ONLINE PERSONAL IDENTIFICATION IN NIGHT USING MULTIPLE FACE REPRESENTATIONS

Ajay Kumar, T. Srikanth
Biometrics Research Laboratory
Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India

ABSTRACT

This paper details a fully automated face authentication system using low-cost near infrared imaging. The image normalization step consists of eye center localization, scale correction and orientation correction. This paper investigates the comparison and combination of four face matchers on the automatically normalized face images: elastic bunch graph matching (EBGM), trace transform, PCA, and LDA. The performance evaluation is presented on the near infrared images acquired from the 102 users in two sessions with different pose and expressions. Our experimental results achieve the best results with the EER of 3.92% from the EBGM matcher while the combination of results from the different matchers can significantly improve the performance and achieve the EER of 2.28 %. The near infrared face database developed in this work is also made publicly available to foster further research.

1. INTRODUCTION

Personal authentication using state-of-art approaches in face recognition literature has reported to be highly sensitive to the illumination and therefore illumination invariant face representations have been suggested [3]. Some defense and security related applications require face identification in dark, *i.e.* in absence of any visible illumination. Thermal infrared (IR) imaging is quite expensive, highly sensitive to ambient temperature and therefore not yet materialized into a commercial for real/online applications. The face images acquired in near IR imaging can capture subsurface facial features that are extremely difficult to modify [4]-[5]. The availability of low-cost near IR webcams, now commonly employed for indoor/video surveillance, has necessitated more efforts to achieve reliable face identification using near IR imaging.

1.1 Prior Work

The face identification using active near IR imaging has generated new interests among researchers and some work has already reported in the literature [1]-[5]. Recently Stan *et al*. [3] have also used near IR imaging to illustrate illumination invariant face identification. Authors in [3] have employed learning-based approach using local binary pattern features and shown promising results. Zou *et al*. [1] have also demonstrated face recognition using near IR imaging. Authors have concluded that near IR face images offer better classification severability than those images acquired under ambient illumination. However the work detailed in [1] is only preliminary on 18 subjects data and require much efforts on performance improvement. Zhao and Giget [2] employed discrete cosine transform based features to examine near IR face recognition on 10 subjects and achieved limiting performance. Chan *et al.* [4] have presented more detailed analysis for face recognition using principal component analysis. However the work detailed in [4] employed thermal IR imaging (7.0 – 14 μm) and quite expensive as compared to near IR imaging focused in this paper.

1.2 Our Work

A summary of prior work presented in previous section suggests that there have been little efforts on *automated* face identification using near IR imaging and with notable exception in [3], none to investigate the utility of low quality near IR images for face identification. The review of prior work also suggests that the face identification with near IR imaging using elastic bunch graph matching (EBGM), trace transform and combination of multiple representations have not yet been investigated. *It is prudent to examine these approaches on near IR imaging as researchers have shown promising results using* the images acquired in visible illumination. There are several publicly available database acquired using visible illumination. However, to the best of our knowledge, there is no publicly available face database using near IR imaging. Therefore the image database acquired in this work is made available [10] to other researchers.

This paper presents a completely automated and accurate system for online face authentication using near IR images. The automated localization of eyes, orientation and scale correction, feature extraction using multiple representations and score combination steps have been employed to achieve promising performance. We perform comparative evaluation from the four matchers based on: EBGM, trace transform, linear discriminant analysis (LDA), and principal component analysis (PCA). Then the combination of these matchers is investigated for further performance improvement.

2. SYSTEM OVERVIEW

The block diagram of the near IR based multi-matcher face identification system is illustrated in figure 1. The frontal images of subjects are acquired using a near IR webcam. The acquired images are firstly subjected to eye center localization which extracts the coordinates of the eye centers. The eye centers serve as references for the face normalization. The face normalization mainly consists of scale correction, orientation correction, and extraction of region of interest, *i.e.*, face region using elliptical masking. The segmented face regions are used to extract features using multiple representations. In this we employed four matchers that are based on EBGM, trace

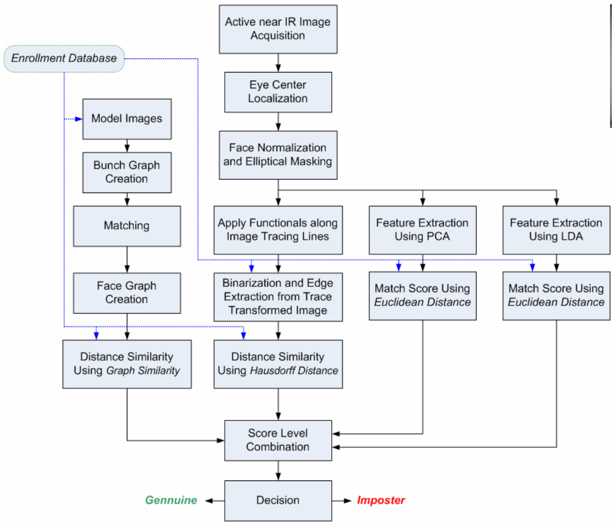

Figure 1: The block diagram of investigated approach

transform, LDA, and PCA. The EBGM requires bunch graph creation using model images which is completed during training phase. The matching scores from the four matchers are combined using score level fusion. The combined matching score is used to authenticate every user into one of the two, *i.e.* genuine or imposter, classes.

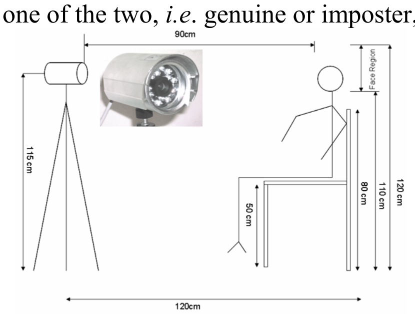

Figure 2: The image acquisition setup

2.1 Image Acquisition

The image acquisition setup used for imaging the subjects is shown in figure 2. The volunteers were requested to sit on the chair, change pose and expression, while imaging. The imaging was done in night at IIT Delhi. The webcam employed for imaging has 18 near IR LEDs whose illumination is centered at 850 nm. The images can be acquired at 30 frames per second and were of size 768 × 576 pixels. The average distance between the camera and face is 90 cm.

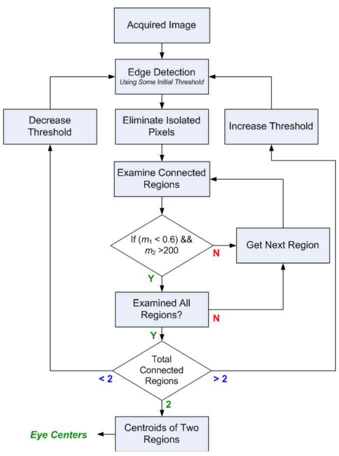

Figure 3: Eye center location algorithm for the acquired images

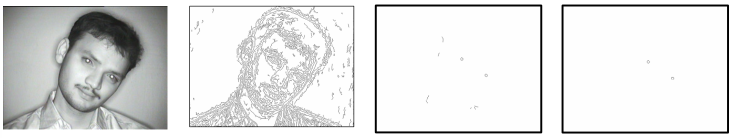

Figure 4: Steps in the automated eye center localization

3. EYE LOCALIZATION

The bright pupil effect generates a nearly perfect circle with white/bright illumination and can be easily detected from the near IR images. The eye detection using bright pupil algorithm has been detailed in [2]. However, as noted in [2], this approach is not robust and does not generate precise localization results. Therefore this algorithm has been modified to reliably locate the eye centers and is shown in figure 3.

The images are firstly subjected to multi-stage canny edge detection which smoothens the image using a Gaussian filter. The derivative of the smoothed image is subjected to thresholding using some initial threshold. The isolated pixels in the resulting binarized images are removed. The connected components (pixels) from this image are labeled and subjected to feature measurements (m_1 and m_2). The measurement m_1 represents convex area of the component while m_2 (e) represents the eccentricity of the connected component and estimated as follows:

$$b = \sqrt{((\rho_{11} - \rho_{12})^2 + 4\rho_{12}^2}\ , \quad a = \sqrt{((\rho_{11} - \rho_{12})^2 - 4\rho_{12}^2}$$

$$e = 2\frac{\sqrt{(b/2)^2 - (a/2)^2}}{b} \qquad (1)$$

where a and b represents the estimated length of semi minor and semi major axis respectively. The ρ_{11}, ρ_{22}, ρ_{12}, represents the normalized second-order moments of pixels in the image. The objective is to search two pupils which resemble circles in the binarized image. The pupil diameter is very small and order of 10 mm, therefore the convex area is chosen smaller (200). The shape of the targeted connected regions is nearly circular and therefore the low value of the eccentricity (0.6) is chosen.

The scale factor is the ratio of distance between located eye coordinate (source) to the destination eye coordinates. The angle between the line joining two pairs of coordinates, *i.e.*, source and destination, is the required rotation angle. Figure 5 illustrates the steps employed to achieve the face normalization using located eye centers.

4. FEATURE EXTRACTION

In this work, four face matchers were investigated for their performance on near IR images; EBGM, PCA, LDA, trace transform.

A. EBGM: The elastic bunch graph matching algorithm is based has been succinctly developed in stages which are detailed in several references [6]-[8]. A set of 40

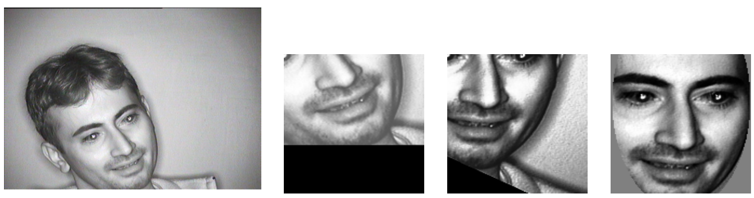

Figure 5: Orientation and scale correction steps for an image

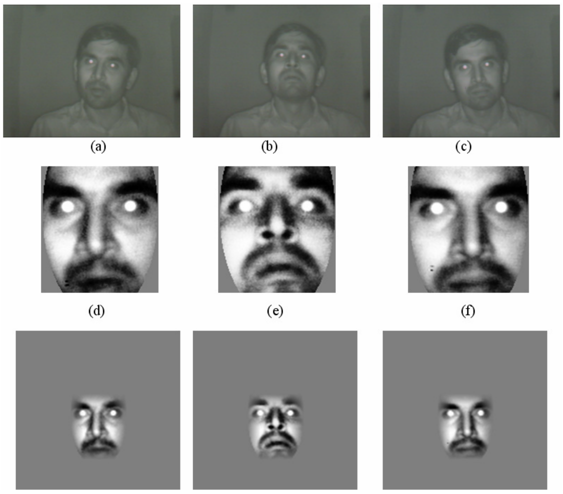

Figure 6: The acquired image samples, normalized images and corresponding normalized images for EBGM

complex coefficients (5 frequencies × 8 orientations) obtained from the convolution of Gabor kernels on an image point is referred as a jet. The phase similarity measure between two jets is computed as follows:

$$M_{\phi}(\psi,\psi') = \frac{\sum_p c_p c'_p \cos(\phi_p - \phi'_p - \vec{d}\vec{k}_p)}{\sqrt{\sum_p c_p^2 \sum_p c'^2_p}} \quad (2)$$

The implementations of image normalization, bunch graph localization, and extraction of similarity scores is detailed in [8] and utilized to generate the matching scores from the acquired images.

B. Trace Transform: The trace transform can be used to extract the facial features are invariant to translation, rotation and scaling. In this work we experimented with several different functions listed in [9] and selected the three functionals to generate the combined matching scores.

$$T(f(x)) = \int_0^\infty r\,|\,f(x)\,|\,dx, \quad T(f(x)) = \int_0^\infty \min\{f(x)\}dx,$$
$$T(f(x)) = \int_0^\infty |\,r\,| * |\max(f(r))|\,dr, \quad (3)$$

where $r = x - c$, and c = median$\{x, f(x)\}$. The similarity between the two trace transform shapes is measured by using Hausdorff distance.

C. PCA: The appearance based approach is detailed in several references and also referred as *eigenfaces*. The projection of training face image on *eigenfaces* is used to compute the characteristic features while Euclidean distance is employed for their classification.

D. LDA: The Linear Discriminant Analysis (LDA) is another appearance based approach which has been quite promising on visible face images. The LDA analysis employed in our work is same as detailed in [11].

5. EXPERIMENTAL RESULTS

The face images using near IR imaging were acquired using the simple setup detailed in section 2.1. The near IR face image database consists of the faces collected from the students and staff at IIT Delhi and the subjects in the

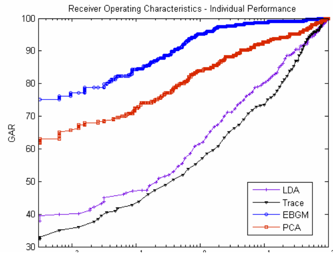

Figure 7: The comparative ROC from the four matchers

Table 1: Performance Indices for Individual Performance

	EER	DI
EBGM	3.9214 %	1.1314
PCA	7.843 %	0.8809
LDA	15.032 %	0.8146
Trace	18.62 %	0.7460

are in the age group 17-50 years. This database has been acquired in our campus during Feb - Jun 2007 using a webcam with near infrared illumination from the 102 users. The database has been mostly acquired in two stages and in each stage three images were acquired for every user. The average interval between the two image acquisitions is four week. The database of 612 images has wide variations in the pose and expression, and is available for the researchers [10]. All the acquired images are subjected to the automated eye localization and face normalization as detailed in section 3.

The figure 5 shows example of the steps employed to achieve face normalization. In this figure the automatically extracted source eye coordinates are (267, 407) and (373, 252) for left and right eyes respectively. The corresponding destination eye coordinates have been fixed to (30, 45) and (100, 45) respectively. Therefore the scaling factor and required angle of rotation is 0.5712 and 27.1924 respectively. The normalization scheme ensures smooth edges and scales down images to 128 × 128 pixels.

The dataset of three images, acquired during the first session, was used for the training and rest three images were used for the testing. The EBGM model images consisted of 50 images that were representative of the variability in the training images. The receiver operating characteristics (ROC) for each of the matcher is used to ascertain the performance. The quantitative performance is ascertained using equal error rate and the decidability index (DI) [3]. The distribution of genuine and imposter scores from the test data obtained using EBGM, trace transform and PCA matcher is illustrated in figure 8. The scores illustrated in figure 8 have been normalized in the range [0, 100] using min-max normalization. The ROC from each of the four matchers is shown figure 7 while the corresponding performance indices are shown in table 1. The experimental results suggest that the EBGM achieves best performance while the performance of trace transform is poor. The stability of trace transform shapes has been ascertained in [9] and it can tolerate reasonable level of pose and expression variations in visible face images. However, the employed near IR face database for performance evaluation has high variations in pose and expressions which is the plausible reason for poor performance from trace transform matcher. The LDA and trace transform matching scores were firstly combined to ascertain the performance improvement. The ROC plot and the performance indices from the average

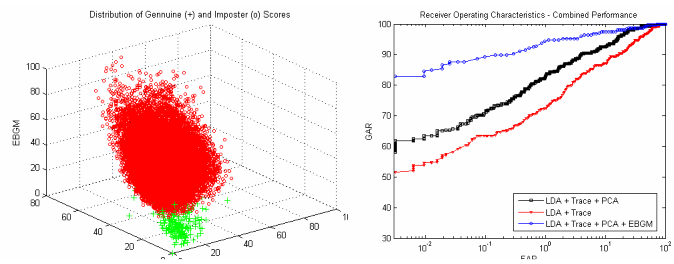

Figure 8: Distribution of genuine and imposter scores (left) and the ROC from the combination of matchers (right)

(sum rule) combination of two matchers is illustrated in figure 8 and table 2 respectively. This figure and the corresponding table also illustrate the ROC and performance indices when the PCA matching scores are combined with the combination of LDA and trace transform matching scores. The results suggest that the combination is useful to further improve the performance. The experimental results from the combination of all the four matchers, *i.e.* EBGM, trace transform, LDA and PCA, show further performance improvement while achieving the EER of 2.28%. The appearance based techniques, *i.e.* PCA and LDA, are very sensitive to the rotational, translational and scale variations in the acquired images. The experimental results suggest that the employed eye localization and face normalization schemes have been quite successful in handling these variations.

The computational complexity is the key for the online success of any combination of matchers for the near IR based face authentication. Therefore the computational time for the eye localization, face normalization and various components in the four matchers was computed and summarized in table 3. The computational time illustrated in table 3 employed GNU C++ compiler for Linux (Fedora Core 5) operating system running on Pentium IV processor (3.4 Ghz) with 2GB of RAM. We did not make any substantial efforts to optimize the code and utilized the implementation for EBGM available in [8] along with the OpenCV library [12]. The computational complexity estimated in table 3 for various implementations suggests that the combined computations can achieve online face authentication.

Table 2: Performance Indices for the Combined Performance

	EER	DI
LDA + Trace + PCA + EBGM	2.2876 %	3.3343
LDA + Trace + PCA	7. 148 %	1.9071
LDA + Trace	11.433 %	1.3093

Table 3: Computational time for different stages of the algorithm

Operation	Time (msec)
Image Acquisition	150
Eye Localization	200
Face Normalization	2.4
EBGM Normalization	1
Fitting Graphs to Novel images + Bunch Graph creation	50
Extracting Face Graph	200
Similarity measurement for EBGM	80
Feature extraction using Trace transform	3
Hausdorff distance calculation	100
Creating *Eigenspace* using Training set	40
Projecting test image onto *Eigenspace*, and also calculating Euclidean distance	10
Creating *Fisherspace* using Training set	60
Projecting test images on to *Fisherspace* and distance calculation	10

6. CONCLOSIONS

This paper has suggested a *completely automated* and reliable system for online face authentication using low quality (cost) near IR images. The automated localization of eyes, orientation and scale correction, feature extraction using multiple representations and score combination steps have been investigated to achieve reliable authentication. The experimental results have suggested that our eye localization method is quite accurate and fast. The comparative performance evaluation from the four matchers (EBGM, trace transform, PCA and LDA) suggests that the best performance is achieved by EBGM matcher while the combination of all matchers can effectively improve the performance.

The primary goal of our work is to build a reliable automated face authentication system that can perform in night, using near IR imaging, and potentially be used for secured access control. Our investigation in this paper has been limited to the problem of authentication and assumed that users are fairly cooperative so that the images without any mask and eyeglasses can be acquired. Further investigation for the recognition problem on variety of images will be useful and is the part of the ongoing work.

7. REFERENCES

[1] T.-Y X. Zou, J. Kittler, and K. Messer, “Face recognition using active near-IR illumination,” *Proc. British Machine Vision Conf.*, pp. 209-219, Sep. 2005.

[2] S. Zhao and R.-R. Grigat, “An automatic face recognition system in the near infrared spectrum,” *Proc. MLDM*, LNAI 3587, pp. 437-444, Jul. 2005.

[3] S. Z. Li, R. Chu, S. Liao and L. Zhang, “Illumination invariant face recognition using near-infrared images,” *IEEE. TPAMI*, vol. 29, no. 4, pp. 627-639, Apr. 2007.

[4] X. Chen, P. J. Flynn, and K. W. Bowyer, “IR and visible light face recognition,” *Computer Vision & Image Understanding*, vol. 99, pp. 332-358, 2005.

[5] S. G. Kong, J. Heo, B. R. Abidi, J. Paik, M. A. Abidi, “Recent advances in visual and infrared face recognition,” *Comp. Vis. & Image Under.*, vol. 97, pp. 103-135, 2005.

[6] L. Wiskott, J.-M. Fellous, N. Krüger, and C. von der Malsburg, “Face recognition by elastic bunch graph matching,” *IEEE TPAMI*, vol. 19, pp. 775–779, Jul. 1997

[7] K. Okada, J. Steffens, T. Maurer, H. Hong, E. Elagin, H. Neven, and C. von der Malsburg, “The Bochum/USC face recognition system and how it fared in the FERET phase III test,” *Face Recognition: From Theory to Applications*, Springer-Verlag, pp. 186-205, 1998.

[8] http://www.cs.colostate.edu/evalfacerec/algorithms5.html

[9] A. Kadyrov and M Petrou, “The Trace transform and its applications” *IEEE TPAMI*, vol. 23, pp 811-828, 2001.

[10] http://web.iitd.ac.in/~ajaykr/FaceIR.htm

[11] P. Belhumeur, J. P. Hespanha, D. J. Kriegman, “Eigenfaces vs. Fisherfaces: recognition using class specific linear projection”, *IEEE TPAMI,* vol. 19, no. 7, pp. 711-720, 1997.

[12] http://sourceforge.net/projects/opencvlibrary