Periocular Recognition Using uMLBP and Attribute Features

Zahid Ali[1], Unsang Park[1], Jongho Nang[1], Jeong-Seon Park[2], Taehwa Hong[3], and Sungjoo Park[4]
[1] Dept. of Comp. Sci. & Eng., Sogang University, Seoul, Republic of Korea
[e-mail: zahid,unsangpark,jhnang@sogang.ac.kr]
[2] Dept. of Multimedia, Chonnam National University, Jeollanam-do, Republic of Korea
[e-mail: jpark@jnu.ac.kr]
[3] Samsung Electronics, Kyungki-do, Republic of Korea
[e-mail: taehwa.hong@samsung.com]
[4] Smart Media Research Center, Korea Electronics Technology Institute, Seoul, Republic of Korea
[e-mail: bpark@keti.re.kr]
*Corresponding author: Unsang Park

Received March 3, 2017; revised May 28, 2017; accepted August 9, 2017; published December 31, 2017

Abstract

The field of periocular biometrics has gained wide attention as an alternative or supplemental means to conventional biometric traits such as the iris or the face. Periocular biometrics provide intermediate resolution between the iris and the face, which enables it to support both. We have developed a periocular recognition system by using uniform Multiscale Local Binary Pattern (uMLBP) and attribute features. The proposed system has been evaluated in terms of major factors that need to be considered on a mobile platform (e.g., distance and facial pose) to assess the feasibility of the use of periocular biometrics on mobile devices. Experimental results showed 98.7% of rank-1 identification accuracy on a subset of the Face Recognition Grand Challenge (FRGC) database, which is the best performance among similar studies.

Keywords: periocular biometrics, uMLBP, periocular attribute classifiers, partial least squares, FRGC

1. Introduction

Biometric trait-based user recognition has long been used for login authentications on PCs and laptop computers [6]. Fingerprints and the face are the two most widely used biometric traits for PC login authentication systems. Even though fingerprints have shown high verification accuracy, their contact-based sensing interface has offered lower user's convenience. This has triggered the development of non-contact user authentication systems; the face is the most popular non-contact biometric trait. However, facial biometrics suffer from lower recognition accuracy due to the large intra-class variation.

On the other hand, periocular biometrics have recently been introduced as a new biometric trait [17]. The increasing demands in preventing unauthorized access on mobile devices have been coupled with the characteristics of periocular biometrics to populate a new possibility of the use of periocular biometrics on mobile platforms. The advantages of using periocular biometrics on smartphones are listed below.

- Periocular biometrics can be captured during the normal use of mobile devices.
- Since the facial image is usually captured within arm's length of mobile devices, there are high perspective distortions and occlusion problems. Variations from wobble and illumination are other problems that degrade the facial recognition accuracy of mobile devices. Periocular Biometrics can mitigate these problems of facial biometrics. **Fig. 1** shows scenarios where Periocular Biometrics can help in identity verification.
- The smartphone industry is trying to incorporate iris recognition systems on smartphones. Periocular Biometrics can supplement the iris-based recognition system with no additional hardware and minimal overhead in software upgrades.
- The rapid increase of recognition accuracy of periocular biometrics to the level of facial biometrics has made the periocular-based login authentication system on mobile platforms more viable.

Fig. 1. Example images with face occlusions and illumination variations where periocular biometrics can help in identity verification. (Photo courtesy of www.freeimages.com).

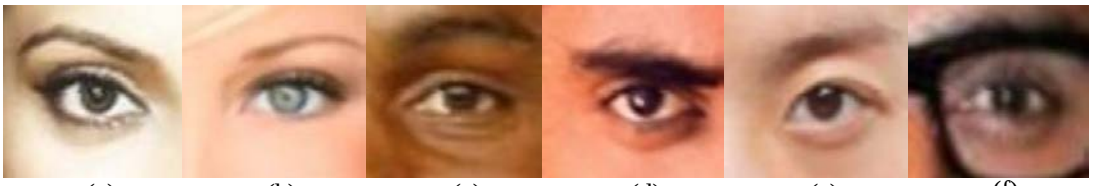
Compared to ocular biometrics, periocular biometrics offer several benefits [16]. One of the most significant benefits of periocular biometrics is its alternative use for person recognition when ocular biometrics fail [4]. Even if ocular biometrics are used

successfully, periocular biometrics can still be added to the ocular biometric system to increase recognition accuracy. It has been shown that the combination of iris and periocular biometrics is more effective than using only one of the two methods [24]. In contrast to ocular biometrics which is mostly used in the near-infrared spectrum, periocular biometrics can be used both in the visible (VS) and the near-infrared (NI) spectrums [16][23]. Using classical encoding and matching techniques, Santos et al. [18] demonstrated the utility of an iris and periocular biometrics by combining the two modalities using Neural Network to overcome the issues associated with mobile environments. Recently, it has been shown that periocular biometrics can also provide a good recognition performance in non-ideal recognition conditions, such as surveillance environments [10]. **Table 1** summarizes recent studies in periocular-based person identification.

Table 1. Literature review on periocular-based person identification. The FRGC database is underlined for comparison.

	Approach	**Database**	**Rank-1 identification accuracy (%)**	**EER (%)**	**d'**
Park et al. 2011 [16]	GO, LBP, SIFT	FRGC ($568_{subjects}$, $2{,}272_{images}$)	87.32	21.78 19.26 6.96	-
Woodard et al. 2010 [23]	uLBP	FRGC ($410_{subjects}$, 4100_{images}), MBGC ($85_{subjects}$, 1600_{images})	FRGC (90.00), MBGC (87.00)	-	-
Woodard et al. 2010 [24]	uLBP	MBGC ($80_{subjects}$ for right, $70_{subjects}$ for left eyes)	92.50	20*	-
Adams et al. 2010 [1]	GEFE feature	FRGC ($410_{subjects}$, 820_{images}), FERET($54_{subjects}$, 162_{images})	FRGC (92.16), FERET (85.06)	-	-
Dozier et al. 2011 [5]	GEFeS feature	FRGC ($410_{subjects}$, 820_{images})	93.09	-	-
Juefei-Xu et al. 2011 [9]	WLBP + UDP	FG-NET ($82_{subjects}$, 1002_{images})	100.00	0.6	-
Joshi et al. 2012 [8]	uLBP	CASIA ($400_{subjects}$, 2400_{images}), UBIRIS ($400_{subjects}$, 2400_{images})	CASIA iris (88.79), UBIRIS periocular (81.03), fusion (96.50)	7.33*	-
Uzair et al. 2013 [21]	LBP + PCA	MBGC ($85_{subjects}$, 3163_{images})	97.70	13.13	-
Jonathon et al. 2013 [7]	PDM + mSIFT	FOCS ($136_{subjects}$, 9581_{images}), UBIPr ($259_{subjects}$, 10252_{images})	FOCS (97.00), UBIPr (99.00)	23.95 6.43	-
Santos et al. 2015 [18]	LBP+HOG+ SIFT+uLBP +GIST	CSIP ($50_{subjects}$, 10 setups, 2004_{images})	Periocular (92.70), Iris (71.10), Fusion (93.40)	0.15 0.71 0.14	-
Bakshi et al. 2014 [28]	PIGP	UBIris v2	82.86	12.69	1.47
Bakshi et al. 2015 [2]	PILP	BATH ($200_{subjects}$, 8000 NI_{images}) CASIA ($411_{subjects}$, 16212 NI_{images}) UBIrisv2 ($241_{subjects}$, 1877 VS_{images}) FERETv4 ($1191_{subjects}$, 14126 VS_{images})	BATH (100) CASIA (100) UBIris (87.62) FERET (85.80)	0.13 0.38 4.49 7.53	3.31 2.90 2.29 2.01
Proposed method	uMLBP + LDA+PA	FRGC ($568_{subjects}$, $2{,}272_{images}$)	98.70	4.89	1.42

* Average EER = (EER of right eye + EER of left eye)/2



(a) (b) (c) (d) (e) (f)

Fig. 2. Example images to illustrate the describable attribute based recognition. (a) Indian Female, (b) Caucasian Female with Thin eyebrows, (c) African American with bags under eyes, (d) Male with thick eyebrows, (e) Asian and (f) Wearing eyeglasses.

In this article, we propose a state-of-the-art periocular recognition system using uniform Multiscale Local Binary Pattern (uMLBP), Linear Discriminant Analysis (LDA) and periocular attribute features. Our contributions differ from previous studies in following respects:

- First, we propose a periocular recognition system using uMLBP and LDA that shows state-of-the-art recognition accuracy on the subset of the FRGC database. We compared the proposed method with state-of-the-art periocular recognition methods and two other state-of-the-art MLPB methods that were applied on facial recognition. LDA has been widely used in facial recognition, but has never been used in periocular recognition. We showed that the combination of uMLBP and LDA provide improved performance in periocular recognition.

- Second, we propose to use the describable attributes extracted from the periocular region to further improve the identification accuracy. **Fig. 2** shows periocular regions that can be differentiated based on their describable facial attributes. Facial attributes are naturally used to recognize a person. For example, we could describe Barak Obama as a Middle aged African American male, with curly short hairs, without eye glasses, oval face shape and thick lips, etc. These traits or facial attributes are helpful to distinguish two persons very quickly. Using similar idea, Kumar et al. [11] proposed a face verification method using describable facial attributes. Kumar et al. [11] used high-level visual features, or attributes, to represent a face image. These describable facial attributes such as, gender, ethnicity, age, etc., are insensitive to illumination conditions, expression, pose and other imaging variations [11]. As the describable attribute features of face image is called facial attributes, we call it for periocular image as periocular attribute (PA) in this article. We also used Partial Least Squares (PLS) method to reduce the feature dimensions, which not only improved the computation speed at post-processing steps but also increased the recognition accuracy. On an embedded platform, these describable periocular attributes can be computed in parallel with the uMLBP+LDA approach. Finally the scores from the two different methods can be fused to provide robust recognition performance. To the best of our knowledge, no existing research utilizes the attribute information along with low-level features in periocular biometric-based recognition.

- Third, to emphasize a real world scenario, a periocular image database is constructed for periocular recognition research in mobile platforms with considering variations in distance and pose captured by a smartphone. The effect of perspective distortion as a function of the capture distance was evaluated to test the feasibility of applying periocular recognition system on mobile platforms.

The rest of the paper is organized as follows. Section 2 describes proposed method and its implementation in detail. Section 3 describes the experimental procedures and their results. Finally, we draw the conclusions and future scope of our research work in Section 4.

2. Proposed Approach

In this section we provide detailed explanations about the proposed periocular recognition method based on uMLBP and LDA subspace learning, the attribute based periocular region matching method, and fusion of uMLBP and periocular attribute (PA) based recognition methods. The overall periocular recognition system based on uMLBP+LDA and PA+PLS is shown in **Fig. 3**. We first explain the workflow of the overall system briefly, and then discuss each component in more details in the following sections.

The overall system diagram shown in **Fig. 3** shows process flows for uMLBP+LDA based and PA+PLS based periocular recognition. The training pipeline (orange color) consists of labelled dataset of periocular attributes. An example of 4 classes i.e., age (old vs young), gender (male vs female), ethnicity (Caucasian vs black, India[1], and Asian), eyebrows (thick vs thin) are shown in **Fig. 3**. For each image in the dataset, STASM landmark detector [13] is applied to acquire facial landmarks. The facial landmarks are used for rotation and scale normalization of the input images and cropping the four sub-regions from the two periocular regions as shown in **Fig. 3**. The low level features from the four regions are extracted and then mapped to the learnt PLS weights to reduce the feature dimensions. These feature are then used to train 15 SVM based PA classifiers. The subject ID is verified by matching the feature vectors with the stored feature vectors in the periocular database. In following section, we give the detail explanation of all the modules shown in **Fig. 3**.

2.1. Periocular recognition system based on uMLBP and LDA

We reviewed representative approaches in periocular-based person recognition systems, as summarized in **Table 1**. The majority of the approaches have used uniform Local Binary Pattern (uLBP) or multiscale Local Binary Pattern (MLBP) due to their superior performance in representing texture features and robustness in scales. LBP was first introduced by Ojala [15] to describe texture-type features. The LBP was later extended to uLBP and MLBP [20] to handle the performance degradations incurred from noise and scale variations. We used the uMLBP as a low-level feature in the proposed system to address both uniformity and multiscale aspects.

[1] Even though Indian is not known as an ethnicity, we used that category following the original facial attribute work of [11].

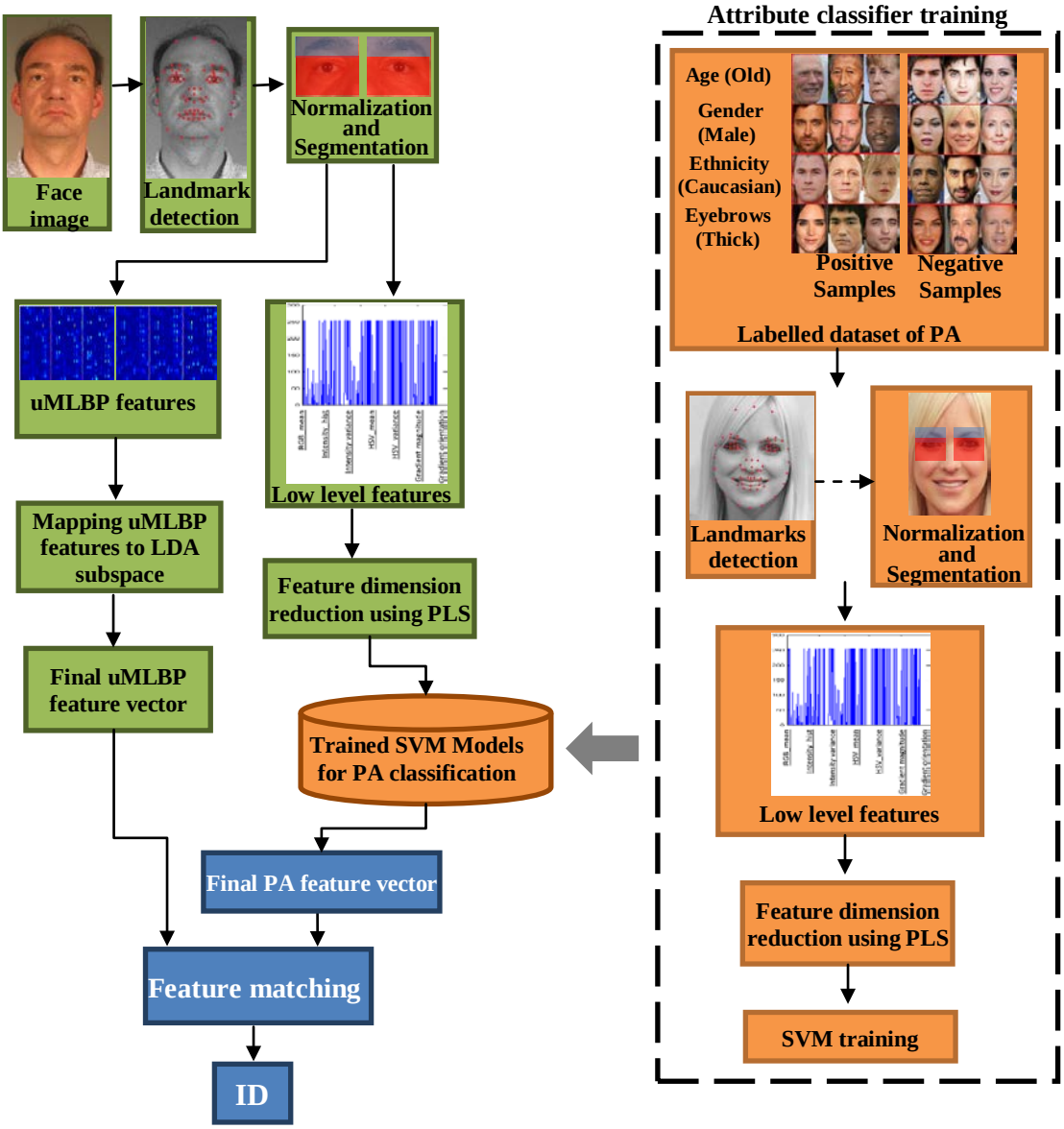


Fig. 3. Overall system diagram.

In order to describe the basic LBP operator, we assume there are NP circular neighborhood pixels ($=g_{np}$) in a circle of radius R centered at C ($=g_C$), then the LBP operator can be defined as:

$$LBP_{NP,R}(x,y) = \sum_{np=0}^{NP} \delta(g_{np} - g_C)2^{NP}, \quad (1)$$

where $\delta(d) = 1\ iff\ d \geq 0$. A subset of these 2^{NP} binary patterns, which contains at most two bitwise transitions (0→1 or 1→0) are called uniform patterns. The measure of uniformity is given by:

$$LBP_{uniform} = |\delta(g_{NP-1} - g_C) - \delta(g_0 - g_C)| + \sum_{np=1}^{NP} |\delta(g_{np} - g_C) - \delta(g_{NP-1} - g_C)| \quad (2)$$

Some example uniform patterns are, 11111111 (0 transitions), 00110000 (2 transitions) and 10111111 (2 transitions), whereas the patterns 10101010 (7 transitions) and 00110011 (3 transitions) are non-uniform. These uniform binary patterns can be used to represent spots, flat areas, edges and corners robustly. The total number of these uniform binary patterns in 2^{NP} patterns are $(NP-1)\times NP + 3$. The uniform LBP (uLBP) operator is then defined as:

$$LBP^{u2}_{NP,R}(x,y) = \begin{cases} L_u(LBP_{NP,R}(x,y) & , \ if\ LBP_{uniform} \le 2 \\ (NP-1)\times NP + 3 & , \ otherwise \end{cases} \quad (3)$$

where *u2* represents $LBP_{uniform}$ is up to 2, L_u assigns a unique label $\in [0, (NP-1)\times NP + 2]$ among the uniform LBP. For example, there are total 59 unique uniform LBP features for $NP = 8$.

In our experiments, we have used another variant of uLBP called uMLBP, which is computed by varying the sampling radius R and concatenating uLBP features at each radius (scale) to achieve a multi-resolution representation of uLBP. We used patch radii of 1, 3, 9 and 12 in our experiments. The uMLBP descriptors at these patch radii are converted to column vectors and concatenated all together.

In addition, a subspace learning method based on LDA [3] was used to construct compact and discriminative features for mobile environment. LDA minimizes the redundant information from uMLBP features, which is inherited by uMLBP features due to resampling the same image at multiple resolutions. LDA maximizes the ratio of between-class scatter ($\mathbf{S_B}$) and within-class scatter ($\mathbf{S_W}$), given as:

$$\mathbf{S_B} = \sum_{i=1}^{NC} ns^i \, (e^i - e)(e^i - e)^T \quad (4)$$

$$\mathbf{S_W} = \sum_{i=1}^{NC} \sum_{j=1}^{ns^i} (fs_j^i - e^i)(fs_j^i - e^i)^T \quad (5)$$

where NC is the total number of classes, ns^i and e^i are the number of samples and the mean of all the face images in i^{th} class, respectively, fs_j^i represents the j^{th} feature sample in i^{th} class, and e is the mean of all the samples. The optimal projection $\mathbf{W_{LDA}}$ is chosen as the matrix that maximizes the expression below.

$$\mathbf{W_{LDA}} = arg\max_{\mathbf{W}} \frac{|\mathbf{W S_B W}^{\mathrm{T}}|}{|\mathbf{W S_W W}^{\mathrm{T}}|} \quad (6)$$

Periocular recognition starts with periocular region cropping from the facial image as shown in **Fig. 3**. We detected facial landmarks using STASM landmarks detector [13], and did the rotation normalization, scale normalization, and cropped with a fixed aspect ratio (241×227 pixels) for both eyes. The uMLBP features are then extracted from the periocular regions. Then, all the uMLBP feature vectors are transformed by LDA. We used 4 fold cross validation method and reported average recognition performance from the four folds. The matching scores of uMLBP+LDA method is computed by:

$$S_{uMLBP+LDA}(P,G) = exp^{-c_m \times dist(P,G)} \quad (7)$$

where c_m is a constant (=0.1) and $dist(P,G) = \sqrt{(F_P - F_G)^2}$. The transformation from distances to scores can be achieved in many different ways. We experimented with inverse and exponential functions and found that the use of exponential function gave a better mapping results. We were able to achieve significant performance improvement by using the LDA for the first time in periocular recognition studies, as it can be seen in **Table 1**. Also, mapping to LDA subspace reduced the size of uMLBP feature dimensions by 92.8% (uMLBP feature dimensions before and after mapping to LDA subspace is equal to 5900 and 425, respectively). On an embedded platform, this reduction in feature dimensions may lead to significant reduction in processing and implementation complexity.

2.2. Periocular Recognition System Based on Periocular Attributes

2.2.1. Database Acquisition and Labelling

In order to have a large dataset of faces with labelled attributes for the training, we used Extreme Picture Finder software (www.exisoftware.com/picture_finder). The software takes keywords, such as Asian people, male, people with thick eyebrows, etc., and searches the corresponding images using several search engines. In total, we downloaded 68,382 images. We used STASM [13] to detect face landmarks and normalize the facial images. After removing images with poor landmark detection, we had 39,612 face images with their 15 describable attributes, such as Ethnicity (African American, Caucasian, Indian and Asian), Eyebrow shape (Arched, Thick, Thin), Eye shape (narrow eyes), Bags under eyes, Eyeglasses, Gender (Male, Female), and having eye makeup. Examples of the positive and negative samples for the four classes i.e., old age, male, Caucasian ethnicity and thick eyebrows are shown in **Fig. 3**.

2.2.2. Feature Extraction and Dimension Reduction using Partial Least Squares

Using the detected landmarks, we cropped the periocular regions of all the face images in our labelled dataset of facial attributes. A set of k low-level feature extractors $\mathbf{f_{Ai}}$ are applied to the input periocular image I to form a feature set $\mathbf{f_A}(I) = \{\mathbf{f_{A1}}(I), \ldots., \mathbf{f_{Ak}}(I)\}$. We describe each extractor $\mathbf{f_{Ai}}$ in terms of different low-level features, which includes the mean and variance of RGB, intensity and HSV values, magnitude and orientations of edge gradients, as shown in **Fig. 3**. For eyebrows related periocular attributes, we extracted the features from the eyebrows region only, represented by two blue boxes in **Fig. 3**.

The use of many types of low-level features increase the size of the overall feature vectors, which resulted in over-fitting of the SVM models, hence in order to reduce the feature dimensions we experimented several methods such as LDA, Principle Component Analysis (PCA), and partial least squares (PLS), out of which PLS performed the best.

PCA, LDA and PLS are representative methods used for feature dimensionality reduction. In case of uMLBP+LDA, our main purpose of using LDA is to reduce feature dimensions, while in case of PA+PLS, our purpose of using PLS is not only to reduce the feature dimensions, but also learn attribute classifiers. The PCA only utilize the variance of all the features without considering the class labels. The LDA takes class labels into account, and it tries to find optimal projection by maximizing the ratio of between-class and within-class scatter. On the other hand, PLS uses the annotation labels not only to maximize inter-class variance, but also maximize the fit and minimize the misfit [29]. This means we take into

account the classes and try to reduce the dimension while maximizing the separation of classes, but at the same time we do not let the misfit to increase, which is ideal in classification scenario. We believe the superior performance of PLS on attribute features compared with those of LDA is due to the low dimensionality of the attribute features.

To understand PLS, let **F** denote a matrix containing *o* samples of *m*-dimensional feature vectors, and similarly let $\boldsymbol{l}$ represent the class labels for *o* samples. In order to build an indirect relationship between **F** and $\boldsymbol{l}$, PLS searches for a set of components (called latent vectors) that performs a simultaneous decomposition of **F** (*o* × *m*) and $\boldsymbol{l}$ (*o* × 1) with the constraint that these components cover the covariance between **F** and $\boldsymbol{l}$ as much as possible. PLS decomposes the zero-mean matrix **F** and zero-mean vector $\boldsymbol{l}$ into:

$$\mathbf{F} = \mathbf{U}\mathbf{O}^T + \mathbf{E}, \qquad \boldsymbol{l} = \mathbf{V}\boldsymbol{q}^T + \boldsymbol{h} \tag{8}$$

where **U** and **V** (*o* × *p* matrices) contains *p* latent vectors, **O** (*m* × *p*) and $\boldsymbol{q}$ (1 × *p*) represent the loadings and **E** (*o* × *m*) and $\boldsymbol{h}$ (*o* × 1) are the residuals. The PLS method, using the nonlinear iterative partial least squares (NIPALS) algorithm [19], constructs a set of weight vectors (or projection vectors) $\mathbf{W}_{\boldsymbol{PLS}} = \{\boldsymbol{w}_{\boldsymbol{PLS\_1}}, \boldsymbol{w}_{\boldsymbol{PLS\_2}}, \ldots, \boldsymbol{w}_{\boldsymbol{PLS\_p}}\}$ such that:

$$[cov(\boldsymbol{u_i}, \boldsymbol{v_i})]^2 = \max_{|\boldsymbol{w}_{\boldsymbol{PLS\_i}}|=|\boldsymbol{c}_{\boldsymbol{PLS\_i}}|=1} [cov(\mathbf{F}\boldsymbol{w}_{\boldsymbol{PLS\_i}}, \boldsymbol{l}\boldsymbol{c}_{\boldsymbol{PLS\_i}})]^2 \tag{9}$$

where $\boldsymbol{u_i}$ is the i^{th} column of matrix **U**, $\boldsymbol{v_i}$ the i^{th} column of matrix **V**, $cov(\boldsymbol{u_i}, \boldsymbol{v_i})$ is the sample covariance between latent vectors $\boldsymbol{u_i}$ and $\boldsymbol{v_i}$, and $|\boldsymbol{w}_{\boldsymbol{PLS\_i}}|$, $|\boldsymbol{c}_{\boldsymbol{PLS\_i}}|$ represents the determinants of $\boldsymbol{w}_{\boldsymbol{PLS\_i}}$ and $\boldsymbol{c}_{\boldsymbol{PLS\_i}}$ respectively. Eq. 9 finds the latent vectors $\boldsymbol{u_i}$ *and* $\boldsymbol{v_i}$ by updating the weights, such that the covariance of **F** and $\boldsymbol{l}$ becomes the maximum. After getting latent vectors $\boldsymbol{u_i}$ *and* $\boldsymbol{v_i}$, **F** and $\boldsymbol{l}$ are deflated (confidence is lowered) by subtracting their rank-1 approximations based on $\boldsymbol{u_i}$ *and* $\boldsymbol{v_i}$. This procedure is iterated until desired number of latent vectors are extracted. The dimensionality reduction for periocular attributes is performed by projecting the low level features onto the weights $\mathbf{W}_{\boldsymbol{PLS}}$, obtaining the latent vector $\boldsymbol{z_i}$ as a result, which are then used to train the linear SVM-based attribute classifiers as shown in **Fig. 3**. To learn the PLS weights, we used all the features of the labelled dataset of PA and performed 10 fold cross validation.

2.2.3. Attribute Based Periocular Region Matching

Consider an attributes function $\mathbf{a_i}$ that maps an input image *I* to a real value a_i. Large positive values of a_i indicate the presence or strength of the i^{th} attribute, while negative values indicate its absence [11]. We trained a total of 15 linear SVM classifiers using vlfeat library [22]. A grid search approach was used to compute the value of cost parameter for training each SVM classifier. When the low-level feature vector of a test image is passed to the trained SVM models for PA classification, the module outputs a 1×15 attribute vector **ac**(*I*), as shown in **Fig. 3**.

In order to find the similarity between two periocular regions based on PA, we defined a periocular region matching method. This method compares the attribute vectors, **ac**(I_1) and **ac**(I_2) of two periocular images I_1 and I_2, and returns the attribute recognition decision vector **av**(I_1, I_2) by computing the similarity between the attributes of the two images, as shown in **Fig. 4**. Let a_i = $\mathbf{ac_i}(I_1)$ and b_i = $\mathbf{ac_i}(I_2)$ be the outputs of the i^{th} attribute (i =1,2,3, …,15) classifiers for each periocular image. Then, the term $\big||a_i| - |b_i|\big|$ represents the dissimilarity of the i^{th}

attribute in those two periocular images. The sign (+, -) of the product $a_i b_i$ signifies the presence or absence of that attribute. The attribute score (S_{PA}) is computed as:

$$S_{PA}(P,\ G) = \ exp^{-\ c_a \times d} + \ MC_{PA}(P, G) \tag{10}$$

where, $d = \left(\sum_i^{15} pa_i\right)\ \ (for\ pa_i \geq 0)$, $\mathrm{MC_{PA}}(P_i,\ G_j) = n/N$ (PA match count matrix), $pa_i = \mathrm{sign}(a_i) * \mathrm{sign}(b_i) * \left||a_i| - |b_i|\right|$, a_i and b_i denote the i^{th} attribute values of subject a and b, P_i and G_j represent probe and gallery indices, n is the number of attributes having same signs and N is total number of attributes. Value of c_a is constant and fixed as 0.25.

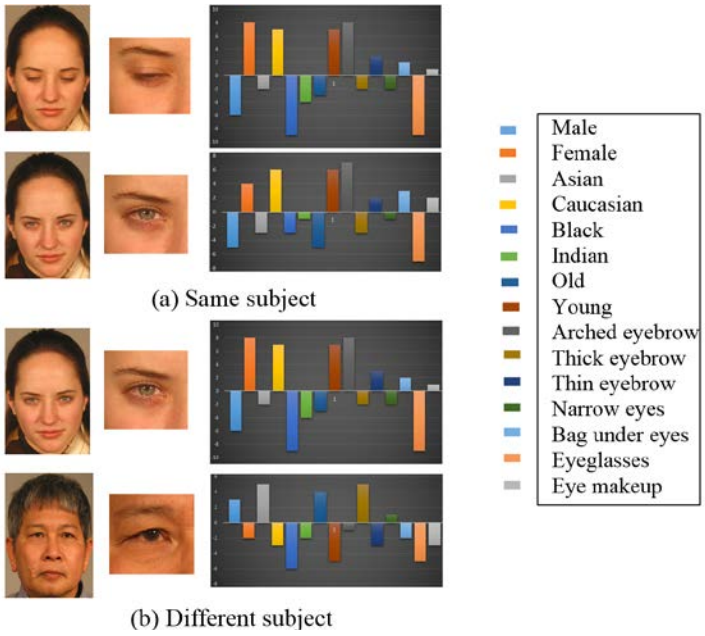


Fig. 4. Example of PA feature vector comparison of the same and different subject pairs

3. Experimental Results

We describe the database used in the experiments and then provide experimental results in three different subsections: (i) the general performance of the proposed system on the FRGC and Sogang Periocular Database (SPDB), (ii) the effect of distance variation on recognition accuracy, and (iii) the performance comparisons between the proposed method and two other state-of-the-art LBP-based methods [12][14].

3.1. SPDB and FRGC Database

Existing periocular databases have mostly addressed challenging factors regarding pose, expression, illumination, and blinking. We address distance as one of the major factors in our periocular database composition. The distance between a subject and a camera plays an important role in the level of perspective distortions, especially on mobile platforms. We have captured facial images at three different distances (0.5m, 1.0m and 1.5m) and five different facial poses (frontal, left, right, up, and down at ±20°) using the built-in

camera of a smartphone (Samsung Galaxy S3). The facial image resolution is 2448×3264 in JPEG format. The database is composed of 15,000 images of 50 different subjects (2 images at 5 different poses at 3 different distances for each of the 50 subjects).

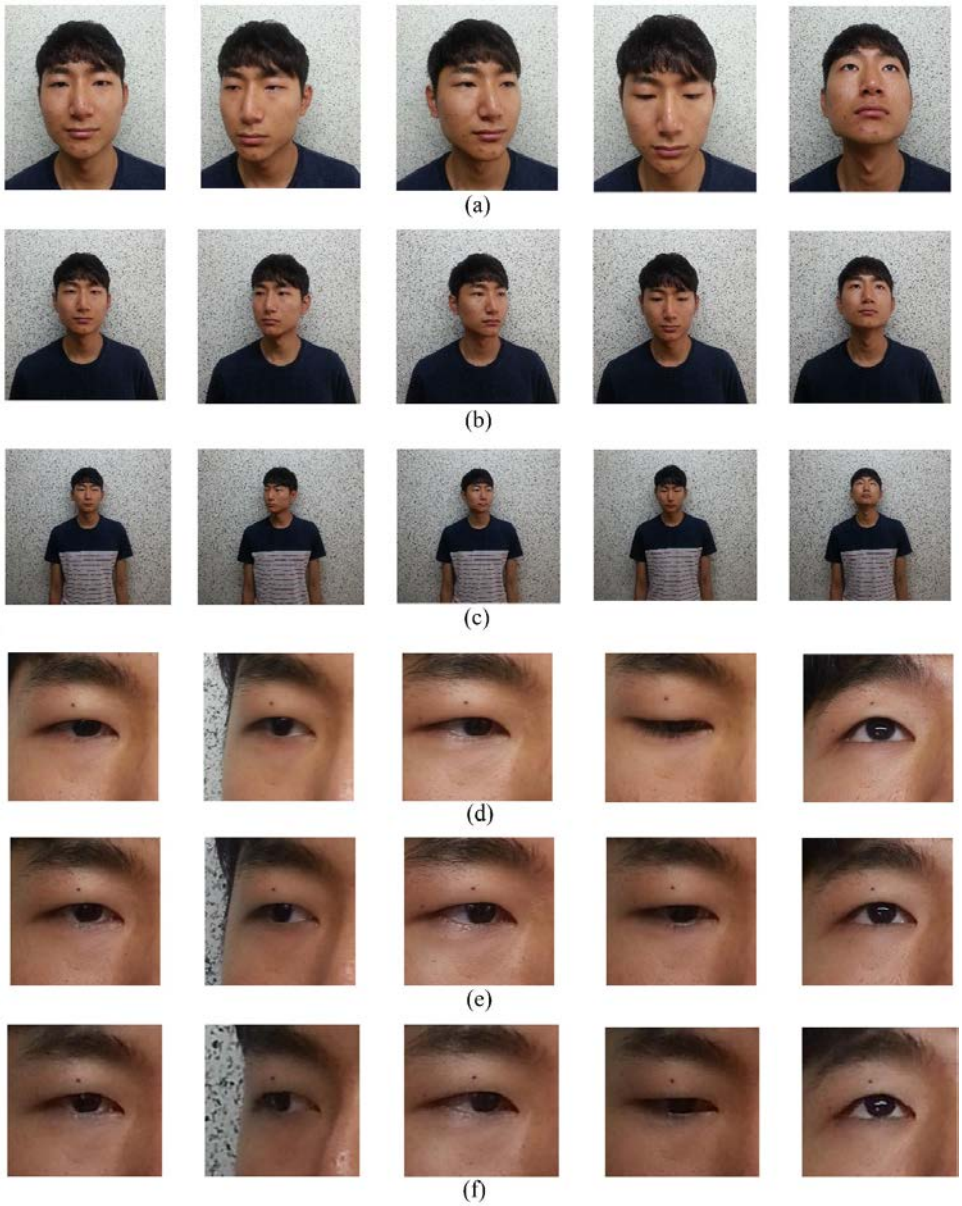


Fig. 5. Example images from the Sogang Periocular Database (SPDB). Facial images of frontal; left; right; up; and down at (a) 0.5m, (b) 1.0m and (c) 1.5m. Periocular images of frontal; left; right; up; and down at (d) 0.5m, (e) 1.0m and (f) 1.5m.

The periocular images captured at three different distances and various facial poses, makes SPDB a special database to study the characteristics of periocular biometrics in smartphone environments. Example images of SPDB and cropped periocular regions of a subject at three different distances are shown in **Fig. 5**. We also used a subset of the FRGC

ver. 2.0 database along with SPDB in the recognition experiments. The subset of FRGC contains 1136 frontal images of 568 subjects (2 frontal images per subject).

3.2. Performance Evaluation

3.2.1. Periocular recognition based on uMLBP and LDA learning

We performed a four-fold cross-validation to evaluate the subspace learning method on FRGC dataset. Moreover, we used the learned model of FRGC dataset and applied it to the SPDB feature space. The recognition accuracy of 98.37% in the FRGC database is the state-of-the-art performance based on the literature survey shown in **Table 1**. The performance on SPDB was evaluated using four different poses in query images (yaw and pitch at ±20°) and only frontal images enrolled in the gallery database (capture distance of 0.5m). Similarly as FRGC, the uMLBP with LDA showed the best performance in the SPDB database. Even though the recognition accuracy with SPDB containing pose variation is much lower than that of FRGC when only low level features were used, their performances are improved to similar level after applying LDA. The recognition performance on the FRGC database using uLBP and uMLBP with or without LDA are summarized in **Table 2**.

We used weighted score sum method to combine the matching scores of the PA with the other four primary methods. For each fusion, a weight $w \in [0,1]$ is assigned to primary method's score matrix and weight '1-w' is assigned to PA's score matrix. Finally, both the score matrices are added to get the final (fused) score (using eq. 7 and eq. 10) as:

$$S_{final} = w \times S_{uMLBP+LDA} + (1 - w) \times S_{PA} \tag{11}$$

Table 2 shows that the state-of-the-art recognition accuracy of 98.7% is achieved when we fuse the uMLBP+LDA method with PA+PLS method, for FRGC database.

Table 2. Recognition accuracies of four different approaches with the FRGC and SPDB databases.

Feature Type	Database			
	FRGC (2,272 images, leave-one-person-out* fashion)		**SPDB** (800 non-frontal images for probe, 200 frontal images for gallery)	
	Primary method	**Fusion with PA+PLS**	**Primary method**	**Fusion with PA+PLS**
uLBP**	92.25%	92.6% (w'=0.88)	86.1%	86.3% (w'=0.85)
uMLBP***	92.21%	92.4% (w'=0.90)	84.1%	85% (w'=0.85)
uLBP+LDA	96.7%	96.83 % (w'=0.98)	95.5%	95.71% (w'=0.99)
uMLBP+LDA	98.37%	98.7% (w'=0.99)	97.5%	97.8% (w'=0.99)

∗ We take one image as a probe from the first to the last in turn, and other images of all other subjects are considered as gallery set.
∗∗ Patch size 74×74 pixels, radius 12 pixels
∗∗∗ Patch size 74×74 pixels, radius 1, 3, 6, and 12 pixels
w' = optimal weight

It can be seen that uMLBP alone showed a little worse performance than uLBP, but the use of LDA subspace learning improved the uMLBP accuracy more than the accuracy of uLBP+LDA. We think this is due to the fact that although uMLBP contains more discriminative texture information than uLBP, it also induce some repetitive information, causing between-to-within class scatter ratio to drop. **Fig. 6** depicts the visual representation of uMLBP features of four periocular regions. We selected two images of two different subjects from SPDB and FRGC dataset. In this example the uMLBP features are computed with patch radius R=3, 12 and 21, and then concatenated. In **Fig. 6** we highlighted some of the prominent features (with green boxes), which are repetitive with little variations. LDA, on the other hand, uses this redundancy information in more effective way to learn the high dimensional reference system or the Eigen vectors, and hence result in better recognition accuracy than other methods with no LDA.

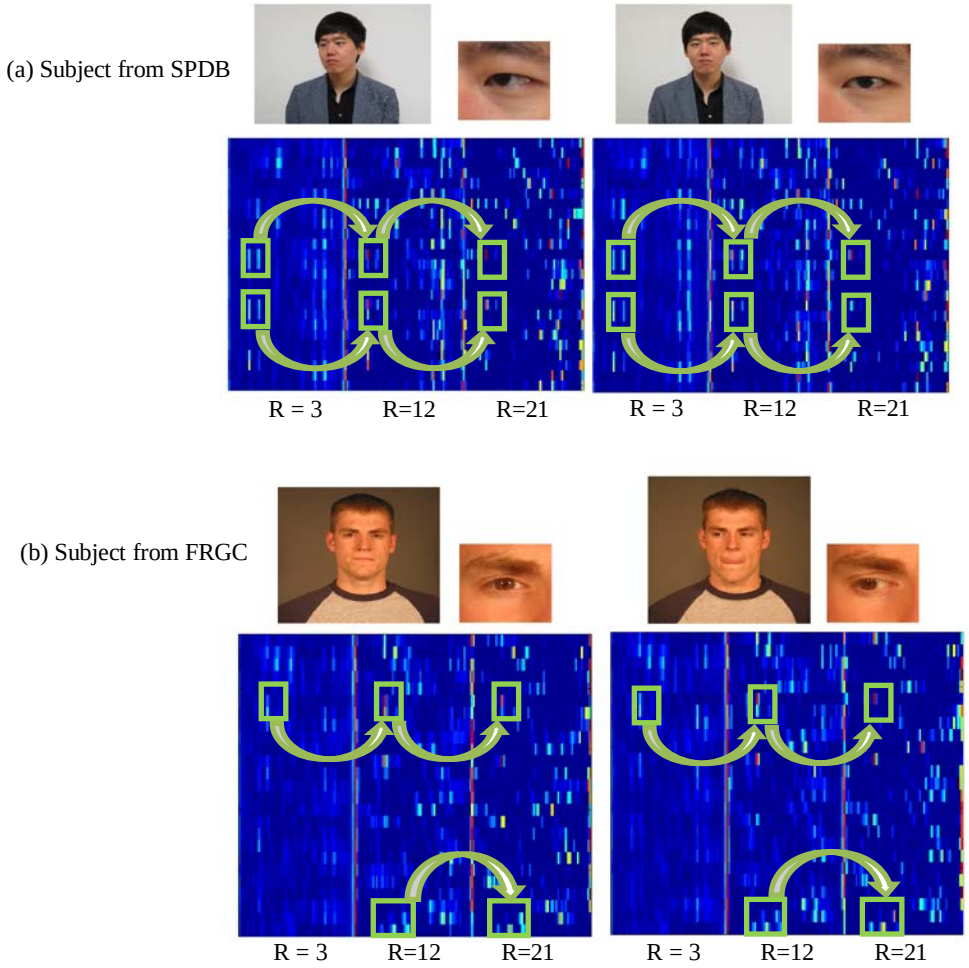


Fig. 6. uMLBP feature visualizations of the periocular images of the two subjects from SPDB and FRGC datasets. The uMLBP feature is computed with patch radii R = 3, 12 and 21.

Table 3. Effect of different features on between-to-within class scatter ratio

Method	$\|S_B\|$	$\|S_W\|$	$S_{BW} = \|S_B\|/\|S_W\|$
uLBP	52.33	73.87	0.71
uMLBP	26.56	40.48	0.66
uLBP+LDA	87.76	6.48	13.53
uMLBP+LDA	88.37	3.81	23.21

To support this hypothesis, we computed the between-class and within-class scatter matrix for the four cases, i.e., uLBP, uLBP+LDA, uMLBP and uMLBP+LDA. In multivariate analysis, it is often required to estimate a single representative value of the complete scatter matrix. Many researchers [25],[26],[27] proposes $S = tr(\mathbf{\Sigma})/m$, where tr operator represent trace operation applied on the scatter matrix $\mathbf{\Sigma}$, and m represents the dimension of the feature vector. Here S is also equal to the sum of Eigen values of the scatter matrix. We calculated the ratio of the sum of the Eigen values of between-class scatter matrix to that of within-class scatter matrix for the four cases, S_{BW}, and compared the results in **Table 3**. In **Table 3**, $|S_B|$ and $|S_W|$ denote the magnitude (single representative value) of the between-class scatter and within-class scatter respectively. It is evident from **Table 3** that S_{BW} of uLBP (S_{BW_uLBP}) is little greater than the S_{BW} of uMLBP (S_{BW_uMLBP}), due to which the rank-1 accuracy of uMLBP is little worse than that of uLBP. However, the S_{Bw} of uMLBP+LDA is significantly greater than that of uLBP+LDA, which causes significant improvement in the rank-1 accuracy of uMLBP+LDA, as shown in **Table 2**.

3.3. Effect of Distance Variation on Periocular Recognition Accuracy

Next, we evaluated the effects of distance on the recognition performance. For this experiment we used the SPDB, having facial images at three different distances (0.5m, 1.0m and 1.5m), both for probe and gallery and evaluated the recognition accuracy in every combination (c.f., **Table 4**). We used uMLBP features combined with LDA for the features, and used only frontal images to evaluate the distance effect separately from the pose variation effect. The recognition results showed similar performances at different combinations of capture distances of probe and gallery images, but the uses of gallery and probe images captured at the same distances showed the best accuracies. Based on the results in **Table 4**, we conclude that the periocular recognition system can be used at closer distance (close to the normal operating distance of a smartphone), and the gallery database can be constructed using images captured at 0.5m distance.

Table 4. Recognition accuracies with respect to the capture distance using the uMLBP+LDA method.

		Probe		
Gallery	**Distance**	**0.5m**	**1.0m**	**1.5m**
	0.5m	99.5%	98.0%	97.5%
	1.0m	98.0%	99.0%	97%
	1.5m	97.5%	97.0%	97.5%

3.4. Comparison with State-of-the-art LBP Based Features

Even though the proposed system showed the best performance on the FRGC database as shown in the previous subsections, the proposed method (uMLBP and LDA) is not much different from the conventional methods. However, as can be seen in **Table 1**, conventional periocular recognitions mostly rely on LBP or uLBP rather than uMLBP, which can better handle the scale variations. Also, no previous approaches utilized LDA. Therefore, we believe the combination of uMLBP and LDA is proposed for the periocular recognition for the first time in this paper, and it is well-suited for periocular recognition.

For additional evaluation of the proposed method, we compared our method with two other state-of-the-art LBP-based features [12][14]. Liao [12] proposed Multi-scale Block Local Binary Pattern (MB-LBP) that extend the MLBP by considering blocks instead of each pixel as the unit values. The MB-LBP is also extended by considering the concepts similar with uniform-LBP, and then combined with AdaBoost to select effective features for the classifier construction, which named as Statistically Effective Multi-scale Block Local Binary Pattern (SEMB-LBP). On the other hand, Nikisins et al. [14] extended the MLBP by applying feature weighting methods.

Table 5. Comparison of the performance in facial recognition between [12] and our implementation of the proposed method in [12]. The FERET database is used for the evaluation.

	Liao [12]	**Our implementation of [12]**
Rank-1 identification accuracy	91.81%	96.34% (*PS*=56×72,*s*=9)*
		97.20% (*PS*=56×56,*s*=9)
		96.99% (*PS*=32×32,*s*=9)
#images and #subjects used in training and testing	12,776 images of 222 subjects for training, 932 images of 466 subjects for testing	3,772 images of 275 subjects for training, 930 images of 465 subjects for testing**

∗ *PS* and *s* represent patch size and the scale of SEMB-LBP, respectively.

∗∗ There is no detailed explanation about the database construction for the experiment in [12]. We selected the number of images and subjects to construct the database similarly with that of [12].

We first implemented the methods of [12] and [14] because their codes are not publically available. We carefully followed the descriptions in those articles and compared the performance of our implementations of [12] and [14] using the same facial image databases (FERET). Then, we compared the methods of [12][14] and the proposed method on the periocular database (FRGC).

Table 5 shows the comparison of the performances between [12] and our implementation of [12]. Since the detailed explanation of the dataset construction is not provided in [12], our dataset is slightly different from that in [12]. However, the number of images and subjects in the test dataset is almost the same. It can be seen that our implementation showed superior identification accuracy than that in [12]. Similarly, **Table 6** summarizes the comparison between [14] and our implementation of [14]. Our implementation showed very similar performances as those of [14], and even showed better performances in some parameter settings.

Table 6. Comparison of the performance in facial recognition between [18] and our implementation of the proposed method in [18]. The FERET database is used for the evaluation.

	Nikisins [14]	Our implementation of [14]	Methods
Rank-1 identification accuracy	96.80%	97.11%	MSLBP*
	97.80%	97.31%	MSLBP+MF
	98.10%	98.50%	MSLBP+MF+FW
	99.20%	98.64%	MSLBP+MF+BF**
#images and #subjects used in training and testing	1,986 images of 993 subjects both for training and testing	2,010 images of 1,005 subjects both for training and testing***	

* The parameters used are L=11, K=8, NR=3 where L, K, and NR represent sampling radius, number of patches, and number of radii, respectively.

** MF, FW, and BW represent mean filter, feature weighting, and block weighting, respectively.

*** There is no detailed explanation about the database construction for the experiment in [14]. We selected the number of images and subjects to construct the database similarly with that of [14].

Finally, **Table 7** shows the periocular recognition performance comparisons between our implementations of [12] and [14] and our proposed methods. It is evident that the proposed method (i.e., uMLBP+LDA+PA+PLS) provides the best recognition accuracy. The recognition performances of uLBP, uMLBP, uLBP+LDA, and uMLBP+LDA are also provided in **Table 7** for the comparison.

Table 7. Comparison of the performances of various approaches in periocular recognition. The FRGC database is used for the evaluation.

	Our implementation of [12]	Our implementation of [14]	Proposed method
Rank-1 identification accuracy	88.86% (PS=56×56,s=9)	96.61% (MSLBP+MF+BF)	98.7% (uMLBP+LDA+PA+PLS)

4. Conclusions and Future Work

In this paper, we proposed a state-of-the-art periocular recognition method using uMLBP, LDA, and periocular attributes. We also collected a new database (SPDB) to investigate the effect of distance and pose variation in periocular recognition on mobile platforms. From the experimental results, we observed that the performance of the periocular recognition system did not degrade much depending on the capture distance. Besides the uMLBP+LDA based periocular recognition method, we also proposed the use of 15 periocular attributes for more robust recognition. We expect that the use of more number of periocular attributes (e.g., colour of the pupil, monolid eyes, almond eyes, upturned eyes, downturned eyes, hooded eyes, wrinkles in periocular region, etc.) can

further improve the recognition accuracy. Future work should also include reducing the negative effects of the pose and distance variations and finding deformation invariant descriptors or learning methods.

Acknowledgment

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT) (2014-0-00077, Development of global multi-target tracking and event prediction techniques based on real-time large-scale video analysis)

References

[1] J. Adams, D. L. Woodard, G. Dozier, P. Miller, K. Bryant, and G. Glenn, "Genetic-based type ii feature extraction for periocular biometric recognition: Less is more," in *Proc. of Int. Conf. on Pattern Recognition*, pp. 205-208, 2010. Article (CrossRef Link).

[2] S. Bakshi, P. K. Sa, and B. Majhi, "A novel phase-intensive local pattern for periocular recognition under visible spectrum," *Biocybernetics and Biomedical Engineering*, vol. 35, no. 1, pp. 30-44, 2015. Article (CrossRef Link).

[3] P. N. Belhumeur, J. P. Hespanha, and D. 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. Article (CrossRef Link).

[4] S. Bharadwaj, H. Bhatt, M. Vatsa, and R. Singh, "Periocular biometrics: When iris recognition fails," in *Proc. of IEEE Int. Conf. on Biometrics: Theory, Applications, and Systems*, pp. 1-6, 2010. Article (CrossRef Link).

[5] G. Dozier, K. Purrington, K. Popplewell, J. Shelton, T. Abegaz, K. Bryant, J. Adams, D. L. Woodard, and P. Miller. 2011. "GEFeS: Genetic & evolutionary feature selection for periocular biometric recognition," in *Proc. of IEEE Workshop on Computational Intelligence in Biometrics and Identity Management*, pp. 152-156, 2011. Article (CrossRef Link).

[6] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 4-20, 2004. Article (CrossRef Link).

[7] S. M. Jonathon, and B. V. K. V. Kumar, "What is a "good" periocular region for recognition?" in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition Workshops*, pp. 117-124, 2013. Article (CrossRef Link).

[8] A. Joshi, A. K. Gangwar, and Z. Saquib, "Person recognition based on fusion of iris and periocular biometrics," in *Proc. of Int. Conf. on Hybrid Intelligent Systems*, pp. 57-62, 2012. Article (CrossRef Link).

[9] F. Juefei-Xu, K. Luu, M. Savvides, T. Bui, and C. Suen. 2011. "Investigating age invariant face recognition based on periocular biometrics," in *Proc. of Int. Joint Conf. on Biometrics*, pp. 1-7, 2011. Article (CrossRef Link).

[10] F. Juefei-Xu and M. Savvides, "Unconstrained periocular biometric acquisition and recognition using COTS PTZ camera for uncooperative and non-cooperative subjects," in *Proc. of IEEE Workshop on Applications of Computer Vision*, pp. 201-208, 2012. Article (CrossRef Link).

[11] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, "Attribute and Simile Classifiers for Face Verification," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pp. 1962-1977, 2011. Article (CrossRef Link).

[12] S. Liao, and S. Z. Li, "Learning multi-scale block local binary patterns for face recognition," in *Proc. of Int. Conf. on Biometrics*, pp. 828-837, 2007. Article (CrossRef Link).

[13] S. Milborrow and F. Nicolls, "Active Shape Models with SIFT Descriptors and MARS," in *Proc. of Int. Conf. on Computer Vision Theory and Application*, pp. 380-387, 2014.
[14] O. Nikisins, "Weighted multi-scale local binary pattern histograms for face recognition," in *Proc. of Int. Conf. on Applied Mathematics and Computational Methods*, pp. 76-81, 2013.
[15] T. Ojala, M. Pietikainen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51-59, 1996. Article (CrossRef Link).
[16] U. Park, R. Jillela, A. Ross, and A. K. Jain, "Periocular biometrics in the visible spectrum," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 1, pp. 96-106, 2011. Article (CrossRef Link).
[17] U. Park, A. Ross, and A. K. Jain, "Periocular biometrics in the visible spectrum: a feasibility study," in *Proc. of IEEE Int. Conf. on Biometrics: Theory, Applications, and Systems*, pp. 1-6, 2009. Article (CrossRef Link).
[18] G. Santos, E. Grancho, M. V. Bernardo, and P. T. Fiadeiro, "Fusing Iris and Periocular Information for Cross-sensor recognition," *Pattern Recognition Letters*, vol. 57, pp. 52-59, 2015. Article (CrossRef Link).
[19] W. R. Schwartz, A. Kembhavi, D. Harwood, and L. S. Davis, "Human Detection Using Partial Least Squares Analysis," in *Proc. of Int. Conf. on Computer Vision*, pp. 24-31, 2009. Article (CrossRef Link).
[20] J. L. Starck, F. Murtagh, and A. Bijaoui, "Image restoration with noise suppression using the wavelet transform," *Astronomy and Astrophysics*, vol. 288, no. 1, pp. 342-348, 1994.
[21] M. Uzair, A. Mahmood, A. Mian, and C. McDonald, "Periocular biometric recognition using image sets," in *Proc. of IEEE Workshop on Applications of Computer Vision*, pp. 246-251, 2013. Article (CrossRef Link).
[22] A. Vedaldi, and B. Fulkerson, "{VLFeat}: An Open and Portable Library of Computer Vision Algorithms." VLFeat Library. http://www.vlfeat.org/.
[23] D. L. Woodard, S. Pundlik, J. Lyle, and P. Miller, "Periocular region appearance cues for biometric identification," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition Workshops*, pp. 162-169, 2010. Article (CrossRef Link).
[24] D. L. Woodard, S. Pundlik, P. Miller, R. Jillela, and A. Ross, "On the fusion of periocular and iris biometrics in non-ideal imagery," in *Proc. of Int. Conf. on Pattern Recognition*, pp. 201-204, 2010. Article (CrossRef Link).
[25] L. Dumbgen, "On Tyler's M-functional of scatter in high dimension," Annals of the Institute of Statistical Mathematics, vol. 50, pp. 471-491, 1998.
[26] E. Ollila, T. P. Hettmansperger, and H. Oja, "Affine equivariant multivariate sign methods," *Preprint, University of Jyvaskyla*, 2004. Article (CrossRef Link).
[27] D. E. Tyler, "Robustness and efficiency properties of scatter matrices," *Biometrika*, vol. 70, pp. 411-420, 1983. Article (CrossRef Link).
[28] S. Bakshi, P. K. Sa, and B. Majhi, "Phase Intensive Global Pattern for Periocular Recognition," in *Proc. of IEEE India Conference (INDICON)*, pp. 1-5, 2014. Article (CrossRef Link).
[29] T. D. Bie, N. Cristianini, and R. Rosipal, "Eigenproblems in Pattern Recognition," *Applications in Pattern Recognition, Computer Vision, Neuralcomputing, and Robotics*, Ch. 5, part II, pp. 129-167, 2005, Springer Berlin Heidelberg.

Zahid Ali received his B.E in Electronic Engineering from NED University in 2009 and M.E in Computer Science and Engineering from Chosun University in 2013. He was awarded Global IT Scholarship by Republic of South Korea in 2013. He is currently a Ph.D. candidate in the department of Computer Science and Engineering at Sogang University, where he is also working as research assistant in Computer Vision and Image Processing lab. His main research interests are at the intersection of computer vision and machine learning, developing techniques for unconstrained face recognition and objects detection using Convolutional Neural Networks (CNN's).

Unsang Park received his B.S. and M.S. degrees from the Department of Materials Engineering, Hanyang University, Seoul, Korea, in 1998 and 2000, respectively. He received his M.S. and Ph.D. degrees from the Department of Computer Science and Engineering, Michigan State University, MI, USA in 2004 and 2009, respectively. He has been an assistant professor in the Department of Computer Science and Engineering at Sogang University since 2012. His research interests include pattern recognition, image processing, computer vision, and machine learning.

Jengho Nang received his Ph.D. and M.S. degrees in Computer Science from Korea Advanced Institute of Science and Technology (KAIST), Daejon, Korea, in 1992 and 1988, respectively, and his B.S. degree in Computer Science from Sogang University, Seoul, Korea, in 1986. He has been a professor of Computer Science and Engineering Department, Sogang University since 1993. His research interests include multimedia system, deep neural network, and internet technology.

Jeong-Seon Park received her B.S. and M.S. degrees from the Department of Computer Science, Chungbuk National University, Chungcheongbuk-do, Korea, in 1994 and 1996, respectively. She received her B.S. Ph.D. degree from the Department of Computer Science from Korean University, Seoul, Korea in 2005. Since 2005, she is a professor in the Department of Multimedia, Chonnam National University, Chonnam, Korea. Her research interests include image processing, pattern recognition and IT convergence application.

Taehwa Hong received the B.S. degree in electrical engineering, in 1997, the M.S. degree in electrical engineering, in 1999, and the Ph.D. degree in electrical engineering, in 2006, from Yonsei University, Seoul, Korea. He currently works as a principal engineer of Mobile Communications Business at Samsung Electronics Co., Ltd. His research interests are in image processing, computer vision, image recognition and machine learning.

Sungjoo Park received the B.S. and M.S. degrees from Kyunpook National University, Daegu, Korea, and Ph.D. degree from Kwangwoon University, Seoul, Korea, from the Electronics & Communications Department. Currently, he is a director at Digital Media Research Center of Korea Electronics Technology Institute. His research interests include A/V signal processing, media service platform, and the convergence of digital broadcasting and telecommunication.