Face Recognition in Non-Uniform Illumination Conditions Using Lighting Normalization and SVM

CHIUNHSIUN LIN
National Taipei University
69 Sec. 2 Chian_Kwo N. Road, Taipei, Taiwan, 10433
TAIWAN

CHING-HUNG SU
The Department of Electronic Engineering,
Lunghwa University of Science and Technology,
No.300, Sec. 1, Wanshou Rd., Gueishan Shiang, Taoyuan County, Taiwan, 33306
TAIWAN

Abstract: - An efficient face recognition scheme is developed to recognize face in color images with non-uniform illumination conditions. The proposed scheme comprises two phases, namely face detection and face recognition. For the face detection phase, a lighting normalization function and an isosceles triangle approach are utilized to detect facial regions accurately. For the recognition phase, a SVM scheme is adopted to uniquely identify facial characteristics. The primary advantage of the proposed face recognition system is its ability to handle different facial image sizes under non-uniform illumination conditions. Moreover, the system performs better than PCA algorithms in term of success rate.

Key-Words: - Face detection; Face recognition; Triangle-based segmentation; Lighting normalization function; SVM; Non-uniform illumination conditions

1 Introduction

Automatic face detection and recognition of human faces is one of the most intricate and important problems in computer vision and pattern recognition. It can be used as the security mechanism to replace metal key, plastic card, and password or PIN number. Above and beyond, it can also be applied in the intelligent systems for human-computer interaction and criminal identification. Abundant of researches have been conducted on human face detection and face recognition. For comprehensive study of previous face detection techniques, please see [1, 2]. Some principal techniques can see color extraction [3, 4, and 5], neural networks [6, 7, and 8], PCA [9, 10], SVM [11, 12], template matching [13, 14, and 15]. For comprehensive survey of previous face recognition techniques, please see [16, 17, 18 and 19]. In general, face recognition systems can be classified into the following two categories [19, 20]. (a) Feature-based: The feature-based matching approach utilize the relationship between facial features such as eyes, mouth and nose [21~26]. (b) Template matching: The template matching approach employs the complete features of face image [27~33]. However, few can achieve a completely reliable performance.

Currently, the topics are still the most arduous tasks for pattern recognition by reason of the large variation on the signal strength associated with unpredictable illumination conditions and other uncertain sizes of faces. Hence, it is necessary for an automatic face recognition system to have an excellent preprocess (a face detection system). In this paper, we focus on face images captured with non-uniform illumination as shown in Fig. 1

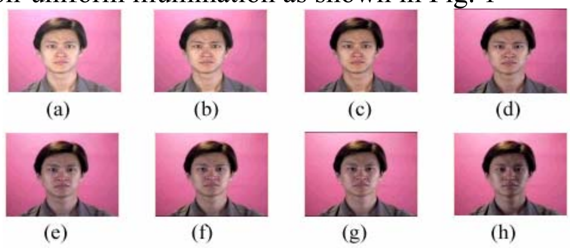

Fig. 1: Face images (before face detection) captured under non-uniform illumination.

The rest of the paper is organized as follows. In section 2, face detection based on a triangle-based approach and a lighting normalizing function is described. In section 3, a SVM function is utilized to recognize human face. Experimental results are demonstrated in section 4 to verify the validity of the proposed face recognition system. Finally, conclusions are given in section 5.

2 Face Detection

There are six steps in the first part of the designed system. The input image can be either a color or a gray level image. First, convert the input image to a binary image. Second, label all 4-connected components in the image to form segments and find out the center of each segment. Third, detect any 3 centers of 3 different segments that form an isosceles triangle. If the triangle *i j k* is an isosceles triangle as shown in Fig. 2(a), then it should possess the characteristic of ″the distance of line *i j* = the distance of line *j k*″. Due to the imaging effect, imperfect binarization result and various poses of human faces, a 25% deviation is given to absore the tolerance. Here, ″abs″ means the absolute value, "D (i, j)" denotes the Euclidean distance between the centers of block i (right eye) and block j (mouth), "D (j, k)" denotes the Euclidean distance between the center of block k (left eye) and block j (mouth), "D (i, k)" represents the Euclidean distance between the centers of block i (right eye) and block k (left eye).The first matching rule can thereby be stated as (abs(D(i, j)-D(j, k)) < 0.25*max(D(i, j), D(j, k))), and the second matching rule is (abs(D(i, j)-D(i, k)) < 0.25*max(D(i, j), D(j, k))). Since the labeling process is operated from left to right then from top to bottom, we can get the third matching rule as ″i < j < k″. For example, as shown in Fig. 2(a), if three points (i, j, and k) satisfy the matching rules, then we think that they construct an isosceles triangle. Assuming that the real facial region should cover the eyebrows, the eyes, the mouth and some area below the mouth, the coordinates can be determined as follows:

$$X1 = X4 = Xi - 1/4*D\ (i, k); \quad (1)$$
$$X2 = X3 = Xk+1/4*D\ (i, k); \quad (2)$$
$$Y1 = Y2 = Yi+1/4*D\ (i, k); \quad (3)$$
$$Y3 = Y4 = Yj -1/4*D\ (i, k); \quad (4)$$

Assuming that (Xi, Yi), (Xj, Yj) and (Xk, Yk) are the three center points of blocks i, j, and k, that form an isosceles triangle. Then (X1, Y1), (X2, Y2), (X3, Y3), and (X4, Y4) are the four corner points of the face region as shown in Fig. 2(b). X1 and X4 locate at the same coordinate of (Xi - 1/4*D (i, k)); X2 and X3 locate at the same coordinate of (Xk + 1/4*D (i, k)); Y1 and Y2 locate at the same coordinate of (Yi + 1/4*D (i, k)); Y3 and Y4 locate at the same coordinate of (Yj - 1/4*D (i, k)); where D (i, k) is the Euclidean distance between the centers of block i (right eye) and block k (left eye).

Fourth, normalizing a potential face region can decrease the effects of variation in the distance and location. Since all potential faces are normalized to a standard size (i.e. 60 × 60 pixels) in this step, the potential face regions selected in the previous section can have different sizes. Herein, the potential facial region is resized by the bicubic interpolation technique as described in the textbook written by Gonzalez R. C. et al. [34]. In this step, we have already removed the restriction of changeable sizes.

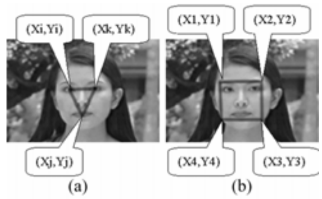

Fig. 2 IF (X*i*, Y*i*), (X*j*, Y*j*) and (X*k*, Y*k*) are the three center points of blocks i(right eye), j(mouth), and k(right eye), respectively. The four corner points of the face region are thus (X1, Y1), (X2, Y2), (X3, Y3), and (X4, Y4).

Fifth, feed every normalized potential face region into a weighting mask function that is applied to decide the actual location of the face region. Fig. 3 illustrates all the procedures of how to obtain a detected face. Fig. 3(a) demonstrates the original image; Fig. 3(b) shows the binary image of the original image; Fig. 3(c) displays the grayscale image of the original image; Fig. 3(d) depicts the isosceles triangle formed by the 3 centers of 3 blocks; Fig. 3(e) shows the best binary potential face region cropped from the binary image covering the isosceles triangle; Fig.3(f) illustrates the best color potential facial region cropped from the original image that covers the isosceles triangle; Fig. 3(g) displays the best binary potential face region which is normalized to a standard size (60 × 60 pixels), and Fig. 3(h) depicts the result of face detection. The detail of face detection can be found in Lin and Fan [15]. After these processes, we utilize a lighting normalization function to standardize the illumination of the detected face.

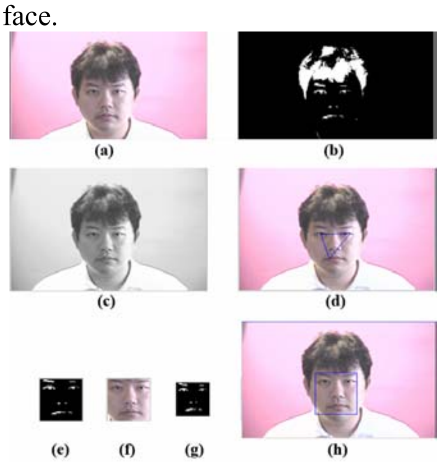

Fig. 3 illustrates all the procedures of how to obtain a detected face.

Sixth, we employ the detected face, such as Fig. 3(f), into next procedure. The detected face, such as Fig. 3(f), is transferred to grayscale image first. Then, we utilize a lighting normalization function to standardize the illumination of the grayscale detected facial region. The illumination of the grayscale detected facial region is modified by regulating the average grayscale value to a constant (which in the proposed system is 135). Next, A simple global thresholding function is then performed with a threshold value of T (in the proposed system, T = 100) to generate a satisfactory binary image. The average grayscale value of the detected facial image is modified to a constant 135, because the average grayscale values of the detected facial images are about 110–160. Hence, the average gray value of 110 and 160 [(110+160)/2 =135] is selected to adjust the average grayscale value of a facial image, which in this case is 135. If the average gray value of the grayscale detected facial region is adjusted to a constant between 125 and 145 should be all right. The average grayscale value of different images seems to be impossible to regulate to a constant. Nevertheless, regulation is possible in the detected face, since the outline of the detected face that all covers the eyebrows, two eyes, and one mouth is outstandingly stable, as demonstrated in Fig. 4. Fig. 4(a) shows the detected face. Fig. 4(b) illustrates the detected face by a simple global thresholding process with a threshold value of T (T = 100), which produces unstable binary images (the block of mouth is almost invisible). Fig. 4(c) depicts the detected face from adjusting the average grayscale value to a constant (constant = 135), and then by simple global thresholding with threshold T (T = 100) to obtain stable binary images. Figs. 4(d), 4(g), 4(j) and 4(m) display the detected face with different intensity of illumination, respectively. Figs. 4(e), 4(h), 4(k) and 4(n) depict the original detected facial region with different intensity of illumination, and then by simple global thresholding with threshold T (T = 100) to acquire unstable binary images (the block of mouth is almost invisible). Figs. 4(f), 4(i), 4(l) and 4(o) show the original detected facial region with different intensity of illumination. The average grayscale value was first adjusted to a constant value of 135, and then converted to a stable binary image by a simple global thresholding process with a threshold value of T (T = 100, and the block of mouth is easy to find)). In this task, the problems of non-uniform illumination conditions can be resolved using the lighting normalization function.

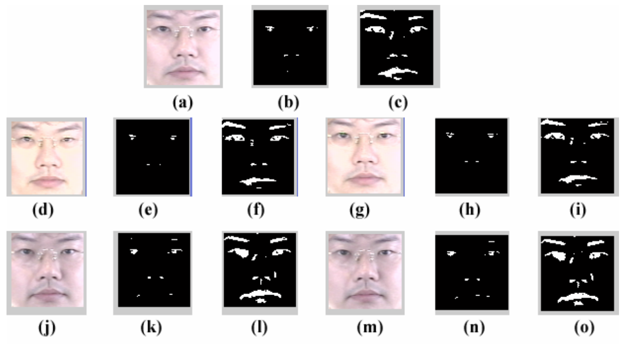

Fig. 4: The process of the original detected facial region by adjusting the average gray-level value to a constant (constant = 135), and then by simple global thresholding with threshold T (T = 100) to acquire stable binary images.

3 Face Recognition Using SVM

Face recognition is a tricky mission because of the changeable illumination conditions. For example, the illumination changes between indoor and outdoor environments are an unsolved problem for face recognition. Moreover, it is a multi-class problem. It can be solved by two familiar approaches. We assume N is the number of classes (e.g. N different individuals). The first proposal is "one against the rest approach". This technique includes N binary classifiers, and each of them separates a single class from all the remaining classes. The final output is the class that corresponds to the binary classifier with the highest output value. The second proposal is "one against one approach". This technique includes $N(N-1)/2$ binary classifiers, and each of them separates a pair of classes. The final output is decided by voting, or decision tree. In this paper, we use the first proposal for our system.

In the previous section, a set of face regions in an image was selected. The success rate of the face detection process as section 2 is assumed to be 100% in this section. Two issues should be considered in face recognition first. (1) Determine the features to be used to represent a face. (2) Determine how to utilize these features for recognition. This section presents an efficient support vector machine classifier for face recognition. SVM is a new and promising classification and regression technique proposed by Vapnik [35] and his group at AT&T Bell Laboratories. SVM based on Structural Risk Minimization (SRM) principle. SVM is superior to those methods based on Empirical Risk Minimization (ERM) principle, such as ANN. SVM minimizes an upper bound on VC dimension as opposed to ERM that minimize the error on the training data. The main idea of SVM comes from (1) a nonlinear mapping of the input space to a high dimensional feature space, and (2) given two linearly separable classes, designs the classifier that leaves the maximum margin from two classes in the feature space. SVM displays good performance, has been applied extensively for pattern classification and handwriting recognition.

Let x_i, $i = 1, 2 \ldots, N$, be the feature vectors of the training set, X. These belong to either of two classes, ω_1, ω, which are assumed to be linearly separable. The goal is to design the decision hyperplane:

$$g(\underline{x}) = \underline{w}^T \underline{x} + w_0 = 0 = w_1 x_1 + w_2 x_2 + \ldots + w_l x_l + w_0 \quad (5)$$

Assume $\underline{x}_1, \underline{x}_2$ on the decision hyperplane

$$0 = \underline{w}^T \underline{x}_1 + w_0 = \underline{w}^T \underline{x}_2 + w_0$$

$$\Rightarrow \underline{w}^T (\underline{x}_1 - \underline{x}_2) = 0 \ \forall \underline{x}_1, \underline{x}_2 \quad (6)$$

Since the difference vector x1 –x2 obviously lies on the decision hyperplane (for any $x_1 - x_2$), it is apparent from equation (6) that the vector ω is orthogonal to the decision hyperplane. The decision hyperplane is also called Optimal Separating Hyperplane (OSH) that minimizes the risk of misclassifying not only the samples in the training set but also the other samples of the test set. For a two class classification problem, the goal is to separate the two classes by the only decision hyperplane that is established by available samples. Consider the samples in Fig. 5(a), where there are many possible linear classifiers (e.g. H1, H2, H3 and H4) that can separate the data, but there is only one decision hyperplane (shown in Fig. 5(b)) that maximizes the margin (the distance between the decision hyperplane and the nearest data point of the two classes).

It sounds fabulous and seems to require advanced techniques like AI searching or some time-consuming complicated computation. However, SVM used some statistical learning theory to solve these problems in reasonable time.

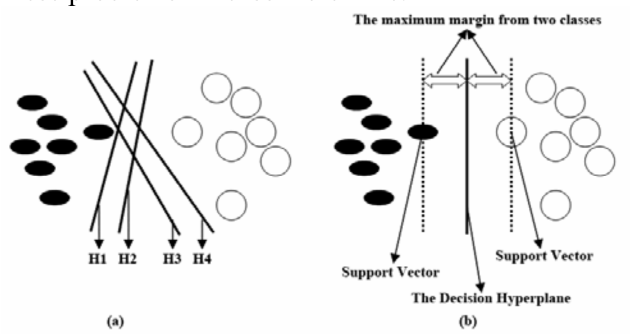

Fig. 5 The only one decision hyperplane maximizes the margin.

4 Experimental Results

This section describes experimental results, and two sets of database are demonstrated to verify the efficiency of the proposed approach. The first image database contains 48 individuals were taken by a digital camcorder with genuine non-uniform illumination conditions. Since we need more images (at least eight images per person) with non-uniform illumination conditions step by step slightly, we produce these images by ourselves at the National Taipei University. The resolution of each color image is 352×240 pixels. The process for taking video is ten photographic lights were turned on first in our laboratory. Then, one photographic light was switched off every second until only three photographic lights were still turned on. We obtain 384 images with non-uniform illumination conditions modified by real illumination changes. Subsequently, 192 images are employed as training images (48 individuals with 4 training images per person). The other 192 images are utilized as test images (48 different persons with 4 test images per person). An example with non-uniform illumination images is displayed in Fig. 6. After the lighting normalization, Figs. 6 (a), (c), (e) and (g) are employed as training images, while Figs. 6 (b), (d) (f) and (h) are adopted as test images. The test and training image sets are distinct from each other. Among them, 176 faces can be recognized correctly. Experimental results indicate that the success rate is 91.67% (176/192 = 91.67%), corresponding to a failure rate is 8.33 % (16/192 = 8.33 %).

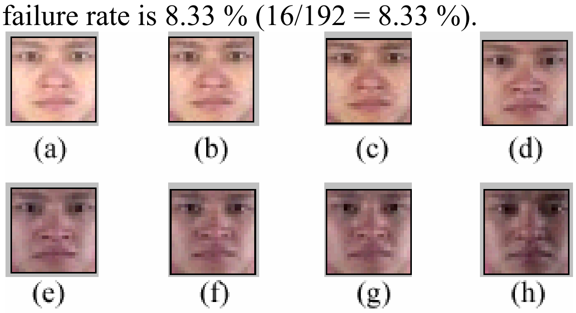

Fig. 6: Examples of the detected faces with non-uniform illumination that are taken by various luminosities.

In order to compare the recognition performances in a benchmark database, the second face database that we employed is some parts of the "AR face database" [36]. This face database was created by Aleix Martinez and Robert Benavente in the Computer Vision Center (CVC) at the U.A.B. It contains over 4,000 color images corresponding to 126 people's faces (70 men and 56 women). Images contain frontal view faces with different facial expressions, illumination conditions, and occlusions (sun glasses and scarf). The pictures were taken at the CVC under strictly controlled conditions. No restrictions on wear (clothes, glasses, etc.), make-up, hair style, etc. were imposed to participants. Each person participated in two sessions, separated by two weeks (14 days) time. The size of the test images range is 640*480 pixels. We select the parts exclude the occlusions images (sun glasses and scarf), and examples of the detected faces with various conditions that are taken from AR database are shown in Fig. 7. After the lighting normalization, the first section, Figs. 7 (a), (c), (e) and (g) are employed as training images, while Figs. 7 (b), (d) and (f) are adopted as test images. After the lighting normalization, the second session, (a), (c), (e) and (g) are employed as testing images, while Figs. 7 (b), (d) and (f) are adopted as training images. The test and training image sets are distinct from each other. Therefore, there are totally 420 faces (include 30 individuals with 7 training images per person and 7 testing images per person) that are used to verify the validity of our system. Among them, 184 faces can be recognized correctly. Experimental results indicate that the success rate is 87.62% (184/210 = 87.62%), corresponding to a failure rate is 12.38 % (26/210 = 12.38 %).

Fig. 7: Examples of the detected faces with various conditions that are taken from AR database.

A summary of the experimental results is exhibited in Table 1 (FD denotes ″face detection″).

Table 1 Comparing Success Rate between Different Databases and Approaches:

Approaches	Success Rate using our Database	Success Rate using AR Database
SVM with FD (Ours)	91.67%	87.62%
PCA with FD	77.67 %	73.68 %
SVM without FD	85.38 %	80.33 %
PCA without FD	72.67 %	65.86 %

5 Conclusion

This investigation proposes an efficient SVM approach to recognize human faces embedded in digital camcorder images. The proposed face recognition system performs admirably on facial images with various sizes under non-uniform illumination conditions and complicated backgrounds. The automatic detection and face recognition system comprises both detection and recognition phases. For the detection phase, the lighting normalization function and the isosceles triangle approach are adopted to detect facial regions accurately. The part of face detection can exterminate the influence on variable face sizes under non-uniform illumination conditions to make the recognition process exceedingly robust. Subsequently, the recognition phase, the SVM classification approach is applied as a unique signature for face recognition. The system performs better than PCA algorithms in term of testing time and success rate. For future direction, the proposed approaches can be extended to exploit face recognition with different poses, expressions and other pattern recognition problems.

References:

[1] E. Hjelmas and B. K. Low, Face detection: A survey, Computer Vision and Image Understanding, Vol. 83, 2001, pp. 236 – 274.

[2] Ming-Hsuan Yang, D.J. Kriegman, N. Ahuja, Detecting faces in images: a survey, IEEE Transactions on PAMI, Vol. 24, Issue 1, Jan. 2002, pp. 34 – 58.

[3] C. Garcia and G. Tziritas, Face detection using quantized skin color regions merging and wavelet packet analysis, IEEE Transactions on Multimedia, Vol. 1, Issue 3, 1999, pp. 264 – 277.

[4] Rein-Lien Hsu, M. Abdel-Mottaleb, A.K. Jain, Face detection in color images, IEEE Transactions on PAMI, Vol. 24, Issue 5, 2002, pp. 696 – 706.

[5] J. Kovac, P. Peer, F. Solina, Illumination independent color-based face detection Proceedings of the 3rd International Symposium on Image and Signal Processing and Analysis, Vol.: 1, 2003, pp. 510 – 515.

[6] H. A. Rowley, S. Baluja, and T. Kanade, Neural network-based face detection, IEEE Transactions on PAMI, Vol. 20, Issue 1, Jan. 1998, pp. 23 – 38.

[7] R. Féraud, O.J. Bernier, J.-E. Villet, and M. Collobert, A Fast and Accuract Face Detector Based on Neural Networks, IEEE Trans. PAMI, Vol. 22, no. 1, Jan. 2001, pp. 42-53.

[8] Ming-Jung Seow, D. Valaparla, V.K. Asari, Neural network based skin color model for face detection, 32nd Applied Imagery Pattern Recognition Workshop, 2003, pp. 141 – 145.

[9] R. Gottumukkal and V.K. Asari, Real time face detection from color video stream based on PCA method, 32nd Applied Imagery Pattern Recognition Workshop, 2003, pp. 146 - 150.

[10] M. Kirby and L. Sirovich, Application of the Karhunen-Loeve procedure and the characterization of human faces, IEEE Trans. on PAMI, Vol. 12, Jan. 1990, pp. 103–108.

[11] J. Fang, G. Qiu, Human face detection using angular radial transform and support vector machines, Int. Conf. on Image Processing, Vol. 1, Sept. 2003, pp. 669-672.

[12] Peichung Shih., Chengjun Liu, Face detection using discriminating feature analysis and Support Vector Machine, Pattern Recognition, Vol. 39, 2006, pp. 260 – 276.

[13] Young Ho Kwon, N. da Vitoria Lobo, Face detection using templates, International Conference on Pattern Recognition, Vol. 1, 9-13 Oct. 1994, pp. 764 – 767.

[14] Haizhou Ai, Luhong Liang, Guangyou Xu, Face detection based on template matching and support vector machines, Int. Conf. on Image Processing, Vol. 1, Oct. 2001, pp. 1006 – 1009.

[15] Chiunhsiun Lin, Kuo-Chin Fan, Triangle-based Approach to the Detection of Human Face, Pattern Recognition, Vol. 34, No. 6, 2001, pp. 1271-1284.

[16] A. Samal, P. A. Iyengar, Automatic recognition and analysis of human faces and facial expressions: A survey. Pattern Recognition Vol.: 25, Issue: 1, 1992, pp. 65-77.

[17] R. Chellappa, C. L. Wilson, and S. Sirohey, Human and machine recognition of faces: A survey, Proc. IEEE, Vol. 83, no. 5, 1995, pp. 705–740.

[18] W. Zhao, R. Chellappa, A. Rosenfeld, P. J. Phillips, Face Recognition: A Literature Survey, ACM Computing Surveys, 2003, pp. 399-458.

[19] Ruiz-del-Solar, J., Navarrete, P., Eigenspace-based face recognition: a comparative study of different approaches, IEEE Transactions on Systems, Man and Cybernetics, Part C, Vol. 35, Issue 3, 2005, pp. 315-325.

[20] R. Brunelli, T. Poggio, Face recognition: features versus templates, IEEE Transactions on PAMI, Vol. 15, Issue 10, 1993, pp. 1042 – 1052.

[21] X. Jia and M. S. Nixon, Extending the feature vector for automatic face recognition, IEEE Transactions on PAMI, Vol.: 17 Issue: 12, 1995, pp. 1167–1176.

[22] L. Wiskott, J. M. Fellous, N. Kuiger, C. von der Malsburg, Face recognition by elastic bunch graph matching IEEE Transactions on PAMI, Vol. 19, Issue 7, 1998, pp. 775 – 779.

[23] T. K. Kim, H. Kim, W. Hwang, J. Kittler, Independent component analysis in a local facial residue space for face recognition, Pattern Recognition, Vol.: 37, Issue: 9, 2004, pp. 1873-1885.

[24] A-Nasser Ansari, Mohamed Abdel-Mottaleb, Automatic facial feature extraction and 3D face modeling using two orthogonal views with application to 3D face recognition, Pattern Recognition Vol. 38, Issue: 12, 2005, pp. 2549-2563.

[25] B. G. Park, K. M. Lee, S. U. Lee, Face recognition using face-ARG matching, IEEE Transactions on PAMI, Vol. 27, Issue 12, Dec. 2005, pp. 1982 – 1988.

[26] X. Xie, K. M. Lam, Face recognition under varying illumination based on a 2D face shape model, Pattern Recognition, Vol.: 38, Issue: 2, February, 2005, pp. 221-230.

[27] Rajkiran Gottumukkal, Vijayan K. Asari, An improved face recognition technique based on modular PCA approach, Patterrn Recognition Letters, Vol. 25, Issue: 4, 2004, pp. 429-436.

[28] Yanwei Pang, Zhengkai Liu, Nenghai Yu, A new nonlinear feature extraction method for face recognition, Neurocomputing Vol. 69, Issue: 7-9, 2006, pp. 949-953.

[29] B. Moghaddam, T. Jebara, A. Pentland, Bayesian face recognition, Pattern Recognition, Vol.: 33, Issue: 11, 2000, pp. 1771-1782.

[30]H. Moon, P.J. Phillips, Computational and Performance aspects of PCA-based Face Recognition Algorithms, Perception, Vol. 30, 2001, pp. 303-321.

[31] B. Draper, K. Baek, M.S. Bartlett, J.R. Beveridge, Recognizing Faces with PCA and ICA, Computer Vision and Image Understanding (Special Issue on Face Recognition), Vol. 91, Issues 1-2, 2003, pp. 115-137.

[32] B.K. Gunturk, A.U. Batur, Y. Altunbasak, M.H. Hayes, III, R.M. Mersereau, Eigenface-domain super-resolution for face recognition, IEEE Transactions on Image Processing, Vol. 12, Issue 5, 2003, pp. 597 – 606.

[33] Hyun-Chul Kim, Daijin Kim, Sung Yang Bang, Sang-Youn Lee, Face recognition using the second-order mixture-of-eigenfaces method, Pattern Recognition, Vol. 37, Issue: 2, 2004, pp. 337-349.

[34] Gonzalez, R. C., Woods, R. E., Digital Image Processing, Addison-Wesley Publishing Company, Reading, MA., 1992.

[35] V. N. Vapnik, An overview of statistical learning theory, IEEE Transactions on Neural Networks, Vol.: 10, Issue:5, 1999, pp. 988-999.

[36] A.M. Martinez and R. Benavente. The AR Face Database, CVC Technical Report #24, June 1998.