heading when you are looking at where you want to go? In B. Bardy, R. Bootsma & Y. Guiard. (Eds) *Proc. of the 8th International Conference on Event Perception and Action.* 171-174

Wann J.P & Turnbull J. (1993) Motor Skill Learning In Cerebral Palsy: Movement, Action and Computer Enhanced Therapy. *Bailliere's Clinical Neurology 2, 15-28*

Warner D. & Jacobson L. (1992) Medical rehabilitation, cyberstyle. *AI Expert* 4, 65-72.

Wing A.M., Lough, S., Turton A., Fraser C. & Jenner J.R. (1990) Recovery of elbow function in voluntary positioning of the hand following hemiplegia due to stroke. *Journal of Neurology, Neurosurgery & Psychiatry* 53, 126-134.

Woodworth RS (1899) The accuracy of voluntary movement. *Psychological Review Monograph Supplements* 3, 3.

Nervous system correlates of virtual reality experience

L Pugnetti[1], L Mendozzi[2], E Barberi[3],
F D Rose[4] and E A Attree[5]

[1,2,3]Scientific Institute S.Maria Nascente
Don Gnocchi Foundation
Milan, Italy

[4,5]Department of Psychology
University of East London
London E15 4LZ, UK

[1]*melau@mbox.vol.it,* [4]*rose2@uel.ac.uk*

ABSTRACT

In recent years several papers have been published in which the effect of exposure to virtual reality (VR) on the activity of the nervous system have been discussed. This area of knowledge is of importance to those interested in VR applications, and especially to those who seek to apply VR to cognitive rehabilitation following brain damage. This paper reviews what is known about the effects on nervous system activity of the interaction with virtual environments, comments on the authors' experience with both normal and neurologically impaired subjects, and outlines a suggested programme for future research.

Keywords: nervous system, psychophysiology, cognition, ARCANA.

1. INTRODUCTION

The search for central nervous system (CNS) correlates of virtual reality (VR) experiences is of considerable importance to those exploring the applicability of this rather new technology to the field of cognitive rehabilitation. The latter is, itself, a rather new clinical discipline that is now emerging as a distinct form of treatment as clinical neuropsychology evolves from its experimental stages. Indeed, the endeavour to integrate a new technology within a new therapeutical methodology that has not yet reached maturity is a typical example of a high risk but potentially highly rewarding operation. Those who seek to do this are well aware of the necessity to obtain as much evidence as possible that VR is psychologically well accepted, physiologically and clinically well tolerated - and cost-effective when applied on a large scale basis.

Though there is little doubt that this evidence will ultimately come from clinical studies, a body of ancillary evidence is available from the results of studies specifically aimed at showing how target brain functions react to this new stimulating technology. Our two centres are making joint efforts to this aim.

2. AN OVERVIEW OF CNS CORRELATES OF VR EXPERIENCES

Besides an indeterminate number of statements made in generic papers in the early days of VR that alluded to the possibility that VR could profoundly affect the brain at the psychological (i.e.) learning and cognition), neurophysiological (i.e. perception) and emotional (i.e. affect, motivation) levels, the bulk of published reports dealing more or less specifically with these issues comes from research undertaken to adapt this technology to work within the range of human sensorimotor constraints - so called 'human factors' (Barfield et al. 1994). This was, and still is, a legitimate preoccupation since early immersive VR systems were shown to cause an excessive workload on human sensory-motor systems, and to produce patterns of stimulation which could not be neurophysiologically coded as 'normal' by the brain. Both factors caused symptoms to appear. The threat to the application of VR was perceived as significant and great efforts have been made to study in greater detail what was causing "cybersickness", - and what should be done to reduce its incidence to a minimum (Biocca, 1992). The issue of VR side-effects has been subject of recent papers and reviews (Pausch et al., 1992) and will not be again addressed here. Rather, we will attempt to underline what has been done to study, by objective means, non-detrimental effects on the brain of VR experiences.

Studies can be classified into two broad categories: 1) physiological studies using VR as a means of putting subjects into a predefined and highly controlled environment, and 2) studies that make use of physiological measurements to assess any bodily reaction to VR. Several pieces of work have been undertaken so far within these two categories. While for the second category of studies the underlying assumption is that the overall pattern of stimulation is still atypical as compared to that provided by the natural environment, the first relies on the closest possible approximation to natural conditions, and this requires a most sophisticated and rare technology. Nonetheless, Eberhart and Kizakevich (1993) have already suggested that VR provides a unique means to undertake physiological studies of human beings under stressful circumstances and have underlined the necessity of taking into serious consideration the biological variability of physiological responses across subjects. This approach has been pursued mostly by military research and training centres, but does not appear to have had yet an impact outside these boundaries. Decety et al. (1994) have used VR to study how the brain is activated during passive observation of complex movements by a virtual hand and by imagination of such movements. They used a PET scan to map functionally activated brain areas. Their study is the best example of how VR can be used to provide highly specified stimulation of a kind that has eluded control so far.

However, VR experiences are not necessarily stressful events. VR can be perceived as almost any kind of real experience depending on a mix of factors (Barfield and Hendrix, 1995). So, the first application of psychophysiological studies is to provide objective data to supplement behavioural data and subjective reports, which are by far the most frequently used measures in VR research. Knowing how and when physiological parameters change during a VR experience can help in the process of tailoring VR applications to an individual's resources (Durlach, 1994). Clearly, this is even more compelling if the user is an impaired individual or if the VR experience must be prolonged in time. For ethical reasons, the assessment of any VR system devoted to people with disabilities or subjects at risk should include an analysis of its effects on target physiological variables (Whalley, 1995).

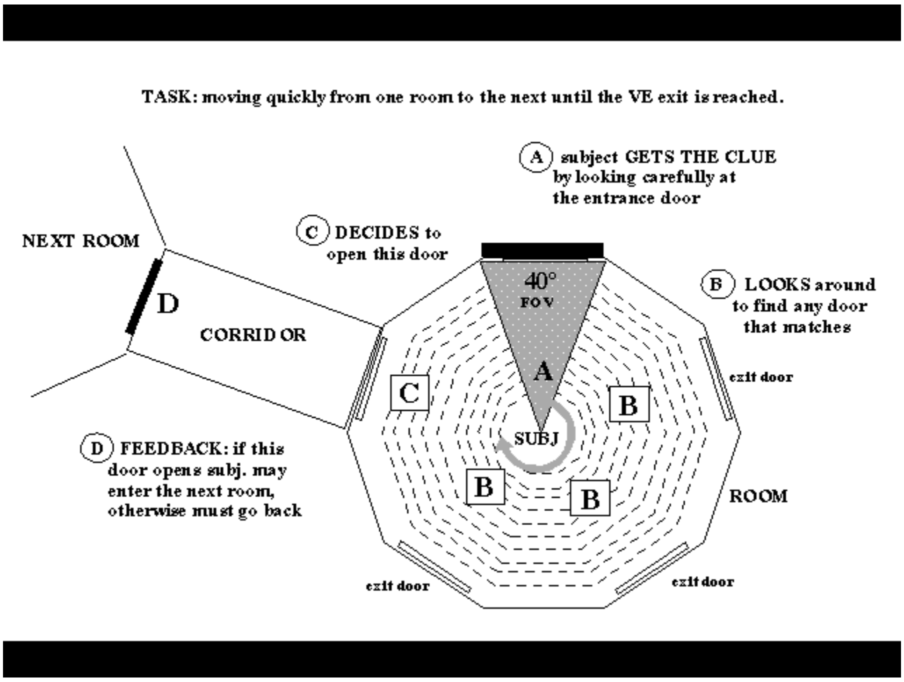


Figure 1. *Top-down schematic view of ARCANA's virtual environment. A sequence of cognitive and behavioural steps involved in the task is also shown.*

We have followed this approach in developing prototype systems devoted to clinical studies. We will focus particularly on the ARCANA project (Pugnetti et al., 1995) as an example of the way VR can be used to implement a specific cognitive paradigm and to study its physiological correlates. ARCANA's virtual environment (VE) features only two very simple internal architectural elements: rooms and connecting corridors. While in a room, subjects are supposed to move about by opening one of several possible doors that either lead to the next room or to a dead end corridor. The subject's task is to proceed by trial and error until he/she finds the door that leads to the outside of the building (figure 1). Cuing is provided by the doors: their variable shapes and colours serve as matching criteria to make correct selections among several alternatives. Each subject is then expected to develop a strategy to avoid the frustration of frequent failure.

Hence, the cognitive requirements of this paradigm appear to be very similar to those of popular tests of categorisation and cognitive flexibility such as the Wisconsin Card Sorting Test (WCST). Our VR setup, however, has important additional requirements because subjects need to explore the virtual space around them, to keep spatial aspects of it within working memory, to master unfamiliar ways to move inside this space that taxes their visuomotor co-ordination, to adapt to a narrowed field of vision which demands an increased reliance upon head rotation to get all the visual information needed, and finally to interpret non-verbal feedback. In addition, information must be handled in a rather abstract form and organised into a strategy which must be checked for correctness at any new choice point, and eventually changed if not successful. This explains why there are really no correlations between the cognitive performance on a battery of standard paper-and-pencil tests and their VR-implemented analogue.

Peculiar to the VR setup for cognitive paradigms is that it allows stimuli to become integral and congruent features of an environment, and that it can be programmed so that specific cognitive events take place within specific contexts and develop on a time scale that is similar, if not identical, to that of normal purposeful behaviour. This puts those interested in assessing brain-behaviour relationships into a very favourable situation. Events flow on a time scale that allows meaningful segmentation of psychophysiological signals, and relationships can be interpreted more confidently because events take place in unambiguous contexts. Because of this, it should be possible to generalise to real life situations with a much greater degree of certainty.

We are interested in the analysis of ongoing EEG and auditory evoked potentials (EP) of our subjects performing our VR test because its underlying cognitive model is fairly well developed (Perrine, 1993) and substantial research has been already been done on its CNS correlates (Silberstein et al., 1995). Moreover, it has also been found to be an outcome predictor for neuropsychological rehabilitation (Berger et al, 1993).

These analyses can give insights into two aspects of CNS function during VR-based cognitive activity: 1) how it changes as a function of time and performance (i.e. long-term variability), and 2) how specific cognitive steps (i.e. stimulus-response sets) are reflected in short-term changes. We report here the results of interim analyses of ongoing studies.

Figure 2. *An example of 15 second polygraph recording from a subject in a virtual environment. Note the changes in EEG and EOG patterns as the subject stops moving straight and begins visual exploratory activity by turning his head right then left.*

Twelve healthy young subjects participated in the first study assessing EEG changes. Polygraphic recordings of behaving individuals were made by means of a portable 12 channel solid-state memory digital recorder (Micromed Brain Spy). A continuous recording was obtained from 7 scalp unipolar positions 1 EKG, 1 EMG, 1 PNG and 2 EOG channels (figure 2). Each recording session lasted about 1 hour, covering an average of 15 minutes both pre and post VR and 30 minutes of VR testing. EEG signals were sampled at 128Hz and 8 bit resolution, and analysed off-line on a Neuroscan v.3 workstation. Ten of 12 recordings were acceptable in terms of EEG signal quality and biological artifacts. Ocular artifacts on EEG channels were either corrected with a covariance algorithm (Semlitsch et al, 1995) or rejected with loss of the affected data epochs. Myographic artifacts were never found to be prominent on EEG channels. Since subjects participated in the study with their eyes open, the amount of alpha activity over occipital and parietal regions was considered as an index of brain activation. Indeed, the plot of mean alpha power of the 10 subjects against time shows a fairly close inverse relationship with several parameters of performance, such as the time elapsed in each room and the number of errors (figure 3). We interpret this finding either as a correlate of a more automatic mode of cognitive processing or as a progressive build up of mental fatigue. This latter hypothesis seems less likely since the post VR resting EEG did not change from the pre VR level. We may tentatively conclude that learning occurs in this VR setting and that it has central neurophysiological correlates.

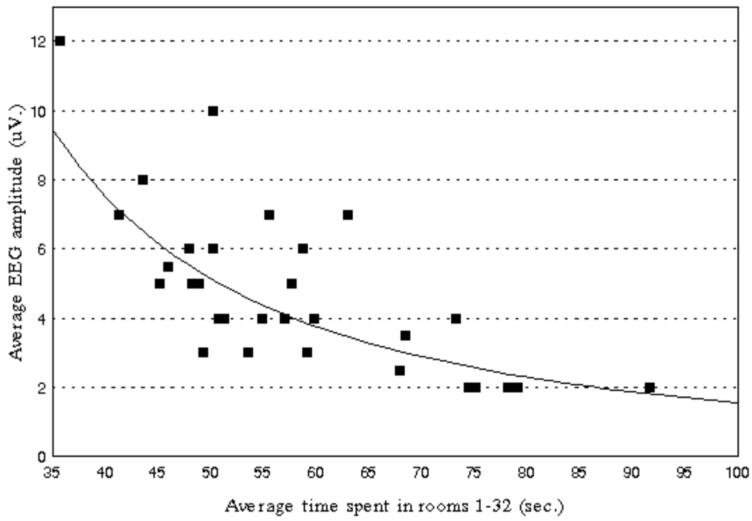


Figure 3. *Plot of EEG alpha amplitude against the time spent in each of 32 rooms. Each point is the average of 10 healthy controls.*

Cortical auditory evoked responses were elicited in another study of 10 subjects by releasing footstep sounds binaurally as they operated a pointer (a virtual key) to move forward in the virtual environment. The subjects spent an average of 40% of the time moving inside the VE. Cortical P1-N1-P2 components were measured from the grand averages of artifact-free epochs spanning the whole duration of VR testing. The latter were compared with similar responses elicited during pre- and post-VR periods. While the latter two conditions yielded comparable results, engaging in the VR task substantially reduced the amplitude of N1-P2 components; no effect was seen on the latency of the responses. This was also an expected result and is interpreted as a correlate of the activation during cognitive processing of cortical areas usually participating in EP production. The ERPs technique we used is similar to the steady state evoked potentials used by Silberstein et al. (1995) and to the ‘tracer strategy’ used by John and Eston (1995) and seems particularly promising for studying CNS correlates of mental workload in subjects. For each subject we can generate a function describing the relationship between sound intensity and ERP parameters (e.g. amplitude) in baseline conditions. By analogy, we can then use this function to measure, with the same units, the variations of ERP parameters that we observe during VR when the subject is supposed to be actively engaged in the primary task and only a fraction of his/her processing abilities is devoted to an auditory input which he/she does not have to process actively. It is expected that a heavily involved subject will produce ERPs of smaller amplitude than one who finds the task easier and can process additional inputs at the same time. We have found this to be true by comparing the ERP waveforms obtained during periods of good performance with those of bad performance (figure 4). On two occasions, subjects began to complain of nausea but were able to terminate their VR session. In both, ERPs disappeared completely after malaise onset. We interpret these findings as further supporting the hypothesis that ERP amplitude can be used as an index of saturation of residual cognitive resources left to process concurrent but irrelevant ambient stimuli.

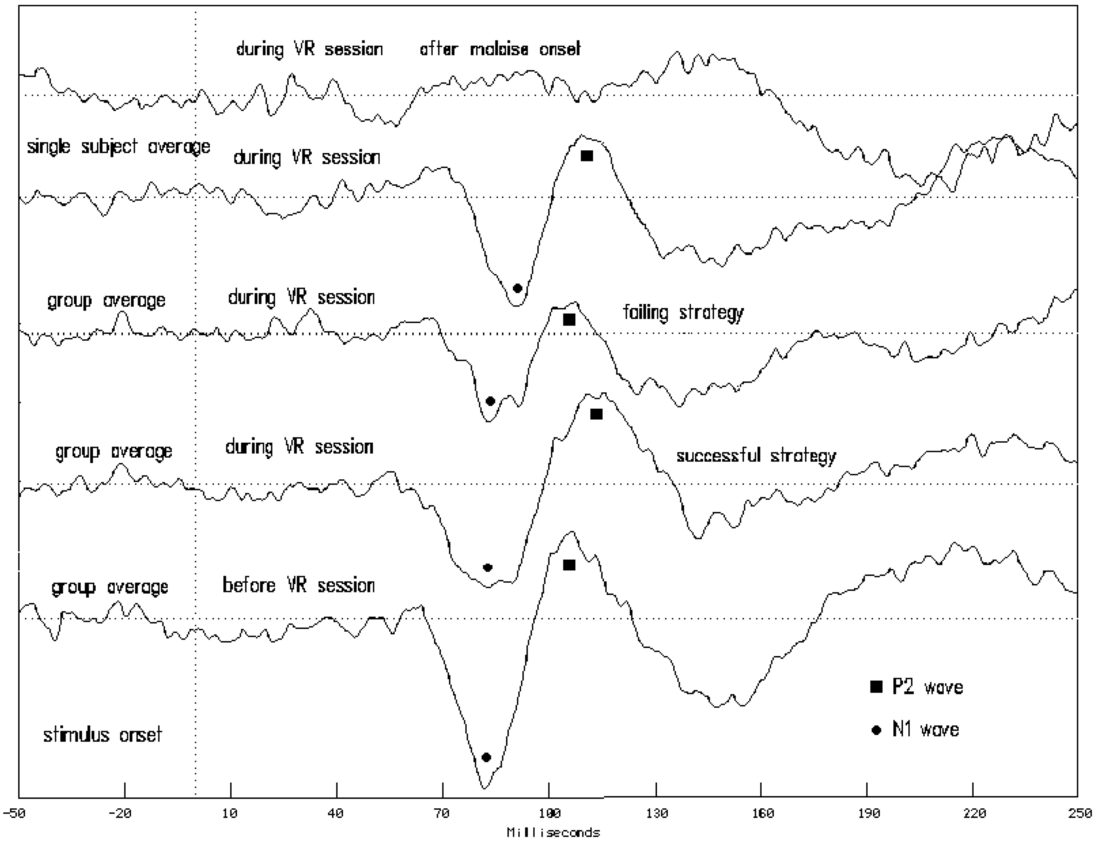


Figure 4. *Averaged evoked responses to auditory stimuli in healthy subjects. Peak amplitude of N1 and P2 components (Cz location is shown) is reduced during VR experience, especially when subjects find it difficult to deal with the cognitive demands of the task and a consolidated strategy is not yet achieved. The appearance of mild nausea in one subject makes the response impossible to recover. (Note: from bottom to top of figure, averages 1-3 and 4-5 have a different display gain).*

ERP mapping could be obtained from a subject wearing a 19 channel headset. For this experiment, 1500 Hz tones were used as stimuli unrelated to VR events. The average ERP recorded during VR shows a decrease in peak amplitude of N1-P2 components especially over anterior regions (figure 5). This observation suggests that the frontal lobes may be more involved than other brain areas. The literature on neurophysiological correlates of the WCST strongly supports this hypothesis (Weinberger et al., 1986; Silberstein et al. 1995).

We are also looking at short term changes in EEG as correlates of specific cognitive events. A critical one is the period lasting 5 to 8 seconds preceding and following feedback. We have programmed this event to take place always in the corridors (figure 1), where no exploratory activity is needed and where important cognitive activities such as anticipation and evaluation of feedback can occur in a predictable sequence. We have made a distinction between pre and post feedback periods and between positive and negative feedback. Preliminary findings pertain only to a subset of the 286 possible periods across the 9 subjects. The amount of alpha (8-12 Hz) and beta (13-20Hz) frequencies were computed for 30 post-negative feedback 8 second periods and 30 post-negative feedback 8 second periods. Beta activity was significantly increased in epochs following negative feedback as compared to epochs following positive feedback over both frontal and parietal electrodes. In the case of negative feedback, EEG findings have a behavioral correlate in an increase in the time spent in the next room, which we have called “negative feedback-evoked time response”’ (figure 6). This response occurs in more than 70% of the time in healthy subjects and reflects the necessity of a re-evaluation of the strategy after a failure. It is well known that many cognitively impaired subjects lack the capacity to make this re-evaluation and to modify their behaviour accordingly. Hence, we expect to find differences between the ability of healthy and impaired subjects to produce this response and its neurophysiological correlates.

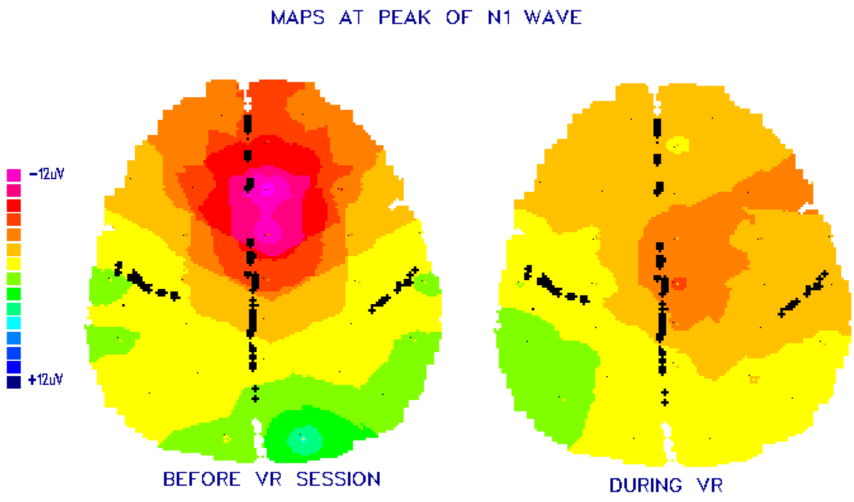


Figure 5. *Colour-coded map showing instantaneous distribution of a subject's peak N1 wave amplitude (uV.) of an auditory ERP recorded from 19 scalp positions before and during a VR session.*

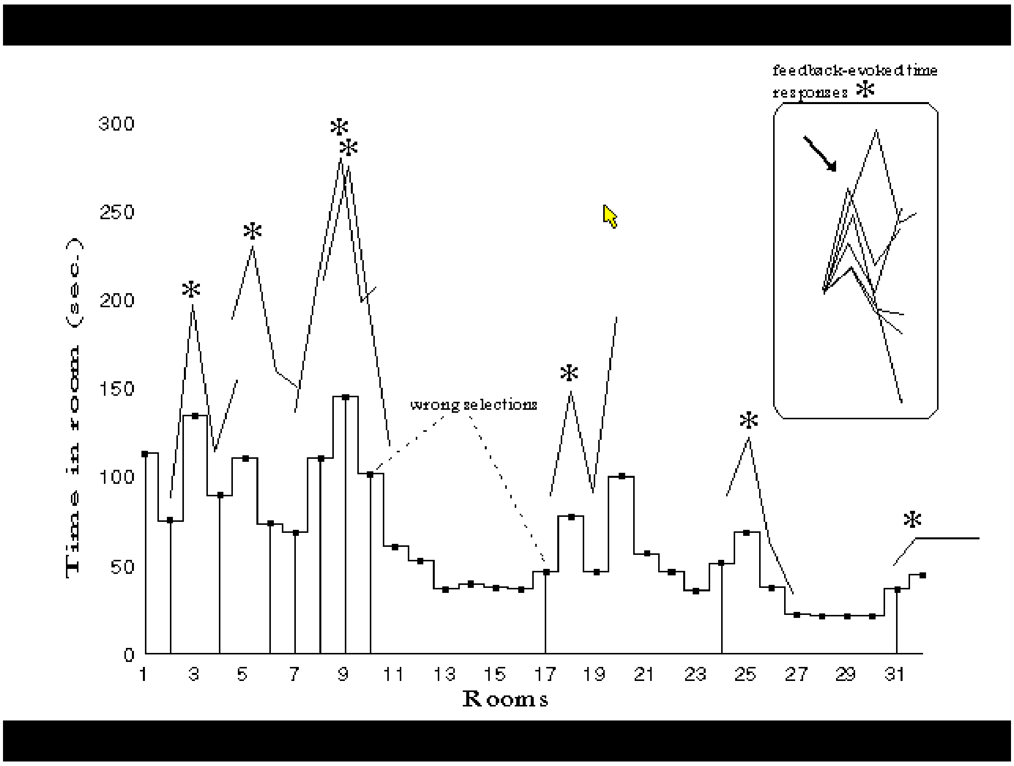


Figure 6. *Plot of time spent in each of ARCANA's consecutive room by a subject. Thin bars indicate when wrong selections occurred. An increase in the time spent in the next room frequently occurs after negative feedback, and is shown in the graph by asterisks.*

In summary, if confirmed and extended by future work and analyses, these findings should have relevance for devising clinically effective VR tools. A similar goal is being pursued by Cole et al. (1995) who have studied multichannel EEG changes prior to and after human interaction with dolphins in order to develop a VR simulator of this therapeutic experience. They have found that EEG spectral peak frequency decreases after real interaction with dolphins and that interhemispheric coherence increases.

Another set of useful data can be obtained by recording EKG changes during a VR experience. As is well known, a tachogram recovered from ECG beat-to-beat analysis is easy to obtain and is useful to assess the degree of physiological stress the subject has experienced (figure 7). This aspect is clearly important to document when dealing with disabilities because we do not want to induce stress for a number of intuitive ethical and medical reasons. Stressful paradigms may affect subjects' performance, acceptance of the technological setup and may delay learning. On the other hand, milder stress may

Proc. 1st Euro. Conf. Disability, Virtual Reality & Assoc. Tech., Maidenhead, UK, 1996

facilitate performance in apathetic individuals. We suggest that heart rate changes also be analysed - with the aim of modifying VR applications to match as closely as needed the real-life situations they intend to simulate. We have found heart rate changes to serve as correlates of psychological frustration in sensitive subjects who were failing on their VR task (Pugnetti and Mendozzi, 1994). Also, heart rate changes can underline differences in psychophysiological stress between real and simulated situations. When VR medical applications induce greater changes in physiological parameters than their real counterparts, they probably need to be re-evaluated in terms of psychophysiological and cognitive demands.

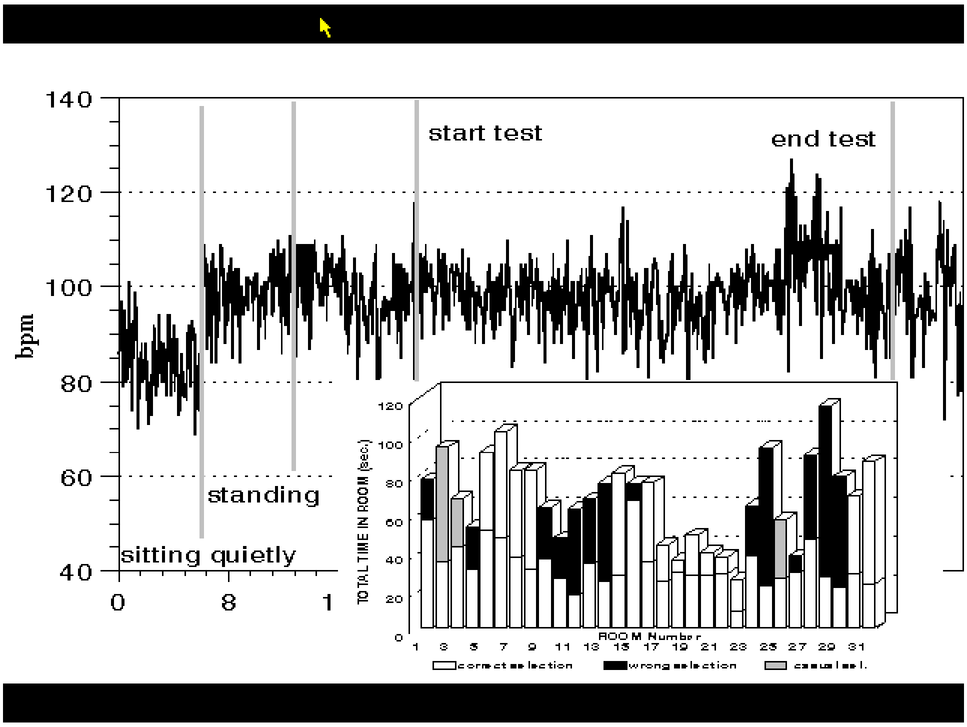


Figure 7. *Plot of mean HR measures taken every 5 seconds in a healthy subject before and during a VR session. The histogram shows time spent in rooms. Black bars indicate wrong selections. Note a HR increase at the times when he was having difficulty finding the right strategy to solve ARCANA's task.*

3. THE FUTURE

A further step in applying VR systems to enhancing learning, enabling communication and assisting cognitive rehabilitation is to bring the experience gained with in-the-field studies with VR into systems capable of modulating critical aspects of the VE in response to biological signal levels. Warner presented both theoretical justification and practical applications of this approach (Warner, 1995) which he calls 'interventional informatics'. Two children with severely disabilities children were enabled to regain control over part of their real environment by means of biosignals and VR associated technologies but, more importantly, their learning abilities were facilitated.

Another most promising rationale for undertaking psychophysiological studies of VR experiences is the implementation of biocybernetic and neurofeedback protocols within VR scenarios. The basic rationale for the former is that by means of on-line biosignal processing (i.e. multichannel EEG) an index of specific mental involvement (e.g. attention) is derived and fed in a closed-loop system to control the amount of items the subject has to pay attention to or control. This approach has being pursued to develop systems such as the EAST, a videogame-based tool intended to demonstrate the concept of rewarding specific brain signal patterns with success at playing. This application is expected to be of value in the treatment of behavioral-cognitive impairments of children with ADD (Pope, 1995). A related application of this concept is expected to emerge within the specific field of neurofeedback to improve the effectiveness of learning voluntary physiological control. According to Budzynski (1995), VR biofeedback has its rationale rooted in the right brain being the mediator of both visuospatial abilities and emotional responses. Bio-VR will act by teaching patients to control emotions via spatial aspects of the VE.

Acknowledgements: This work is being supported by the Italian Ministry of Health, contract n.ICS 49.6/RF 93.25 to the Scientific Institute S. Maria Nascente, and by a grant of the 'Gruppo Volontari per la Ricerca sulla Sclerosi Multipla' to E Barbieri.

4. REFERENCES

W Barfield and C Hendrix (1995), The effect of update rate on the sense of presence within virtual environments, *Virtual Reality*, **1**, 1, pp3-16.

W Barfield, C Hendrix, O Bjorneseth, K A Kaczmarek, W Lotens (1995), Comparison of human sensory capabilities with technical specifications of virtual reality equipment, *Presence*, **4**, 4, pp.329-356.

H J C Berger, K P M van Spaedonck, M W I M Horstink, E L Buytenhuijs, P W J M Lammers and A R Cools (1993), Cognitive shifting as a predictor of progress in social understanding in high-functioning adolescents with autism: a prospective study, *J. Autism Develop. Dis.*, **23**, 2, pp341-359.

F Biocca (1992), Will simulation sickness slow down the diffusion of virtual environment technology? *Presence*, **1**, 3, pp. 334-343.

T H Budzynski (1995), Virtual reality biofeedback: a brief concept paper, *Biofeedback*, **23**, 3, pp.12-13.

D M Cole (1995), Emergent technologies from human-dolphin interaction research, In *Interactive Technology and the New Paradigm for Healthcare* (R M Satava, K Morgan, H S Sieburg, R Mattheus, and J P Christensen, Eds), IOS Press and Ohmsha, pp. 72-78.

J Decety, D Perani, M Jeannerod, V Bettinardi, B Tadary, R Woods, J C Mazziotta, F Fazio (1994), Mapping motor representations with positron emission tomography, *Nature*, **371**, pp.600-602.

N I Durlach (1994), Psychophysical considerations in the design of human-machine interfaces for teleoperator and virtual-environment systems, *Proc. Med. Meets VRII*, San Diego, pp.45- 47.

R C Eberhart and P N Kizakevich (1993), Determining physiological effects of using VR equipment, *Proceedings of the First Annual Conference, Virtual Reality and Persons with Disabilities*, CSU Northridge, pp.47-49.

E R John and P Eston (1995), Quantitative electrophysiological studies of mental tasks, *Biol. Psychol.*, **40**, pp.101-113.

E M Kolasinski (1995, *Simulator sickness in virtual environments. Technical report 1027*, United States Army Research Institute for the Behavioural and Social Sciences, Virginia.

R Pausch, T Crea and M Conway (1992), A literature survey for virtual environments: military flight simulator visual systems and simulator sickness, *Presence*, **1**, 3, pp.344-363.

K Perrine (1993), Differential aspects of conceptual processing in the category test and Wisconsin card sorting test, *J. Clin. Exp. Neuropsychol.*, **15**, 4, pp.461-473.

A T Pope (1995), Psychophysiology and human engineering methods research, *Biofeedback*, **23**, 3, pp.4-6

L Pugnetti and L Mendozzi (1994), Psychopysiological monitoring and models for the remediation of cognitive deficits, *Proceedings of the European Conference on Virtual Reality in Education, Training and Disability,* It. VR Soc., pp41-56.

L Pugnetti, L Mendozzi, A Motta, A Cattaneo, E Barbieri, A Brancotti (1995), Evaluation and retraining of adults' cognitive impairment: which role for virtual reality technology? *Comp. Biol. Med.*, **25**, 2, pp.213-228.

C Regan, (1995), An investigation into nausea and other side-effects of head-coupled mmersive virtual reality, *Virtual Reality*, **1**, 1, pp17-32.

H V Semlitsch, P Anderer, P Schuster, O Presslich (1986), A solution for reliable and valid reduction of ocular artifacts applied to the P300 ERP, *Psychophysiol.*, **23**, pp.695-703.

R B Silberstein, J Ciociari, A Pipingas (1995), Steady-state visually evoked potential topographyduring the Wisconsin card sorting test, *Electroencephal. Clin. Neurophysiol.*, **96**, pp.24-35.

D Warner, J Sale (1995), Intervention informatics: healing with information, In *Interactive Technology and the New Paradigm for Healthcare* (R M Satava, K Morgan, H S Sieburg, R Mattheus, and J P Christensen, Eds), IOS Press and Ohmsha, pp.399-405.

D R Weinberger, K F Berman, R F Zeck (1986), Physiologic dysfunction of dorsolateral prefrontal cortex in schizophrenia, *Arch.Gen.Psychiat.* **43**, pp.114-124.

L J Whalley (1994), Ethical issues in the application of virtual reality to medicine, *Comp. Biol. Med.*, **25**, 2, pp.107-114.

Virtual reality enriched environments, physical exercise and neuropsychological rehabilitation

D. A. Johnson[1], S. Rushton[2] and J. Shaw[3]

[1]Neuropsychology Department, Astley Ainslie Hospital,
Grange Loan, Edinburgh, SCOTLAND

[2,3]Department of Psychology, University of Edinburgh,
7 George Square, Edinburgh, SCOTLAND

[1]*David.Johnson@ed.ac.uk*, [2]*Simon.Rushton@ed.ac.uk*

ABSTRACT

This paper presents preliminary data on the effects of physical exercise and virtual reality upon mood and cognition in severely brain injured adults. The work draws upon two established fields of experimental psychology, enriched environments and physical exercise, in order to propose a new orientation for neurological rehabilitation. The results will be discussed in relation to organisation and delivery of clinical services, and to the theoretical model of von Steinbuchel and Poppel (1993).

Keywords: head injury, activation, virtual reality, rehabilitation, physical exercise, enriched environments

1. INTRODUCTION

Neurological rehabilitation lacks a coherent theoretical basis. Whilst rehabilitation has some intuitive appeal there is little evidence that it produces any significant benefit in the long term. The practice of neurological rehabilitation has seen little significant change for the better since the turn of the century and there remains an overall scepticism of therapeutic efficacy. To some, rehabilitation is little more than well intentioned chalantry (Bell 1992). Despite the well meaning and very professional efforts of rehabilitation staff there remains a very substantial degree of handicap in many brain injured patients in the long term..

Consideration of these problems leads us to suggest that the general approach of neurological rehabilitation must be changed. We suggest that, in order to achieve any significant advances, neurological rehabilitation must firstly delineate the “what-” from the “how-” functions of the brain (Von Steinbuchel & Poppel 1993) and regard the brain as a dependent variable (Bakker 1984). To adopt this altered perspective requires little more than was advocated by Zangwill in 1947: “...the more we find out about the brain and its functions the more we are likely to discover methods of rehabilitation based on strictly scientific principals...”. We suggest that experimental neuroscience has provided us with sufficient knowledge about the brain and its functions to enable the adjustment required to achieve a conceptual rethink in rehabilitation. Further, that our neuroscience colleagues have investigated techniques in both intact and brain damaged subject to the extent that there exists sufficient data to make the clinical application of their work an urgent priority.

In this paper we present our attempts to apply two principle areas of experimental data to a clinical brain damaged population with the aim of effecting a significant and beneficial change in the “how-functions” of the brain.

2. HEAD INJURY

Head Injury, which causes traumatic brain injury, is the major cause of disability in handicap across the entire age spectrum. Post traumatic impairments appear in a wide range of functions including arousal-activation, general levels of activity, motivation, mood and cognition. Either singly or, as is more usual, in combination these post traumatic

 ISBN 0 7049 1140 X

impairments persist indefinitely and adversely affect inter-personal relationships, social adjustment, academic and vocational status. Persistent impairments give rise to suboptimal physiological status, social withdrawal and isolation.

Rehabilitation has generally failed to address the pathophysiological response to trauma, thus neglecting a potentially important vehicle for effecting neurological change. Rather than focusing efforts on the expected underlying and general depression of the cerebral activation, specific problem areas such as memory or activities of daily living are pursued without adequate attention to the likely contribution from underlying neuropathophysiological factors. Attempts at cognitive remediation, for example, may not produce improvements in the specific areas targeted, but benefit may arise in an improved general state of arousal-activation, effort, or motivation of the individual recipients. Depression of cerebral arousal-activation is a common pathophysiological response to head injury. It is reasonable to suggest, therefore, that the first step in any neurological rehabilitation programme should be to facilitate this fundamental process, thereby increasing the brain's ability to receive, process and act upon information. A more activated brain may be better able to benefit from more traditional rehabilitation interventions. Attempts to improve cerebral arousal-activation by the use of pharmacological agents have generally failed. There is no convincing evidence that stimulant medication has any significant and long-lasting impact upon the functioning of the brain damaged individual. There is also the problem of potentially serious and adverse side effects produced by the drugs used. A more easily accessible route to improve arousal-activation may be by way of physical exercise, which has unequivocal benefits upon physical and mental status.

Inactivity is a common complaint after head injury, even in the absence of physical disability. Patients may demonstrate impaired levels of aerobic fitness and complain of fatigue, both of which can impair the rate of progress within rehabilitation. Mental and physical fatigue are common complaints after head injury. Fatigue restricts the level of effort, the duration and quality of performance in all activities. Loss of control over activity generally creates a state of dependence upon others. The sudden and often dramatic imposition of that state is unacceptable to many, giving rise to significant and persistent dissonance. Motivation becomes eroded and a sense of hopelessness pervades the individual's life. The state of behavioural inactivity may exacerbate any endogenous changes in neurotransmitter activity underlying aberrations of mood caused directly by brain injury. Those changes may compound the perceived lack of control, negative feelings and retardation. Acute recovery from head injury includes a hyperdynamic cardiovascular response and significant disruption of central neurotransmitter activity. Is is likely that the early traumatic disruption of the biochemical equilibrium persists after head injury and underlies the chronic post traumatic abnormalities in mood, cognition and behaviour (Johnson, Roethig-Johnston & Richards 1993).

3. PHYSICAL ACTIVITY

Inactivity accelerates the rates of decline of major physiological adaptive systems which eventually reach the point at which the individual's ability to prevent or recover from acute stresses is impaired. The individual's ability to cope with such stresses and preserve subsequent function depends upon the maintenance of adequate physiological reserves, particularly neurological control, mechanical performance and energy metabolism. It is also assisted by such modifying factors as positive affect (e.g. self-confidence). Activity and exercise are increasingly recognised as a way of improving mental health, with reports of reductions in depression and anxiety, and increased perceptions of self control (e.g. Blackburn & Jacobs 1988). The most likely basis for these benefits is a change in central monoaminergic activity. For example, it is suggested that the antidepressant effect of physical exercise shifts monoaminergic function back to normal in clinically depressed subjects. Brown et al (1992) reported improvements in depression, anxiety, hostility, confused thinking, vigour, self efficacy and fatigue for psychiatric institutionalised adolescents. There are similarities between structural and functional declines associated with ageing and the effects of enforced inactivity, such as arises following severe head injury. It is reasonable to draw parallels between severe head injury and normal ageing in which there is a diffuse loss of physiological capacity and reserve and a reduced ability to adapt to changes.

Results from animal experimental studies suggest that chronic exercise may also result in permanent structural changes in the brain. For example, physical exercise in rats improves vascularisation in the cerebellar cortex, while a combination of motor learning with physical activity results in a greater communication network within the brain (e.g. Black et al 1990). Fordyce & Farrar (1991) investigated the effects of physical activity upon hippocampal cholinergic function. Rats demonstrated improved performance on spatial learning and memory tasks following a sustained treadmill activity programme. Neeper et al (1995) reported evidence that physical exercise can increase brain-derived neurotrophic factor in the hippocampus and neocortex. Neeper et al stated that "... interestingly, the greatest effects of exercise upon brain derived neurotrophic factor occurred in highly plastic areas, responsive to environmental stimuli".

4. ACTIVITY AFTER HEAD INJURY

Brain injury may severely disrupt normal patterns of interaction with the environment (Tinson 1989). Moreover, this disruption occurs in a variety of ways. To the extent that the injury entails motor or sensory impairment the opportunities for active interaction are inevitably diminished. Interaction is reduced still further as a consequence of impaired motivation, fatigue, hypo-arousal and impaired concentration and memory, all of which are common sequelae of brain injury. Further, such problems may increase as the time since injury lengthens and the head injured person becomes more withdrawn and isolated in all spheres of activity. Clinicians agree that to increase levels of interaction between the head injured person and their environment is a vital part of any rehabilitation process.

The importance of interaction with the environment is not simply a matter of clinical experience however. Within the animal literature there is strong empirical support for enforced environmental interaction producing clear beneficial effects upon both brain and behaviour (Renner & Rosenzweig 1987). In this context enforced interaction with the environment is referred to as "Environmental Enrichment" and is produced by group-housing of laboratory animals in large, complex and stimulating environments containing an array of manipulable objects. The effects of an enriched environment upon brain and behaviour are complex. In problem solving tasks enriched animals are superior to impoverished or isolation-housed counterparts, and the more complex the task then the greater the advantage. At the same time, enrichment leads to a variety of structural and functional changes in the brain, including a greater mass of cerebral cortex, greater cellular connectivity within the cortex, increased glial activity, a higher cortical metabolic rate and a range of neurochemical changes. In reviewing the progress in enrichment research Will & Kelche (1992) concluded that the enriched environment for brain damaged subjects constitutes a potentially powerful tool which combines the additive effects of its various social and physical components. However, Will & Kelch called for the therapeutic efficacy of enrichment, whether it reflects true recovery or compensation, to be assessed by clinicians.

5. VIRTUAL REALITY

Virtual Reality (VR) provides a powerful means of increasing levels of environmental interaction in a highly controlled and structured manner. The vital characteristic of VR is that it is interactive. Within the virtual world created by the computer every response that the user makes has a consequence to which he must adapt in terms of mental processes and behaviour. Moreover, since interaction with the virtual environment can be made contingent upon whatever motor capacity the patient has then this technology is particularly well suited to applications in neurological rehabilitation. The application of VR to brain damaged patients offers a unique and powerful way of increasing the quantity and quality of direct interaction with the environment and of reducing cognitive and behavioural impairments. A VR enriched environment offers the potential for significant gains in physical and mental function at all levels. Whilst the real or ideal situation (e.g. cycling in the outside world) would necessarily involve the head injured person using community resources on a regular or frequent basis, these naturalistic situations require a high degree of motivation, physical ability and social contact than is usually present in acute rehabilitation patients. There are also the safety concerns associated with impaired balance and motor control, for example. Thus, the VR enriched environment offers a unique and safe opportunity for intervention.

6. PROPOSAL

This preliminary study aims to combine two different experimental approaches to facilitating recovery after brain damage. Patients will be introduced to a graded physical exercise programme, within or outwith an enriched VR environment, with the aim of increasing physical and mental parameters. The anticipated improvements are in general physical, cognitive emotional and behavioural functioning. Any such improvements may have knock on effects in other areas, such as adjustment and outcome. Generalisation from the training programme will be addressed by the incorporation of lifestyle exercise programmes which teach the individual to integrate multiple short bouts of physical activity into the course of their daily lives.

7. METHOD

7.1 Virtual Environment System

The system consists of an instrumented Monarch exercise bike, a 120Mhz Pentium PC with an analogue to digital converter and a 29 inch NEC 4PG monitor. The handle bar rotated and was connected to a potentiometer thus providing the heading angle. A tacho-generator was attached to the wheel providing a measure of cycle speed. The system ran at 20Hz and the sampled values from the bike produced contingent changes in heading direction and speed in the virtual environment. The virtual environment was written using Renderware 2.0 and consisted of a ground plane on which where scattered a series of coloured objects with a mountain range in the distance. The subject's task was to cycle over to the various objects in turn. Several levels of complexity increased the memory and motor requirements. The former was achieved by having differing colours of objects and then requesting that the objects be taken in specific order (e.g. all the reds, followed by all the greens; or red, green, red, green etc.), and the latter by changing the size of the objects and also introducing "puddles" that had to be careful steered between so as to reach the target object. A time course record held the trajectory taken and was available for assessment of improvement in performance time and accuracy.

8. DESIGN

A multiple subjects cross-over design was chosen. The reasons for this are several-fold: in normal clinical studies a commonly employed design in the multiple baseline design. This attempts to demonstrate an effect of an intervention by contrasting the improvement on one set of tasks against another. For example, if a memory training task is being assessed then it would be predicted that after intervention memory tests would reveal an improvement but visuo-spatial tests **would not**. If such a dissociation is found then it can be inferred that the improvement in memory is due to the intervention rather than the endogenous changes. Another design often employed is between groups comparison which assigns participants to either the experimental group or a control group. If there is a statistically significant difference in the improvement between the two groups then it is inferred that the intervention is responsible for this discrepancy. These two designs are unfortunately inappropriate for a small scale investigation of the efficacy of the VR/PE intervention. This is because it would be predicted that any effect would be across the board. Additionally, patients with head injuries are very heterogeneous. Given a large number of patients it would be possible to match them and hope that the confounding factor of individual variations would at attenuated. However, the current study it too modest in size to allow this.

A cross-over design allows the change during an intervention period to be contrasted against the change during a control period. As the results both come from the same patient it removes individual variation as a confounding variable. As endogenous factors will underlie changes, especially during the early stages of recovery it is necessary to compare the rate of increase of change between conditions and also to counter balance the ordering of conditions between patients so that 50% start with VE/PR and 50% start with the control.

9. CONTROL CONDITIONS

The control condition consists of use of standard cognitive rehabilitation software using a BBC Master pc and established training programmes. This is considered to be the one of the most widely used and available interventions. Although no intervention whatsoever might appear to be the best control, ethical considerations require that the best current intervention be used during this time.

10. SUBJECTS

Twenty subjects, aged between 16 and 45 years, will be recruited from patients admitted to the Scottish Brain Injury Rehabilitation Service, Astley Ainslie Hospital, Edinburgh, with a primary diagnosis of head injury. Those patients with a history of previous neurological insult will be excluded. Subjects will be randomly assigned to either experimental or control conditions. Each subject will undergo assessments of cognition, mood and behaviour, physical status and . All assessments will be blind.

11. INITIAL RESULTS

Two early pilot studies (Scott 1995, Shaw 1996) have demonstrated changes on a number of the measures. However, those studies were significantly constrained by their design and the availability of patients so it is not possible to infer more than "promise". At the time of writing a further, more intensive study is underway and the results will be available at the time of presentation to Conference.

©1996 ECDVRAT and University of Reading, UK ; ISBN 0 7049 1140 X

12. CONCLUSIONS

Our aim in this work is to combine two different experimenatl approaches from experimental psychology and apply them to a major clinical population which suffers chronic and severe neurological handicap. Following the approach of von Steinbuchel and Poppel (1993) we seek to tackle rehabilitation at a different, lower, level of analysis than that typically employed in traditional rehabilitation units. The potential benefits from this approach may be substantial and far reaching for the individual patient and for the practice of neurological rehabilitation.

Acknowledgements: We are indebted to Martyn Smyth for the programming, to the technicians in the University Psychology Department, and to continuing collaboration from Dr John Wann and Professor David Rose.

13. REFERENCES

D.J. Bakker (1984). The brain as the dependant variable. *Journal of Clinical Neuropsychology*,**6**, pp. 1-16.

D.S Bell (1992). *Medico-Legal Assessment of Head Injury.* Springfield, Illnois: Charles C. Thomas.

J. E Black, K.R Isaacs, B. J Anderson et al. (1990). Learning causes synaptogenesis, whereas motor activity causes angiogenesis, in cerebellar cortex of adult rats. *Proc. Natl. Acad. Sci. USA.* **87**, pp. 5568-72.

H Blackburn and D.R. Jacobs (1988). Physical activity and the risk of coronary heart disease. *New England Journal of Medicine*, **319**, pp. 1217-1219.

S.W. Brown, M.C Welsh, E.E. Labbe, W.F. Vitulli and P. Kulkarni. (1992). Aerobic exercise in the psychological treatment of adolescents. *Perceptual and Motor Skills,* **74**, pp. 555-560.

D.E Fordyce, R.P Farrar (1991). Physical activity effects on hippocampal and parietal cortical cholinergic function and spatial learning in F344 rats, *Behavioural Brain Research,* **43** pp.115-123.

D.A. Johnson, K. Roethig-Johnston and D. Richards. (1993). Biochemical parameters of recovery in acute severe head injury. *British Journal of Neurosurgery.* **7**, pp.53-60.

S.A Neeper, F Gomez-Pinilla, J Choi and C Cotman (1995). Exercise and neurotrophins. *Nature,***373**, pp.109.

M.J Reener and M.R Rosenzweig (1987). *Enriched and impoverished environments: effects on brain and behaviour.*New York: Springer Verlag.

N.V. Steinbuchel and E. Popper (1993). Domains of rehabilitation: a theoretical perspective. *Behavioural Brain Research.* **56**, pp. 1-10.

D. Tinson. (1989). How stroke patients spend their days. *International Disability Studies*, **11**, pp. 45-49.

B. Will and C. Kelche (1992). Environmental approaches to recovery of function from brain damage: a review of animal studies (1981-1991). In *Recovery from brain damage* (Eds. F.D Rose & D.A Johnson). New York: Plenum Press, pp. 79-104.

O.L Zangwill (1947). Psychological aspects of rehabilitation in cases of brain injury. *British Joutnal of Psychology*,**37**,pp. 60-69.

ECDVRAT '96

Keyword Index

Keyword Index

Author Index

Author Index