PAPER

Phase-Based Periocular Recognition with Texture Enhancement

Luis Rafael MARVAL-PÉREZ[†], *Nonmember*, **Koichi ITO**[†a)], *and* **Takafumi AOKI**[†], *Members*

SUMMARY   Access control and surveillance applications like walking-through security gates and immigration control points have a great demand for convenient and accurate biometric recognition in unconstrained scenarios with low user cooperation. The periocular region, which is a relatively new biometric trait, has been attracting much attention for recognition of an individual in such scenarios. This paper proposes a periocular recognition method that combines Phase-Based Correspondence Matching (PB-CM) with a texture enhancement technique. PB-CM has demonstrated high recognition performance in other biometric traits, e.g., face, palmprint and finger-knuckle-print. However, a major limitation for periocular region is that the performance of PB-CM degrades when the periocular skin has poor texture. We address this problem by applying texture enhancement and found out that variance normalization of texture significantly improves the performance of periocular recognition using PB-CM. Experimental evaluation using three public databases demonstrates the advantage of the proposed method compared with conventional methods.

key words: *periocular recognition, phase-only correlation, phase-based image matching, phase features, texture enhancement, biometrics*

1. Introduction

Reliable authentication of individuals in unconstrained scenarios is increasingly required for applications, such as immigration control, entrance-exit management, surveillance, law enforcement, forensics, etc. [1]. Capturing high-resolution images in less-constrained environments with relaxed cooperation is crucial for user experience in terms of convenience and acceptability [2], and it is relatively easy using the state-of-the-art imaging technology. However, accurate person authentication requires careful consideration in the choice of an adequate biometric trait. In general, iris and face have been used in unconstrained scenarios. Face recognition performance has greatly improved in the last decades [3], and iris recognition has arguably the highest performance in controlled settings [4]. On the other hand, the applicability of face recognition is limited, since face recognition methods have to deal with factors such as facial expressions, lighting variations, and occlusions in order to achieve accurate authentication. Iris recognition methods also have to deal with different factors, especially, partial occlusions due to specular reflections and eyelashes, non-frontal gaze, motion blur, and defocus blur. Such impairments degrade recognition performance or sometimes prevent recognition at all.

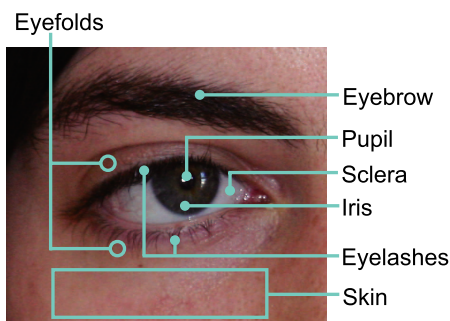Over the last years, the *periocular region* — the ex-

Fig. 1   Discriminative features included in the periocular region, where the image sample is from the UBIPr database.

tended region around the eye — has received considerable attention [5]–[7]. Periocular region includes many discriminative components, such as iris, sclera, skin, eyefolds, eyelashes and eyebrows [6]. These components are indicated in Fig. 1 for an image sample from the UBIPr database [8]. They allow highly accurate recognition comparable with iris recognition in non-controlled scenarios [9]–[11]. In addition, existing sensing setups for face recognition and iris recognition can be used also for periocular recognition.

In order to achieve efficient performance in periocular recognition, highly accurate image matching techniques are indispensable, since images captured from periocular region usually exhibit nonlinear deformation due to variations in head pose and facial expression as well as partial occlusion by eyeglasses, hairs, etc. This paper proposes a periocular recognition algorithm using Phase-Based Correspondence Matching (PB-CM) [12], which has demonstrated efficient performance in face, palmprint and finger knuckle recognition [13]. PB-CM employs local block matching using *phase features* obtained from 2D Discrete Fourier Transform (DFT) of image blocks. It combines phase features with image resolution pyramids to deal with deformation caused by variations in head pose and facial expression. The final matching score is calculated by summarizing results of local block matching in PB-CM, resulting in having the robustness against partial occlusion. The above characteristics of PB-CM are useful also in periocular recognition because of its robustness against partial occlusion and non-linear image deformation.

A major problem of PB-CM is that its performance is significantly degraded when it is applied to regions with poor texture such as the skin under the eye. Addressing this problem, we combine phase-based correspondence match-

ing with a texture enhancement technique to make a highly robust recognition algorithm for periocular images. Experimental evaluation using three public databases demonstrates the efficient performance of the proposed algorithm in periocular recognition compared with conventional algorithms.

Contributions of this paper are summarized as follows: (i) a new periocular image recognition algorithm using phase-based correspondence matching, (ii) a technique for improving its recognition performance through variance normalization, and (iii) systematic experimental evaluation of the algorithms using three public databases.

2. Related Work

Let us start by reviewing a simplified flow diagram in Fig. 2 of a biometric system that performs periocular recognition. The recognition process comprises sensing, periocular region extraction, feature extraction and matching. During sensing, the system captures a raw image which includes the periocular region in the visible or near-infrared spectrum. Once the image is captured, the system normalizes the periocular region using eye detection [9], [14]–[16] or eye corners detection [8]. Normalization involves scaling, rotation and cropping of the captured image. From the normalized periocular image, the system extracts discriminative feature vectors. Then, the similarity between a probe image and a registered image is evaluated by comparing these features.

Previous works on periocular recognition applied traditional features used in biometric recognition, especially, in face recognition. Examples of works employ Histograms of Oriented Gradients (HOGs) [17]–[19], Local Binary Patterns (LBPs) [8], [17], [19]–[22], Principal Component Analysis (PCA) [22] and Scale-Invariant Feature Transform (SIFT) [8], [17], [18], etc. The recognition methods using these features are relatively robust against imperfect alignment and changes in facial expression. However, their performance is limited since they do not fully exploit the texture information within the periocular region.

Among recent periocular recognition methods, we find those based on Convolutional Neural Networks (CNNs) [23]–[25] and those based on a correlation filter known as Fusion Optimal-Trade-off Synthetic Discriminant Function (FOTSDF) [18], [26]–[28]. For comparison purpose, we put our attention on two of these methods: Semantic-Assisted CNN (SCNN) [23] and Periocular Probabilistic Deformation Model (PPDM) [27]. One major disadvantage of these state-of-the-art methods is that they require training data. CNN-based approaches require large training datasets for optimization which are not usually available, and FOTSDF-based approaches rely on image samples of target users for training and parameter selection.

In this paper, we propose Phase-Based Correspondence Matching (PB-CM) with texture enhancement for periocular recognition. PB-CM handles various nonlinear transformation by comparing local blocks at their corresponding location in a similar way to SIFT feature matching. Com-
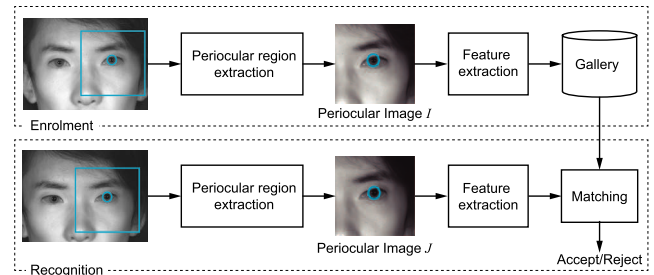

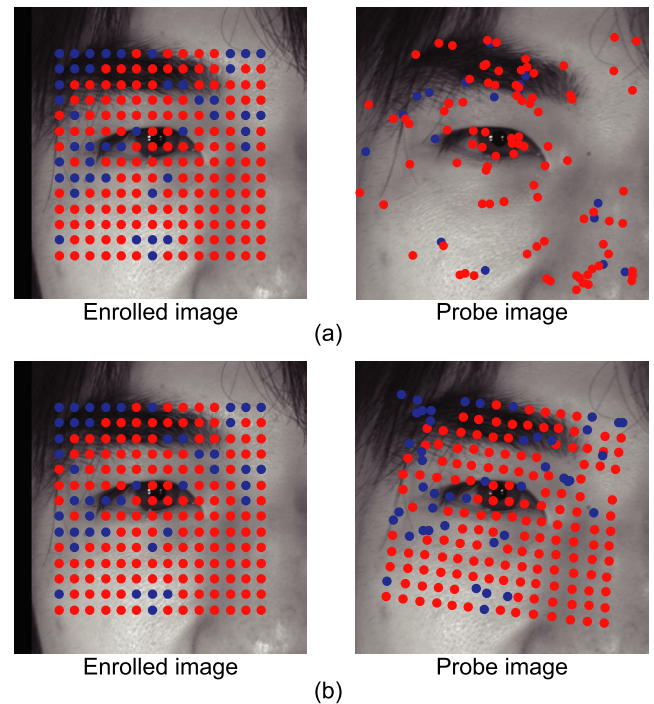**Fig. 2**   Flow diagram of a periocular recognition system.

Fig. 3   Correspondence matching examples for two periocular images of the same person (genuine pair) using m-SIFT features matching (a) and using PB-CM (b). Red dots indicate corresponding locations successfully estimated, and blue dots indicate failed estimations.

paring with SIFT feature matching (Fig. 3(a)), PB-CM utilizes precise corresponding locations for accurate similarity evaluation (Fig. 3(b)). PB-CM employs phase features which have shown to be effective for representing various biometric texture, such as fingerprint, iris, face, palmprint and finger-knuckle-print [29]. However, the skin below the eyes has usually a weak texture and does not yield the same recognition performance as other parts of the periocular region. We found out that we can enhance the skin texture with variance normalization and improve the discriminative capacity of the phase features in periocular recognition. In this manner, the novel combination of PB-CM with texture enhancement makes an intensive use of the available periocular texture. This allows our method to compete favorably against advanced periocular recognition approaches such as FOTSDF-based [27] and CNN-based [23] methods.

3. Fundamentals for Phase-Based Periocular Recognition

Let us formalize the image matching problem. The problem is to accurately measure the similarity between two periocular images, a given enrolled image I (registered in the gallery) and a probe image J. In a real situation, the periocular regions in these images are not perfectly normalized, and may contain occlusions and global deformation, which must be addressed in the image matching task. The block-wise image matching, which compares multiple local image blocks extracted from the given pair of images, could handle these problems, since occlusions and deformation are less pronounced at the scale of small blocks. For these comparisons to be accurate, it is necessary for the local blocks to be aligned, i.e., to be at their corresponding locations between the two images, like the locations indicated by red dots in Fig. 3(b). We use Phase-Only Correlation (POC) to find the corresponding locations and to measure the collective similarity between the groups of blocks.

We first present fundamentals of POC for robust image matching. Second, we introduce a texture enhancement technique, which significantly improves matching performance for periocular images. Third, we present a technique for phase-base correspondence search and possible similarity measures for periocular recognition. We leave the overall design of the periocular recognition algorithm to the next section.

3.1 Phase-Based Image Matching

Let us consider two $N_1 \times N_2$ local blocks, $f(n_1, n_2)$ and $g(n_1, n_2)$, extracted from the image I with the center coordinate $\boldsymbol{p}$ and from the image J with the center coordinate $\boldsymbol{q}$, respectively. The ranges of image coordinates are given by $n_1 = -N_1^-, \cdots, N_1^+$ and $n_2 = -N_2^-, \cdots, N_2^+$, where $N_1^- = \lceil (N_1 - 1)/2 \rceil$, $N_1^+ = \lfloor (N_1 - 1)/2 \rfloor$, $N_2^- = \lceil (N_2 - 1)/2 \rceil$ and $N_2^+ = \lfloor (N_2 - 1)/2 \rfloor$. In order to extract the local phase features, we calculate the 2D DFTs of the blocks with the following equations:

$$
\begin{aligned}
F(k_1, k_2) &= \sum_{n_1, n_2} \left\{ f(n_1, n_2) - f^{DC} \right\} w(n_1, n_2) \\
&\quad \times W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} \\
&= A_F(k_1, k_2) e^{j \theta_F (k_1, k_2)}, \quad (1) \\
G(k_1, k_2) &= \sum_{n_1, n_2} \left\{ g(n_1, n_2) - g^{DC} \right\} w(n_1, n_2) \\
&\quad \times W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} \\
&= A_G(k_1, k_2) e^{j \theta_G (k_1, k_2)}, \quad (2)
\end{aligned}
$$

where $\sum_{n_1, n_2}$ denotes $\sum_{n_1=-N_1^-}^{N_1^+} \sum_{n_2=-N_2^-}^{N_2^+}$, the twiddle factors are given by $W_{N_1} = e^{-j \frac{2\pi}{N_1}}$ and $W_{N_2} = e^{-j \frac{2\pi}{N_2}}$, and the ranges of the transformed blocks are $k_1 = -N_1^-, \cdots, N_1^+$, $k_2 = -N_2^-, \cdots, N_2^+$. $A_F(k_1, k_2)$ and $A_G(k_1, k_2)$ are the amplitude

components of the two image blocks, and $\theta_F (k_1, k_2)$ and $\theta_G(k_1, k_2)$ are the phase components.

We modify the original POC-based matching [31], [32] as follows. In order to address the wrap around effect of the 2D DFT, we apply a window function $w(n_1, n_2)$, i.e., Hanning window, to Eqs. (1) and (2) in our case. As observed above, we subtract the DC components of $f(n_1, n_2)$ and $g(n_1, n_2)$, which are denoted by f^{DC} and g^{DC}, since they do not contribute to the texture description.

We define the local phase features at the coordinates $\boldsymbol{p}$ and $\boldsymbol{q}$ as

$$
X(k_1, k_2) = \frac{F(k_1, k_2)}{|F(k_1, k_2)|}, \quad (3)
$$

$$
Y(k_1, k_2) = \frac{G(k_1, k_2)}{|G(k_1, k_2)|}. \quad (4)
$$

The normalized cross-power spectrum $R(k_1, k_2)$ is computed from the phase features $X(k_1, k_2)$ and $Y(k_1, k_2)$ as

$$
\begin{aligned}
R(k_1, k_2) &= \overline{X(k_1, k_2)} Y(k_1, k_2) \\
&= e^{j \{ \theta_G (k_1, k_2) - \theta_F (k_1, k_2) \}}, \quad (5)
\end{aligned}
$$

where $\overline{X(k_1, k_2)}$ is the complex conjugate of $X(k_1, k_2)$. Then, the POC function $r^{POC}(n_1, n_2)$ is defined as the 2D Inverse DFT (2D IDFT) of $R(k_1, k_2)$:

$$
\begin{aligned}
r^{POC}(n_1, n_2) &= \frac{1}{N_1 N_2} \sum_{k_1, k_2} R(k_1, k_2) \\
&\quad \times W_{N_1}^{-k_1 n_1} W_{N_2}^{-k_2 n_2}, \quad (6)
\end{aligned}
$$

where $\sum_{k_1, k_2}$ denotes $\sum_{k_1=-N_1^-}^{N_1^+} \sum_{k_2=-N_2^-}^{N_2^+}$.

The POC function $r^{POC}(n_1, n_2)$ exhibits a sharp peak when the two image blocks are similar. The location of this correlation peak indicates the translational displacements $\boldsymbol{\delta} = (\delta_1, \delta_2)$ between the two image blocks, while the peak height α serves as a measure of their similarity [30]. In this way, we can compare two image blocks by computing their POC function. However, the correlation peak of the POC function can be very noisy due to the low signal-to-noise ratio in the high-frequency domain. In computer vision applications, this problem has been addressed with a spectral weighting function to decrease the influence of high frequency components [31], while in biometrics a simpler strategy to omit the high frequency components is adopted. This approach is called Band-Limited POC (BLPOC) [32], since it limits the frequency band to a smaller size $K_1 \times K_2$ ($K_1 < N_1, K_2 < N_2$) in the 2D IDFT computation of Eq. (6). More precisely, the frequency range is restricted as $k_1 = -K_1^-, \cdots, K_1^+$ and $k_2 = -K_2^-, \cdots, K_2^+$, where $K_1^- = \lceil (K_1 - 1)/2 \rceil$, $K_1^+ = \lfloor (K_1 - 1)/2 \rfloor$, $K_2^- = \lceil (K_2 - 1)/2 \rceil$ and $K_2^+ = \lfloor (K_2 - 1)/2 \rfloor$. Then, the BLPOC function $r(n_1, n_2)$ is defined by rewriting Eq. (6) as

$$
r(n_1, n_2) = \frac{1}{N_1 N_2} \sum_{k_1, k_2}{}' R(k_1, k_2) W_{N_1}^{-k_1 n_1} W_{N_2}^{-k_2 n_2}, \quad (7)
$$

where $\sum'_{k_1,k_2}$ denotes $\sum_{k_1=-K_1^-}^{K_1^+}\sum_{k_2=-K_2^-}^{K_2^+}$. The peak location of the BLPOC function $\left(\delta_1^{\mathrm{BL}}, \delta_2^{\mathrm{BL}}\right)$ indicates the translational displacement $\boldsymbol{\delta} = (\delta_1, \delta_2)$ between the image blocks such that $\delta_1 = \delta_1^{\mathrm{BL}} N_1/K_1$ and $\delta_2 = \delta_2^{\mathrm{BL}} N_2/K_2$. The BLPOC function is particularly effective for matching a variety of biological textures and is useful for biometric authentication [29].

We refer to the phase features $X(k_1, k_2)$ and $Y(k_1, k_2)$ as band-limited functions. That is to say, they are defined for $k_1 = -K_1^-, \cdots, K_1^+$ and $k_2 = -K_2^-, \cdots, K_2^+$. Note that $r(n_1, n_2)$ is also band-limited, and hence it is defined for $n_1 = -K_1^-, \cdots, K_1^+$ and $n_2 = -K_2^-, \cdots, K_2^+$.

3.2 Texture Enhancement

Phase-based image matching has shown to be effective for recognizing popular biometric traits such as fingerprint [32], iris [33], palmprint [12], [34], and finger knuckle [35], [36], which contain homogeneous texture components. When we apply phase-based image matching to periocular recognition, however, we have to deal with heterogeneous texture components contained in periocular images. For example, while the appearance of eyebrows have very distinctive textures, the skin around the eyes might have poor texture, which deteriorates the recognition performance. It is necessary to enhance the periocular images so as to make the poor texture more visible and make the whole image more homogeneous. Therefore, we employ variance normalization of texture, also known as contrast normalization [37], which adjusts the local variance of pixel intensity across the whole image to the same level. The phase-only image, which, for example, is obtained from IDFT of X in Eq. (3), corresponds to the image having an equalized variance of pixels. This fact indicates that phase features in the frequency domain corresponds to the image in the spatial domain after variance normalization. Therefore, the important information of phase features can be enhanced by applying variance normalization in the spatial domain. Variance normalization consists of three steps as illustrated in Fig. 4.

(i) Extract local patches: Extract a small (e.g. 7×7 pixels) local patch from the original image for every pixel.

(ii) Normalize patches: For all the patches, eliminate their DC components and normalize their variance. A given patch $v(n_1, n_2)$ is normalized as follows:

$$\hat{v}(n_1, n_2) = \frac{v(n_1, n_2) - v^{DC}}{\max\left(\eta, \sqrt{\sum_{\{n_1', n_2'\}\in C}\{v(n_1', n_2') - v^{DC}\}^2}\right)},$$

(8)

where C is a set of patch coordinates, η is a small constant to cancel noise and v^{DC} is the DC component which is given by

$$v^{DC} = \sum_{\{n_1, n_2\}\in C}\frac{v(n_1, n_2)}{|C|}.$$
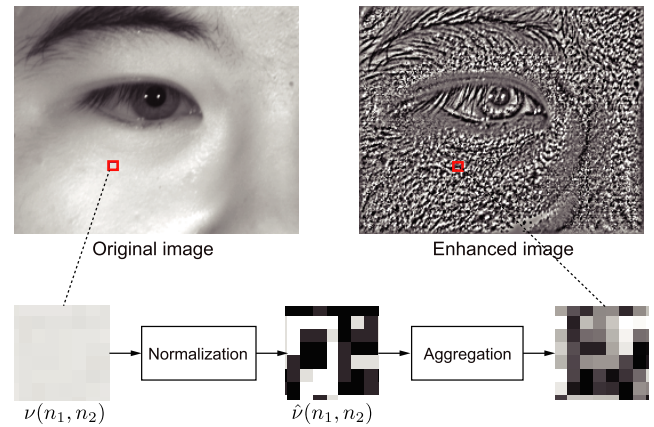
(9)

Fig. 4 Texture enhancement with variance normalization, which consists of three steps: (i) extract one patch per pixel, (ii) normalize the variance of each patch and (iii) combine them into a new image.

(iii) Aggregate normalized patches: Construct a new image by aggregating the patches using their original pixel coordinates. The intensity of each pixel in the new image is given by adding overlapping neighbor patches.

The general formulation of the steps (i) and (ii) is called Divisive Normalization Transform (DNT), and it is effective to reduce the statistical dependencies [38]. DNT has been applied to image processing tasks such as image compression [39] and contrast enhancement [40] among others.

In order to illustrate the effect of texture enhancement on the BLPOC function, Fig. 5 depicts an example of BLPOC matching without texture enhancement and another example with texture enhancement. Both examples show the image blocks $f(n_1, n_2)$ and $g(n_1, n_2)$, the phase features $X(k_1, k_2)$ and $Y(k_1, k_2)$ and the respective BLPOC functions $r(n_1, n_2)$. The BLPOC function without enhancement exhibits multiple peaks, and the location of the highest peak does not correspond to the correct translational displacement between the two image blocks. On the other hand, the BLPOC function with texture enhancement produces a single peak, which location indicates the correct translational displacement between the two image blocks.

3.3 Phase-Based Correspondence Search

As mentioned above, accurate authentication in biometrics cannot be performed unless dealing with various transformation between the images, for example, those caused by variations in facial expression and head pose. To address the global image transformation appeared in periocular images, we combine phase-based block matching with a coarse-to-fine search strategy using multi-scale image pyramids [41]. This technique is known as Phase-Based Correspondence Matching (PB-CM) and it consists of two main steps: correspondence search and similarity evaluation. Ito et al. [12] proposed this technique for palmprint recognition and it was later extended to other biometric traits [13], [42]. We modified this technique for periocular recognition and demonstrated that the designed method can achieve the highest
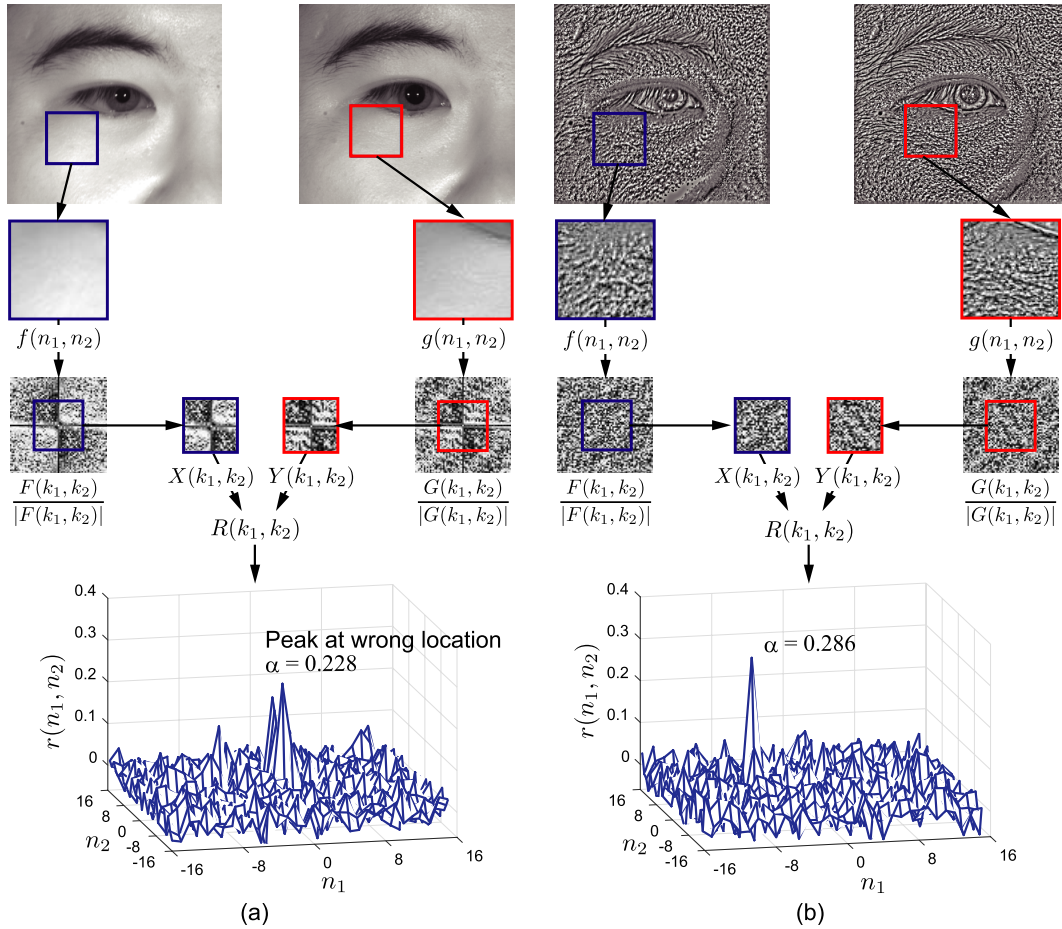
Fig. 5 Effect of texture enhancement on the BLPOC function for comparison of a genuine block pair with poor texture (left: without texture enhancement, right: with texture enhancement).

level of performance on matching periocular features textures when compared with the other methods reported so far.

Given a pair of periocular images, the problem considered here is to find a set of corresponding block pairs between the two images and evaluate their similarity on the basis of block-wise correlation computation. Let P be a set of N_b local block locations (i.e., coordinates) on the enrolled image I, where $P = \{\boldsymbol{p}_1, \cdots, \boldsymbol{p}_{N_b}\} \subseteq \mathbb{Z}^2$ determined in advance. The correspondence search problem is to find the set of corresponding block locations $Q = \{\boldsymbol{q}_1, \cdots, \boldsymbol{q}_{N_b}\} \subseteq \mathbb{Z}^2$ on the probe image J. The search of Q comprises three main steps:

(i) Generate multi-resolution image pyramids

We indicate the resolution layer of the image pyramids by the superscript l, where $l = 0, \cdots, l_{max}$. That is I^l for the enrolled image and J^l for the probe image. We generate the image pyramids by setting $I (= I^0)$ and $J (= J^0)$ and applying the following equations:

$$I^l(n_1, n_2) = \frac{1}{4} \sum_{j_1=0}^{1} \sum_{j_2=0}^{1} I^{l-1}(2n_1 + j_1, 2n_2 + j_2), \quad (10)$$

$$J^l(n_1, n_2) = \frac{1}{4} \sum_{j_1=0}^{1} \sum_{j_2=0}^{1} J^{l-1}(2n_1 + j_1, 2n_2 + j_2), \quad (11)$$

for all $l = 1, \cdots, l_{max}$. After generating image pyramids, texture enhancement described in Sect. 3.2 is carried out for all the resolution layers.

(ii) Determine a set of local block locations in the enrolled image pyramid

Let $P^l = \{\boldsymbol{p}_1^l, \cdots, \boldsymbol{p}_{N_b}^l\}$ denote a set of local block locations on the the l-th layer enrolled image I^l. A set of block locations $P^0 = \{\boldsymbol{p}_1^0 \cdots, \boldsymbol{p}_{N_b}^0\}$ on the original enrolled image I^0 are given in advance as $P^0 = P$ and $\boldsymbol{p}_t^0 = \boldsymbol{p}_t$. Then, the lower-resolution coordinates can be automatically computed as

$$\boldsymbol{p}_t^l = \left\lfloor \frac{1}{2^l} \boldsymbol{p}_t^0 \right\rfloor, \quad (12)$$

where $t = 1, \cdots, N_b$ and $l = 1, \cdots, l_{max}$.

(iii) Estimate the corresponding block locations

Let $Q^l = \{\boldsymbol{q}_1^l, \cdots, \boldsymbol{q}_{N_b}^l\}$ denote a set of corresponding block locations on the l-th layer probe image J^l. These coordinates are estimated using a coarse-to-fine strategy. The l-th
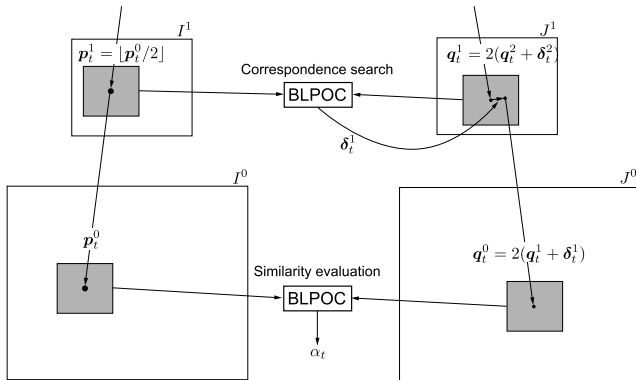
Fig. 6   Sketch of phase-based correspondence matching.

layer corresponding block location q_t^l ($\in Q^l$) is derived from the $(l+1)$-th layer correspondence (p_t^{l+1}, q_t^{l+1}) recursively by

$$q_t^l = 2\left(q_t^{l+1} + \delta_t^{l+1}\right), \tag{13}$$

where $t = 1, \cdots, N_b$, $l < l_{max}$ and δ_t^{l+1} denotes the translational displacement between p_t^{l+1} and q_t^{l+1} estimated by the BLPOC function. This recurrence starts from the coarsest layer $l = l_{max}$, where we assume the simplest approximation:

$$q_t^{l_{max}} = p_t^{l_{max}}. \tag{14}$$

The recurrence ends at the original resolution layer $l = l_0$. As a result, we obtain a set of coordinates $Q = Q^0$ on the probe image that corresponds to a set of coordinates P on the enrolled image. Fig. 6 depicts the coarsest-to-fine strategy for estimating the corresponding points q_t^1 on J^1 and q_t^0 on J^0. The result of the original layer l^0 is used for similarity evaluation as it is explained below in Sect. 3.4.

We modify the original PB-CM [12], [13], [42] in order to dedicate PB-CM to periocular image matching as summarized below. The original PB-CM uses three level resolution pyramids, that is, $l_{max} = 3$ [13]. In the case of periocular images, this is challenging due to the presence of occlusions and variations in gaze. Nonetheless, detection of eyes and facial landmarks is very mature and can ensure reliable global alignment of periocular images. Our experimental observation shows that correspondence search with two-level resolution pyramid, $l_{max} = 1$, should suffice to a wide range of periocular recognition applications. We also introduce the new similarity evaluation metric for periocular recognition described in the next subsection.

3.4   Similarity Evaluation

We evaluate the similarity between I and J through block-wise matching on the original resolution using the corresponding locations Q obtained from the correspondence search. This is to compute the BLPOC functions between a set of blocks at locations P on I and a set of blocks at locations Q on J so as to find the correlation peak values.
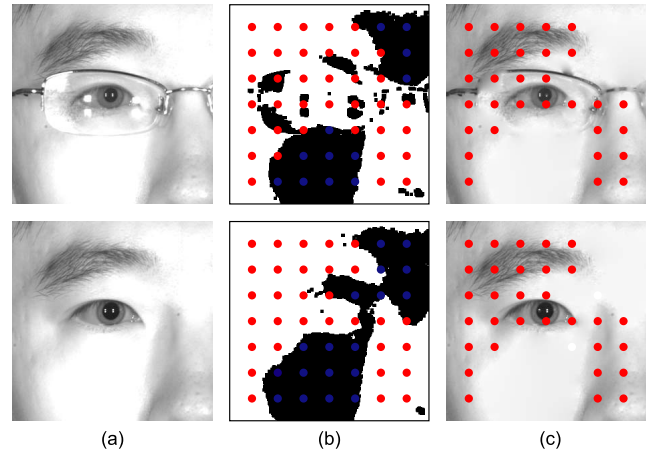

Fig. 7   Block center selection using thresholding to detect reflections: (a) the original image pair, (b) reflection masks and block locations with selected locations indicated by red dot and the discarded locations indicated by blue dots and (c) images after filling the areas that have reflections and the final block locations used for matching.

So far, we assumed that all the blocks in $Q = q_1, \cdots, q_{N_b}$ are valid in that they precisely correspond to the blocks in $P = p_1, \cdots, p_{N_b}$.

In a real situation, however, we have to consider that some blocks in Q are meaningless, since periocular images can present different kinds of occlusions such as glasses, hair, hats and specular reflections, which significantly disturb the matching operation. In particular, specular reflections have a considerable impact on recognition performance [43], which we actually confirmed through experiments. To address this problem, we detect the specular reflections with simple thresholding operation and discard the blocks (in P and Q) that are mostly covered by reflections. To be precise, in our experiments, a block is discarded when reflections occupies more than 50% of the block area.

We first define reflection masks M_I and M_J for the enrolled image I and the probe image J, respectively, through thresholding:

$$M_I(n_1, n_2) = \begin{cases} 1 & \text{if } I(n_1, n_2) < 252 \\ 0 & \text{otherwise} \end{cases}, \tag{15}$$

$$M_J(n_1, n_2) = \begin{cases} 1 & \text{if } J(n_1, n_2) < 252 \\ 0 & \text{otherwise} \end{cases}, \tag{16}$$

where 1 indicates valid pixels and 0 indicates possible specular reflections. In order to weaken the effect of abrupt intensity changes caused by specular reflections, we fill the area with reflections by interpolating inward from the pixel values on the outer boundary of such area[†].

Using the reflection masks, we can select valid blocks for which more than half of the pixels are valid. Fig. 7 depicts the block selection for two images with considerable specular reflections (Fig. 7(a)). We can observe how the effect of the specular reflection on the glasses is reduced in

---

[†]For example, MATLAB provides the *regionfill* function for this purpose.
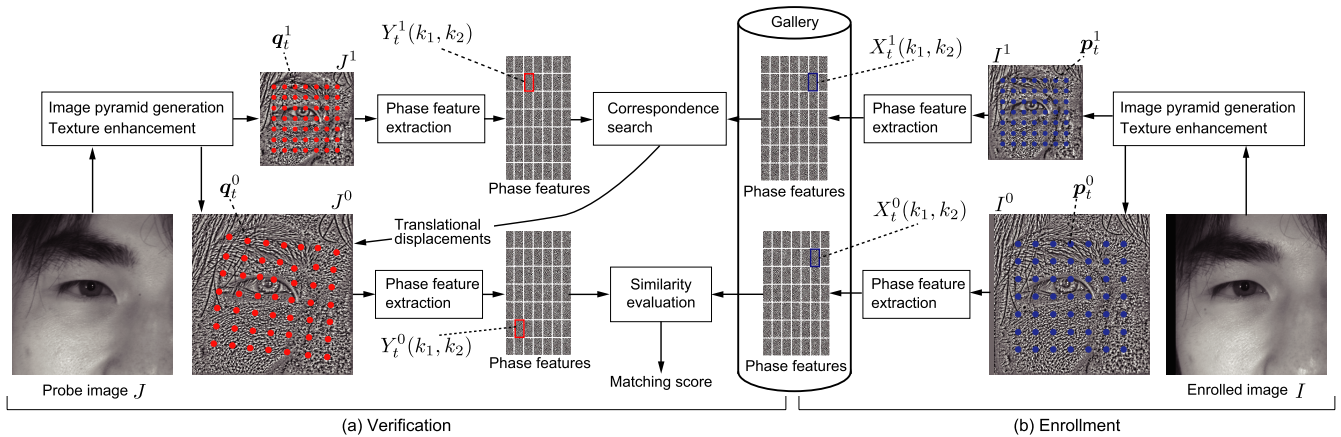
Fig. 8 Overall of the proposed periocular recognition method: (a) verification procedure of phase-based correspondence matching with texture enhancement and (b) enrolment procedure of phase-based correspondence matching with texture enhancement.

Fig. 7(c). After the block selection, we can determine the valid block pairs (p_t, q_t), where both of two blocks are valid in I and J, respectively. Let V be the set of all the indices t for the valid block pairs (p_t, q_t), i.e., $V = \{t | (\boldsymbol{p_t}, \boldsymbol{q_t})\}$ is a valid block pair, where $\boldsymbol{p_t} \in P$ and $\boldsymbol{q_t} \in Q$. Then, the correlation peak value α_t of the BLPOC function between the block pair $(\boldsymbol{p_t}, \boldsymbol{q_t})$ is said to be valid if and only if the block pair is valid, i.e., $t \in V$.

We consider three measures for similarity (i.e., matching scores) between I and J using the valid correlation peak values. The first measure is the straightforward average of the valid peak values α_t as

$$S_{\text{ave}} = \frac{1}{|V|} \sum_{t \in V} \alpha_t. \tag{17}$$

The second measure is the n-rank of peak values ordered from highest to lowest as

$$S_{\text{rank}} = \alpha_{t_n}, \tag{18}$$

where the valid peak values are sorted as $\alpha_{t_1} \geq \cdots \geq \alpha_{t_n} \geq \cdots \geq \alpha_{t_{|V|}}$. The third measure is the number of peak values that are greater than a threshold value thr as

$$S_{thr} = \sum_{t \in V} h(\alpha_t), \tag{19}$$

where

$$h(x) = \begin{cases} 1 & \text{if } x \geq Thr \\ 0 & \text{otherwise.} \end{cases}.$$

4. Periocular Recognition Algorithm

We describe our proposed algorithm for periocular recognition using PB-CM with texture enhancement. Fig. 8 shows the overall of the recognition method. The enrolment procedure is depicted in Fig. 8(b) and described by Algorithm 1. During enrolment, we register a given periocular image as

---

Algorithm 1 Enrolment procedure

Require: An image I to be enrolled
Ensure: A set of phase features $X_t^l(k_1, k_2)$ for I
  Generate the reflection mask $M_I(n_1, n_2)$ by Eq. (15)
  Fill the area with specular reflections on I as described in Sect. 3.4
  Generate a two-layer resolution pyramid and apply texture enhancement so as to obtain I^0 and I^1 as described in Sect. 3.3 and Sect. 3.2
  for $t \in \{1, 2, \cdots, N_b\}$ **do**
    if the block at location $\boldsymbol{p_t^0}$ has more than 50% of valid pixels according to $M_I(n_1, n_2)$ **then**
      Extract a block from I^0 at $\boldsymbol{p_t^0}$
      Compute the phase feature $X_t^0(k_1, k_2)$ by Eq. (3)
      Extract a block from I^1 at $\boldsymbol{p_t^1}$
      Compute the phase feature $X_t^l(k_1, k_2)$ by Eq. (3)
    end if
  end for
  Store the computed phase features in the gallery

---

an array of local phase features extracted from its multi-resolution image pyramid with texture enhancement. A phase feature extracted from the enhanced enrolled image I^l at position $\boldsymbol{p_t^l}$ is denoted by $X_t^l(k_1, k_2)$.

The verification procedure is depicted in Fig. 8(a) and described by Algorithm 2. During verification, given a probe image J, we extract a set of phase features from its multi-resolution image pyramid with texture enhancement. We compare these phase features with the phase features in the gallery using correspondence search and similarity evaluation. $Y_t^l(k_1, k_2)$ denotes a phase feature extracted from the enhanced probe image J^l at a position $\boldsymbol{q_t}$.

In addition, we reduced the registered-data size of the phase features stored in the gallery. We stored only a reduced portion of the total phase information by taking advantage of the spectrum symmetry and including only the band required for the calculation of the BLPOC functions. Also, we applied quantization to the phase angles (see [13], [34] for details). Phase quantization has shown slight degradation of the recognition performance by carefully considering the number of bits used.

Algorithm 2 Verification procedure

Require: A probe image J
Ensure: A matching score between J and the enrolled image I
  Generate the reflection mask $M_J(n_1, n_2)$ by Eq. (16)
  Fill the area with specular reflections on J as described in Sect. 3.4
  Generate a two-layer resolution pyramid and apply texture enhancement
  so as to obtain J^0 and J^1 as described in Secs. 3.3 and 3.2
  Initialize the corresponding points q_t^1 ($t = 1, 2, \cdots, N_b$) on J^1 by
  Eq. (14)
  Initialize the set V, i.e., $V = \{\emptyset\}$
  for $t \in \{1, 2, \cdots, N_b\}$ **do**
    if there is a phase feature $X_t^1(k_1, k_2)$ in the gallery **then**
      Extract a corresponding block from J^1 at q_t^1
      Compute the phase feature $Y_t^1(k_1, k_2)$ by Eq. (4)
      Compute the BLPOC function between $X_t^1(k_1, k_2)$ and $Y_t^1(k_1, k_2)$
      by Eq. (5) and Eq. (7)
      Estimate the translational displacement δ_t^1
      Determine q_t^0 from q_t^1 and δ_t^1 by Eq. (13)
      Extract a corresponding block from J^0 at q_t^0
      if the block at q_t^0 has more than 50% of valid pixels according to
      $M_J(n_1, n_2)$ **then**
        Compute the phase feature $Y_t^0(k_1, k_2)$ by Eq. (4)
        Compute the BLPOC function between $X_t^0(k_1, k_2)$ and
        $Y_t^0(k_1, k_2)$ by Eq. (5) and Eq. (7)
        Obtain the peak value α_t and add t to the set V
      end if
    end if
  end for
  Compute the matching score between I and J from the values α_t ($t \in V$)
  using one of the similarity measures described in Sect. 3.4

---

5. Experimentals and Discussion

This section describes performance evaluation of the proposed method and two baseline methods, which are one based on LBP [44] and the other based on SIFT [45], using three publicly available databases: CASIA-Iris-Distance in the CASIA Iris Image Database Version 4.0 (CASIA) [46], UBIPr database (UBIPr) [8], and ocular still challenge of the NIST FOCS database (FOCS) [47]. We also compare the recognition performance of our method with the performance reported in the literature for advanced periocular recognition methods [23], [27].

For the comparative evaluation, we present the Receiver Operating Characteristic (ROC) curve, which is a plot of the False Rejection Rate (FRR) against the False Acceptance Rate (FAR). In all the experiments, we measured the verification performance using the Equal Error Rate (EER), which is the error rate when FAR [%] is equal to FRR [%]. We compute the EER [%] values by linear interpolation of the ROC curves.

5.1 Databases

The images in the CASIA database were collected by the Institute of Automation, Chinese Academy of Sciences [46]. This database consists of 2,567 partial face images taken from 142 subjects under near infrared illumination, where the subjects stand at a distance around 3m from the camera. The size of this images is $2,352 \times 1,728$ pixels (width×height)

and they cover from the mouth to the forehead of the subjects. The images contain small variations in head pose and occlusions due to hair, eyeglass, and specular reflections. When the tilt (or rotation of the face) is more than six degrees, we correct it by aligning the eyes position horizontally. We scaled the images to one-fourth of the original image resolution and cropped two periocular regions of 300×300 pixels for each eye.

The University of Beira Interior Periocular (UBIPr) database [8] consists of 10,950 periocular images captured in the visible spectrum from 342 subjects [47]. Each subject has images at five different resolutions: 501×401 pixels at 8m (stand-off distance), 561×441 pixels at 7m, 651×501 pixels at 6m, 801×651 pixels at 5m and $1,001 \times 801$ pixels at 4m. We scale all these images to a common size of 240×300 pixels and transform them into gray-scale.

The Face and Ocular Challenge Series (FOCS) database [47] consists of 9,581 periocular images of 136 subjects where 4,792 images are from left eyes and 4,789 images are from right eyes. Images are captured with a near-infrared camera at an image resolution of 750×600 pixels. These periocular images were extracted from video sequences of subjects while walking. This database contains images with blur, occlusion and gaze deviation. The images also exhibit drastic variations in illumination and sensor noise. We scale these images to a size of 300×240 pixels.

5.2 Baseline Methods

In order to compare the recognition performance of the proposed method, we implemented two baseline recognition methods, which employ well-known image descriptors: Local Binary Patterns (LBP) [48] and modified Scale-Invariant Feature Transform (m-SIFT) [18]. The first method used the LBP operator [49], which assigns a label to every pixel of an image by thresholding the neighborhoods of each pixel with the center pixel value and considering the result as a binary number. Then, the histogram of the labels is used as a texture descriptor. Periocular recognition is performed by block-wise comparison of the local histograms. We tested two block sizes for the histogram computation, which are 30×30 pixels for small size and 50×50 pixels for large size. In the experiments with the CASIA database, images are divided into 10×10 (column×row) for small blocks and 6×6 for large blocks. In the case of FOCS and UBIPr databases, we divided images into 10×8 for small blocks and 6×5 for large blocks. We use MATLAB implementation of LBP[†]. The modified SIFT (m-SIFT) method [18] is a biometric recognition method based on SIFT features [45]. In our experiments, m-SIFT is implemented using the VLfeat library[††] like Ross et al. [18]. For a fair comparison, we also applied the reflection mask described in Sect. 3.4 to select or discard features for the two baseline methods. If the circular region around a SIFT keypoint has less than 70% of valid

---

[†]http://www.cse.oulu.fi/CMV/Downloads/LBPSoftware
[††]http://www.vlfeat.org/

Table 1 Parameters for PB-CM and texture enhancement. Parameters indicated with ∗ are used in the comparative evaluations.

PB-CM	
Parameter	Value
Block size	48×48 pixels*
Horizontal/vertical spacing between blocks	16 pixels*
Number of blocks	14×14 blocks, where we removed 5 blocks corresponding to the upper eyelid
Band limitation for BLPOC	50% bandwidth for 24×24 phase components
	67% bandwidth for 32×32 phase components

Texture enhancement	
Parameter	Value
Noise rejection parameter η	$0.0004 \times \|C\| \times 255$*, where $\|C\|$ is the size of the patch
Patch size	3×3 pixels, 5×5 pixels* and 7×7 pixels

Similarity measure	
Parameter	Value
Threshold of S_{thr}	0.30 for 50% bandwidth
	0.28* for 67% bandwidth
Rank of S_{rank}	12 for 50% bandwidth
	8* for 67% bandwidth

area, that keypoint is discarded, and if an LBP block has less than 70% of valid area, that block is discarded.

5.3 Impact of Texture Enhancement on Recognition Accuracy

We employed the CASIA database to evaluate the effect of texture enhancement, since CASIA images are larger than those of the other two databases and display richer skin texture. The parameters for PB-CM and the texture enhancement are presented in Table 1. Texture enhancement affects the appearance of the images with a whitening-like effect, which intensifies the texture representation in the higher frequency bands. Hence, the use of wider BLPOC bandwidth compared with the usual 50% bandwidth [13] can improve verification performance. For this reason, we consider two band limitation setups for BLPOC and three patch size for texture enhancement. Table 2 presents the EERs for the aforementioned configurations. We observe consistent improvement using texture enhancement for both eyes. The introduction of texture enhancement and 67% BLPOC bandwidth can reduce EER to less than one third of the original case (50% bandwidth) and no enhancement. Considering these results, in the following experiments, we employ a patch size of 5×5 pixels and a bandwidth of 67%. For this configuration, the similarity measures S_{thr} and S_{rank} consistently outperform S_{ave}, and hence we omit S_{ave} in the following experiments.

5.4 Comparative Performance Evaluation

We compared our proposed method with the baseline methods for the three databases. We adopt the parameters marked with ∗ in Table 1. The number of blocks for PB-CM is changed depending on the database used, i.e., 14×14 for CASIA and 14×11 for FOCS and UBIPr. Additionally, we employ 4-bit quantization for phase angle representation (see Appendix for further on the effect of quantization). We also compare the performance of the proposed method without texture enhancement in order to demonstrate the effectiveness of texture enhancement in periocular recognition. Table 3(a), (b) and (c) summarize the verification performance for the CASIA, UBIPr, and FOCS databases, respectively. Fig. 9 compares the ROC curves of the proposed method with S_{rank}, the proposed method with texture enhancement and S_{rank} and the conventional methods using m-SIFT and LBP (the best case) for the three databases. The proposed method with texture enhancement outperforms m-SIFT, while m-SIFT outperforms LBP for the three databases. The proposed method with texture enhancement outperforms that without texture enhancement in all the cases. Hence, we discuss the experimental results only for the proposed method with texture enhancement in the following.

As for the CASIA database shown in Table 3(a), the proposed method exhibits EERs one order of magnitude lower than those of m-SIFT. As observed at FAR=0.01% in Fig. 9(a) and (b), the proposed method has a significantly low FRR, i.e., 0.23% for left eye and 0.09% for right eye, compared with m-SIFT, i.e., 9.6% for left eye and 6.4% for right eye. The CASIA database has an advantage for the proposed method due to its relatively good quality images with some skin texture. This database is also acceptable for m-SIFT, since the images have around 400 m-SIFT keypoints on average.

As for the UBIPr database shown in Table 3(b), the proposed method and m-SIFT performed relatively close, while LBP performed poorly. Unlike the CASIA database, the UBIPr database contains images with significant large head-pose variation specially in yaw rotations, which are difficult to address. This is a relatively advantageous scenario for m-SIFT due to its robustness against image deformation. Also, the UBIPr images contain a significant number of SIFT keypoints, where around 1,200 keypoints in average per image are detected. Nonetheless, as observed at FAR=0.1% in Fig. 9(c) and (d), the proposed method exhibits significantly low FRRs, i.e., 9.1% for left eye and 6.5% for right eye, with respect to m-SIFT, i.e., 16.7% for left eye and 10.29% for right eye.

As for the FOCS database shown in Table 3(c), all the methods performed poorly. The proposed method performs slightly better than m-SIFT in terms of EER. This database is highly challenging, since it contains images with substantial blur and noise.

5.5 Comparison with Methods Based on Trained Features

In Sect. 5.4, we demonstrated that the proposed method achieves higher performance on periocular recognition compared with other well-established biometric recognition

Table 2 EERs [%] of phase-based image matching with and without texture enhancement. Bold fonts indicate the best EER per similarity measure.

| Bandwidth	Texture enhancement	Left eye			Right eye		
		S_{ave}	S_{rank}	S_{thr}	S_{ave}	S_{rank}	S_{thr}
50%	without enhancement	0.415	0.451	0.472	0.165	0.313	0.351
	3×3-pixel patch	0.237	0.197	0.210	0.094	0.116	0.128
	5×5-pixel patch	0.237	0.179	0.227	0.089	0.098	0.108
	7×7-pixel patch	0.250	0.205	0.237	0.089	0.102	0.123
67%	without enhancement	0.442	0.624	0.484	0.179	0.487	0.272
	3×3-pixel patch	0.235	0.170	0.180	**0.085**	0.103	0.109
	5×5-pixel patch	**0.223**	**0.157**	**0.145**	**0.085**	**0.080**	0.077
	7×7-pixel patch	0.250	0.165	0.178	**0.085**	0.081	**0.075**
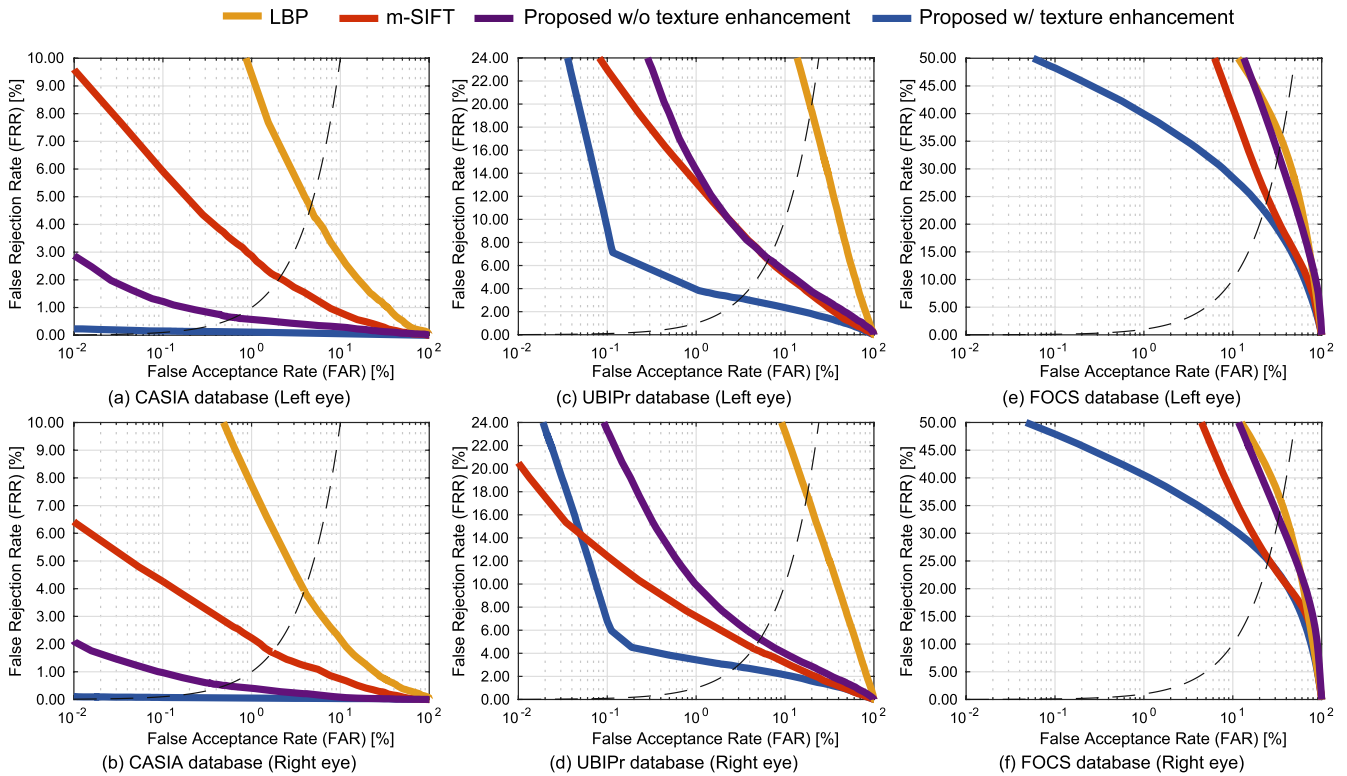
Fig. 9 Comparison of ROC curves of the proposed method and conventional methods.

methods, which employ highly robust feature descriptors: m-SIFT and LBP. The major difficulty of periocular recognition is that the skin under the eyes has considerably weak texture and hence conventional methods of biometric feature representation could not capture its inherent features. Another possibility to address this problem would be to use state-of-the-art machine learning techniques with training data set. In order to enhance the credibility of this paper, in this section, we additionally consider two methods: one is based on Convolutional Neural Networks (CNNs), which are gaining traction also in biometric applications [50], [51], and the other is based on correlation filters, which have been studied on periocular recognition [9], [18], [26]–[28], [52]. Specifically, we compared with "Semantic assisted CNN (SCNN)" proposed by Zhao et al. [23] and "Periocular Probabilistic Deformation Model (PPDM)" proposed by Smereka

et al. [27].

We prepared a subset of the CASIA database that imitates the setup used in [23] for SCNN, although we used different image segments. Table 4 presents the parameters of this setup, and Table 5 shows the resulting EERs of our method and SCNN implementation†. The error rates of our method are one order of magnitude lower than those of SCNN. We assumed that our segments were favorable to SCNN, since in our experiments SCNN yielded EER=4.32%, which is lower than EER=6.61% reported in the literature [23]. Therefore, the comparison in Table 5 should be fair. Note that the EERs of the proposed method in Table 5 differ from those in Table 3 due to the difference in experimental setup.

We also compare our method with PPDM considering

---

†http://www4.comp.polyu.edu.hk/ csajaykr/scnn.rar

Table 3 EERs [%] of the proposed method and conventional methods on CASIA database, UBIPr database and FOCS database, where TE indicates texture enhancement.

(a) CASIA database

Method	Left eye	Right eye
LBP (Block size: 30×30 pixels)	6.067	4.877
LBP (Block size: 50×50 pixels)	4.595	3.920
m-SIFT	2.065	1.710
Proposed with S_{rank}	0.640	0.510
Proposed with S_{thr}	0.610	0.290
Proposed with TE and S_{rank}	0.143	0.075
Proposed with TE and S_{thr}	0.150	0.078

(b) UBIPr database

Method	Left eye	Right eye
LBP (Block size: 30×30 pixels)	19.59	17.70
LBP (Block size: 50×50 pixels)	30.10	29.88
m-SIFT	5.57	4.15
Proposed with S_{rank}	6.55	5.25
Proposed with S_{thr}	5.19	4.05
Proposed with TE and S_{rank}	3.16	2.87
Proposed with TE and S_{thr}	3.47	3.17

(c) FOCS database

Method	Left eye	Right eye
LBP (Block size: 30×30 pixels)	35.30	35.45
LBP (Block size: 50×50 pixels)	46.60	43.03
m-SIFT	24.69	25.26
Proposed with S_{rank}	32.64	32.79
Proposed with S_{thr}	32.02	32.46
Proposed with TE and S_{rank}	22.46	25.08
Proposed with TE and S_{thr}	24.67	26.46

Table 4 Experimental setup for comparing with SCNN.

Parameter	Value
Image selection	As provided in the webpage[†]
Image resolution	0.294 times the original resolution, which is estimated from image samples provided in the webpage[†]
Image size	240×240 pixels
Right eye images	Right eye images are flipped horizontally, since SCNN is trained for left eye images
Number of blocks for PB-CM	10×10 blocks

Table 5 EERs [%] on a subset of CASIA database.

Method	Left eye	Right eye
SCNN	4.32	4.25
Proposed with S_{rank}	0.682	0.237
Proposed with S_{thr}	0.735	0.253

the EERs reported in the literature for left eye images. The EERs of PPDM are 10.47% and 7.67% for CASIA database [23] and UBIPr database [27], respectively. These EERs are more than double of those of our method for the respective databases, i.e., 0.68% and 3.16%. We do not consider that the difference in test protocols can be the reason of such pronounced difference in performance. Therefore, we conclude that our method has clear advantage over PPDM in periocular image matching.

In the case of FOCS database, for left eye images, the EER of SCNN reported in [23] is 10.47%, and the EER of PPDM reported in [27] is 22.44%. Their experimental

evaluation differs in term of preprocessing and selection of test images. However, we can infer that the performance of our method, i.e., EER=22.19% is close to PPDM, while it is clearly worse than the performance of SCNN. This is because our method did not deal with such a high level of noise and blur as is observed in the FOCS images, while CNNs are able to manage a certain degree of feature ambiguity due to image quality degradation. We will consider introducing deblurring and denoising based on spatiotemporal analysis of video sequences in addition to texture enhancement so as to realize accurate periocular recognition for walking persons in future work.

6. Conclusion

We proposed a periocular recognition algorithm using phase-based correspondence matching with texture enhancement. We demonstrated that the phase-based approach also exhibits impressive performance in periocular recognition compared with conventional approaches. By using texture enhancement, we improved the recognition performance of phase-based image matching on the skin texture of periocular images. Experimental evaluation using three public databases demonstrated a clear advantage of the proposed method matching on periocular images over conventional methods.

Acknowledgments

References

[1] M. De Marsico, M. Nappi, and M. Tistarelli, Face Recognition in Adverse Conditions, IGI Global, 2014.

[2] K. Ricanek, M. Savvides, D.L. Woodard, and G. Dozier, "Unconstrained biometric identification: Emerging technologies," Computer, vol.43, no.2, pp.56–62, Feb. 2010.

[3] S.Z. Li and A. Jain, Handbook of Face Recognition, Springer, 2011.

[4] M.J. Burge and K. Bowyer, Handbook of Iris Recognition, Springer, 2013.

[5] G. Santos and H. Proença, "Periocular biometrics: An emerging technology for unconstrained scenarios," Proc. IEEE Symp. Computational Intelligence in Biometrics and Identiry Management, pp.14–21, April 2013.

[6] I. Nigam, M. Vatsa, and R. Singh, "Ocular biometrics: A survey of modalities and fusion approaches," Information Fusion, vol.26, pp.1–35, Nov. 2015.

[7] F. Alonso-Fernandez and J. Bigun, "A survey on periocular biometrics research," Pattern Recogn. Lett., vol.82, no.2, pp.92–105, Oct. 2016.

[8] C.N. Padole and H. Proença, "Periocular recognition: Analysis of performance degradation factors," Proc. Int'l Conf. Biometrics, pp.439–445, March 2012.

[9] V.N. Boddeti, J.M. Smereka, and B.V.K.V. Kumar, "A comparative evaluation of iris and ocular recognition methods on challenging ocular images," Proc. Int'l Joint Conf. Biometrics, pp.1–8, Oct. 2011.

[10] S. Bharadwaj, H.S. Bhatt, M. Vatsa, and R. Singh, "Periocular biometrics: When iris recognition fails," Proc. IEEE Int'l Conf. Biometrics: Theory, Applications and Systems, pp.1–6, Sept. 2010.

[11] P.E. Miller, J.R. Lyle, S.J. Pundlik, and D.L. Woodard, "Performance

evaluation of local appearance based periocular recognition," Proc. IEEE Int'l Conf. Biometrics: Theory, Applications and Systems, pp.1–6, Sept. 2010.

[12] K. Ito, S. Iitsuka, and T. Aoki, "A palmprint recognition algorithm using phase-based correspondence matching," Proc. Int'l Conf. Image Processing, pp.1977–1980, Nov. 2009.

[13] S. Aoyama, K. Ito, and T. Aoki, "Similarity measure using local phase features and its application to biometric recognition," Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition Workshops, pp.180–187, June 2013.

[14] D.S. Bolme, B.A. Draper, and J.R. Beveridge, "Average of synthetic exact filters," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp.2105–2112, June 2009.

[15] A. Uhl and P. Wild, "Combining face with face-part detectors under Gaussian assumption," Lecture Notes in Computer Science (Proc. Int'l Conf. Image Analysis and Recognition), vol.7325, pp.80–89, June 2012.

[16] F. Alonso-Fernandez and J. Bigun, "Near-infrared and visible-light periocular recognition with Gabor features using frequency-adaptive automatic eye detection," IET Biometrics, vol.4, no.2, pp.74–89, June 2015.

[17] U. Park, R.R. Jillela, A. Ross, and A.K. Jain, "Periocular biometrics in the visible spectrum," IEEE Trans. Inf. Forensics Security, vol.6, no.1, pp.96–106, March 2011.

[18] A. Ross, R. Jillela, J.M. Smereka, V.N. Boddeti, B.V.K.V. Kumar, R. Barnard, X. Hu, P. Pauca, and R. Plemmons, "Matching highly non-ideal occular images: An information fusion approach," Proc. Int'l Conf. Biometrics, pp.446–453, April 2012.

[19] S.S. Behera, M. Gour, V. Kanhangad, and N. Puhan, "Periocular recognition in cross-spectral scenario," Proc. Int'l Joint Conf. Biometrics, pp.681–687, Oct. 2017.

[20] P.E. Miller, A.W. Rawls, S.J. Pundlik, and D.L. Woodard, "Personal identification using periocular skin texture," Proc. ACM Symp. Applied Computing, pp.1496–1500, March 2010.

[21] D.L. Woodard, S.J. Pundlik, J.R. Lyle, and P.E. Miller, "Periocular region appearance cues for biometric identification," Proc. IEEE Conf. Computer Vision and Pattern Recognition Workshops, pp.162–169, June 2010.

[22] M. Uzair, A. Mahmood, A. Mian, and C. McDonald, "Periocular region-based person identification in the visible, infrared and hyperspectral imagery," Neurocomputing, vol.149, pp.854–867, Feb. 2015.

[23] Z. Zhao and A. Kumar, "Accurate periocular recognition under less constrained environment using semantics-assisted convolutional neural network," IEEE Trans. Inf. Forensics Security, vol.12, no.5, pp.1017–1030, May 2017.

[24] H. Proença and J.C. Neves, "Deep-PRWIS: Periocular recognition without the iris and sclera using deep learning frameworks," IEEE Trans. Inf. Forensics Security, vol.13, no.4, pp.888–896, April 2018.

[25] A.A. Alahmadi, M. Hussain, H. Aboalsamh, and M. Zuair, "ConvSRC: Smartphone based periocular recognition using deep convolutional neural network and sparsity augmented collaborative representation," CoRR, vol.abs/1801.05449, pp.1–26, Jan. 2018.

[26] J.M. Smereka and V.B.K.V. Kumar, "What is a "good" periocular region for recognition?," Proc. IEEE Conf. Computer Vision and Pattern Recognition Workshops, pp.117–124, June 2013.

[27] J.M. Smereka, V.N. Boddeti, and B.V.K.V. Kumar, "Probabilistic deformation models for challenging periocular image verification," IEEE Trans. Inf. Forensics Security, vol.10, no.9, pp.1875–1890, Sept. 2015.

[28] J.M. Smereka, B.V.K.V. Kumar, and A. Rodriguez, "Selecting discriminative regions for periocular verification," Proc. IEEE Int'l Conf. Identity, Security and Behavior Analysis, pp.1–8, Feb. 2016.

[29] K. Ito and T. Aoki, "Phase-based image matching and its application to biometric recognition," Proc. Asia-Pacific Signal and Information Processing Association Annual Summit and Conf., pp.1–7, Oct. 2013.

[30] C.D. Kuglin and D.C. Hines, "The phase correlation image alignment method," Proc. Int'l Conf. Cybernetics and Society, pp.163–165, 1975.

[31] K. Takita, T. Aoki, Y. Sasaki, T. Higuchi, and K. Kobayashi, "High-accuracy subpixel image registration based on phase-only correlation," IEICE Trans. Fundamentals, vol.E86-A, no.8, pp.1925–1934, Aug. 2003.

[32] K. Ito, H. Nakajima, K. Kobayashi, T. Aoki, and T. Higuchi, "A fingerprint matching algorithm using phase-only correlation," IEICE Trans. Fundamentals, vol.E87-A, no.3, pp.682–691, March 2004.

[33] K. Miyazawa, K. Ito, T. Aoki, K. Kobayashi, and H. Nakajima, "An effective approach for iris recognition using phase-based image matching," IEEE Trans. Pattern Anal. Mach. Intell., vol.30, no.10, pp.1741–1756, Oct. 2008.

[34] S. Iitsuka, K. Ito, and T. Aoki, "A practical palmprint recognition algorithm using phase information," Proc. Int'l Conf. Pattern Recognition, pp.1–4, Dec. 2008.

[35] S. Aoyama, K. Ito, and T. Aoki, "A finger-knuckle-print recognition algorithm using phase-based local block matching," Information Sciences, vol.268, pp.53–64, June 2014.

[36] D. Kusanagi, S. Aoyama, K. Ito, and T. Aoki, "A practical person authentication system using second minor finger knuckles for door security," IPSJ Trans. Comput. Vis. Appl., vol.9, no.8, pp.1–13, March 2017.

[37] J. Mairal, F. Bach, and J. Ponce, Sparse Modeling for Image and Vision Processing, Now Publishers, 2014.

[38] S. Lyu, "Divisive normalization as an efficient coding transform: Justification and effectiveness," Neural Comput., vol.23, no.11, pp.2942–2973, Nov. 2011.

[39] J. Malo, I. Epifanio, R. Navarro, and E.P. Simoncelli, "Nonlinear image representation for efficient perceptual coding," IEEE Trans. Image Process., vol.15, no.1, pp.68–80, Jan. 2006.

[40] S. Lyu and E.P. Simoncelli, "Nonlinear image representation using divisive normalization," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp.1–8, June 2008.

[41] K. Takita, M.A. Muquit, T. Aoki, and T. Higuchi, "A sub-pixel correspondence search technique for computer vision applications," IEICE Trans. Fundamentals, vol.E87-A, no.8, pp.1913–1923, Aug. 2004.

[42] L.R. Marval-Perez, S. Aoyama, K. Ito, and T. Aoki, "Score level fusion of multibiometrics using local phase array," Lecture Notes in Computer Science (Proc. Pacific-Rim Conf. Multimedia), vol.9315, pp.215–224, Sept. 2015.

[43] S. Crihalmeanu, A. Ross, and R. Derakhshani, "Enhancement and registration schemes for matching conjunctival vasculature," Lecture Notes in Computer Science (Proc. Int'l Conf. Biometrics), vol.5558, pp.1240–1249, June 2009.

[44] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," Pattern Recogn., vol.29, no.1, pp.51–59, Jan. 1996.

[45] D. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. Comput. Vision, vol.60, no.2, pp.91–110, Jan. 2004.

[46] "CASIA Iris Image Database V4.0," http://www.idealtest.org/dbDetailForUser.do?id=4

[47] K.W. Bowyer and M.J. Burge, Handbook of Iris Recognition, Springer, 2016.

[48] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with local binary patterns: Application to face recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol.28, no.12, pp.2037–2041, Dec. 2006.

[49] M. Pietikäinen, A. Hadid, G. Zhao, and T. Ahonen, Computer Vision Using Local Binary Patterns, Springer, 2011.

[50] B. Bhanu and A. Kumar, Deep Learning for Biometrics, Springer, 2017.

[51] M. Vatsa, R. Singh, and A. Majumdar, Deep Learning in Biometrics, CRC Press, 2018.

[52] F. Juefei-Xu and M. Savvides, "Subspace-based discrete transform encoded local binary patterns representations for robust periocular

matching on NIST's face recognition grand challenge," IEEE Trans. Image Process., vol.23, no.8, pp.3490–3505, Aug. 2014.

Appendix: Quantization of Phase Features

We evaluated the effect of phase quantization on the recognition performance using the CASIA database. We used the parameters shown in Table 1 and selected a patch size of 5×5 pixels for texture enhancement and 67% bandwidth for BLPOC. Table A·1 shows the EERs for different quantization levels of phase angles: 2bits, 3bits and 4bits. We did not observe significant degradation of recognition performance due to quantization. However, in the case of 2-bit quantization, the results are not consistent for both eyes. We observe little degradation of the recognition performance for left eye, while slight improvement is observed for right eye. We consider 4-bit quantization to be a good balance between degree of compression and recognition performance. Also, 3-bit quantization is a valid consideration when the registered data size is the main concern.

Table A·1   EERs [%] of phase-based image matching with and without quantisation of phase angles. Bold fonts indicate the best EER per similarity measure.

Phase quantization	Left eye			Right eye		
	S_{ave}	S_{rank}	S_{thr}	S_{ave}	S_{rank}	S_{thr}
No	0.223	0.157	0.145	0.085	0.080	0.077
4 bits	**0.219**	**0.143**	**0.150**	**0.087**	0.075	**0.078**
3 bits	0.223	**0.143**	0.165	0.089	0.089	0.084
2 bits	0.257	0.179	0.210	0.080	**0.071**	0.088

Takafumi Aoki   received the BE, ME, and DE degrees in electronic engineering from Tohoku University, Sendai, Japan, in 1988, 1990, and 1992, respectively. He is currently a professor in the Graduate School of Information Sciences (GSIS) at Tohoku University. Since April 2012, he has also served as the Vice President of Tohoku University. His research interests include theoretical aspects of computation, computer design and organization, LSI systems for embedded applications, digital signal processing, computer vision, image processing, biometric authentication, and security issues in computer systems. He received more than 20 academic awards as well as distinguished service awards for his contributions to victim identification in the 2011 Great East Japan Disaster.

Luis Rafael Marval-Pérez   received the B.E. degree in electronic engineering from Simon Bolivar University, Venezuela, in 2006, and the M.S. degree in information sciences from Tohoku University, Sendai, Japan in 2012. He is currently a Ph.D. candidate at Tohoku University. His research interests include pattern recognition, sparse modeling, machine learning and computer vision.

Koichi Ito   received the B.E. degree in electronic engineering, and the M.S. and Ph.D. degree in information sciences from Tohoku University, Sendai, Japan, in 2000, 2002 and 2005, respectively. He is currently an Associate Professor of the Graduate School of Information Sciences at Tohoku University. From 2004 to 2005, he was a Research Fellow of the Japan Society for the Promotion of Science. His research interest includes signal and image processing, computer vision and biometric authentication.