Robustness of Face Recognition to Variations of Illumination on Mobile Devices Based on SVM

Gi Pyo Nam[1], Byung Jun Kang[2] and Kang Ryoung Park[1]
[1]Division of Electronics and Electrical Engineering, Dongguk University, Seoul, Rep. of Korea
Biometrics Engineering Research Center (BERC)
[e-mail: {oscar1201, parkgr}@dongguk.edu]
[2]ETRI, Daejeon, Rep. of Korea,
[E-mail: kangbyj@etri.re.kr]
*Corresponding author: Kang Ryoung Park

Received November 9, 2009; revised January 17, 2010; accepted January 27, 2010; published February 27, 2010

Abstract

With the increasing popularity of mobile devices, it has become necessary to protect private information and content in these devices. Face recognition has been favored over conventional passwords or security keys, because it can be easily implemented using a built-in camera, while providing user convenience. However, because mobile devices can be used both indoors and outdoors, there can be many illumination changes, which can reduce the accuracy of face recognition. Therefore, we propose a new face recognition method on a mobile device robust to illumination variations. This research makes the following four original contributions. First, we compared the performance of face recognition with illumination variations on mobile devices for several illumination normalization procedures suitable for mobile devices with low processing power. These include the Retinex filter, histogram equalization and histogram stretching. Second, we compared the performance for global and local methods of face recognition such as PCA (Principal Component Analysis), LNMF (Local Non-negative Matrix Factorization) and LBP (Local Binary Pattern) using an integer-based kernel suitable for mobile devices having low processing power. Third, the characteristics of each method according to the illumination variations are analyzed. Fourth, we use two matching scores for several methods of illumination normalization, Retinex and histogram stretching, which show the best and 2nd best performances, respectively. These are used as the inputs of an SVM (Support Vector Machine) classifier, which can increase the accuracy of face recognition. Experimental results with two databases (data collected by a mobile device and the AR database) showed that the accuracy of face recognition achieved by the proposed method was superior to that of other methods.

Keywords: Face recognition, illumination normalization, SVM, mobile device

This work was supported by the Korea Science and Engineering Foundation (KOSEF), the Korean government through the Biometrics Engineering Research Center (BERC) at Yonsei University [R112002105070020(2009)].

DOI: 10.3837/tiis.2010.01.002

1. Introduction

Recently, mobile devices such as smart phones, UMPCs (ultra-mobile PCs) and net-books have become increasingly popular. However, the security issues of mobile devices have been brought into question, since it is often necessary to protect private information and content. Another issue is that internet and banking transactions via mobile devices have also increased. In conventional systems, PINs (Personal ID Numbers), passwords or secure USB (Universal Serial Bus) keys have been used for authenticating users. However, these methods are risky, since keys and passwords can be stolen or forgotten. Biometric technology has been used to overcome these problems.

Biometric systems identify users by using unique characteristics such as fingerprints, irises, palm veins and faces. Nowadays, biometric systems are used for door access control, immigration administration and user authentication of ATMs (Automatic Teller Machines). Biometric technologies have been also adopted on mobile devices [1][2][3][4][5]. Face recognition on mobile devices has been demonstrated by Qian *et al.* [1] and Hadid *et al.* [2]. Czyz *et al.* [3] showed the performance degradations of face and speech identification on mobile devices. Dae Jong Lee *et al.* presented research about faces and signatures recognition on mobile devices [4]. In police applications, mobile biometric systems using faces or fingerprints are actually being used in the Los Angeles police department [5].

Among biometrics, face recognition has been widely used for authentication, because it can be easily implemented using a built-in camera of a mobile device, while providing user convenience. Though fingerprints, palm veins and iris recognition show better accuracy than face recognition, they have shortcomings in that they are difficult to apply to small sized mobile devices, because they require additional image capturing devices. Though face recognition can be easily used on mobile devices, accuracy is often degraded by illumination variations. Especially, because mobile devices can be used in both indoor and outdoor environments, there can be many illumination changes, which can reduce accuracy. To overcome these problems, we propose a new face recognition method robust to illumination variations on mobile devices.

There has been a lot of previous researches about illumination normalization of face images. Xie *et al.* proposed local illumination normalization based on the Lambertian model [6]. However, their method of local region segmentation generally takes too much processing time for the CANDIDE model. Qing *et al.* improved face recognition accuracy under varying lighting conditions by using the Gabor phase probabilistic model [7]. Shan *et al.* proposed GIC (Gamma Intensity Correction) with HE (Histogram Equalization) and QIR (Quotient illumination Relighting) [8]. They used the GIC method to normalize the overall image intensity and eliminated the side light by using HE. In addition, the QIR method normalized illumination with Quotient illumination obtained by the bootstrap method from face images captured under various lighting conditions. Nam *et al.* proposed a method using adaptive filters such as the Retinex filter, histogram equalization filter and the ends-in contrast stretching filter [9]. Wang *et al.* proposed the SQI (Self-Quotient Image) method to normalize illumination and eliminate shadows [10]. However, all of these methods take too much processing time to be applied to mobile devices with low processing speeds. In addition, the performance enhancement achieved by using the single illumination normalization method is limited. Therefore, we propose a new face recognition method on mobile device robust to the illumination variations by combining two methods of illumination normalization based on an

SVM (Support Vector Machine). We propose a method of combining two matching scores (Retinex + PCA + LBP and histogram stretching + PCA + LBP) using an SVM, which can increase the accuracy of face recognition. By using two methods of illumination normalization (Retinex and histogram stretching), the proposed method can fully cope with various illumination changes on face images. Considering mobile platforms of low processing power, we used the integer-based Retinex and PCA methods, which have faster processing speeds compared to methods based on floating point operations.

The rest of this paper is organized as follows. In section 2, we explain the proposed method. Experimental results and conclusions are described in sections 3 and 4.

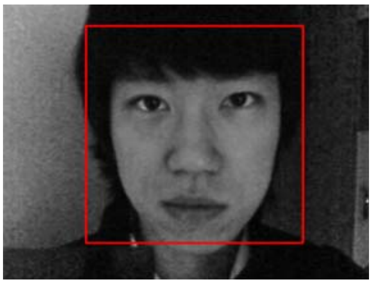
2. Proposed System

2.1 Acquiring face images

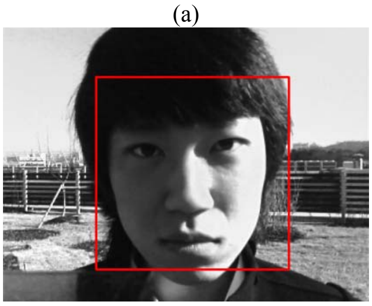
First, we capture a facial image with a built-in camera of an UMPC (Ultra-Mobile Personal Computer). Most mobile systems include a built-in visible camera and face images can be obtained without using additional camera devices or hardware, which is a great advantage of face recognition on mobile devices. Then, the face and eye regions are detected using the Adaboost face detector. Details are provided in the next section.

2.2 Face and eye detection based on Adaboost

To detect the face and eye regions, we use the Adaboost algorithm. The Adaboost algorithm generally constructs a strong classifier by combining several weak classifiers [11] [12]. This algorithm takes a long time to train weak classifiers, but it is very fast in case of testing. Based on size filtering, the facial regions that are incorrectly detected are removed. **Fig. 1** shows some examples of the detected facial regions. Because face recognition on mobile devices can be used in both indoor and outdoor environments, some complicated backgrounds may be included in the detected facial box shown in **Fig. 1**, which can greatly reduce face recognition accuracy.



(a)



(b)

Fig. 1. The detected facial regions using the Adaboost face detector (a) indoors (b) outdoors

To overcome this problem, after detecting the face region, two eye points are located by using an eye detection method based on the Adaboost algorithm in the restricted area of the face region. Based on the detected eye positions, the face region is redefined, as shown in **Fig. 2**.

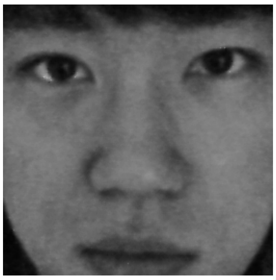


Fig. 2. The redefined facial region based on the detected two eye positions

2.3 Preprocessing

2.3.1 Size normalization

While capturing the face images of the same person, there are size variations of each captured facial region due to differences in Z distances between each user's face and the camera lens. Therefore, the size of the redefined facial region is normalized as 32×32 pixels, as shown in **Fig. 3**.

Fig. 3. The facial region after size normalization

2.3.2 Illumination normalization

Because mobile devices can be used in both indoor and outdoor environments, there can be large illumination variations, which can reduce overall face recognition accuracy. To overcome these problems, we compared face recognition performance with illumination variations on mobile devices for several illumination normalization procedures suitable for mobile devices with low processing power. These include the Retinex filter, histogram equalization and histogram stretching, as follows.

1) Retinex filter

The Retinex filter is generally used for illumination normalization and improves the contrast and brightness of an image [13]. A detailed explanation is as follows. The image intensity ($L(x, y)$) can be represented as the product of the illumination value ($I(x, y)$) and the reflectance ratio ($r(x, y)$) [14]:

$$L(x,y) = I(x,y) \times r(x,y) \tag{1}$$

$I(x, y)$ means the amount of incident light at the (x, y) position. $r(x, y)$ represents the reflectance ratio for the objects with respect to the incident light at the (x, y) position, representing a face image without illumination variations. Based on Eq. (1), illumination variations are reduced by eliminating the illumination ($I(x, y)$) from the image intensity ($L(x, y)$). We use the logarithm described by Eq. (2) and Eq. (3).

$$\log L(x,y) = \log(I(x,y) \times r(x,y)) \tag{2}$$

$$\log r(x,y) = \log L(x,y) - \log I(x,y) \tag{3}$$

The illumination ($I(x, y)$) is estimated by the convolution operation of the Gaussian filter ($F(x, y)$) and an image ($L(x, y)$), as described by Eq. (4) and (5) [9].

$$R(x,y) = \log r(x,y) = \log L(x,y) - \log(L(x,y) * F(x,y)) \tag{4}$$

$$F(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{5}$$

Consequently, $R(x, y)$ ($\log r(x, y)$) represents the Retinex image. **Fig. 4** shows the resultant image of **Fig. 2** by Retinex filtering.

Fig. 4. The resultant image of **Fig. 2** by Retinex filtering

To reduce the processing time, we use the integer-based normalization algorithms for Retinex filtering. That is, in case of Retinex filtering, the original coefficients of the Gaussian filter by Eq. (5) are the floating-point values. So, we converted all the floating-point coefficients of the Gaussian filter by a 10-bit left shift (*'<<10' in C program code)* and this is a similar concept to multiplying by 1000 (about 2^{10})). We used these integer coefficients for the Gaussian convolution operation. Then, a 10-bit right shift was done for each of the pixel values by Gaussian filtering (*'>>10' in C program code*) in order to prevent an increase of amplitude by Gaussian filtering with the coefficients multiplied by 2^{10}.

2) Histogram equalization

Due to the low processing time and the ability to enhance images, histogram equalization has been widely used [9][14]. However, additional noises such as camera thermal noises can be generated in face images while increasing the contrast and brightness of the images. as shown in **Fig. 5**.

3) Histogram stretching

Histogram stretching improves the contrast and brightness values in an image based on a linear equation while histogram equalization does so in a nonlinear manner. **Fig. 6** shows the illumination normalized image obtained by histogram stretching.

Fig. 5. The illumination normalized image of **Fig. 2** obtained by histogram equalization

Fig. 6. The illumination normalized image of **Fig. 2** obtained by histogram stretching

2.4 Face recognition

While illumination normalization is performed in the preprocessing step, local shading or thermal noise factors still remain in the face images, as shown in **Figs. 4**, **5**, and **6**. To overcome this problem, we compared the performances of global and local methods of face recognition such as PCA (Principal Component Analysis), LNMF (Local Non-Negative Matrix Factorization) and LBP (Local Binary Pattern). Especially, we use the integer-based PCA, LNMF and LBP methods for considering mobile environments with low processing speed. In detail, we converted the floating-point eigenfaces into integer-based ones by using a 20-bit shift. This means that the floating-point values of the eigenfaces are converted by a 20-bit left shift (*'<<20' in C program code)* and this is a similar concept to multiplying by 1,000,000 (about 2^{20}) or to be precise 1,048,576. In addition, integer operations are performed by a 7-bit shift when calculating the eigen-coefficients. This means that the floating-point eigen-coefficients are converted by a 7-bit left shift (*'<<7' in C program code)* and this is a similar concept to multiplying by 100 (about 2^7) or to be precise 128. Here, the optimal numbers of bit shifts were determined by considering both the processing time and accuracy of face recognition [17]. The reason we used the bit shift operation is that its processing speed is faster than that of multiplication.

1) PCA method

The PCA method is the most popular face recognition method, because it is very good at representing and analyzing face data [15][16]. However, the performance of PCA is generally sensitive to illumination variations, because it uses global face information [6].

For the minimum error of face recognition (see **Table 1**), the proposed method selected 140 eigenfaces. In general, the PCA method generates floating-point eigenfaces and corresponding floating-point eigenvalues, which require a large processing time on mobile devices. So, we convert the floating-point eigenfaces into integer-based ones by using a 20-bit shift. In addition, integer operations are performed by a 7-bit shift when calculating the eigen-coefficients. Here, the optimal number of bit shifts was determined by considering both the processing time and accuracy of face recognition [17]. Consequently, 140 integer eigen-coefficients were required for representing one face image of 32×32 pixels, as shown in **Fig. 3**.

2) LNMF method

NMF (Non-negative Matrix Factorization) generally trains a parts-based face representation by using the non-negativity constraints instead of the orthogonality of PCA [18][19]. However, there is a problem in that the NMF has difficulty producing a factorization including local features. So, instead we used LNMF for obtaining the local face features. Unlike NMF, LNMF produces kernels with orthogonality [18]. For the minimum error of face recognition, LNMF used 100 kernels. Consequently, 100 eigen-coefficients of integer-type are required for representing one face image of 32×32 pixels (see **Fig. 3**).

3) LBP method

The LBP method generally obtains binary patterns from a non-parametric (3×3) pixels kernel [20]. As shown in **Fig. 7** and Eq. (6), (7), the binary patterns are determined by comparing the gray values of the eight surrounding pixels (“95”, “93”, “95”, “116”, “133”, “132”, “135”, “111”) with that of the center pixel (“112”). If the gray value of a surrounding pixel is greater than that of the center pixel, the binary value is “1”, otherwise it is “0”. In **Fig. 7**, the calculated binary patterns are “00011110”.

$$LBP(x_c, y_c) = \sum_{0}^{7} s(i_n - i_c)2^n \tag{6}$$

$$s(x) = \begin{cases} 1 & if \quad x \geq 0 \\ 0 & if \quad x < 0 \end{cases} \tag{7}$$

The (3×3) pixel kernel is moved in the face region of **Fig. 3**, sequentially as in the convolution operation. At each kernel position, eight bits are obtained. In detail, the binary patterns of a total of 7,200 bits (30 in the horizontal direction × 30 in the vertical direction × 8 bits) are obtained in one face image of 32×32 pixels, as shown in **Fig. 3**.

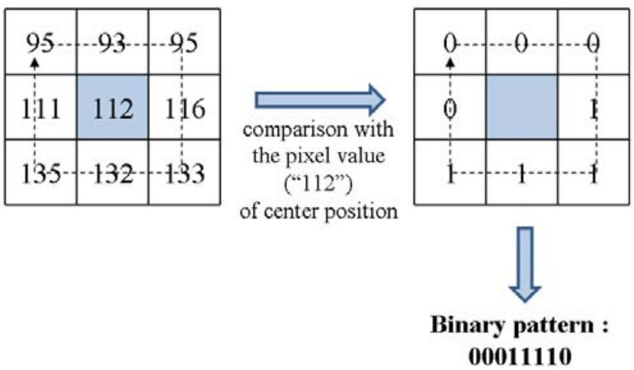


Fig. 7. The LBP method

4) Calculating matching scores

The Euclidian distance or the Hamming distance is used for calculating the dissimilarity between an input face and an enrolled one. Because the eigen-coefficients of the PCA and LNMF methods are represented as integers, the Euclidean distance (ED) is used for measuring dissimilarity. The Euclidean distance can be expressed as:

$$ED = \sqrt{\sum_{i=1}^{n} (A_i - B_i)^2} \tag{8}$$

where A_i and B_i are the eigen-coefficients of the enrolled image and the input image, respectively. n is the total number of eigen-coefficients of the PCA and LNMF methods. Although the Cosine distance can be also used, experimental results show that face recognition accuracy for the Euclidean distance is better than that for the Cosine distance.

Because the features of the LBP methods are shown as binary numbers, the Hamming distance *(HD)* is used for measuring dissimilarity, as shown in Eq. (9).

$$HD = \frac{\| BPA \otimes BPB \|}{T} \tag{9}$$

where $\otimes$ denotes the Boolean Exclusive-OR operator between the corresponding pairs of bits. T denotes the total number of corresponding pairs. BPA and BPB represent the whole binary pattern strings of an input image and the enrolled one, respectively.

2.5 Combining methods by using SVM

Since it is very difficult to estimate illumination changes accurately, there is no universal method of illumination normalization which can cope with all cases, so we propose a method which combines two methods of illumination normalization. For that, two matching scores by two methods of illumination normalization are used as the inputs of an SVM (Support Vector Machine) classifier.

The SVM discriminates two or more classes by using support vectors. It is a pattern-matching method used to solve two-class problems by determining the optimal linear decision hyper-plane. It is based on the concept of structural risk minimization, which determines the maximum distance between two classes [34]. In general, the SVM can be represented as (2) [34].

$$f(x) = \operatorname{sgn}\left(\sum_{i=1}^{k} \alpha_i y_i K(x, x_i) + b \right) \tag{10}$$

where k represents the number of data points and $y_i \in \{-1, 1\}$ represents the class label of training point x_i. The coefficients α_i can be found by solving a quadratic programming problem with linear constraints, and b is the bias. The SVM was then extended to non-linear decision surfaces by using a kernel function.

By training, it selects the support vectors and then decides the hyper-plane which discriminates two or more classes based on those support vectors. For the experiments, we used the winSVM program [33], which is based on mySVM [35]. One of the following five kernel functions is mainly used for the SVM classifier [21].

$$\text{dot kernel}: K(x, y) = x * y \tag{11}$$

$$\text{polynomial kernel}: K(x, y) = (x * y + 1)^d \tag{12}$$

$$\text{RBF kernel}: K(x, y) = \exp(-\gamma \| x - y \|^2) \quad \gamma > 0 \tag{13}$$

$$\text{neural kernel}: K(x, y) = \tanh(ax * b) \tag{14}$$

$$\text{anova kernel}: K(x, y) = \left(\sum_i \exp(-\gamma(x_i - y_i)) \right)^d \tag{15}$$

The dot kernel is used for linear classification and the others are used for nonlinear classification. The optimal kernel is selected by training for the minimum classification error; the polynomial kernel was selected. We classified the genuine and imposter distribution as "-1" and "1" for the output values by the SVM, respectively. As shown in **Fig. 8**, the two matching scores are calculated from the two images normalized by Retinex and histogram stretching, respectively. These two scores are used as the inputs of the SVM and the final decision is based on its output. Since the accuracies of face recognition by Retinex and histogram stretching with PCA and LBP are higher than the other cases, as shown in **Table 2** and **6**, we use these two methods of normalization with PCA and LBP for the SVM classifier, as shown in **Fig. 8**.

2.6 Overall procedure of proposed method

Fig. 8 is the overall procedure of the proposed method. First, we capture a facial image with a built-in camera of an UMPC (Ultra-Mobile Personal Computer). Next, the face region is detected using the Adaboost face detector, as shown in **Fig. 8**.

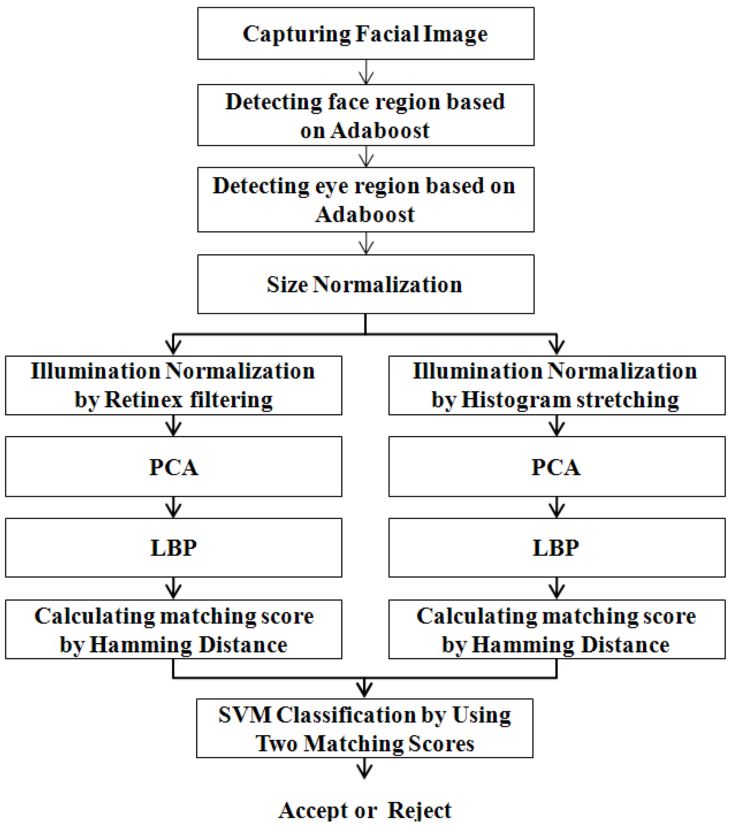


Fig. 8. Flowchart of the proposed method

After detecting the face region, two eye points are located by using an eye detection method based on the Adaboost algorithm in the restricted area of the face region. Based on the detected eye positions, the face region is re-defined as shown in Fig. 2.

Then, the size of the redefined face region is normalized as 32×32 pixels. The illumination variance is reduced by the illumination normalization method. We use Retinex filtering and histogram stretching for illumination normalization, respectively. From the two normalized images, facial features are extracted by using PCA (Principal Component Analysis) and LBP (Local Binary Pattern) and two matching scores are obtained. These two matching scores are used as the inputs of an SVM (Support Vector Machine) classifier and the final decision on face recognition is based on its output.

3. Experimental Results

We collected face images with the built-in camera of a commercial UMPC (Ultra Mobile PC) and tested our algorithm on the same UMPC. The machine featured 1.2 GHz CPU, 512 MB memory, 30 GB HDD, 4.5 inch LCD display and the Windows-XP operating system. The face database consisted of 2,868 face images from 192 classes. The image resolution is 640 × 480 pixels. Half the images were used for training and the other ones were used for testing. The database included face images from indoor environments with fluorescent lamps or without light. Also, it included face images from outdoor environments with sunlight. In this case, there were four directional sources of sunlight, which we referred to as frontal, left side, right side and back lighting.

The objective of our research is to propose algorithms for illumination normalization and face recognition, which show robustness to illumination variations irrespective of whether the environment is indoors or outdoors. So, for training, we used face images that were illuminated uniformly with fluorescent light. For testing, we used face images that showed various kinds of illumination variations in both indoor and outdoor environments. Because the proposed method focuses on robustness to illumination variations, all the images were captured with neutral expressions.

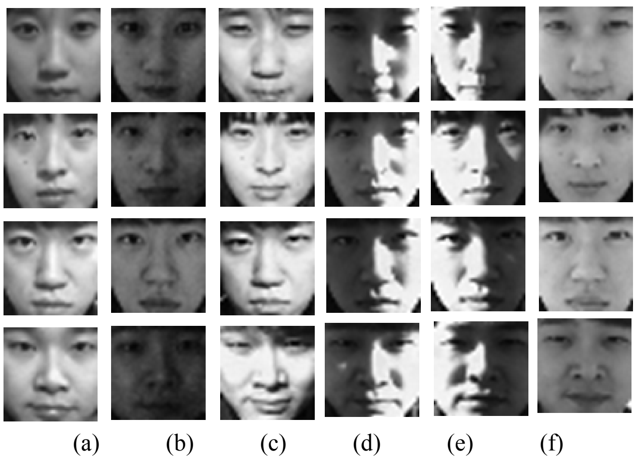


Fig. 9. Face database images (a) indoors with fluorescent lamp (b) indoors without light (c) outdoors with frontal sunlight (d) outdoors with right sunlight (e) outdoors with left sunlight (f) outdoors with back sunlight

However, the images included a few narrowed eyes when the subjects were faced with frontal sunlight. Fig. 9 shows some examples of the face database images.

In the first experiment, we compared the face recognition accuracy under various kinds of illumination conditions without illumination normalization. As shown in **Table 1**, we measured the EER (Equal Error Rate) in order to obtain the face recognition accuracies with the PCA, LNMF, LBP, PCA+LBP and LNMF+LBP methods. The EER means the error rate when the FAR (False Acceptance Rate) is equal to the FRR (False Rejection Rate) [22]. The FAR denotes the error rate when an unenrolled user (an imposter) is accepted as an enrolled user (a genuine user). The FRR denotes the error rate when an enrolled user is rejected as an un-enrolled one.

After obtaining the coefficients of the input face image, the distance between these coefficients and those of the enrolled one are calculated. If the distance is greater than the predetermined threshold, the input face image is rejected as unenrolled, otherwise, it is accepted as enrolled. To determine the optimal threshold with the training data, we used a Bayesian classifier based on a Bayesian rule [32]. That is, after obtaining the genuine and imposter distributions from the training data, the optimal threshold was determined for the minimum EER of face recognition.

For the distance measurements, various methods such as the Euclidean, Cosine and Hamming distances were compared. Since the coefficients of PCA and LNMF are real, the performances of the Euclidean and Cosine distances could be compared. Experimental results showed that the performance for the Euclidean distance was better than that for the Cosine distance. Since the coefficients of LBP are represented as binary codes, as shown in **Fig. 7**, the Hamming distance is used.

Since the obtained coefficients of PCA and LNMF are formulated as feature vectors rather than matrices, we performed the following procedure for PCA+LBP and LNMF+LBP. Using the example with PCA, since the size of the face image is 32×32 pixels, as shown in **Fig. 3**, the size of the obtained eigenface is also 32×32 pixels and the total number of eigenfaces is 1024(32×32) not considering the dimension reduction [16]. It follows that the total number of obtained eigen-coefficients is also 1024, because one eigen-coefficient is calculated from one eigenface [16]. Then, we reduce the 1024 eigen-coefficients to 32×32 matrix form in order to apply LBP. That is, the first 32 eigen-coefficients become the first row of the matrix and the next 32 eigen-coefficients become its second row. The next 32 eigen-coefficients also become its third row. The procedure is iterated and we can obtain a 32×32 matrix from the 1024 eigen-coefficients. With this matrix, we can apply the LBP method as shown in **Fig. 7**. The same procedure is applied for LNMF+LBP.

Table 1. Face recognition accuracy without illumination normalization

Face Recognition Method	PCA	LNMF	LBP	PCA+LBP	LNMF+LBP
EER (%)	31.606	26.437	26.836	29.638	34.239

For various kinds of illumination variations, the EER was greater than 26% in all methods of face recognition, as shown in **Table 1**. The LBP and LNMF methods using more local face information show better recognition accuracy than the others. Also, the PCA method shows a higher EER under various illumination conditions. That means that only the global representations of the face images are more susceptible to illumination variations. In addition, the LNMF+LBP method shows the lowest recognition accuracy, because unnecessary local face information such as noise is extracted by using the binary patterns of local face information obtained from the LNMF method. For the PCA+LBP and the LNMF+LBP methods, LBP was applied to the facial coefficients, which were obtained by the PCA or LNMF methods. For face images without illumination changes, the face recognition accuracies of the PCA, LNMF, LBP, PCA+LBP and LNMF+LBP methods were 2.251%, 2.807%, 3.129%, 2.485% and 5.468%, respectively.

In the second experiment, we compared the accuracy of the face recognition algorithms with illumination normalization. **Fig. 10** shows some examples of face images with illumination normalization.

Fig. 10. The face images with illumination normalization (a) original image (b) with histogram equalization (c) with histogram stretching (d) with the Retinex filter

For histogram equalization, the EERs were slightly higher than for other illumination normalization methods, as shown in **Table 2**, since the quality of the face images was degraded in indoor environments without light and outdoor environments with back sunlight, while performing histogram equalization, as shown in **Fig. 10** (b) and **Fig. 11**. The recognition accuracies of almost all the methods were reduced due to partial noise factors generated by histogram equalization.

Table 2. Face recognition accuracy without illumination normalization

	Face Recognition Method	PCA	LNMF	LBP	PCA+LBP	LNMF+LBP
EER (%)	Histogram Equalization	31.763	29.601	27.114	29.082	35.942
	Histogram Stretching	31.643	26.498	27.440	27.548	33.853
	Retinex Filter	16.896	21.196	21.123	**16.280**	23.068

When using histogram stretching, the EERs of the PCA+LBP and LNMF+LBP methods became slightly smaller than those of **Table 1**. When using the Retinex filter, all the EERs were lower than those for histogram equalization and stretching, as shown in **Table 2**. Also, the EERs with the Retinex filter were also lower than those shown in **Table 1**. Among them, the PCA+LBP method with Retinex filter showed the best recognition accuracy.

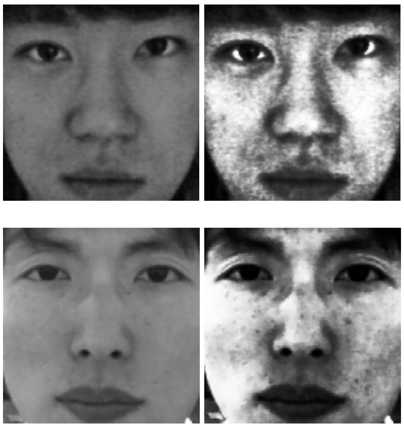


Fig. 11. Image degradation by histogram equalization

In the next test, we used the outputs of the Retinex filter and histogram stretching with PCA and LBP as the inputs of the SVM classifier, as shown in **Fig. 8**. For SVM training, we used half the images in database, and the other half were used for testing. Among the five kernels of Eq. (10) ~ (14), the polynomial kernel was selected for the minimum classification error by training. The accuracies of face recognition with the test data are shown in **Table 3**. By comparing **Table 2** and **3**, we can know that the accuracy of the proposed method is higher than those of the other cases.

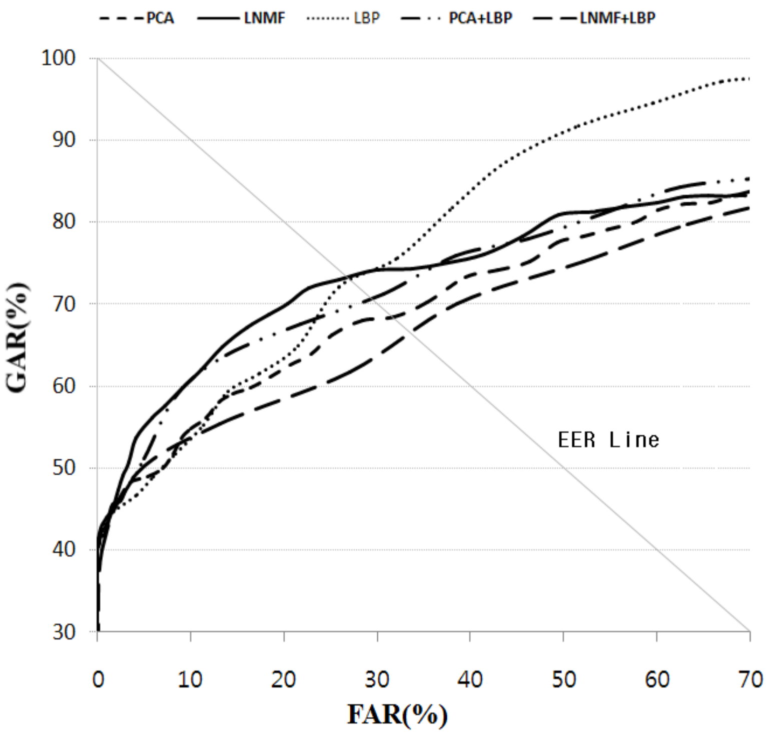
Fig. 12 shows the ROC (Receiver Operating Characteristic) curves of face recognition in case of using each illumination normalization method. **Fig. 13** shows the ROC curves to compare the result of the proposed method of **Fig. 8** to the result of the Retinex filter with PCA+LBP, which shows the best accuracy when using single illumination normalization, as shown in **Table 2**. In **Fig. 12** and **13**, the horizontal axis means FAR (False Accept Rate), and the vertical axis means GAR (Genuine Acceptance Rate). GAR is 100 - FRR (False Rejection Rate) and the points which intersect with EER (Equal Error Rate) Line mean the EER.

The reason the accuracy of the proposed method was better than that of the other cases is that illumination changes can include local and global variations in general. So, the proposed method, which combines two methods of illumination normalization, shows better performance compared to that for single illumination normalization.

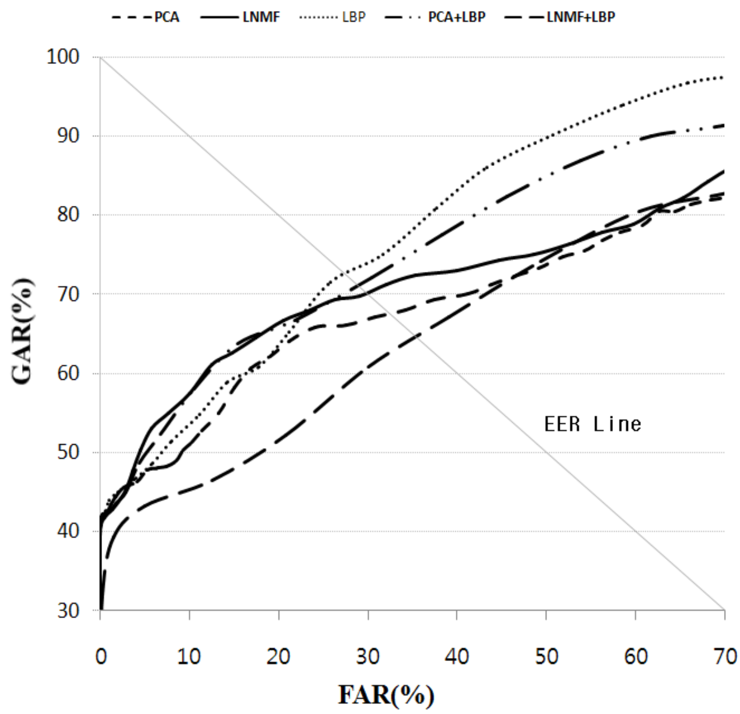
Table 3. The result of the proposed method of **Fig. 8** (unit: %)

FAR	FRR	EER
13.67	14.52	**14.095**

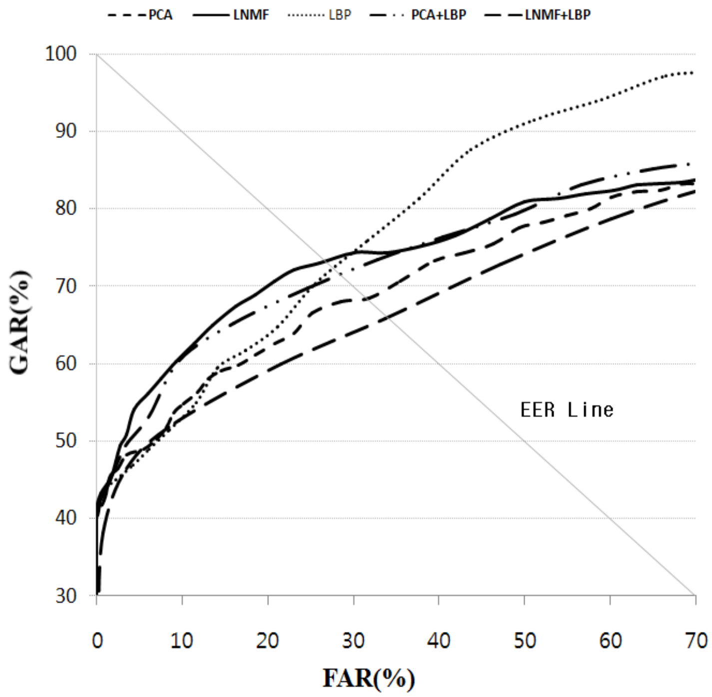
In the next experiment, we measured the processing time of all the steps shown in **Fig. 8** (See **Table 4**). The total processing time of the proposed method including face and eye detections, size normalization, Retinex+PCA+LBP, histogram stretching+PCA+LBP, calculating Hamming distance two times and classification by SVM was about 56.7 ms. Consequently, the proposed algorithm can be operated at a real-time speed of about 17.6 frames/sec. By comparing the processing time of the Retinex filter with PCA+LBP, which shows the best performance for single illumination normalization, as shown in **Table 2**, the increased processing time of the proposed method is as small as 0.072 ms, because the additional procedures of the proposed method include only histogram stretching, PCA+LBP, calculating Hamming distance and SVM classification.



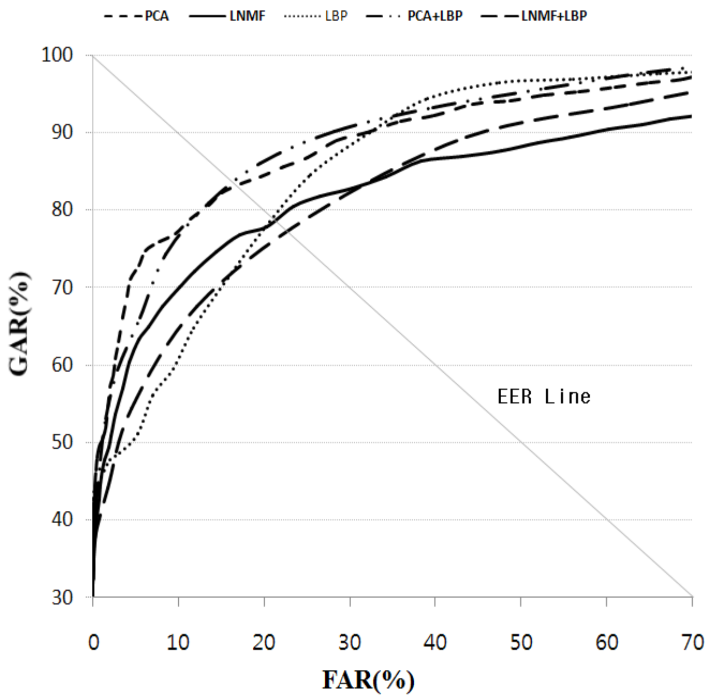
(a)



(b)



(c)



(d)

Fig. 12. ROC curves (a) without illumination normalization (b) with histogram equalization (c) with histogram stretching (d) with Retinex filter

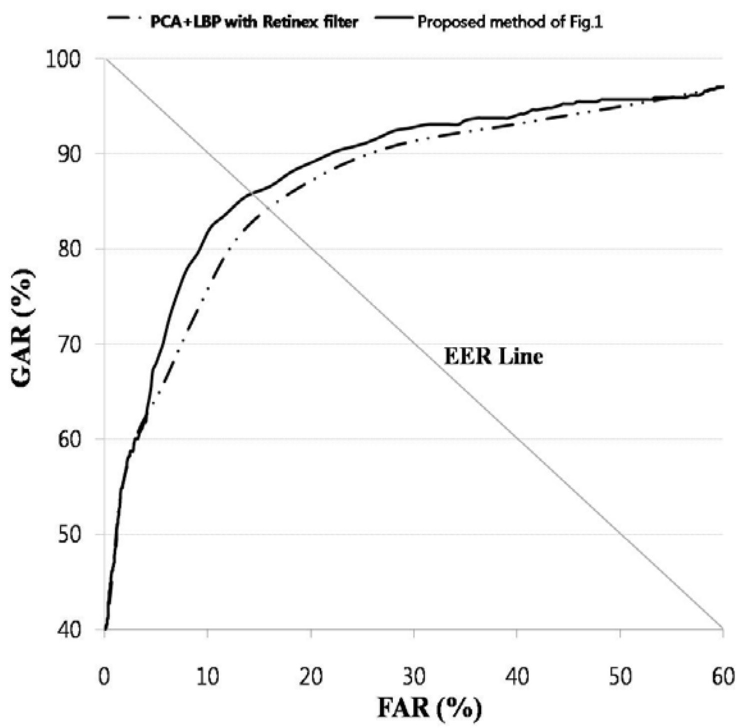


Fig. 13. ROC curves to compare SVM combining to Retinex with PCA+LBP

Table 4. Processing time of the proposed method of **Fig. 8**

All Steps of Fig. 8		Processing Time (ms)
Face detection by Adaboost		31
Eye detection by Adaboost		24.9
Size Normalization		4.17E-4
Illumination Normalization	Histogram Equalization	0.02
	Histogram Stretching	0.04
	Retinex Filter	0.66
Face Recognition	PCA	0.001
	LNMF	0.018
	LBP	0.016
	PCA+LBP (x 2)	0.031
	LNMF+LBP	0.034
Calculating Dissimilarity	Euclidean Distance	4.91E-6
	Hamming Distance (x 2)	3.02E-5
SVM Classification		0.001

In the final experiments, we compared the accuracy with another open face database called the AR face database [23]. The image resolution of the AR database is 768×576 pixels. There are 1,783 images from 136 subjects (including 76 men and 60 women) in the first and second session. For each subject, 7 ~ 14 images were captured under neutral and different expressions, different illuminations and occlusion conditions. Because our research objective is to make the face recognition algorithm robust to illumination variation, face images with different expressions and occlusion conditions were not used for training and testing. Only face images with or without illumination changes were used, as shown in **Fig. 14**. Consequently, a total of 1,008 face images (135 subjects x 4 ~ 8 images in the first and second sessions) from the AR database were used.

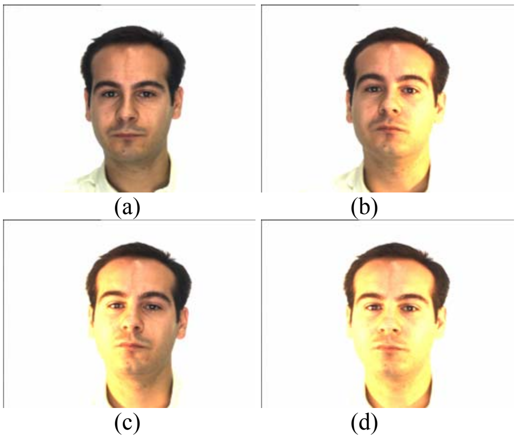


Fig. 14. Sample images from the AR database (a) without illumination (b) with right light (c) with left light (d) with lights from both sides

Without illumination normalization, the face recognition accuracy for LBP was best, as shown in **Table 5**. The accuracy was much enhanced with illumination normalization. As shown in **Table 6**, PCA+LBP with Retinex filter shows the best accuracy, as in the results of **Table 2**. The result of the proposed method of **Fig. 8** is shown in **Table 7**. By comparing **Table 6** and **7**, we can know that the accuracy of the proposed method is higher than those of the other cases.

Table 5. Face recognition accuracy without illumination normalization

Face Recognition Method	PCA	LNMF	LBP	PCA+LBP	LNMF+LBP
EER (%)	28.325	32.780	**19.963**	20.372	32.395

Table 6. Face recognition accuracy with illumination normalization

	Face Recognition Method	PCA	LNMF	LBP	PCA+LBP	LNMF+LBP
EER (%)	Histogram Equalization	33.161	36.321	19.821	36.855	36.968
	Histogram Stretching	28.285	32.818	19.963	21.592	32.745
	Retinex Filter	15.419	29.407	16.345	**14.849**	25.167

Table 7. The result of combining Retinex and histogram stretching by using SVM

FAR	FRR	EER
10.09 %	10.27 %	**10.18 %**

Fig. 15 shows some examples of using the proposed face recognition system on the UMPC.

(a) (b)

Fig. 15. Examples of using the proposed face recognition system on UMPC (a) indoors (b) outdoors

4. Conclusions

In this paper, we propose a new face recognition method robust to variations of illumination on mobile devices based on an SVM. The accuracy of face recognition under various illumination conditions was best when two matching scores from the Retinex filter + PCA + LBP and histogram stretching + PCA + LBP methods were used as the inputs of the SVM. In future work, we plan to study local shade elimination by combining illumination normalization and de-noising filtering.

References

[1] Q. Tao, R. N. J. Veldhuis, "Biometric Authentication for a Mobile Personal Device," in *Proc. of International Conference on Mobile and Ubiquitous Systems*, pp. 1-3, 2006.
[2] A. Hadid, J. Y. Heikkila, O. Silven, M. Pietikainen, "Face and Eye Detection for Person Authentication in Mobile Phones," in *Proc. of International Conference on Distributed Smart Cameras*, pp. 101-108, 2007.
[3] J. Czyz, S. Bengio, C. Marce, L. Vandendorpe, "Scalability Analysis of Audio-visual Person Authentication," in *Proc. of International Conference on Audio and Video Based Biometric Person Identification*, pp. 752-760, 2003.
[4] D. J. Lee, K. C. Kwak, J. O. Min, M. G. Chun, "Multi-modal Biometrics System Using Face and Signature," *LNCS,* vol. 3043, pp. 635-644, 2004.
[5] T. J. Smith, "LA County Sheriff's Department (LASD) Mobile ID Project : From Pilot to Countrywide roll-out," *Biometrics 2008*, 2008.
[6] X. Xie, K. M. Lam, "An Efficient Illumination Normalization Method for Face Recognition," *Pattern Recognition Letters*, vol. 27, pp. 609-617, 2006.
[7] L. Qing, S. Shan, X. Chen, W. Gao, "Face Recognition under Varying Lighting Based on the Probabilistic Model of Gabor Phase," in *Proc. of 18th Int'l Conf. Pattern Recognition, ICPR 2006*, vol. 3, pp. 1139-1142, 2006.
[8] S. Shan, W. Gao, B. Cao, D. Zhao, "Illumination Normalization for Robust Face Recognition Against Varying Lighting Conditions," in *Proc. of IEEE Int'l Workshop on Analysis and Modeling of Faces and Gestures*, 2003.
[9] M. Y. Nam, P. K. Rhee, "An Efficient Face Recognition for Variant Illumination Condition," *ISPACS2005*, vol. 1, pp. 111-115, 2004.

[10] H. Wang, S. Z. Li, Y. Wang, "Face Recognition under Varying Lighting Conditions Using Self Quotient Image," in *Proc. of IEEE Int'l Conf. on Automatic Face and Gesture Recognition*, 2004.

[11] P. Viola, M. Jones, "Robust Real-time Face Detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp.137-154, 2004.

[12] OpenCV, Available on: http://opencvlibrary.sourceforge.net/ (accessed on January 2, 2010).

[13] G. D. Hines, Z. Rahman, D. J. Jobson, G. A. Woodell, "Single-scale Retinex Using Digital Signal Processors," in *Proc. of Global Signal Processing Conference*, Sep. 2004.

[14] R. C. Gonzalez, R. E. Woods, "Digital image processing," Addison-Wesley, MA, 1992.

[15] L. I. Smith, "A Tutorial on Principal Components Analysis", 2002.

[16] M. Turk, A. Pentland, "Eigenfaces for Recognition," *J. Cognitive Neuroscience*, vol. 3, no. 1, pp. 71-86, 1991.

[17] S. Han, H. A. Park, D. H. Cho, K. R. Park, S. Y. Lee, "Face Recognition Based on Near-Infrared Light Using Mobile Phone," in *Proc. of 8th ICANNGA 2007, Lecture Notes in Computer Science* , vol. 4432, pp. 440-448, 2007.

[18] S. Z. Li, X. W. Hou, H. J. Zhang, Q. S. Cheng, "Learning Spatially Localized, Parts-based Representation," in *Proc. of IEEE Int. Conf. on Computer Vision and Pattern Recognition*, Hawaii, Dec. 11-13, 2001.

[19] D. D. Lee, H. S. Seung, "Learning the Parts of Objects by Non-negative Matrix Factorization," *Nature*, vol. 401, pp. 788-791, 1999.

[20] Y. Rodriguez, S. Marcel, "Face Authentication Using Adapted Local Binary Pattern Histograms," *LNCS*, vol. 3954, pp. 321-332, 2006.

[21] H. A. Park and K. R. Park, "Iris Recognition based on Score Level Fusion by using SVM," *Pattern Recognition Letters*, vol. 28, no. 1, pp. 94-104, 2007.

[22] N. A. Macmillan, C. D. Creelman, "Detection Theory: A User's Guide," *Cambridge University Press*, NY, 1991.

[23] A. M. Martinez, R. Benavente, The AR face database, *CVC Tech*, Report #24, 1998.

[24] E. H. Land, J. J. McCann, "Lightness and Retinex Theory," *J. Opt. Soc. Am.*, vol. 61, pp 1-11, 1971.

[25] A. Moore, J. Allman, R. M. Goodman, "A Real-time Neural System for Color Constancy," *IEEE Trans. Neural Networks*, vol. 2, pp. 237-247, Mar. 1991.

[26] B. D. Thompson, Z. Rahman, S. K. Park, "Retinex Preprocessing for Improved Multi-spectral Image Classification," *Visual Information Processing VIII*, Proc. SPIE 3716, 1999.

[27] Sony UMPC, Available on: http://b2b.sony.com (accessed on January 2, 2010).

[28] R. P. Wildes, "Iris Recognition: An Emerging Biometric Technology," *Proc. IEEE*, vol. 85, no. 9, pp. 1348-1363, 1997.

[29] E. Land, "An Alternative Technique for the Computation of the Designator in the Retinex Theory of Color Vision," in *Proc. Nat. Acad, Sci.*, vol.83, pp. 3078-3080, 1986.

[30] D. J. Jobson, Z. Rahman, and G. A. Woodell, "Retinex Processing for Automatic Image Enhancement," *Human Vision and Electronic Imaging VII, SPIE Symposium on Electronic Imaging*, vol. 4662, 2002.

[31] http://www.soe.ucsc.edu/classes/ee264/Winter02/yan.ppt (accessed on 2010.02.08).

[32] R. O. Duda, and P. E. Hart, "Pattern Classification and Scene Analysis," Wiley-Interscience, USA.

[33] http://www.cs.ucl.ac.uk/staff/M.Sewell/winsvm/ (accessed on 2010.02.08)

[34] V. Vapnik, "Statistical Learning Theory," John Wiley & Sons, NY, USA, 1998.

[35] http://www-ai.cs.uni-dortmund.de/SOFTWARE/MYSVM/index.html (accessed on 2010.02.08).

Gi Pyo Nam received a B.S. degree in Digital Media from Sangmyung University, Seoul, South Korea, in 2009. He is currently pursuing a M.S. degree in Electronics and Electrical Engineering at Dongguk University. He is also a research member of the Biometrics Engineering Research Center. His research interests include biometrics and image processing.

Byung Jun Kang received his B.S. and M.S. degrees in Software and Computer Science from Sangmyung University, Seoul, South Korea, in 2004 and 2006, respectively. In 2009, he also earned a Ph.D degree in Computer Science from the Department of Computer Science, Sangmyung University, Seoul, South Korea. He is currently a Researcher at the Electronics and Telecommunications Research Institute. His research interests include biometrics and image processing

Kang Ryoung Park received his BS and MS degrees in Electronic Engineering from Yonsei University, Seoul, Korea, in 1994 and 1996, respectively. He also received his PhD degree in Computer Vision from the Department of Electrical and Computer Engineering, Yonsei University, in 2000. He was an assistant professor in the Division of Digital Media Technology at Sangmyung University from March 2003 to February 2008. He has been an assistant and associate professor in the Division of Electronics and Electrical Engineering at Dongguk University since March 2008. He is also a research member of BERC. His research interests include computer vision, image processing and biometrics.