Hindawi Publishing Corporation
Journal of Electrical and Computer Engineering
Volume 2016, Article ID 6390741, 9 pages
http://dx.doi.org/10.1155/2016/6390741

Research Article

Improving Robustness of Biometric Identity Determination with Digital Watermarking

Juha Partala, Angelos Fylakis, Anu Pramila, Anja Keskinarkaus, and Tapio Seppänen

Physiological Signal Analysis Team, Center for Machine Vision and Signal Analysis, University of Oulu, Oulu, Finland

Correspondence should be addressed to Juha Partala; juha.partala@ee.oulu.fi

Received 17 June 2016; Revised 7 September 2016; Accepted 5 October 2016

Academic Editor: Isao Echizen

The determination of an identity from noisy biometric measurements is a continuing challenge. In many applications, such as identity-based encryption, the identity needs to be known with virtually 100% certainty. The determination of identities with such precision from face images taken under a wide range of natural situations is still an unsolved problem. We propose a digital watermarking based method to aid face recognizers to tackle this problem in applications. In particular, we suggest embedding multiple face dependent watermarks into an image to serve as expert knowledge on the corresponding identities to identity-based schemes. This knowledge could originate, for example, from the tagging of those people on a social network. In our proposal, a single payload consists of a correction vector that can be added to the extracted biometric template to compile a nearly noiseless identity. It also supports the removal of a person from the image. If a particular face is censored, the corresponding identity is also removed. Based on our experiments, our method is robust against JPEG compression, image filtering, and occlusion and enables a reliable determination of an identity without side information.

1. Introduction

An identity is naturally captured using the physical characteristics of a person. Those characteristics can be measured using different biometric traits such as face, iris, voice, or fingerprint that have been successfully used in many security oriented applications such as identification, authentication, and access control. Biometrics has enabled the development of user-friendly security mechanisms and, for certain applications, has enabled us to get rid of cryptographic keys and identifiers and to exchange those with biometric ones. Fuzzy identity-based encryption [1] enables the application of public biometric identifiers for public key encryption without the hassle of certificates. In contrast to traditional public key encryption, where a certificate is needed to ensure the identity of the recipient, identity-based encryption [2] derives public keys directly from the public identity. In fuzzy identity-based encryption, biometric templates can be used with a certain threshold of error tolerance without any secrecy requirements on the biometric features.

In order to be conveniently applicable for fuzzy identity-based encryption, the biometric modality has to be public and easy to acquire by others. Face is a natural public biometric trait that we all use to identify individuals in everyday life. It is also one of the most studied traits regarding biometric identification and authentication. Therefore, it is a natural biometric modality to derive identities for identity-based encryption and other identity-based schemes. However, to achieve good accuracy, face recognition techniques often assume that the faces are aligned and photometrically and geometrically normalized. In addition, occlusion and non-frontal facial poses create significant challenges even to state-of-the-art face recognition algorithms [3]. These changes in the biometric acquisition process cause intrasubject variations that are common for images taken under natural conditions and posted to social networks or public photo-sharing services such as Facebook or Instagram. These variations lead to a large number of errors in the extracted biometric features and render their use in many applications impossible.

In this paper, we suggest a digital watermarking based solution to this problem. While our approach works for any biometric modality, we concentrate on the determination of biometric identities from an image taken under natural conditions. We boost the performance of *image independent*

biometrics, the determination of identity in any condition, with *image dependent* biometrics, features computed from a single image, combined with digital watermarking. Image dependent biometrics are subjective to variations caused by image processing such as filtering and compression. However, since the pose, lighting, and occlusion are fixed for a fixed image, image dependent biometrics are spared from a large class of intrasubject variations. Therefore, we are able to improve the performance of applications that need robust biometrics such as fuzzy identity-based encryption and other privacy related applications.

Our solution is based on the embedding of supplementary information about the identities in an image to the image itself. We apply a print-scan robust embedding algorithm and tie its payload directly to the biometric features of the people in the image. Thus, the biometric features of a person in that image work as a key to the supplementary information. If the amount of errors in biometric feature extraction is greater than the threshold of the application, the watermark can be used in correcting those errors. Note that the biometric features for the key are computed from *that particular image* which means that biometric feature extraction is extremely robust compared to unconstrained situations. In addition, since the key is based on biometrics, the removal of a person from an image also removes his or her identity.

The paper is organized as follows. In Section 2, we present the background on face biometrics and digital watermarking relevant to the paper. Section 3 is devoted to the details of our scheme. The experiments are described in Section 4. Finally, Section 5 provides the discussion and the conclusion.

2. Background

2.1. Face Biometrics. Biometrics refers to the automatic recognition of individuals based on their characteristics [3]. They are often used in verifying the claimed identity of the subject (one-to-one matching) or to identify individuals (one-to-many matching). In a biometric system, raw biometric data is first acquired using a sensor. This raw data can be in the form of an image, audio, or a physiological signal. *Feature extraction* is applied on the raw data to extract an identifying set of features into a *biometric template* that should be unique to each individual. For certain applications, such as fuzzy identity-based encryption and fuzzy extractor [4] based schemes, we are not interested in classification but directly apply the templates.

Automatic computation of 2D face biometrics consists of several subtasks such as face detection, alignment, normalization, and description. Face detection methods attempt to detect and indicate face regions in arbitrary images. Face alignment refers to the geometric normalization of the face region and is often based on eye locations. In addition to alignment, the image is photometrically normalized to remove lighting variations. Finally, face description refers to the process of feature extraction and can be based on principal component analysis [5], linear discriminant analysis [6], local binary patterns (LBP) [7], or deep learning [8].

Acquisition of biometrics (with different sensors for instance) induces variations into the extracted features. These are called intrasubject variations that can be caused, among other things, by differences in sensors, pose and expression changes, illumination, and occlusion [3]. Intrasubject variations lead to errors in the extracted biometric template. These errors in turn affect the true and false acceptance rates (TAR and FAR). According to a test conducted by NIST in 2012, state-of-the-art face recognition algorithms reach TAR of approximately 96% at FAR 0.1% on frontal images in controlled conditions. On a challenging dataset with a larger intraclass variation, such as the People In Photo Albums set [9], the performance is much worse. For example, the deep learning based DeepFace [8] has an overall accuracy of 46.66%. Performance in uncontrolled conditions is too low for applications such as fuzzy identity-based encryption where the identity of the person has to be certain.

2.2. Watermarking. Watermarking is often used in biometric systems to add another layer of security, generally either by directly embedding a biometric template into the host data or by protecting the biometric data with a watermark [10, 11]. Hämmerle-Uhl et al. [11] discuss different applications for biometric watermarking. These include steganographic approaches, multibiometric recognition, two-factor authentication, sample replay prevention, and sensor and sample authentication.

In steganography, biometric data is hidden in arbitrary data for transmitting so that the attacker is unaware that data is being transferred. In multibiometric recognition, biometric data is embedded into the biometric sample. The advantage of such method is that two different modalities can be transmitted at the same time and/or recognition performance is increased. Somewhat similar application is two-factor authentication in which authentication data is put, for example, on a smart card. The smart card can contain, for example, a fingerprint of the person [12] and the fingerprint image is watermarked with the face of the person therefore enabling a second layer of authentication. The idea of sample replay prevention is to prevent the use of sniffed sample data to fool the sensor. The sensor embeds a watermark to the sample image before transmitting it for feature extraction. An attacker tries to remove the watermark in order to be able to use sniffed data for replay attacks or as fake traits and consequently the watermark must be robust. Sensor and sample authentication are very similar except that the attacker aims at inserting a watermark in order to mimic correctly acquired sensor data. This can be prevented with semifragile watermarking [11].

Using the biometrics as a key, however, is a more recent concept. Dutta et al. [13] proposed a method for applying iris biometrics as a key in audio to prove the ownership of a piece of music. They argued that a random pseudorandom sequence is not enough for proving ownership unless the sequence is uniquely mapped to an entity that is logically or physically owned by the claimant. They proceeded to extract features from an iris image and used these features as seed of the audio watermark.

An image can have multiple faces and therefore our method should support multiple watermarking. Sheppard et al. [14] divided the multiple watermarking methods into rewatermarking, segmented watermarking, and composite

watermarking. We are the most interested in a special case of segmented watermarking in which the multiple watermarks are embedded by interleaving the separate watermarks instead of embedding watermarks on top of each other.

3. Suggested Scheme

Our goal is to embed supplementary information into an image to aid in automatic face recognition and identity determination. Since embedding capacity is a precious resource, we want to minimize the size of this supplementary information. Therefore, instead of real valued feature vectors, we apply binary biometric templates. In this section, we describe our proposal in detail. We first give an overview and then proceed to explain our face feature extraction and the generation of the binary biometric templates. Finally, we give the details on embedding and extraction.

3.1. Overview

3.1.1. Embedding. Given an image $\mathcal{I}$ our method first detects the faces inside it. Let the correct identities for those faces and their corresponding binary biometric template vectors $B_1, B_2, \ldots, B_m$ be given. These correct identities are given as expert knowledge, for example, by humans tagging people in those images on a social network. For that particular image $\mathcal{I}$, our scheme computes the face features of the detected faces and converts them into binary biometric template vectors $B'_{\mathcal{I},1}, B'_{\mathcal{I},2}, \ldots, B'_{\mathcal{I},m}$ which we call the *image dependent templates*. Note that these templates might contain a significant amount of noise compared to the true identity templates B_i due to intrasubject variations such as pose changes, make up, or lighting. To aid in any identity determination task that later uses this image $\mathcal{I}$, our scheme computes supplementary information $I_{\mathcal{I},i} = B_i \oplus B'_{\mathcal{I},i}$ for all $i \in \{1, 2, \ldots, m\}$, where $\oplus$ is bitwise addition, and embeds $I_{\mathcal{I},1}, I_{\mathcal{I},2}, \ldots, I_{\mathcal{I},m}$ as a payload into $\mathcal{I}$ and thus creates a watermarked image $\mathcal{I}_w$. Note that $B'_{\mathcal{I},i}$ is effectively used as a key which can be later used to derive the correct template B_i from the supplementary information $I_{\mathcal{I},i}$. In contrast to direct embedding of identities into the image, in our scheme the removal of the face of person i from the image renders the computation of $B'_{\mathcal{I},i}$ impossible and thus prevents the determination of B_i and the identification of that person. The embedding process has been depicted in Figure 1.

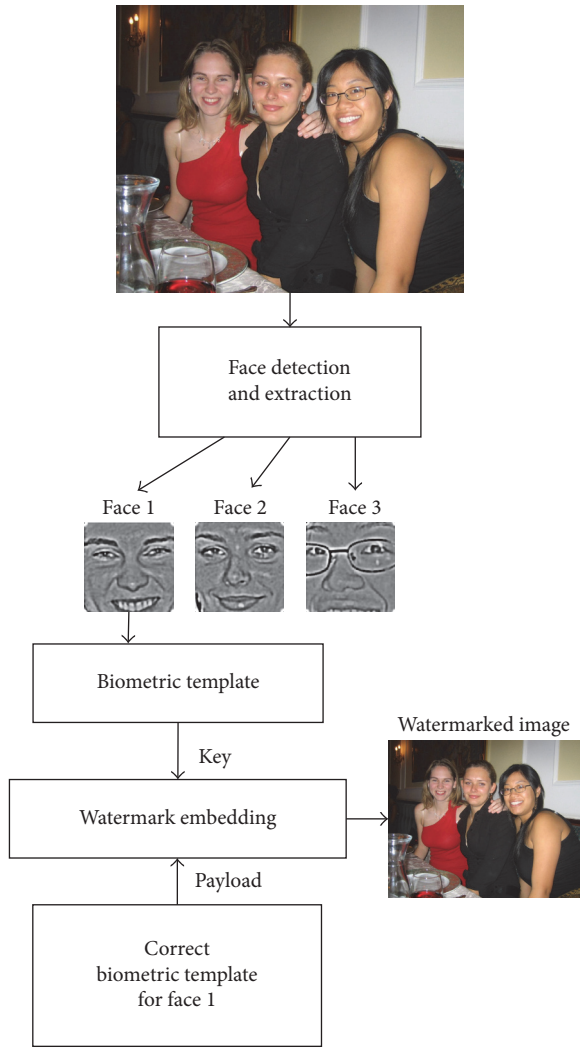

Figure 1: General overview of the embedding process of our proposed scheme. The same process is repeated for all of the faces in the image.

3.1.2. Identity Determination. To determine the identities of people in a watermarked image $\mathcal{I}_w$, faces are first detected. Image dependent templates $\overline{B'_{\mathcal{I},i}}$ are then computed for those faces and the supplementary information $\overline{I_{\mathcal{I},i}}$ extracted for every $i \in \{1, 2, \ldots, m\}$. We have denoted by overlining that errors might have been introduced to both the image dependent features and the supplementary information due to image processing such as compression. That is, $\overline{B'_{\mathcal{I},i}} = B'_{\mathcal{I},i} \oplus e_i$ and $\overline{I_{\mathcal{I},i}} = I_{\mathcal{I},i} \oplus e'_i$ for every $i \in \{1, 2, \ldots, m\}$, where e_i and e'_i are some error vectors. The final template

$$\begin{aligned} \overline{B_i} &= \overline{I_{\mathcal{I},i}} \oplus \overline{B'_{\mathcal{I},i}} = B_i \oplus B'_{\mathcal{I},i} \oplus e'_i \oplus B'_{\mathcal{I},i} \oplus e_i \\ &= B_i \oplus e_i \oplus e'_i \end{aligned} \tag{1}$$

is computed for every face in the image. Compared to the original template B_i, the computed template $\overline{B_i}$ may contain errors e_i due to the extraction of the image dependent template and errors e'_i due to the extraction of the supplementary information. However, being image dependent, $\overline{B'_{\mathcal{I},i}}$ and $B'_{\mathcal{I},i}$ differ only by the amount of errors introduced by image processing variations such as color changes and scaling. Compared to nonrestricted face biometrics, the image dependent features of a person are robust yielding a significant boost to identity determination.

3.2. Face Feature Extraction.

For face biometrics, we apply local binary patterns (LBP). LBP based descriptors are computationally efficient compared with deep learning based methods and suitable even for real-time applications and constrained devices. For details on LBP, its variants, and applications, see, for example, [15]. Face recognition based on LBP was pioneered by Ahonen et al. [7] and we follow

the same methodology. In their work, the (extended) LBP operator computes a label for each pixel by thresholding P sampling points in an evenly spaced circle of radius R. The label is considered as a binary number which means that each pixel is mapped to a bit vector of length P. The pattern of this bit vector is called uniform if there are at most two transitions from 0 to 1 or from 1 to 0 when considering the vector circular. Even in face images, most of the patterns are uniform [7]. Given a face image, it is partitioned into $n \times n$ blocks and a histogram of uniform patterns is computed for each block. This spatially enhanced histogram is used as a face descriptor and classification is based on the Chi squared histogram dissimilarity measure. For details, see [7].

Following the general subtask structure of face recognition, our method consists of face detection, alignment, normalization, and face description. For face detection, we apply the method of Zhu and Ramanan [16] that applies a unified approach to face detection, pose estimation, and landmark point extraction. It reliably estimates head pose and facial landmarks such as eye locations in unconstrained situations and is robust to background clutter. We refer to [16] for details on the algorithm. To extract the face region, the image is first converted into grayscale and rotated based on the eye locations. Next, the image is scaled and a 128×128 pixels' face region is extracted with eyes fixed at coordinates $(27, 37), (101, 37)$. The resulting face region is smoothed with median filtering using 3×3 neighborhood and the lighting is normalized using the method of Tan and Triggs [17].

For face description, we follow Ahonen et al. [7]. The extracted face region is divided into 7×7 rectangular regions and the extended LBP operator is applied on each of those regions separately. Following [7], we chose $R = 2$ and $P = 8$ with uniform patterns giving a total of 49 histograms of length 59 for a single face. These histograms are combined into a single spatially enhanced histogram X (subsequently referred to as a histogram). The process has been depicted in Figure 2.

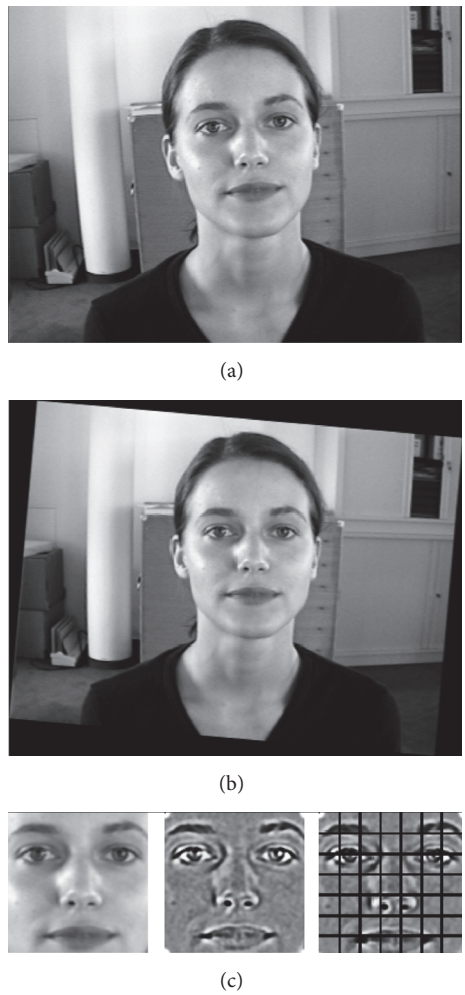

Figure 2: Example image from the BioID face database [18]. The original image (a) is first rotated based on the eye locations (b). The resulting image is scaled, eye locations are fixed, and a 128×128 pixel face region is extracted ((c) left). The lighting is normalized ((c) middle) and the face is divided into 7×7 rectangular regions ((c) right). LBP histograms are computed for each region separately and combined into a spatially enhanced histogram.

3.3. Biometric Template Generation. Based on the spatially enhanced histogram, we generate a binary biometric template for each person using a dictionary of training images. For the training set, we apply the FaceScrub database containing unconstrained face images of 530 celebrities (265 male and 265 female) [19]. A single facial image from each person in the dataset is used. The set of $k = 530$ training images is processed with the feature extraction described above to form a dictionary D of spatially enhanced histograms $D = \{X_1, X_2, \ldots, X_k\}$.

For each pair $X, Y \in D$, weighted Chi square distance

$$\chi^2_W(X, Y) = \sum_{j,i} w_j \frac{\left(x_{i,j} - y_{i,j}\right)^2}{x_{i,j} - y_{i,j}} \tag{2}$$

is computed, where i runs over the histogram bins and the weight terms w_j are determined by the region j in the 7×7 grid. Based on the work of Ahonen et al. [7], the weight terms given by the matrix

$$W = \begin{pmatrix} 2 & 1 & 1 & 1 & 1 & 1 & 2 \\ 2 & 4 & 4 & 1 & 4 & 4 & 2 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix} \tag{3}$$

are chosen. Those terms emphasize the eyebrows, eyes, mouth, and temples. Cheeks and the nose are given zero

weight. For every $X_i \in D$ a bound c_i is computed by taking the median of the set

$$\left\{\chi_W^2(X_i, Y) : Y \in D, Y \neq X_i\right\} \tag{4}$$

for every $i \in \{1, 2, \ldots, n\}$.

To compute an n-bit binary biometric template $B = (b_1, b_2, \ldots, b_n)$ for a given face, we extract its spatially enhanced histogram Z, its distance $\chi_W^2(Z, X_i)$ to a selected set $\{X_1, X_2, \ldots, X_n\} \subseteq D$ of dictionary histograms, and set

$$b_i = \begin{cases} 0, & \chi_W^2(Z, X_i) < c_i, \\ 1, & \chi_W^2(Z, X_i) \geq c_i \end{cases} \tag{5}$$

for every $i \in \{1, 2, \ldots, n\}$. We prune the histogram dictionary for those histograms that produce the most consistent templates for frontal face images. As a second set, we apply the BioID face database (https://www.bioid.com/About/BioID-Face-Database) that consists of 1521 grayscale images of 23 persons captured in realistic settings [18].

To select n histograms, we first generate templates of length k for every image of every person in the test set. For a single person j, each of these templates should be identical. However, due to variations some of the bits have flipped. Let $z_{j,i}$ denote the amount of 0's divided by the amount of 1's for bit position i in the templates of person j. A *consistency vector* $P_j = (p_{j,1}, p_{j,2}, \ldots, p_{j,k})$ and a *bit majority vector* $Q_j = (q_{j,1}, q_{j,2}, \ldots, q_{j,k})$ are computed where

$$p_{j,i} = \max\left(z_{j,i}, 1 - z_{j,i}\right),$$
$$q_{j,i} = \begin{cases} 0, & z_{j,i} \geq 1 - z_{j,i}, \\ 1, & z_{j,i} < 1 - z_{j,i} \end{cases} \tag{6}$$

for every $i \in \{1, 2, \ldots, k\}$. The value $p_{j,i}$ measures the probability of bit flipping on position i for person j. The value $q_{j,i}$ represents the most probable bit in this position.

Let z_i' denote the amount of 0's divided by the amount of 1's in the set $\{q_{j,i} : j \in \{1, 2, \ldots, T\}\}$, where T is the number of persons (for our test set $T = 23$). To pick histograms into the pruned dictionary D', we first compute a *differentiation vector* $R = (r_1, r_2, \ldots, r_k)$, where

$$r_i = \max\left(z_i', 1 - z_i'\right) \tag{7}$$

that measures the amount of variation in bits for position i across persons. In addition, a *total consistency vector* $P = (p_1, p_2, \ldots, p_k)$ is computed by summing the consistency values on bit position i across persons: $p_i = \sum_{j=1}^{T} p_{j,i}/T$. Based on the difference values $R = (r_1, r_2, \ldots, r_k)$, bit indices $i \in \{1, 2, \ldots, k\}$ are sorted into descending order. A total of n dictionary histograms X_i are selected into a pruned dictionary D' based on this ordering provided that the total consistency value p_i is over a predefined bound δ. For our experiments, we chose $\delta = 0.85$. The pruned dictionary D' is used for template generation.

The template length n needs to be chosen based on the capacity of the watermarking scheme. There should be a unique biometric template for every person from the world population so theoretically the threshold of 33 bits has to be exceeded. However, a larger template size provides better resiliency for misidentification. Based on the capacity of our watermarking scheme detailed in Section 3.4, we chose $n = 40$. It gives us enough capacity to host the supplementary information for at least one biometric template even in the smallest images (close to 500 × 500 pixels) found in social networks.

3.4. Watermarking. In our use-case scenario, watermarking is used as a second layer of the biometric feature acquisition process. This layer will provide the supplementary information $I_{\mathcal{F},i}$ embedded into the structure of the image itself. This allows the detection and correction of errors in identity-based schemes. The first requirement of this use-case is to have enough capacity. Specifically, each face in the image corresponds to a payload of 40 bits because of the biometric template size. The second watermarking requirement includes robustness to certain transformations which are common on images that are usually uploaded to social networks. Such transformations include JPEG compression with different compression ratios, color filtering, and possibly removal or modifications of some of the faces. Furthermore, high fidelity is required so that there will be no easily distinguishable artifacts on the images under normal view on a computer monitor. For this requirement, peak signal-to-noise ratio (PSNR) values over 40 dB are generally considered to be acceptable.

The watermarking method used in this scheme is a modification of the multibit watermarking technique proposed by Keskinarkaus et al. [20]. It was designed to be robust to print-scan attacks; that is, the host image can be printed and scanned and the watermark should still be reliably extractable. In the current scheme's case, this satisfies the robustness requirements with some margin so that the strength of the watermark can be lowered to meet the fidelity requirements.

The principal idea is that a message sequence, that is, the payload, is mapped to a directional angle of periodic patterns which are scattered and embedded in using triangular masks placed in permuted locations. The permutations are pseudo-randomly generated. The permuted triangles are grouped in polygon sets. Next, the polygons are grouped in segments which are 40 bit bins, the size of the biometric templates. Each one of the bins is capable of hosting the supplementary biometric information for one face.

Note that the first polygon of permuted triangles is always reserved to host the information about the number of initial faces detected in the image, unlike the original watermarking method where this polygon was reserved for synchronization purposes. This feature enables the proper extraction of information even when some faces have been removed from the image. This is because the number of biometric templates that have to be extracted is always known. Further modification on this method is the adaptability to size. The number of the triangular sequences is proportional to the image size. Thus, bigger images can host a larger number

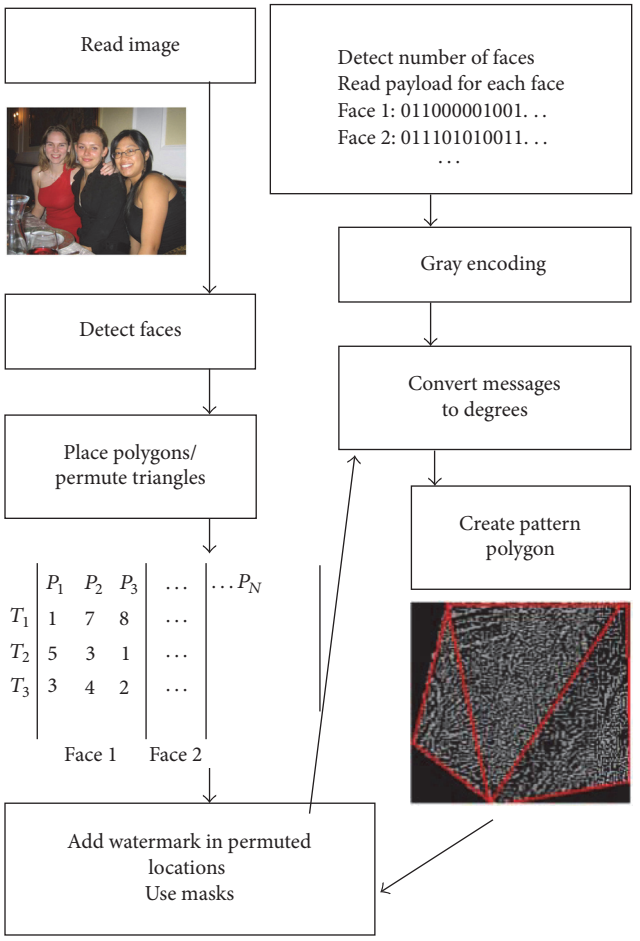

FIGURE 3: Overview of the watermarking process. T_1–T_3 refer to the three triangles acquired using the Delaunay triangulation. P_1–P_N refer to the new polygons being sets of three scattered triangles.

of faces. Compared with the original method, lower watermarking strength settings have been tested to improve fidelity while still maintaining high enough levels of robustness to withstand transformations such as JPEG compression and filter applications which are very common for images that are uploaded on social networks. Also the feature point detection scheme has been removed from the original method [20]. Instead faces can be considered as feature points in order to align an image in its proper position before extraction.

The watermarking technique based on [20] including the modifications to fit this use-case follows the steps below. The process is also depicted in Figure 3.

Embedding Steps

(1) Apply Zhu and Ramanan's face detection method and calculate the number of faces N_F as well as the locations of their feature points in the coordinate system.

(2) Determine the number of convex polygons k that fit to tile the input image $\mathcal{I}$. More information follows in the capacity subsection.

(3) Use Delaunay triangulation to divide each polygon into triangle areas.

(4) Produce a fixed pseudorandom permutation of the triangles that form the polygons. Thus, we create a vector of polygons which are actually sets of 3 scattered triangles. Segment the set vector so that it forms 40 bit bins. That will be 8 polygons. Every face is assigned to one bin. Note that the first polygon is reserved to store the number of detected faces.

(5) Embed the watermarks using the directional patterns in each polygon for each polygon that is now a set of scattered triangles. Use masks to restrict the marked region. Return the watermarked image $\mathcal{I}_w$.

Extracting Steps

(1) Detect faces in input image $\mathcal{I}_w$ using the same method. If the image has been rotated face features can be used in rotating it back to its original angle.

(2) Determine the number of convex polygons k that fit to tile the image.

(3) In the same way divide each polygon into triangle areas.

(4) Use the same seed number to permute the triangles and assign the faces to them.

(5) By detecting the angles of the periodic patterns extract the number of faces from the first polygon and continue by extracting the supplementary biometric information from the 40 bit bins, each one being a set of polygons.

The periodic directional patterns are generated in the same way as described in the original paper [20]. Changes focus on the adaptability of the number of polygons to the image size, the removal of reference points, and the use of number of faces and face locations to allow the embedding of multiple independent watermarks. Last the scaling factor λ is multiplied by a strength varying from 0.4 to 0.6 to reduce the watermarking strength.

Capacity and thus the maximum number of faces of which the information can be embedded in an image is dynamic as it is determined by the image size. In our case since we used a fixed polygon that is circumscribed in a 162×164 rectangle, an image of size $M \times N$ is tiled using $k = \lfloor N/162 \rfloor \times \lfloor M/164 \rfloor$ convex polygons. Because of the fact that each polygon is able to host 5 bits of data and one block is used for storing the initial number of faces, the total capacity is $C = (\lfloor N/162 \rfloor \times \lfloor M/164 \rfloor - 1) \times 5$ bits of data. Each face required $C_F = 40$ bits of space. In a use-case scenario where there are N_F faces there is $C_r = C - N_F C_F$ capacity left which is available to host extra information, for example, error correcting codes or a checksum. To give an example a typical image of size 1280×960 has a capacity of 170 bits and thus it is able to host the biometric templates of 4 faces.

4. Experiments

Both the biometrics and the watermarking algorithm were implemented in Matlab (trademark: The Mathworks Inc.).

Original Watermarked

Figure 4: Comparison of an unwatermarked and a watermarked image with embedding strength of 60%.

Figure 5: The three tested scenarios. An original image, an image with sepia filter applied, and an image with face removed.

Table 1: Mean bit error rate of image dependent template under JPEG compression.

	JPEG 80	JPEG 70	JPEG 60	JPEG 50
Mean BER	0.08	0.08	0.14	0.20

4.1. Test Dataset. For the experiments, we chose 10 images from the Annotated Face in-the-Wild (AFW) dataset [16]. The dataset contains 205 images with a total of 468 faces with large variations in face orientation, appearance, and background clutter. Resolutions range from 1024×768 up to 2606×1733. To capture a wide range of situations, we chose images containing one to four faces in different viewpoints. No annotations were used. Face landmark points were determined by the algorithm of Zhu and Ramanan [16] as a part of our scheme.

4.2. Image Dependent Templates under JPEG Compression. First, the image dependent biometric template performance was measured under JPEG compression. Four JPEG quality factors were evaluated: 80, 70, 60, and 50. First, a reference template was computed for each face in the test set. Then, the image was compressed and the template was recomputed and compared to the reference template. The mean bit error rate (BER) under compression was computed. The experimental results have been collected into Table 1.

4.3. Watermarking Quality Assessment and Performance. To evaluate the combination of biometrics and watermarking, for each individual j, a random identity template B_j of 40 bits representing the true errorless template was generated. To evaluate robustness against JPEG compression, we tested quality factors 100, 90, and 80. Three watermark embedding strengths were evaluated: 40%, 50%, and 60%. An example is depicted in Figure 4. For a single image, all combinations of these experiments were performed resulting in a total of 252 tests. Based on these tests, the mean PSNR and the mean BER of the extracted templates were evaluated.

To measure the robustness to image filtering, as demonstrated in Figure 5, we tested the effect of sepia filtering which is a common image effect. In addition, to measure the integrity of multiple identities in the image, we tested the case of face removal. In this test, one of the faces was selected and the face region determined by the method of Zhu and Ramanan [16] was rendered completely black. The above experiments were repeated for both of these modified images resulting in two additional sets of mean PSNR and mean BER measurements.

The experimental results have been collected into Table 2, where Orig stands for the original image, Sepia stands for the sepia filtered image, and FR stands for the image where a single face has been removed. Finally the parentheses (p%) denote the embedding strength of the watermark.

5. Discussion and Conclusion

The exact determination of an identity is hard in challenging conditions even if we have millions of training samples available [3]. Our suggestion provides a method of improving

Table 2: The mean bit error rate of the final template and the mean peak signal-to-noise ratio [dB] of the watermarked image for different JPEG quality levels.

	JPEG 100		JPEG 90		JPEG 80	
	BER	PSNR	BER	PSNR	BER	PSNR
Orig (40%)	0.08	45.6	0.06	45.6	0.06	45.6
Sepia (40%)	0.16	45.6	0.13	45.6	0.19	45.6
FR (40%)	0.10	45.6	0.11	45.6	0.26	45.0
Orig (50%)	0.06	43.5	0.07	43.5	0.07	43.5
Sepia (50%)	0.13	43.5	0.12	43.5	0.12	43.5
FR (50%)	0.06	43.4	0.10	43.4	0.11	43.4
Orig (60%)	0.08	42.2	0.07	42.2	0.07	42.2
Sepia (60%)	0.17	42.2	0.11	42.2	0.11	42.2
FR (60%)	0.06	42.0	0.08	42.0	0.11	42.0

the performance of identity-based schemes especially in challenging conditions. Our method provides this resilience without side information; merely the watermarked raw data suffices. Furthermore, our method supports manual removal of a person from an image. If the face of a particular person is censored from the image, his or her identity is also removed. Based on local binary patterns, our face biometrics scheme is efficient and suitable for constrained devices. However, our method is not restricted to LBP. Our methodology can be applied with any face recognizer. We believe that the biometric performance can be greatly increased with a state-of-the-art face descriptor such as DeepFace [8] in exchange of computational performance. The classic LBP is not particularly robust to noise [21] which is also seen in our experiments (Table 1). The periodic pattern induced by the watermarking scheme also increases the BER of image dependent templates. Performance could be increased, for example, by masking faces in the watermarking algorithm to prevent the periodic pattern from affecting those regions. There are also LBP variants more robust to such noise [22]. However, based on our experiments, even with the classic LBP, the performance is acceptable for high quality JPEG and embedding strength $\leq 75\%$.

Two datasets were used in the computation of the dictionary for template generation: FaceScrub [19] and BioID [18]. Both sets contain variations in lighting, expression, and gender. In addition, FaceScrub contains a lot of variations in size, compression, and noise. We have not explicitly tested the performance of template generation with regard to individual characteristics such as gender. However, the dictionary generation we applied maximizes the consistency of a template across all of the variations present in the data. Naturally, there are limits to these datasets. They are not able to capture all of the variations encountered in the wild. For instance, we are unsure how the template generation works, for example, for elder people or children since they are largely missing from the training data. To the best of our knowledge, there are no biometric face databases containing in-the-wild variations. Such considerations are left for future work.

By our experiments, our proposal is robust against JPEG compression down to quality level 80. Mean BER of 0.08, 0.07, and 0.07 were reached for JPEG quality levels 100, 90, and 80, respectively, for the original image with embedding strength 60%. Sepia filtering causes an increase in the mean BER. Since this value is not decreased when embedding strength is increased, it seems that the LBP based face descriptor is sensitive to sepia filtering. For robustness against sepia filtering and face removal, embedding strength less than 50% should not be used. In general, face removal does not affect the mean BER for the remaining faces provided that an adequately high embedding strength is used. For all of our tests, the PSNR is high meaning that watermarks were imperceptible.

For extremely short template lengths, the achieved BER may lead to false identification. However, even under sepia filtering the BER ≤ 0.2 is within the applicable range of existing fuzzy identity-based schemes provided that the template length is adequately large. To increase the template length from 40, we suggest choosing a less robust watermarking scheme and to cut down the number of attacks the scheme needs to withstand. For certain applications, zero BER is required. In such a case, we suggest applying fuzzy randomness extractors such as the fuzzy extractor [4] or related schemes [23–25]. These methods apply error correcting codes to bring the BER down to zero in exchange of the final template length. Such errorless determination of the template enables new applications that are not possible when the possibility of false identification is present. For example, traditional errorless identity-based encryption could be used if the template can be extracted without errors.

The payload in our method is derived from the correct identity in the image combined with image dependent features. It is possible for an adversary to exchange the payload of a particular person in the image with his or her chosen payload. To counter such manipulation attacks, the extractor needs to check that the extracted identity corresponds to the one in the image. One possibility is to use the extracted biometric template as a link to a frontal face image of that person taken in good conditions. Tampering of the payload is easy to detect provided that the system shows such a high quality image to the user whenever the identity is requested. However, we note that our proposal does not guarantee cryptographic security against such tampering or cryptographic protection for an identity when a face is removed. Such considerations are left for future work.

The watermarking method we applied is highly robust. It has been shown to perform well even in the challenging print-scan scenario with correctly chosen parameters [20]. The main drawback is the relatively low capacity. In particular, we would want to increase the biometric template size to increase its accuracy and robustness in applications. Capacity increase is naturally possible by sacrificing robustness.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This research was supported by Tekes, the Finnish Funding Agency for Technology and Innovation in VitalSens and INKA projects. This work was also supported in part by Infotech Oulu.

References

[1] A. Sahai and B. Waters, "Fuzzy identity-based encryption," in *Advances in Cryptology—EUROCRYPT 2005*, R. Cramer, Ed., vol. 3494 of *Lecture Notes in Computer Science*, pp. 457–473, Springer, Berlin, Germany, 2005.

[2] A. Shamir, "Identity-based cryptosystems and signature schemes," in *Advances in Cryptology—Proceedings of CRYPTO 84*, G. R. Blakley and D. Chaum, Eds., vol. 196 of *Lecture Notes in Computer Science*, pp. 47–53, Springer, Berlin, Germany, 1985.

[3] A. K. Jain, K. Nandakumar, and A. Ross, "50 years of biometric research: accomplishments, challenges, and opportunities," *Pattern Recognition Letters*, vol. 79, pp. 80–105, 2016.

[4] Y. Dodis, L. Reyzin, and A. Smith, "Fuzzy extractors: how to generate strong keys from biometrics and other noisy data," in *Advances in Cryptology—EUROCRYPT 2004*, C. Cachin and J. L. Camenisch, Eds., vol. 3027 of *Lecture Notes in Computer Science*, pp. 523–540, Springer, Berlin, Germany, 2004.

[5] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[6] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, 1997.

[7] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with local binary patterns: application to face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037–2041, 2006.

[8] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: closing the gap to human-level performance in face verification," in *Proceedings of the 27th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '14)*, pp. 1701–1708, June 2014.

[9] N. Zhang, M. Paluri, Y. Taigman, R. Fergus, and L. Bourdev, "Beyond frontal faces: improving person recognition using multiple cues," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '15)*, pp. 4804–4813, IEEE, Boston, Mass, USA, June 2015.

[10] J. Dong and T. Tan, "Effects of watermarking on iris recognition performance," in *Proceedings of the 10th International Conference on Control, Automation, Robotics and Vision (ICARCV '08)*, pp. 1156–1161, December 2008.

[11] J. Hämmerle-Uhl, K. Raab, and A. Uhl, "Watermarking as a means to enhance biometric systems: a critical survey," in *Information Hiding*, pp. 238–254, Springer, 2011.

[12] A. K. Jain and U. Uludag, "Hiding biometric data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 11, pp. 1494–1498, 2003.

[13] M. K. Dutta, P. Gupta, and V. K. Pathak, "Biometric based unique key generation for authentic audio watermarking," in *Pattern Recognition and Machine Intelligence*, vol. 5909 of *Lecture Notes in Computer Science*, pp. 458–463, Springer, Berlin, Germany, 2009.

[14] N. P. Sheppard, R. Safavi-Naini, and P. Ogunbona, "On multiple watermarking," in *Proceedings of the Workshop on Multimedia and Security: New Challenges*, pp. 3–6, ACM, October 2001.

[15] M. Pietikäinen and G. Zhao, "Two decades of local binary patterns: a survey," in *Advances in Independent Component Analysis and Learning Machines*, E. Bingham, S. Kaski, J. Laaksonen, and J. Lampinen, Eds., pp. 175–210, Elsevier, New York, NY, USA, 2015.

[16] X. Zhu and D. Ramanan, "Face detection, pose estimation, and landmark localization in the wild," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 2879–2886, Providence, RI, USA, June 2012.

[17] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1635–1650, 2010.

[18] O. Jesorsky, K. J. Kirchberg, and R. W. Frischholz, "Robust face detection using the hausdorff distance," in *Proceedings of the International Conference on Audio-and Video-Based Biometric Person Authentication (AVBPA '01)*, pp. 90–95, Springer, Halmstad, Sweden, 2001.

[19] H. W. Ng and S. Winkler, "A data-driven approach to cleaning large face datasets," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '14)*, pp. 343–347, 2014.

[20] A. Keskinarkaus, A. Pramila, and T. Seppänen, "Image watermarking with feature point based synchronization robust to print-scan attack," *Journal of Visual Communication and Image Representation*, vol. 23, no. 3, pp. 507–515, 2012.

[21] J. Chen, V. Kellokumpu, G. Zhao, and M. Pietikäinen, "Rlbp: robust local binary pattern," in *Proceedings of the British Machine Vision Conference (BMVC '13)*, pp. 1–10, Bristol, UK, 2013.

[22] G. Kylberg and I.-M. Sintorn, "Evaluation of noise robustness for local binary pattern descriptors in texture classification," *EURASIP Journal on Image and Video Processing*, vol. 2013, article 17, 2013.

[23] Y. Dodis and Y. Yu, "Overcoming weak expectations," in *Theory of Cryptography*, A. Sahai, Ed., vol. 7785 of *Lecture Notes in Computer Science*, pp. 1–22, Springer, Berlin, Germany, 2013.

[24] G. Cohen, R. Raz, and G. Segev, "Nonmalleable extractors with short seeds and applications to privacy amplification," *SIAM Journal on Computing*, vol. 43, no. 2, pp. 450–476, 2014.

[25] X. Li, "Non-malleable condensers for arbitrary min-entropy, and almost optimal protocols for privacy amplification," in *Theory of Cryptography*, vol. 9014 of *Lecture Notes in Computer Science*, pp. 502–531, Springer, 2015.