Robust multimodal face and fingerprint fusion in the presence of spoofing attacks

Pattern Recognition 50 (2016) 17–25

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Robust multimodal face and fingerprint fusion in the presence of spoofing attacks

Peter Wild *, Petru Radu, Lulu Chen, James Ferryman

University of Reading, School of Systems Engineering, Whiteknights, Reading RG6 6AY, United Kingdom

ARTICLE INFO

Article history:
Received 28 February 2015
Received in revised form
15 June 2015
Accepted 7 August 2015
Available online 19 August 2015

Keywords:
Multibiometrics
Anti-spoofing
1-Median

ABSTRACT

Anti-spoofing is attracting growing interest in biometrics, considering the variety of fake materials and new means to attack biometric recognition systems. New unseen materials continuously challenge state-of-the-art spoofing detectors, suggesting for additional systematic approaches to target anti-spoofing. By incorporating liveness scores into the biometric fusion process, recognition accuracy can be enhanced, but traditional sum-rule based fusion algorithms are known to be highly sensitive to single spoofed instances. This paper investigates 1-median filtering as a spoofing-resistant generalised alternative to the sum-rule targeting the problem of partial multibiometric spoofing where m out of n biometric sources to be combined are attacked. Augmenting previous work, this paper investigates the dynamic detection and rejection of liveness-recognition pair outliers for spoofed samples in true multi-modal configuration with its inherent challenge of normalisation. As a further contribution, bootstrap aggregating (bagging) classifiers for fingerprint spoof-detection algorithm is presented. Experiments on the latest face video databases (Idiap Replay-Attack Database and CASIA Face Anti-Spoofing Database) and fingerprint spoofing database (Fingerprint Liveness Detection Competition 2013) illustrate the efficiency of proposed techniques.

1. Introduction

Fingerprint and face biometrics as most widely adopted traits are being exposed to an increasing threat of presentation attacks. Consequently, there are numerous studies [1,2] and open challenges [3,4] on anti-spoofing techniques assessing the spoofing detector's ability to distinguish between genuine and fake attempts for especially these two traits. Recently, the integration of anti-spoofing scores with recognition scores has received considerable attention [5–7]. The standard approach, as outlined in [5], has been to reject spoofed samples before comparing them against the gallery template. However, recognition scores can be helpful in the probe-attack spoofing detection problem and liveness scores can impact on the recognition task. Considering imposters with access to fake fingers or face photographs reveals an impact on overall accuracy (shifted imposter score distribution for non-zero-effort attempts [7]) and assuming a correlation between successful spoofs achieving a higher score and their corresponding liveness score is likely (and shown) to help in the final judgment of the decision task, especially in an ensemble of classifiers where this paper looks for outliers. It is therefore useful to investigate the benefits of dealing with a holistic (liveness and verification) multi-class problem rather than two separate classification problems (live vs. fake and genuine vs. impostor). If a system involves multiple modalities there is an even larger variety of different ways to treat the problem of combining liveness and recognition scores. Multibiometrics using face and fingerprint biometrics comes with many benefits including expected increased accuracy, higher universality (absence of single characteristics), efficiency (fast indexing), but its robustness to spoofing attempts has been shown to be compromised [8,9]. Furthermore, with the inclusion of multiple modalities the attacker has an even more extended choice to select the easiest modality to be attacked. It is therefore desirable to find new techniques coping with spoofing attacks, which are subject to investigation in this paper. The paper focuses on three objectives: (1) investigation of spoofing robustness in multibiometrics; (2) development of novel methods towards anomaly detection for increased systematic anti-spoofing; and (3) proposition of a novel bootstrap aggregating (bagging) of classifiers method combining features in fingerprint counter-spoofing.

With regard to the first topic on spoofing robustness in multibiometrics, the paper tests degradation in accuracy for the "partial multibiometric spoofing" scenario, where m out of n samples are spoofed, highlighting the tradeoff between accuracy and security for different fusion methods. Fig. 1 illustrates this concept. The

* Corresponding author. Tel.: +44 118 378 7633; fax: +44 118 975 1994.
E-mail addresses: p.wild@reading.ac.uk (P. Wild), p.radu@reading.ac.uk (P. Radu), l.chen@reading.ac.uk (L. Chen), j.m.ferryman@reading.ac.uk (J. Ferryman).

http://dx.doi.org/10.1016/j.patcog.2015.08.007

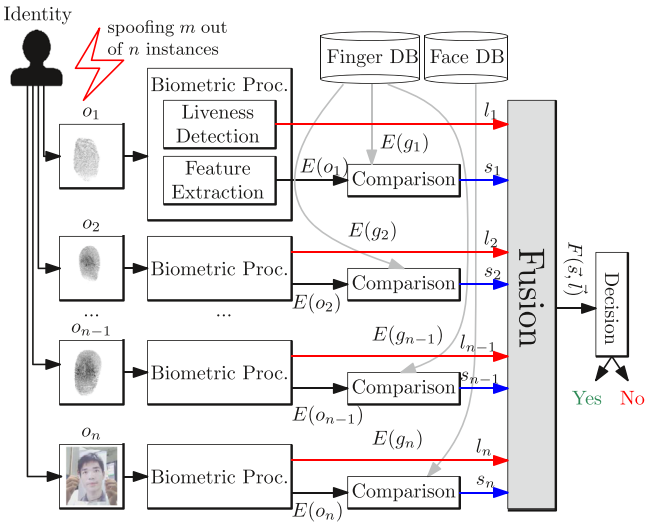

Fig. 1. Partial multibiometric spoofing of observations o_i given templates g_i fusing scores s_i and liveness values l_i.

sensitivity of a recognition and liveness fusion method with regard to spoofing is especially interesting in multimodal configuration, where scores originate from different underlying distributions and multiple traits facilitate a selection of the modality to be attacked. The paper analyses the impact of the number of spoofed fingers or spoofed face on accuracy using the latest biometric datasets. The relative robustness of several score-level fusion rules can be used to choose the most robust fusion rule [9].

As a second outlined contribution, this paper presents a novel multibiometric spoofing-aware fusion method following the idea of anomaly detection and extending research in [10] to multiple modalities. This paper investigates 1-median-based fusion using outlier detection applied in a multibiometric setup. Note that the extension to multiple modalities raises further questions with regard to normalisation. For different modalities, scores generally follow different distributions. Therefore, counter-spoofing is much more challenging than for single-modality approaches, including multi-instance or multi-algorithm approaches. Further, this work presents further theoretical considerations and discusses parameter choice in detail. Recognition scores and liveness scores are likely to be dependent, as spoofing tries to achieve a high recognition score in order to successfully claim the alien (spoofed) identity. In partial multibiometric spoofing this information can be used to further discriminate between genuines and impostors. Despite spoofing sensitivity of traditional fusion techniques, it is a reasonable assumption to claim a higher difficulty for attackers to spoof multiple modalities at the same time or even to obtain the necessary samples to produce a fake fingerprint or face mask. On the other hand, special spoofing-robust fusion schemes might exhibit a reduced level of accuracy. This trade-off between cost and security to limit drawbacks [5,11] is investigated.

Third and last, as a by-product of evaluations the paper further presents a novel spoofing detector again employing a fusion principle: bootstrap aggregating (bagging) of classifiers. This technique is employed in combining the decision outcome of multiple different classifiers. Using also multiple features to be more robust vs. changes in materials (see [12]), the paper aims at investigating this technique in the employed system as an anti-fingerprint spoofing technique towards integral fusion concepts in robust anti-spoofing. Bagging is shown to outperform state-of-the-art detectors on the most challenging LivDet 2013 Crossmatch subset database.

The remainder of this paper is organised as follows: Section 2 introduces the problems of anti-spoofing and spoofing-aware fusion in biometrics. The proposed methods of bagging for spoof-detection and 1-median filtering for spoofing-resistant multibiometric fusion are outlined in Section 3. Section 4 highlights experimental results with regard to the proposed and investigated techniques. This includes a discussion of methods towards anomaly detection in multibiometrics, highlighting parameter choice and optimisation for the proposed 1-median filtering. Section 5 concludes this paper with an outlook on future work.

2. Related work

There are several anti-spoofing or liveness detection algorithms extracting features (usually trained for modality, sensor, material, etc.), in order to determine whether a biometric sample is either *live* or *fake*. For evaluation purposes, *ferrlive* (rate of misclassified live samples) and *ferrfake* (rate of misclassified fake samples) are employed. Whereas for individual modalities the anti-spoofing problem is well defined and evaluated separately from biometric system performance, research on fusion between match scores and liveness factors is still in its infancy [13]. Recently, [14] suggested a framework for verification systems under spoofing attacks. Within the framework [8] adopted in this paper, liveness and recognition scores are combined considering the scenario of probe-spoofing only (i.e. no gallery-spoofing, enforced by e.g., attended enrolment). Formally, given a vector of biometric observation (units, e.g. fingers, eyes) $\vec{o} = (o_1, \ldots o_n)$ from one or more modalities, and corresponding claimed identity template $\vec{g} = (g_1, \ldots, g_n)$, the task of the fusion module F is to compute a unified decision score, using comparison scores $\vec{s} = (s_1, \ldots, s_n)$ and (probe) liveness values $\vec{l} = (l_1, \ldots, l_n)$, so that the verification task V (authentication based on threshold η) can be formulated as follows:

$$V(\vec{o}, \vec{g}) := \begin{cases} \text{accept} & \text{if } F(\vec{s}, \vec{l}) \geq \eta; \\ \text{reject} & \text{else.} \end{cases} \quad (1)$$

Let i be the current index and $E(o_i), E(g_i)$ refer to extracted (modality-specific) features of samples, then $s_i = C(E(o_i), E(g_i)) \in [0, 1]$ is used to denote the normalised comparison result of o_i, g_i and $l_i = L(o_i) \in [0, 1]$ denote the likeliness of a genuine (live) sample. Clearly, it is desirable to find a method F unaffected in performance if m out of the n elements of $\vec{o}$ are spoofed. This testing setup is referred to as "partial multibiometric spoofing", introduced in [10] and extended in this work towards multiple modalities. Note that this notion of live or spoofed probes vs. always-live enrolled gallery samples (assuming attended enrollment) leads to a simpler modelling (2 classes distinguishing live probe from spoof or live, but different sources) than in the general asymmetric case (8 classes based on live/spoof probe, live/spoof gallery sample, and same/different source) or symmetric case (6 classes) [7], fully concentrating on a dichotomous authentication task, which can be evaluated in the traditional way using receiver operating characteristics.

2.1. On combining anti-spoofing and recognition

Marasco et al. [5] are among the first considering fusion of liveness with recognition scores separately for each modality, using simple rejection of spoofed samples. If a spoofing attempt is indicated, the current modality matching score is ignored. This initial study is extended in [15] evaluating sequential fusion, classifier fusion, and Bayesian Belief Networks for combining match scores and liveness measures, highlighting the superiority of the latter method for the LivDet2009 dataset but also that accuracy is decreased when taking liveness detection into account. Chingovska et al. [6] evaluate binary decision rules and Logistic Regression (LR) as decision and score-level fusion techniques combining face recognition and liveness scores addressing the integration (but neglecting the partial spoofing problem) of liveness. They report higher resistance to spoofing attacks (91.54% vs. 10%) but are outperformed by

LR approaches achieving both, high verification accuracy and good spoofing detection. Recently, Poh et al. [7] have targeted the problem of integrating spoofing and matching scores in a probe and gallery-spoofing scenario, investigating Gaussian Copula-based Bayesian classifiers and a mixture of linear classifiers for this task. While their method outperforms classical Support Vector Machine (SVM) based techniques, the approach needs training with regards to the full range of attacks.

The assessment of traditional fusion rules (this work is using Kittler et al.'s classical framework [16]) in the presence of spoofing attacks is a further relevant sub-problem and addressed in this work. Rodrigues et al. [8,17] first addressed this security issue of spoofing attacks against a multimodal biometric system. They presented two methods, one using likelihood ratio and another employing fuzzy logic, both exceeding the accuracy of traditional fusion rules. Also Akhtar et al. [18] studied the impact of spoofing on parallel and serial fusion rules for face and fingerprint reporting that score-level fusion methods from the literature are not robust to spoofing attacks and that serial fusion gave better results for an overall assessment of performance, verification time, user acceptability and robustness.

2.2. Anti-spoofing in fingerprint and face recognition

In fingerprint recognition, there are two general ways to address the spoofing problem: either by actively assessing the liveness (e.g. by measuring pulse, perspiration patterns, or blood pressure), or by passively analysing patterns of spoofed materials (e.g. lack of detail, pattern differences). The latter type, which is the subject of interest in this paper, reveals high risk of material and sensor-dependence [12]. An excellent recent survey of spoofing methods in fingerprint recognition can be found in [2]. Among the most common techniques for static (extracted from single image) texture-based anti-spoofing methods are statistical features [19], Power Spectrum Fourier analysis [20], Ridge Frequency Analysis [19], Local Binary Patterns (LBP) [21] and Local Phase Quantisation [22]. However, recent developments towards material-independent static anti-spoofing suggest to combine multiple features and probably even detectors. Fumera et al. [13] give a good introduction into the problem of combining multiple liveness detectors for a single modality, fusion of liveness detector and matcher for a single modality, and anti-spoofing capabilities of ad hoc fusion rules combining multiple comparison scores.

Face spoofing counter-measures can broadly be classified into texture-based and motion-based counter-measures. A good overview on face counter-spoofing may be found in [23]. The first category assessing textural properties is the more widespread group with approaches like LBP [24], or statistical features [25] exploiting the observation that images/videos with spoofed faces (printed or replayed) do not exhibit the same noise-level like genuine samples. The second type of motion-based approaches targets the reproduction of (flat) printed photographs or re-display of faces on tablets exploiting the difference in 3D appearance of spoofed approaches. For fusion purposes this paper focuses on the first type and employs an existing anti-spoofing system [26].

3. Proposed methods

In order to solve the problem of robust face and fingerprint fusion in the presence of spoofing attacks, this work proposes 1-median filtering for enhanced tolerance with regards to a number of attack-outliers in the ensemble of score-liveness tuples, and bagging of classifiers for enhanced (fused) spoofing resistance. Both methods are described in detail in the next subsections.

3.1. Score and liveness fusion using 1-median filtering

Aiming to overcome the limitations of traditional sum-rule based techniques, which are known to be very susceptive towards outliers and thus easy to be attacked in multibiometric configuration, where an attacker can target the weakest link in the chain of combined biometric units to be attacked (e.g. using a particular available latent fingerprint), 1-median filtering [10] is investigated as a method for joint face and fingerprint score-and-liveness fusion. The motivations in the definition of 1-median filtering are (1) an extension towards a hybrid between sum rule and median rule as in Kittler et al.'s classical fusion methods [16] to find an optimal compromise between (0-spoof) accuracy and (m-spoof) robustness performance; and (2) an incorporation of high-dimensional information to be combined (score $\vec{s}$ and liveness $\vec{l}$ pairs as introduced in Section 2).

Based on the median rule's property to be less affected by outliers (which is very beneficial for spoofing resistance), this fusion method can be formulated as follows:

$$F_{mf}(\vec{s}) := \frac{1}{\sum_{i=1}^{n} M(\vec{s}, s_i)} \sum_{i=1}^{n} M(\vec{s}, s_i) s_i. \tag{2}$$

$$M(\vec{s}, s_i) := \begin{cases} 1 & \text{if } \left| s_i - \underset{j=1}{\overset{n}{med}}\, s_j \right| < \phi; \\ 0 & \text{else.} \end{cases} \tag{3}$$

Note that parameter ϕ limiting the zone of influence allows for an arbitrary tradeoff between the sum rule ($\phi = \infty$ results in $F_{mf}(\vec{s}) = \frac{1}{n}\sum_{i=1}^{n} s_i$, the classical sum rule) and the median rule (for ϕ sufficiently small, the definition becomes $F_{mf}(\vec{s}) = med_{j=1}^{n} s_j$, the median rule). One of the important tasks is to find a suitable (trained) parameter ϕ, which can be either fixed or a function of the scores. The choice of ϕ is not straightforward and should rely on the underlying distribution's properties. Parameter selection is discussed in Section 4.7.

As a motivation for the selection of the median, consider the following theoretical considerations: as observed in [6] unimodal non-zero-effort imposter score distributions (comparing a spoofed sample with a genuine reference) are shifted towards the genuine distribution (comparing two live genuine samples) compared to zero-effort imposters (comparing two live samples from different identities). In an equally weighted mixture-model for random variables X_i, we have $E(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \mu$ assuming independent, normalised (same mean μ and variance σ^2) distributions. However, for the variance, we get $Var(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n^2} Var(\sum_{i=1}^{n} X_i) = \frac{1}{n}\sigma^2$. While this illustrates the positive effect of fusion on imposter scores (narrowing the variance), it also clearly illustrates that if one of the random variables follows a degraded spoof-imposter distribution with lower mean, this is likely leading to a bimodal distribution (especially if n is large). Assuming the distributions can be modelled by Gaussians, a mixture of two normal distributions with highly unequal means has a positive kurtosis, as the smaller distribution lengthens the tail of the more dominant one. While there are exceptions to the rule [27], as a rule of thumb, it is generally suggested, that in skewed distributions, the mean is farther out the longer tail than the median [28], therefore a better representative in the filtering process (which is likely to succeed as can be seen from theoretical considerations if the number of spoofed modalities is low compared to n). However, note that the crucial pre-assumption is a proper normalisation, which ideally should leave genuine score distribution almost unaffected. Further, note that median filtering is not to be mixed up with image- or kernel-based combination and works on score (and liveness-) values to be combined.

Algorithm 1. Median filtering.

Data: comparison pairs' scores $\vec{s} = \{s_1, s_2, \dots s_n\}$ and probe-liveness $\vec{l} = \{l_1, l_2, \dots l_n\}$, trained parameters ϕ, Ψ

Result: verification decision $p \in [0,1]$ indicating live and genuine match

$\mathrm{m} \leftarrow \underset{j=1}{\overset{n}{med}} \begin{bmatrix} s_j \\ l_j \end{bmatrix}$;

for *each score pair* (s_i, l_i) **do**

  if $\left\| \begin{bmatrix} s_i \\ l_i \end{bmatrix} - m \right\| < \phi$ **then**

    reject sample from $\vec{s}, \vec{l}$ and update $n \leftarrow n-1$;

  end

end

$\mathrm{f} \leftarrow \frac{1}{n} \sum_{i=1}^{n} \begin{bmatrix} s_i \\ l_i \end{bmatrix}$; $\mathrm{p} \leftarrow \mathrm{logreg}(f, \Psi)$;

The outlined technique can easily be extended to 2D for combining points (s_i, l_i) of recognition and liveness scores, using the geometric-median (1-median). This is the point minimising the sum of distances to the sample points using score s_i and liveness l_i as coordinate values.

$$F^2_{mf}(\vec{s}, \vec{l}) := \frac{1}{\sum_{i=1}^{n} M\left(\begin{bmatrix} \vec{s} \\ \vec{l} \end{bmatrix}, \begin{bmatrix} s_i \\ l_i \end{bmatrix}\right)} \sum_{i=1}^{n} M\left(\begin{bmatrix} \vec{s} \\ \vec{l} \end{bmatrix}, \begin{bmatrix} s_i \\ l_i \end{bmatrix}\right) \begin{bmatrix} s_i \\ l_i \end{bmatrix}. \quad (4)$$

$$M\left(\begin{bmatrix} \vec{s} \\ \vec{l} \end{bmatrix}, \begin{bmatrix} s_i \\ l_i \end{bmatrix}\right) := \begin{cases} 1 & \text{if } \left\| \begin{bmatrix} s_i \\ l_i \end{bmatrix} - \underset{j=1}{\overset{n}{med}} \begin{bmatrix} s_j \\ l_j \end{bmatrix} \right\| < \phi; \\ 0 & \text{else.} \end{cases} \quad (5)$$

Fig. 2 illustrates how median filtering uses the median as a seed point to select all points in a local neighbourhood, computing the centroid of the set of filtered points as an even better local representative. As the median is less affected by outliers (left example) it is beneficial in the presence of outliers, whereas in case samples are less scattered (right example), no samples are rejected. Algorithm 1 illustrates all the steps. Note that the additional processing time needed for the comparison should have a negligible impact, as n is traditionally rather small.

Note that the 1-median is not necessarily an input point and for performance reasons an approximation (e.g., coordinate-wise median) might be sufficient. As the task of the fusion module is to come to a final single decision score, a further mapping to a single scalar is necessary. For this task, LR or SVMs can be employed to find the hyperplane $\Psi : \vec{w} \cdot \vec{x} - \vec{b} = 0$ optimally separating the sets of genuine and zero-/m-spoof impostors, where m is the number of spoofed samples in the joint fusion scheme

$$F_{mf}(\vec{s}, \vec{l}) := dist(F^2_{mf}(\vec{s}, \vec{l}), \Psi). \quad (6)$$

Separability becomes more difficult for larger values of m (spoofed samples), the presented implementation uses $m = \lfloor n/2 \rfloor$. Threshold variation is equal to moving the hyperplane separating the two (genuine and impostor) joint score- and liveness-distributions. Training of Ψ is discussed in Section 4.5.

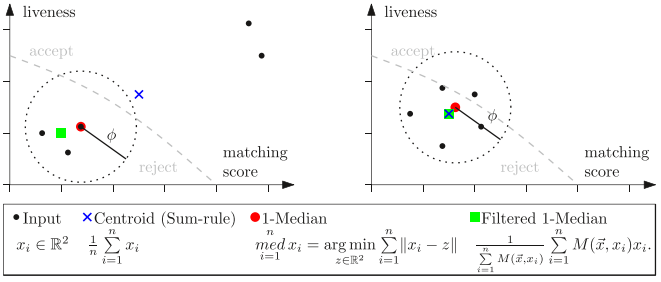

Fig. 2. 1-Median filtering vs. sum-rule in 2D with outliers (left) and without (right).

3.2. Bagging-based fingerprint liveness detection

For fingerprint anti-spoofing relatively poor performance compared to other test sets is reported for the Crossmatch subset of LivDet 2013 fingerprint database [4]. In order to improve those results, this paper proposes the following novel spoof detection algorithm. The employed setup follows a three-stage architecture with preprocessing, feature extraction, and classifier fusion. Fig. 3 illustrates the processing chain. In the preprocessing stage, the fingerprint image is segmented and aligned. The background of the image is removed using Otsu's thresholding [29] and the region of interest is automatically cropped at a dimension of 248 by 256 pixels. Feature extraction extracts global properties and local texture details using three methods selected as representative methods (wavelet-based, statistical and frequency-based) to make maximal use of the fusion technique:

1. *2D Gabor filters* [30]: These filters as a product of a Gaussian and a sinusoid capturing local details are parameterised by Gaussian space constants δ_x and δ_y, frequency f of the modulating sinusoid and orientation θ

$$G(x,y,f) = \frac{1}{2\pi\delta_x\delta_y} e^{-1/2(x'^2/\delta_x^2 + y'^2/\delta_y^2)} \cos(2\pi f x')$$
$$x' = x \sin\theta + y\cos\theta \quad y' = x\cos\theta - y\sin\theta \quad (7)$$

   Similar to [31] θ is set to $0^\circ, 45^\circ, 90^\circ$ and 135° at frequency $f = 0.1$, corresponding to 10 pixels (typical inter ridge distance). The filtered region of interest is divided in blocks of 20×20 pixels. For each block, mean and standard deviation are computed leading to a total of 880 features.
2. *Gray level co-occurrence matrix* (*GLCM*): 6 features each were extracted from 20 by 20 sized pixel-blocks (leading to a total feature vector size of 1100 components) computing local characteristics following [31]: maximum probability, entropy, contrast, energy, homogeneity and inverse difference moment of order k.
3. *Fourier Transform* (*FT*) *based features*: As global features on the Fourier-transformed image, the sum of absolute differences between pairs of concentric circles (at distance 1–3 pixels, evaluated at 25 locations) are computed, similar to [32], yielding 180 components.

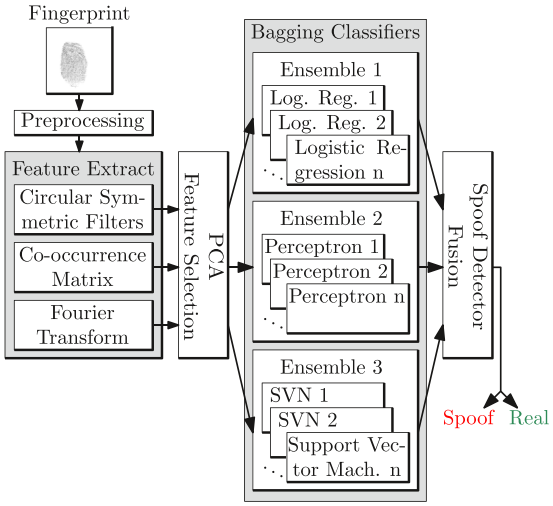

Fig. 3. Proposed fingerprint counter-spoofing based on bagging classifiers.

The final feature of 2160 components is obtained by concatenating the 3 individual feature vectors described above and Principal Component Analysis (PCA) is applied as a feature selection procedure retaining 99% of the variance of the data with 80 PCA components. Besides a fusion of features, the suggested anti-spoofing algorithm employs a multiple classifier framework, which distinguishes itself from a standard multiple classifier system by applying a bagging [33] technique for its component (base) classifiers.

The Bootstrap AGGregatING (bagging) method [33] is used to add base classifiers to the base ensemble using bootstrap replicates on the training set. The bootstrap method facilitates determining the probability distribution of the data without using the Central Limit Theorem [34]. The idea behind bootstrap sampling is to create an artificial random list of the labelled training set by picking some labels more than once. One classifier is trained on this random list and is added to the base ensemble. In the operational phase, the base classifiers are applied to the input feature and their outputs are combined at the decision level by using majority vote. To benefit from the variations of the training set, it is better if the base classifiers are unstable (e.g. neural networks and tree classifiers). In the present work, three base classifiers are employed: (1) regularised LR; (2) single layer perceptron, and; (3) SVM. The three base classifiers are trained $n=100$ times on different bootstrap replicates of the training data. In the operational phase, the 300 classifiers decisions are recorded as 0 or 1, where 1 indicates that the classifier believes that the image is spoofed. The final spoofing score s_f for one test fingerprint image is given as

$$s_f = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} D_j^i}{mn} \tag{8}$$

where m is the number of base classifiers types, n is the number of bootstrap replicates and D_j^i is the decision of base classifier of type I trained on the bootstrap replicate number j. In the proposed implementation, $m=3, n=100$. The value of s_f is in the range $[0,1]$. By setting and adjusting a threshold $t \in [0,1]$, different operating points of the anti-spoofing approach are configurable.

4. Experiments and discussion

In order to evaluate the suggested 1-median filtering and bagging approaches, a modularised setup is employed, using state-of-the-art feature extraction, recognition and spoofing algorithms described in the following sections. After an introduction into database, metrics and employed reference setup, this section concentrates on questions related to baseline performance of spoofing detectors evaluating the bagging classifier approach and inter-relation of recognition and anti-spoofing. Then, partial multibiometric spoofing in multibiometric face and fingerprint context is assessed, considering recognition-only and joint recognition and liveness fusion techniques.

4.1. Setup: database and metrics

As in many other approaches assessing face and fingerprint fusion [18,35], also this work builds on a chimeric dataset pairing face and fingerprints originally originating from different people. This approach is justifiable, since fingerprints and faces as biometric modalities can be assumed to be independent. While also true multi-biometric databases exist with multiple traits collected from the same person, this does not extend to spoofing datasets. Further, spoofing datasets are created for liveness detection purposes and therefore usually do not have to come with a large number of genuine samples, which are needed for the intended recognition-based assessment. In contrast to spoofing evaluations assessing *ferrfake* and *ferrlive* (see Section 2) measures, this evaluation refers to Receiver Operating Characteristics (ROC). Note that for $m>0$ spoofed samples these refer to pairs of Spoof False Acceptance Rate (SFAR) and Genuine Acceptance Rate (GAR, the percentage of genuine users being accepted), rather than False Acceptance Rate (FAR) and GAR pairs, see [36]. For comparing recognition performance (S)EER is employed as the (Spoof) Equal Error Rate where GAR=(S)FAR and decidability index (d-Prime) as $d' = |\mu_1 - \mu_2| / \sqrt{(\sigma_1^2+\sigma_2^2)/2}$ measuring the separation of distributions with mean μ_i and standard deviation σ_i.

Paired data originate from the following datasets (Table 1):

- *LivDet 2013 CrossMatch* [4]: The 4500 images of 99 users support up to 3 genuine samples per finger and a varying number of spoofed samples made from BodyDouble, Latex, Playdoh and WoodGlue.
- *Idiap ReplayAttack* [37]: The counter-spoofing video database of 1300 clips of 50 clients with 320 × 240 pixels resolution provides 8 genuine and 40 attack samples per user offering a *controlled* (homogeneous background) and challenging *adverse* recording setup. There are 4 mobile attacks using iPhones, 4 high-resolution iPad replays, and 2 hard-copy prints.
- *CASIA AntispoofingFace* [38]: This database of 600 clips of 50 clients with 640 × 480 pixels resolution comes with 3 genuine and 9 fake samples per user, offering low, medium and high quality setups and 3 fake attacks.

A chimeric dataset is compiled, combining faces and fingers from the databases above, forming a new set of 85 classes. Note the number of classes, 85, is due to the restriction of LivDet to right-hands only and guaranteeing a minimum number of genuine and spoof fingers to simulate the selection of n out of m spoofed samples of different fingers in a random way. Testing uses right hands only to allow for a training of the employed counter-spoofing detector and learning-based parameters of median filtering. Spoofing attempts are simulated by randomly replacing m out of n samples (4 fingers and 1 face) with spoofs.

Table 1
Employed test databases.

Mod.	Database/set	Images	Users	Training	Testing
FP	LivDet2013 Crossmatch	4500	99	Left hand	Right hand
Face	ReplayAttack	1300	50	*train* set	*devel, enroll, test*
Face	AntispoofingFace	600	50	–	*train, test*

4.2. Setup: baseline system

For experiments, the following baseline system is employed:

- *NIST Biometric Image Software* [39]: For fingerprint feature extraction (using minutiae detection *mindtct*) and comparison (using *bozorth3* in 1:1 verify mode). While the final score is not normalised, (capped) min–max normalisation is used to map scores to the unit interval $[0,1]$.
- *Neurotechnology VeriLook 5.5* [40]: For processing video/still image face samples. The off-the-shelf software extracts a 4–35 kB template via facial reference points and is able to account for off-axis registration (using 15 degrees roll, pitch and yaw parameters). Note that quality assurance (ISO/IEC 19794-5:2005) was deactivated to account for low-quality samples in the dataset. The setup employed in this work used the *low matching speed* setting with switched-off threshold (such that final scores could be obtained).
- *LBP-TOP Face-Liveness* [24]: Using the open-source implementation in [26] for face liveness-detection. The LBP-TOP operator calculates LBP features at three orthogonal planes that intersect

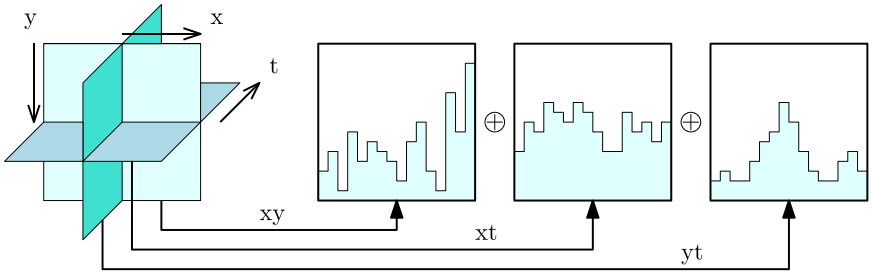

Fig. 4. LBP-TOP process [24]: for each of three planes intersecting at one pixel, LBP histograms are computed and concatenated.

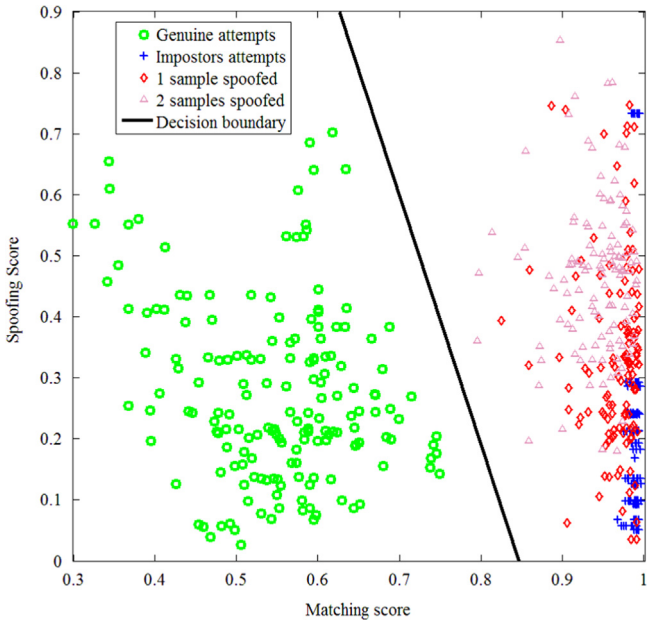

Fig. 5. Fused liveness-comparison scores with trained decision boundary for genuine, impostor, 1-spoof and 2-spoof pairs for finger-and-face fusion.

in the centre pixel. The features are extracted from each separate plane and then concatenated together. A multi-resolution description is then generated such that the histograms along the time domains (*XT* and *YT*) are concatenated for different values of time *t*. Fig. 4 illustrates the process. Compared with traditional 2D LBP features, LBP-TOP can capture spatio-temporal features combining information from both image and time domains. SVMs are then applied for classification.

- *Bagging Fingerprint-Liveness*: As introduced in Section 3.

4.3. Baseline performance of spoofing detectors

In a first experiment, the detection performance of employed spoofing detectors is investigated. For anti-spoofing performance assessment, *ferrfake* and *ferrlive* rates are computed, using the underlying LivDet 2013 Crossmatch dataset for the fingerprint modality, and the ReplayAttack database for face. The implementation of the face spoofing-detection algorithm based on the open source package provided from the original work [26] yielded an accuracy of 85% on ReplayAttack. A similar accuracy is obtained for the presented fingerprint detection scheme, however the rate is much more remarkable given the high quality of the underlying spoofing dataset. Counter-spoofing using bagging classifiers yielded an accuracy of 84% on the Crossmatch set (the method is trained using a distinct subset of the test database using fingers from left hands), a value which significantly outperforms the best accuracy reported (68.8%) in the results of the LivDet liveness detection competition [4] for this subset. Given that integrated feature-level and classifier fusion in spoofing detectors is not a common practice, recognition rates are very promising and suggest to explore this topic even further in the future.

4.4. On the mutual impact of liveness and recognition

Spoofing systems are usually evaluated on their own without taking recognition into consideration [6]. However, the joint operation of liveness and recognition systems in practice raises a series of questions, most notably how to combine recognition and liveness information. This section highlights that not only liveness values are useful in justifying the authenticity of an identification claim, but also vice versa: the recognition score of a template is actually helpful to judge the presence of a spoofing attack. In LivDet, counter-spoofing performance is measured in terms of *ferrfake* and *ferrlive* rates referring to finger images as inputs, not comparisons. If recognition scores are to be considered in the evaluation of spoofing detection, it is important not to assume specific properties of the comparison. In real-world applications, however, a fake fingerprint will be employed to fake the originating identity causing its corresponding score to be distributed according to the spoof-imposter score distribution, whereas a live fingerprint is to originate from either genuine or zero-impostor distributions. The impostor score distribution is likely to shift towards the genuine score distribution in evaluations considering non-zero-effort impostors [6], whereas the genuine score distribution remains unaffected. Since spoofs are unlikely to be perfect, scores are typically degraded, which can be used to judge the presence of a fake sample using, e.g. fuzzy logic as employed in [8]. Simple fuzzy-rule based (inverting the recognition score contribution after threshold $t=0.6$) product fusion of recognition and liveness scores using randomly determined identity claims in experiments increased the spoofing detection capability from 18.36% *ferrfake*=*ferrlive* to 13.86% for the LivDet CrossMatch set. The integration of spoofing scores into recognition accuracy can further increase recognition accuracy when considering 1-spoof impostors in evaluations, e.g. by simply rejecting the sample [15]. Therefore, the real challenge is to find a suitable tradeoff between recognition accuracy and spoofing robustness, subject to investigation in the next sections.

4.5. Combining liveness and recognition scores with logistic regression

With the positive impact of recognition on spoofing detection and vice versa, it is reasonable to proceed towards a holistic framework integrating both evidence as discussed in Sections 2 and 3 and illustrated in Fig. 1. When combining both sets of scores individually using sum rule, logistic regression can be used to learn a more robust boundary. This boundary is trained (and then outliers eliminated using the presented 1-median filtering approach) in already combined (fused) space, which is able to clearly distinguish between different zero-effort and spoof detection. The 1-spoof and 2-spoof examples are used to account for the median tolerating a number of

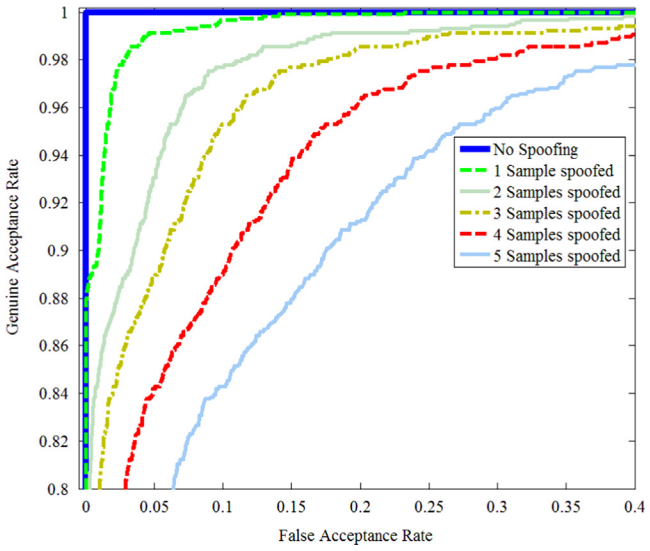

Fig. 6. ROC for partial multibiometric spoofing using sum rule.

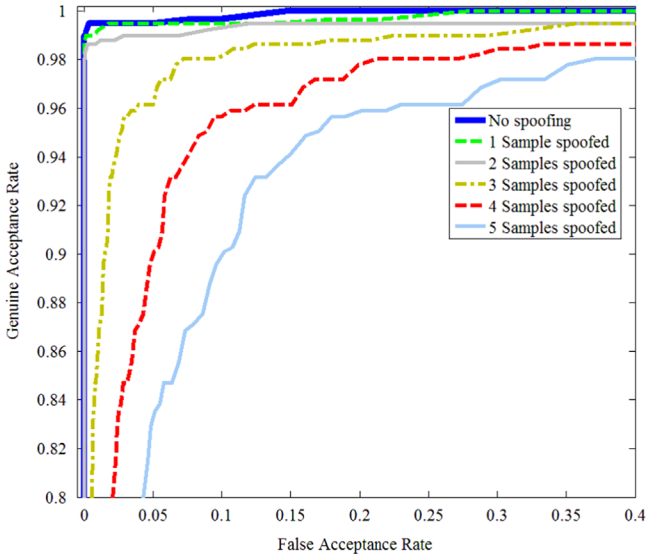

Fig. 7. ROC for partial multibiometric spoofing using median rule.

outliers up to half of the samples. The trained decision hyperplane in Fig. 5 has the following form:

$$\Psi : y = -13.52x + 12.4257. \quad (9)$$

The experiment also clearly illustrates how m-spoof distributions are shifted towards the genuine distribution with increased m.

4.6. Classical fusion in partial multibiometric spoofing

In partial multibiometric spoofing, m out of n biometrics samples of an identity are spoofed ($n=5$ with 4 fingerprints and 1 face in experiments, using equal probabilities). That is, an attacker is assumed to have access to $m = 0, 1, \ldots, n$ latent fingerprints or face masks/print-outs for the attempt to spoof the system, referred to as an *m-spoof* attack. Recognition rates for the common fusion rules sum, product and median are evaluated in this scenario and results listed in Table 3. Confirming assumed behaviour, experiments in this work show that spoofing clearly impacts on recognition. For simple sum rule fusion it is evident that recognition is affected if even a single sample is spoofed, EER is degraded from 0% to 2.32% (d-Prime from 2.91 to 2.62). While perfect separation in the first

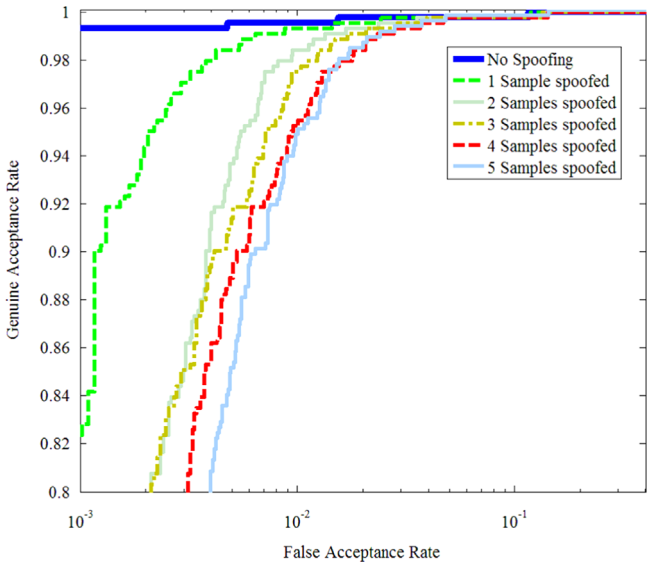

Fig. 8. ROC for partial multibiometric spoofing using 1-median filtering + LR.

Table 2
Average mean vs. median of m-spoof imposter distribution.

Unit	0-spoof	1-spoof	2-spoof	3-spoof	4-spoof	5-spoof
Median	0.991	0.989	0.984	0.965	0.935	0.921
Mean	0.989	0.964	0.940	0.914	0.889	0.873

case is certainly also attributable to the size of the dataset, the degradation is clearly visible, taking also the large number of 5 combinations of 2 modalities in a single authentication attempt in our particular setup into account. Further m-spoofing of faces and fingerprints suggest an increase of abs. 2–3% using sum rule for every additional spoofed sample (EERs of 2.32–12.24% for 1–5 spoofs), which is even more pronounced compared to previous fingerprint-only experiments [10] and spoofing of all samples did not always lead to acceptance (due to degraded accuracy of spoofed samples). The ROCs in Figs. 6 and 7 illustrate the degradation for sum and median rules, respectively.

In contrast to previous experiments on fingerprints only, median rule has shown to be even more successful in combining results. There is little difference between 0-spoof, 1-spoof, and 2-spoof samples (EERs of 0.42%, 0.87%, and 1.20%), suggesting better tolerance vs. spoofing attempts clearly outperforming the sum-rule. However, this comes at a price of clearly degraded initial performance. Further, results indicated that the underlying distributions have a huge impact on the performance of the median rule (reported results refer to min–max-normalised scores). It is therefore important to consider learning distributions beforehand and employ a proper normalisation method. If median rule is extended to median filtering, the low 0% EER of 0-spoofs can be retained and still a better spoofing resistance than sum rule is observed. The product rule performed slightly worse than sum rule with 1–5 spoof EERs of 3.35–15.72%.

As an interesting side-aspect of the evaluation conducted in this paper, the effect of face-only vs. fingerprint-only spoofing is investigated. If always a specific modality is spoofed, a clear discrepancy of spoofing success can be observed: for 1-spoof spoofing restricted to spoofing the face image only (without spoofing detection) 7.33% EER is obtained (using sum rule), whereas fingerprint-only spoofing results in 1.19% EER. This underlines the difficulty of finding suitable counter-spoofing fusion methods being able to tackle these different shifts in distributions. Note that however the quality of the employed

Table 3
EER/SEER (in %) and *d*-Prime results of face-and-finger fusion on the test set varying the number m of spoofed samples.

Method	(S)EER						d-Prime					
	$m=0$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=0$	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$
Sum rule	**0**	2.32	5.71	7.52	10.38	12.24	2.91	2.62	2.27	2.11	1.89	1.67
Product rule	**0**	3.35	7.61	9.48	12.64	15.72	6.44	3.40	2.31	2.03	1.69	1.42
Median rule	0.42	0.87	1.20	3.91	6.74	10.03	2.38	2.37	2.31	2.14	1.89	1.68
Median filtering ($\phi=1\sigma$)	0	1.71	3.65	5.22	8.02	10.60	2.66	2.45	2.19	2.06	1.87	1.69
Median filtering ($\phi=0.5\sigma$)	0	1.03	2.04	3.50	6.23	9.53	2.29	2.17	2.01	1.90	1.75	1.69
1-Median filter	0.47	**0.81**	**1.16**	**1.39**	**1.71**	**1.81**	3.18	3.16	3.12	3.11	3.10	3.08

face database was also very challenging and thus impacting on this result. Experiments continue with the setup using random selection of spoofing samples (i.e. 20% probability to spoof face and $4\times 20\%$ probability to spoof fingerprint).

4.7. Median filtering

In order to evaluate the suggested 1-median filtering an experiment using LR on the joint liveness-and-score pairs is conducted, incorporating liveness information into decision. Obtained EER for the method show much more stable results and also ROC curves are much flatter, see Fig. 8 plotted using log-scale. Median-filtering on probabilities of already combined scores and liveness measures is able to retain EERs below 2% over all spoofing attempts, within a narrow band (0.81–1.81%), and at the same time retains a very high 0-spoof accuracy (0.47%). Results refer to using a filter radius of $\phi=3\sigma$, i.e. relative to the standard deviation of fused samples. Also corresponding *d*-Prime clearly illustrate the much better separation of genuine and m-impostor distribution. However, the tradeoff is a slightly degraded initial 0-spoof performance of 0.47% (which however, is much better than previously published results in [10] due to better spoofing detection).

In a further experiment we verified the superiority of the median compared to the mean when looking at simple recognition-score m-spoof imposter distributions. From Table 2 we can see that (1) mean recognition scores decrease with the number of spoofed samples indicating an effect of the spoofing effort; and (2) median consequently delivered better performance suppressing the negative impact of spoofed samples (note the small changes especially for up to 2-spoof in contrast to the mean) confirming theoretical considerations. Finally, we also tested the effect on variance and found that variance is increasing for a larger number of spoofed samples in multibiometric configuration (e.g., $\sigma^2=0.00012$ for 0-spoof, 0.0222 for 3-spoofed and 0.0293 for the 5-spoof case).

Finally, a critical task in setting up parameters for the median filtering is a suitable choice for the filter radius ϕ. The introduction of filter parameter ϕ results in a tradeoff permitting for better choices between the median rule ($\phi\approx 0$), which is better for suppression of spoofing attempts, and the sum rule ($\phi=\infty$), which delivers the best zero-spoof performance. While one method is to employ static values of ϕ, in order to avoid over-fitting with regards to the training set, dynamic selection methods of ϕ are investigated, based on the scattering of input scores (e.g., as a factor of σ being the standard deviation of the n scores/tuples to be combined). As can be seen from filter evaluations on score-only combinations in Table 3, small values of ϕ (0.5) deliver a closer performance to median filtering, whereas larger values of ϕ are closer to the performance of the sum rule. This way an arbitrary compromise between classical accuracy using the sum rule and potentially slightly degraded 0-spoof performance but higher spoofing resistance, as for the median rule, can be obtained. While the paper does not aim to provide an assessment of computational cost, note that the overhead introduced by median filtering is minimal, as the number n of employed features is typically a fixed and low number.

4.8. Rejection of spoofing samples

Finally, the classical alternative in many implementations implementing counter-spoofing functionality is to reject samples during quality assurance, if they appear to originate from a spoofed source. Unfortunately spoofing detector errors in multibiometric configuration add up and cause a high number of falsely rejected genuine attempts. In order to virtually compare this method with the presented integrated approach, the final score is set to 1, if and only if one or more of the spoofing scores were greater than a threshold ($t=0.44$ is used). The resulting (virtual, as normally this would result in a Failure to Acquire error) EER of approx. 22% clearly illustrates the superiority of integrated recognition-and-spoofing fusion. Further, especially for commercial applications falsely rejected users are considered critical, whereas any threshold is typically set at a very conservative level which limits the use of employed techniques. Compared to techniques integrating liveness results, an advantage is that no information is lost and the overall scores can be taken into account.

5. Conclusion

Experiments in this paper show that 1-spoofing in face and fingerprint fusion can successfully be targeted by employing median instead of sum rule for combinations using its property to be less affected by a certain number of outliers. However, this comes at the cost of a reduced 0-spoof performance. Its extension to 1-median filtering is able to find arbitrary trade-off points between sum and median rule, allowing for better flexibility in choosing the right tradeoff between accuracy and security. The paper investigated how spoofing detection and recognition can mutually benefit from each other and evaluated 1-median filtering as a novel multibiometric fusion method integrating liveness and recognition scores. Results yielded more stable results for this method in partial multibiometric spoofing configuration, where m out of n samples of an identity are spoofed (EERs 0.47–1.81% vs. 0–12.24% for the sum rule). The paper presented an analysis of the filter radius in median filtering and investigated the impact of face vs. fingerprint spoofing. Finally, bootstrap aggregating (bagging) classifiers were proposed for anti-spoofing and shown to deliver highly accurate results (84% accuracy) on the challenging LivDet2013 crossmatch dataset. We believe the following remaining questions should be further investigated in future work: a closer investigation of score normalisation issues for median filtering; an extension towards user-adaptive anti-spoofing and recognition fusion, and; integration of extrinsic factors (e.g. acquisition conditions) and/or quality-related measurements into the fusion scheme.

Conflict of interest

None declared.

Acknowledgements

This work was supported by the EU FASTPASS project under Grant agreement 312583.

References

[1] A. Anjos, J. Komulainen, S. Marcel, A. Hadid, M. Pietikainen, Face anti-spoofing: visual approach, in: S. Marcel, et al., (Eds.), Handbook of Biometric Anti-Spoofing, Springer, London, 2014, pp. 65–82.
[2] E. Marasco, A. Ross, A survey on antispoofing schemes for fingerprint recognition systems, ACM Comput. Surv. 47 (2) (2014) 28:1–28:36.
[3] M. Chakka, et al., Competition on counter measures to 2-d facial spoofing attacks, in: Proceedings of the International Joint Conference on Biometrics, 2011, http://dx.doi.org/10.1109/IJCB.2011.6117509.
[4] L. Ghiani, et al., Livdet 2013 fingerprint liveness detection competition, in: Proceedings of the International Conference on Biometrics, 2013, http://dx.doi.org/10.1109/ICB.2013.6613027.
[5] E. Marasco, P. Johnson, C. Sansone, S. Schuckers, Increase the security of multibiometric systems by incorporating a spoofing detection algorithm in the fusion mechanism, in: C. Sansone, et al. (Eds.), Proceedings of the International Workshop on MCS, Lecture Notes in Computer Science, vol. 6713, Springer, Berlin, 2011, pp. 309–318, http://dx.doi.org/10.1007/978-3-642-21557-5_33.
[6] I. Chingovska, A. Anjos, S. Marcel, Anti-spoofing in action: Joint operation with a verification system, in: Proceedings of the International Conference on Computer Vision and Pattern Recognition, Workshop, 2013, pp. 98–104, http://dx.doi.org/10.1109/CVPRW.2013.22.
[7] N. Poh, R. Wong, G.-L. Marcialis, Toward an attack-sensitive tamper-resistant biometric recognition with a symmetric matcher: a fingerprint case study, in: Proceedings of the Symposium on Computational Intelligence in Biometrics and Identity Management, 2014, pp. 1–6.
[8] R.N. Rodrigues, L.L. Ling, V. Govindaraju, Robustness of multimodal biometric fusion methods against spoof attacks, J. Vis. Lang. Comput. 20 (3) (2009) 169–179.
[9] Z. Akhtar, G. Fumera, G. Marcialis, F. Roli, Evaluation of multimodal biometric score fusion rules under spoof attacks, in: Proceedings of the International Conference on Biometrics, 2012, pp. 402–407, http://dx.doi.org/10.1109/ICB.2012.6199784.
[10] P. Wild, P. Radu, L. Chen, J. Ferryman, Towards anomaly detection for increased security in multibiometric systems: spoofing-resistant 1-median fusion eliminating outliers, in: Proceedings of the International Joint Conference on Biometrics, 2014, http://dx.doi.org/10.1109/BTAS.2014.6996293.
[11] A. Rattani, N. Poh, A. Ross, A Bayesian approach for modeling sensor influence on quality, liveness and match score values in fingerprint verification, in: Proceedings of the International Workshop on Information Forensics and Security, 2013, pp. 37–42, http://dx.doi.org/10.1109/WIFS.2013.6707791.
[12] A. Rattani, A. Ross, Automatic adaptation of fingerprint liveness detector to new spoof materials, in: Proceedings of the International Joint Conference on Biometrics, 2014, http://dx.doi.org/10.1109/BTAS.2014.6996254.
[13] G. Fumera, G. Marcialis, B. Biggio, F. Roli, S. Schuckers, Multimodal anti-spoofing in biometric recognition systems, in: S. Marcel, et al., (Eds.), Handbook of Biometric Anti-Spoofing, Springer, London, 2014, pp. 165–184.
[14] I. Chingovska, A. Rabello dos Anjos, S. Marcel, Biometrics evaluation under spoofing attacks, IEEE Trans. Inf. Forensics Secur. 9 (12) (2014) 2264–2276. http://dx.doi.org/10.1109/TIFS.2014.2349158.
[15] E. Marasco, Y. Ding, A. Ross, Combining match scores with liveness values in a fingerprint verification system, in: Proceedings of the International Conference on Biometrics: Theory, Applications and Systems, 2012, pp. 418–425, http://dx.doi.org/10.1109/BTAS.2012.6374609.
[16] J. Kittler, M. Hatef, R.P. Duin, J. Matas, On combining classifiers, IEEE Trans. Pattern Anal. Mach. Int. 20 (3) (1998) 226–239. http://dx.doi.org/10.1109/34.667881.
[17] R. Rodrigues, N. Kamat, V. Govindaraju, Evaluation of biometric spoofing in a multimodal system, in: Proceedings of the International Conference on Biometrics: Theory, Applications and Systems (BTAS), 2010, pp. 1–5, http://dx.doi.org/10.1109/BTAS.2010.5634531.
[18] Z. Akhtar, G. Fumera, G. Marcialis, F. Roli, Evaluation of serial and parallel multibiometric systems under spoofing attacks, in: Proceedings of the International Conference on Biometrics: Theory, Applications and Systems, 2012, pp. 283–288, http://dx.doi.org/10.1109/BTAS.2012.6374590.
[19] A. Abhyankar, S. Schuckers, Fingerprint liveness detection using local ridge frequencies and multiresolution texture analysis techniques, in: Proceedings of the International Conference on Image Processing, 2006, pp. 321–324, http://dx.doi.org/10.1109/ICIP.2006.313158.
[20] P. Coli, G. Marcialis, F. Roli, Power spectrum-based fingerprint vitality detection, in: Proceedings of the Workshop on Automatic Identification Advanced Technologies, 2007, pp. 169–173. http://dx.doi.org/10.1109/AUTOID.2007.380614.
[21] S. Nikam, S. Agarwal, Fingerprint liveness detection using curvelet energy and co-occurrence signatures, in: Proceedings of the International Conference on Computer Graphics, Imaging and Visualisation, 2008, pp. 217–222, http://dx.doi.org/10.1109/CGIV.2008.9.
[22] L. Ghiani, G. Marcialis, F. Roli, Fingerprint liveness detection by local phase quantization, in: Proceedings of the International Conference on Pattern Recognition, 2012, pp. 537–540.
[23] A. Anjos, M.M. Chakka, S. Marcel, Motion-based counter-measures to photo attacks in face recognition, IET Biometrics 3 (3) (2013) 147–158.
[24] G. Zhao, M. Pietikainen, Dynamic texture recognition using local binary patterns with an application to facial expressions, IEEE Trans. Pattern Anal. Mach. Int. 29 (6) (2007) 915–928. http://dx.doi.org/10.1109/TPAMI.2007.1110.
[25] X. Tan, Y. Li, J. Liu, L. Jiang, Face liveness detection from a single image with sparse low rank bilinear discriminative model, in: K. Daniilidis, et al. (Eds.), ECCV 2010, Lecture Notes in Computer Science, vol. 6316, 2010, pp. 504–517.
[26] T. de Freitas Pereira, A. Anjos, J.M. de Martino, S. Marcel, LBP-TOP based countermeasure against face spoofing attacks, in: Proceedings of the Workshop on Computer Vision, Springer, 2013, pp. 121–132, http://dx.doi.org/10.1007/978-3-642-37410-4_11.
[27] P.T. von Hippel, Mean, median, and skew: correcting a textbook rule, J. Stat. Educ. 13 (2) (2005) 1–13.
[28] D. Moore, G. McCabe, Introduction to the Practice of Statistics, Freeman, New York, 2003.
[29] N. Otsu, A threshold selection method from gray-level histograms, IEEE Trans. Syst. Man Cybern. 9 (1) (1979) 62–66.
[30] J. Daugman, Complete discrete 2-d Gabor transforms by neural networks for image analysis and compression, IEEE Trans. Acoust. Speech Signal Process. 36 (7) (1988) 1169–1179. http://dx.doi.org/10.1109/29.1644.
[31] S.B. Nikam, S. Agarwal, Gabor filter-based fingerprint anti-spoofing, in: J. Blanc-Talon, et al. (Eds.), Proceedings of the International Conference on ACIVS, Lecture Notes in Computer Science, vol. 5259, Springer, Berlin, 2008, pp. 1103–1114, http://dx.doi.org/10.1007/978-3-540-88458-3_100.
[32] H. Choi, R. Kang, K. Choi, A. Jin, J. Kim, Fake-fingerprint detection using multiple static features, Opt. Eng. 48 (2009) 047202–047202–13.
[33] L. Breiman, Bagging predictors, Mach. Learn. 24 (2) (1996) 123–140. http://dx.doi.org/10.1023/A:1018054314350.
[34] B. Efron, R.J. Tibshirani, An Introduction to the Bootstrap, Chapman & Hall, New York, 1994.
[35] N. Poh, A. Merati, J. Kittler, Heterogeneous information fusion: a novel fusion paradigm for biometric systems, in: Proceedings of the International Joint Conference on Biometrics, 2011, pp. 1–8, http://dx.doi.org/10.1109/IJCB.2011.6117494.
[36] P. Johnson, B. Tan, S. Schuckers, Multimodal fusion vulnerability to non-zero effort (spoof) imposters, in: Proceedings of the International Workshop on Information Forensics and Security, 2010, pp. 1–5, http://dx.doi.org/10.1109/WIFS.2010.5711469.
[37] I. Chingovska, A. Anjos, S. Marcel, On the effectiveness of local binary patterns in face anti-spoofing, in: Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG), 2012, pp. 1–7.
[38] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, S. Li, A face antispoofing database with diverse attacks, in: Proceedings of the International Conference on Biometrics, 2012, pp. 26–31, http://dx.doi.org/10.1109/ICB.2012.6199754.
[39] NIST, Biometric Image Software V4.2, 2013, ⟨http://www.nist.gov/itl/iad/ig/nbis.cfm⟩.
[40] Neurotechnology Inc, VeriLook SDK, 2014, ⟨http://www.neurotechnology.com/verilook.html⟩.

Peter Wild received his Ph.D. degree in computer science with “sub-auspiciis-praesidentis” honours (2013) from the University of Salzburg, where he held positions as Research Assistant (2007–2012) and Postdoc (2013). Currently, he is Postdoc in computer vision at the University of Reading. His research was awarded the European Biometrics Industry Award 2013.

Petru Radu is a Research Associate at the University of Reading, UK. In 2013, he obtained his Ph.D. in iris recognition at the University of Kent, where he continued as a Postdoc in Image Processing and Document Analysis. Petrus research interests include statistical pattern recognition, image processing and multiple classifier systems.

Lulu Chen received her Ph.D. degree in computer science from the University of Reading in 2013, and currently works as a Postdoc in the University of Reading. Her research interests include computer vision, image processing, biometric recognition for automated border control, biometric counter-spoofing and human action recognition with sensor fusion.

James M. Ferryman (M96) received his Ph.D. degree in computational vision in 2000 from the University of Reading and became a Full Professor in 2014. Prof. Ferryman leads the Computational Vision Group. His current research interests include biometrics for border control, multi-sensor surveillance, behavioral analysis, cognitive systems, and performance evaluation.