Recognition of Quantized Still Face Images

Tao Wu and Rama Chellappa

Abstract— **In applications such as document understanding, only binary face images may be available as inputs to a face recognition (FR) algorithm. In this paper, we investigate the effects of the number of grey levels on PCA, multiple exemplar discriminant analysis (MEDA) and the elastic bunch graph matching (EBGM) FR algorithms. The inputs to these FR algorithms are quantized images (binary images or images with small number of grey levels) modified by distance and Box-Cox transforms. The performances of PCA and MEDA algorithms are at 87.66% for images in FRGC version 1 experiment 1 after they are thresholded and transformed while the EBGM algorithm achieves only 37.5%. In many document understanding applications, it is also required to verify a degraded low-quality image against a high-quality image, both of which are from the same source. For this problem, the performances of PCA and MEDA are stable when the images were degraded by noise, downsampling or different thresholding parameters.**

I. INTRODUCTION

As an important biometric feature, face images have been used widely to identify humans. In most applications, face images to be identified are grey or colored, which have sufficient information to extract good features, especially when the face images are obtained under controlled conditions. A review of face recognition research conducted before 2003 may be found in [1].

In some special cases, however, face images are collected by scanning a passport or a drivers license. In these scanned images, pixels in the face image may not have a wide range of possible values. When face images are photocopied or faxed, the grey levels may also be distorted or lost. The most commonly used photocopiers and fax machines work in black and white mode only. They scan the original pages and binarize the images using a global threshold. Some techniques have been developed to locate faces using edge information on black and white images [2].

In recent years it has become common to collect biographical or descriptive information about humans by scanning documents. Technologies to analyze scanned documents, such as page segmentation, optical character recognition etc. have been developed over the last 20 years. However, the information collected from different sources may be inconsistent. It is desired to validate and verify the face images collected from these low-quality sources. In these cases, both the gallery and probe set may consist of binary face images only. Thus, face recognition techniques that work on binary images are needed.

Tao Wu and Rama Chellappa are with the Department of Electrical & Computer Engineering and the Center for Automation Research, UMIACS, University of Maryland, College Park, MD 20742, USA taowu@umiacs.umd.edu; rama@cfar.umd.edu

Generally, face recognition task includes two steps. First, some features are extracted from a face image. Then identification or classification is accomplished using the extracted features. When face images are binarized using a global threshold, significant information is lost. Traditional intensity-based recognition algorithms would not work well on binary face images. The loss of information will lead to a reduction in recognition accuracy of almost all popular face recognition algorithms.

Turk and Pentland [3] proposed an eigen face method which used principal component analysis (PCA) to identify face images. This method has become a common method for face recognition on greyscale face images [4]. When the number of grey levels decreases, the performance of PCA could be impacted. Independent component analysis (ICA) was introduced by Jutten et al [5] and Comon et al [6]. This method tries to find statistically independent components lying in the observed data. ICA may have a better performance than PCA under some conditions [7]. In practice, the probability distribution of the observed data is usually unavailable. ICA derives independent components through numerical methods, which are not always guaranteed to yield strictly statistically independent components. Schein etc. proposed a logistic version of PCA for analyzing binary data [8]. Although logistic PCA can achieve less reconstruction error than PCA applied to binary data, it has neither a closed form of computation, nor a unique solution. This restricts the application of logistic PCA to pattern recognition problems. Tang and Tao proposed a binary PCA method [9] by combining the PCA algorithm with Haar-like binary box functions. This method actually was designed to decompose intensity images into a linear combination of binary box functions. Maver and Leonardis proposed a PCA-based method to recognize binary images using grey-level parametric eigenspaces [10]. Their methods attemptted to match binary face images against a greyscale image by estimating the information lost during the binarization using eigenvectors obtained from greyscale images. They introduced an assumption that the true values of edge pixels in the binary image equal to the threshold value. Generally, this assumption does not hold. Thus, their method cannot estimate the lost information effectively. Belhumeur etc. [11] proposed Fisherfaces, which adopted linear discriminant analysis (LDA), to analysis the feature extracted from face images. This method can be related to the Bayesian rule when the features obey Gaussian distribution, which is not true for binary images.

Other popular face recognition methods, based on elastic bunch graph matching (EBGM) [12], active appearance

model (AAM) [13], active shape model (ASM) [14] and albedo estimation [15], will also face some difficulties, since the intensive feature cannot be extracted from binary images. Although local binary pattern (LBP) based methods [16] extract features through local binarization, it also encounters problems on a binary image obtained by a global threshold

In this paper, we first investigate the performance of PCA-based face recognition algorithm under different number of grey levels. Then we process the binary face images with distance and Box-Cox transforms, which make the probability distribution of pixels much more Gaussian-like, and analyze the performance of the combination of PCA and MEDA [17] algorithms on the transformed face images.

The rest of the paper is organized as follows: In section 2, we review some face recognition algorithms suitable for binary images. In section 3, experimental results using the proposed methods are presented. The influence of different numbers of grey levels, the effects of different global binarization criterion and the performance of the method under different degradation models are discussed. A summary and conclusions are given in section 4.

II. ALGORITHMS

A. Review of Principal Component Analysis

Let C denotes the covariance matrix of images, then the PCA basis vectors are obtained by solving the eigenvalues and eigenvectors of C:

$$C = \frac{1}{N} \sum_{i=1}^{N} (x_i - m)(x_i - m)^T \quad C = V \Sigma V^T \quad (1)$$

where N is the number of images, x_i is the ith image, m is the mean of all images, Σ is a diagonal matrix of all eigenvalues, $V = [v_1...v_i...v_N]$. v_i is the eigenvectors corresponding to the ith eigenvalue λ_i. Assume $\lambda_1 \geq ...\lambda_i \geq ...\lambda_N$. The PCA coefficients are obtained by:

$$y_i = W^T x_i \quad (2)$$

where $W^T = [v_1...v_i...v_n]$, n is the number of desired principal components.

Though PCA does not make any assumption on the probability distribution of the pixel values, when the values of pixels obey a Gaussian distribution, PCA coefficients are independent and obey a Gaussian distribution. The eigenvalue λ_i is proportional to the energy, of the image along the direction of v_i. So the number of desired principal components n could be determined by specifying the percentage of energy to be preserved:

$$n = \arg\min \frac{\sum_{i=1}^{n} \lambda_i}{\sum_{i=1}^{M} \lambda_i} > p \quad (3)$$

where M is the number of pixels of an image, p is the specified percentage of energy to be preserved. When the input to PCA does not obey a Gaussian distribution, λ_i still represent the energy, since $x_i - m$ is a zero-mean random vector. The PCA coefficients are neither necessarily independent, nor Gaussian, however.

B. Multiple-Exemplar Discriminant Analysis

Linear discriminant analysis (LDA) is widely used in the field of pattern recognition. MEDA [17] uses several exemplar or even the whole sample set to represent each class. Its effectiveness was demonstrated in [18]. MEDA is an extension of LDA, but has different definitions of within-class and between-class scatter matrices. The MEDA method is summarized as follows. Let x_j^i denotes the jth sample in the ith class. The within-class scatter matrix Σ_W and between-class matrix Σ_B are defined by

$$\Sigma_W = \sum_{i=1}^{C} \frac{1}{N_i} \sum_{j=1}^{N_i} \sum_{k=1}^{N_i} (x_j^i - x_k^i)(x_j^i - x_k^i)^T \quad (4)$$

$$\Sigma_B = \sum_{i=1}^{C} \sum_{j=1;j\neq i}^{C} \frac{1}{N_i N_j} \sum_{k=1}^{N_i} \sum_{l=1}^{N_j} (x_k^i - x_l^j)(x_k^i - x_l^j)^T \quad (5)$$

where C and N_i are the number of all samples and the number of samples in the $i$the class respectively. Then the projection matrix W is obtained by maximizing the function

$$J_W = \frac{\det\left(W^T \Sigma_B W\right)}{\det\left(W^T \Sigma_W W\right)} \quad (6)$$

Generally, MEDA requires multiple samples for each subject in the gallery set. For the face recognition problem, because of the symmetry of human faces, images in the gallery set could be mirrored vertically so as to increase the number of samples.

III. RECOGNITION OF QUANTIZED FACE IMAGES

A. Quantization and Binarization Method

In our experiments, images with different numbers of grey levels were obtained by quantizing the original images using the minimum mean square error (MMSE) criterion [20]. Some examples of normalized face images quantized by the MMSE quantizer are given in Fig. 1.

(a) (b) (c) (d)

Fig. 1. Face images quantized by the MMSE quantizer. (a) The original grey image with 256 grey levels. (b) The original image is quantized into 8 grey levels. (c) The original image is quantized into 4 grey levels. (d) The original image is quantized into 2 grey levels.

Although the quantizer is able to quantize an image into a specified number of grey levels using the MMSE criterion, it is sensitive to illumination variations if it is used to binarize images. When recognizing faces from binary images, regions around eyes, nose and mouth are very important. Since the illumination on the whole face is generally not evenly-distributed, the global binarization threshold could be affected even by slight shadows in the above-mentioned regions. In Fig. 1(a), the surrounding region of the original

face image is slightly darker than the center region. So the MMSE quantizer assigned darker values to the pixels in the surrounding regions, as shown in Fig. 1(d). The shapes of facial organs were submerged in this result, though the quantization error is minimized. We adopted the MMSE quantizer in the experiments to investigate the effect of the number of grey levels on the performance of face recognition algorithms.

For the binary face recognition problem, we adopted a contrast based quantizer. Regions, such as eyes, nose and mouth, on a face image are generally much darker than other regions. The shapes of these regions could be better preserved when the original image is binarized under a contrast criterion, which requires that the bright pixels in the binary result account for a specified percentage. An example is given in Fig. 2. Compared to the MMSE quantizer, this method could not only preserve the shapes of the organs better in binary result, but also can be easily implemented in an ordinary fax machine.

Fig. 2. The original image in Fig. 1 is binarized under the contrast criterion. The percentage of the bright pixels is 80%.

B. Dataset

The performance of PCA, MEDA and the elastic bunch graph matching (EBGM) algorithms was investigated using the Face Recognition Grand Challenge (FRGC) database version 1 [19]. In this database, experiment 1 is focused on the recognition of still frontal face images obtained under controlled illumination. In the gallery set there is only one controlled still image for each subject. Totally, there are 152 images in the gallery set for 152 subjects, 608 images in the probe set. Images were mirrored vertically in order to meet the requirements of MEDA. Similar to experiment 1, experiment 2 is focused on the face recognition problem under controlled illumination, but has 4 different images for each subject in the gallery set. There are totally 608 images in the gallery set for 152 subjects and 2432 images in the probe set. Experiment 4 concerns with face recognition under different illumination conditions, where the images in the gallery set were obtained under controlled conditions and those in the probe set were obtained under uncontrolled conditions. We investigated the performance of the algorithms on experiments 1, 2 and 4 of FRGC.

C. Effect of the number of grey levels

In order to investigate the effect of the number of grey levels, we quantized the images in version 1 experiment 1 database using the MMSE quantizer. The mean peak signal noise ratio (PSNR) of the output of the quantizer when the original images are quantized into different number of grey

levels is plotted in Fig. 3. The quantization PSNR drops as the number of grey levels decreases. When there are less than 8 grey levels, the PSNR is less than 35db. A significant amount of information is lost.
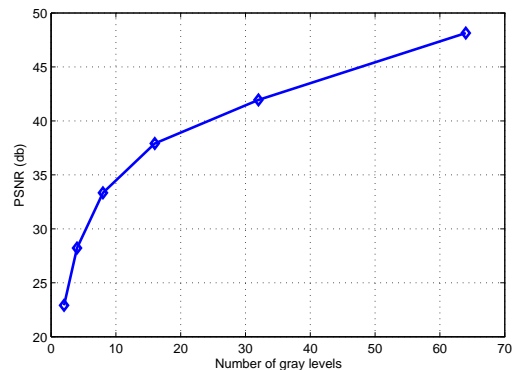

Fig. 3. The mean PSNR when the images were quantized into different number of grey levels .

The eigenvalues of the PCA covariance matrix estimated from the images with different number of grey levels are plotted in Fig. 4. The eigenvalues of the covariance matrix were sorted in the descending order. It shows that for the
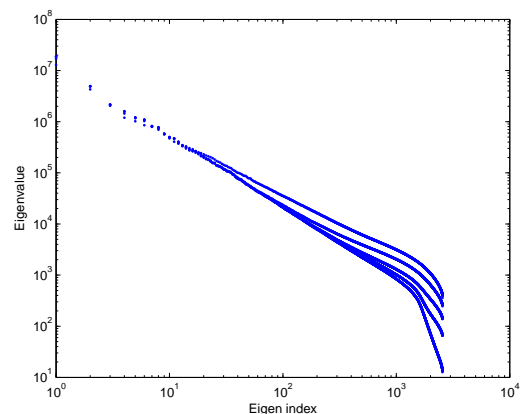

Fig. 4. Eigenvalues of the covariance matrix of 2, 4, 8, 16, 256 grey levels images. Their spectrums almost overlap in low order eigenvalues, and have high order tails. The tails from up to down are from 2, 4, 8, 16, 256 grey levels images.

top n eigenvalues, the percentage of the energy they account for decreases as the number of grey levels decreases. As the number of grey levels decreases, the values of the middle and high orders eigenvalues of the covariance matrix increases, and the energy tends to spread to high order eigenvalues. According to Fig. 3, when images are quantized into fewer greylevels, the power of quantization noise increases. This might be a major reason for the energy spread phenomenon.

We quantized all the images in training, gallery and probe sets of experiment 1 and tested the performance of PCA, MEDA and EBGM algorithms. The recognition accuracies of the algorithms using the Euclidean distance or the cosine

of the angles between two representations [19] as the metric on rank 1 when the images have different number of grey levels are plotted in Fig. 5. It appears that the combination
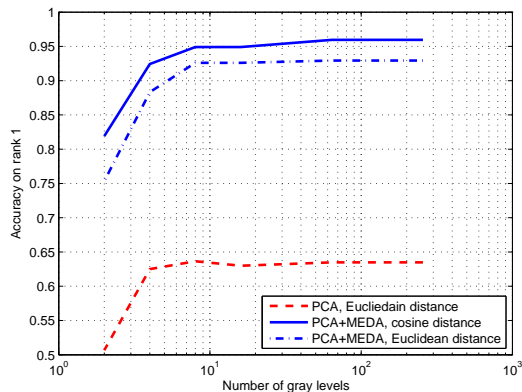

Fig. 5.   Recognition accuracies of PCA and PCA+MEDA methods on rank 1 when using different distance metrics and the images have different number of grey levels on FRGC version 1 experiment 1. The x-axis is logarithmically scaled. The performance of PCA using the cosine distance and Euclidean distance are almost the same, so only the one using Euclidean distance is plotted.

of PCA and MEDA always has better performance than the PCA only method. The cosine distance metric yields a better result than the Euclidean distance metric. The performance of each algorithm is almost the same, when the number of grey levels is greater than 8. But the recognition accuracies drop severely when the images have fewer grey levels. Since the quantization noise increases as the number of grey levels decreases, according to Fig. 3, much information could be lost or corrupted when the images are nearly binary. This might be an leading to the reduction in performance.

D. Performance comparison on binary images

The distribution of the values of pixels in a binary image is much more like a Bernoulli distribution, rather than a Gaussian distribution. It has been suggested in [8] that the performance may drop for non-Gaussian data. We tested the performance of PCA-based methods after the distribution of the pixels were transformed to be more Gaussian.

We tried two different ways to transform the binary images. One is to convolute the images with a Gaussian kernel so that the distribution of pixels will be more like Gaussian than Bernoulli. The other way is to perform a distance transform (DT)[21] on the binary images first, in order to extend the support of the pixels values. Each pixel is assigned a value of the Euclidean distance from it to its nearest edge point. In our experiment, the pixels were assigned positive values if they were white in the binary images. Otherwise, they were assigned negative values. Then the Box-Cox transform[22] is adopted to convert the images processed by DT to be more like Gaussian. The eigenvalues, in descending order, of the covariance matrices estimated from images binarized by the contrast criterion, and from the transformed binary images are plotted in Fig. 6. After Gaussian or distance and Box-Cox transformations, the energy of
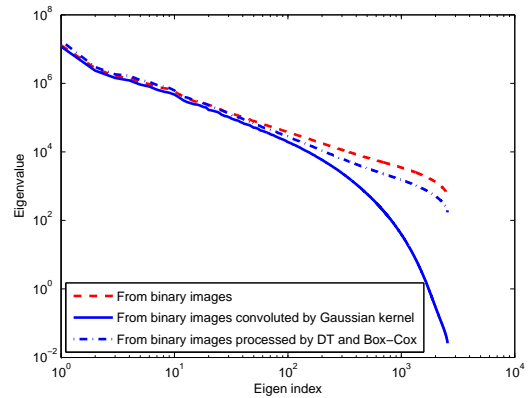

Fig. 6.   Eigenvalues of the convariance matrices estimated from binary images and transformed binary images.

the covariance matrix concentrates more in the low to middle order eigenvalues. Altough Gaussian kernel could make most of the energy concentrate in low order eigenvalues, it blurs the image and suppresses the high frequency components, which generally carry significant discriminant information. The distance and Box-Cox transformations do not damp these components. A cumulative match curve (CMC) comparison of the PCA + MEDA method on binary images, images processed by Gaussian convolution and images processed by distance and Box-Cox transformation is shown in Fig. 7.
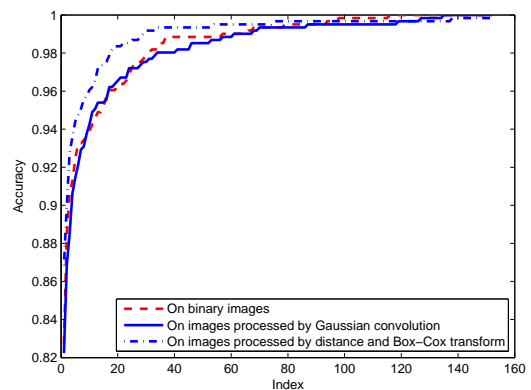

Fig. 7.   Performance on binary FRGC version 1 experiment 1 images. The accuracies on rank 1 for binary images, images processed by Gaussian convolution and images processed by distance and Box-Cox transformation are 82.73%, 82.24% and 87.66%, respectively.

We binarized the images in FRGC experiments 1, 2 and 4 by the contrast criterion, and then processed them with distance and Box-Cox transformations. The CMC curve of PCA + MEDA method is shown in Fig. 8, which shows that with the help of distance and Box-Cox transforms, the accuracy of PCA + MEDA method on rank 1 is 87.66%, which is about 10% lower than the performance on 256 grey levels image (which is 95.96%, as shown in Fig. 5). In experiment 4, the original images were obtained under uncontrolled illumination conditions. The illumination on the

face region varies significantly. The shadows on the face images severely corrupt the output of the global binarization, thus the performance in this situation is very low.
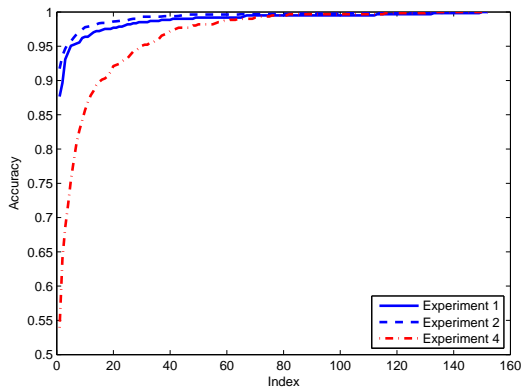

Fig. 8. Performance of PCA+MEDA on binarized images from experiment 1, 2, and 4 of FRGC. The accuracies on rank 1 are 87.66%, 91.74% and 53.95%, respectively.

In addition, we tested the EBGM method implemented by CSU [23] on the same binary dataset obtained from experiment 1 so as to investigate whether binary image data could have any impact on feature-based algorithm. Fig. 9 shows the best result we got. The reason for the poor performance may be that the bunch graph cannot be fit precisely and the features cannot describe the images effectively. This implies that the lose of intensity information could severely reduce the performance of intensity-based methods.
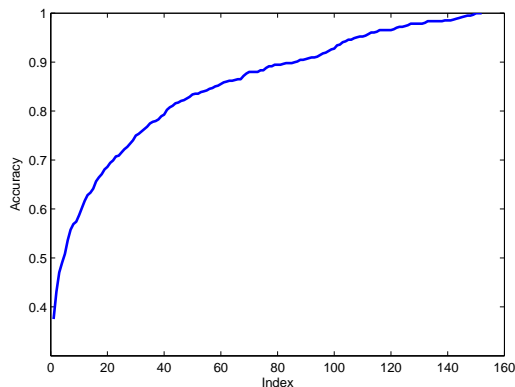

Fig. 9. Performance of EBGM on binarized images from experiment 1.

E. Face verification under noise, down sampling and different binarization threshold

In the application of binary face image recognition, the problem of verifying a degraded low-quality image against a high-quality image is also of interest. When processing documents containing binary text and face images information about one's identity, it is required to verify if the text and face images are consistent. In this case, a high-quality binary image may be obtained by searching a database using the text information, and cropping the low-quality binary face images from the document. Actually, the binary high-quality and low-quality images are from the same source, but the low-quality images are degraded. We used the images in FRGC version 1 experiment 1 to simulate this situation. The gallery set is constructed from binary face images with 80% contrast. An example is shown in Fig. 2. The probe set is binarized with different global threshold from the same source images, or degraded by adding random noise or downsampling. Fig. 10 shows some examples of the images in the gallery set and the degraded images in the probe set. The performance of the
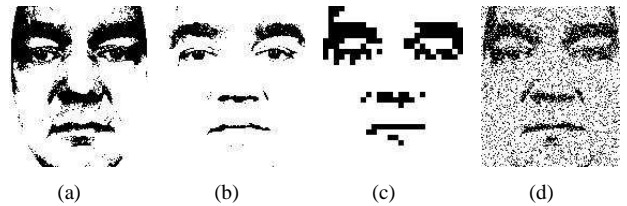

Fig. 10. Degraded binary face images. (a) A binary face image obtained from the original grey image with a contrast of 70%. (b) A binary face image obtained from the original grey image with a contrast of 90%. (c) A binary face image with a contrast of 80% was downsampled to 20% of the original size. (d) A binary face image degraded by additive random noise, PSNR=8db.

PCA+MEDA method on noisy images, downsampled images and images binarized with different parameter under the contrast criterion is shown in Fig. 11, 12 and 13, respectively. According to these results, for the same source images verification problem, the algorithm is able to maintain a high accuracy rate on rank 1 when the PSNR is greater than 7 db, or the images are downsampled to not less than 20% of the original size, or the parameters in the contrast based binarization varies no more than 10%.
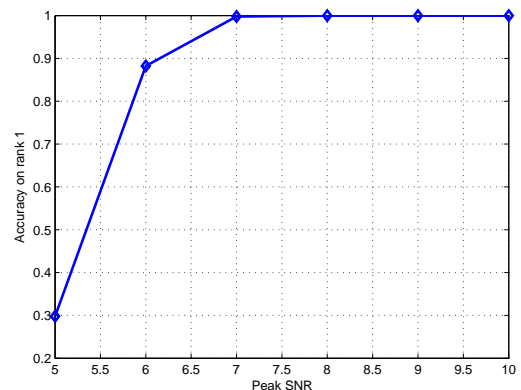

Fig. 11. Same source verification rate under noise.

IV. CONCLUSIONS

In this paper, we investigated the effect of the number of grey levels on the performance of PCA, MEDA and EBGM methods. When there are more than 8 grey levels, the performance is only slightly affected. Otherwise, the performance
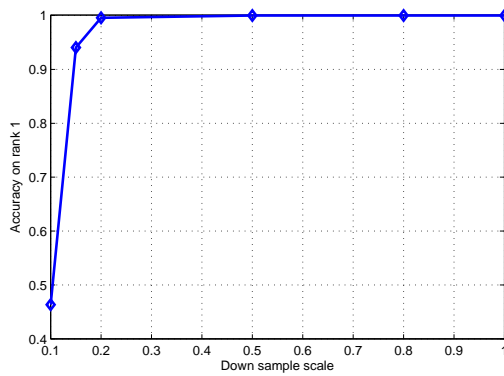
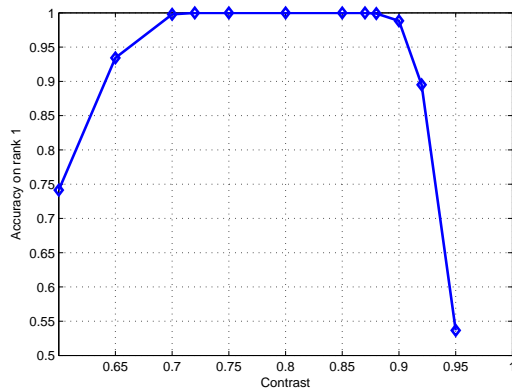Fig. 12.   Same source verification rate on down sampled images.

Fig. 13.   Same source verification rate on binary images obtained by different threshold.

drops severely. The output of MMSE quantizer could be affected by uneven illumination if it is used for binarization. A contrast-based quantizer can convey shape information about facial organs more precisely. With the help of distance and Box-Cox transforms, which make the distribution to be more Gaussian-like and concentrate the energy in the lower orders of eigenvalues, the performance of PCA + MEDA method achieved an accuracy of 87.66% on rank 1 in FRGC version 1 experiment 1. Compared to accuracies of 95.96% on 256 grey levels images, the performance decreases about 10% for binary images. The EBGM method did not perform well. For the same source verification problem, that arises in many document understanding applications, the performance of the algorithm is stable. Variations in illumination, however, will lead to a severe performance drop when the images are globally binarized.

V. ACKNOWLEDGMENT

REFERENCES

[1] W. Zhao, R. Chellappa, P. J. Phillips and A. Rosenfeld, Face Recognition: A Literature Survey, *ACM Computing Surveys*, vol. 35, 2003, pp. 399-458

[2] V. Govindaraju, D.B. Sher, R.K. Srihari, S.N. Srihari, Locating Human Faces in Newspaper Photographs, *In Proceedings of the 1989 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'89)*, 1989, pp. 549-554

[3] M. Turk and A. Pentland, Eigenfaces for Recognition, *J. Cogn. Neurosci.*, vol. 3, 1991, pp. 72-86

[4] G. Shakhnarovich and B. Moghaddam, Face Recognition in Subspaces, *Handbook of Face Recognition*, Springer, 2004, pp. 141-168

[5] C. Jutten and J. Hrault, Blind Separation of Sources, part I: An adaptive algorithm based on neuromimetic architecture, *Signal Processing*, vol. 24, 1991, pp. 1-10

[6] P. Comon, Independent Component Analysis, A New Concept? *Signal Processing*, vol. 36, 1994, pp. 287-314

[7] M. Bartlett, H. Lades, and T. Sejnowski, Independent Component Representations for Face Recognition, *In Proceedings of the SPIE: Conference on Human Vision and Electronic Imaging III*, vol. 3299, 1998, pp. 528-539

[8] A. Schein, L. Saul and L. Ungar, A Generalized Linear Model for Principal Component Analysis of Binary Data, *In Proceedings of the 9th International Workshop on Artificail Intelligence ans Statistics*, 2003, pp. 546431

[9] F. Tang and H. Tao, Binary Principal Component Analysis, *In Proceedings of British Machine Vision Conference (BMVC)*, vol. 1, 2006, pp. 377-386

[10] J. Maver and A. Leonardis, Recognizing 2-tone Images in Grey-level Parametric Eigenspaces, *Pattern Recognition Letters*, vol. 23, 2002, pp. 1631-1640

[11] P. Belhumeur, J. Hespanha and D. Kriegman, Eigenfaces vs. Fisherfaces: Recognition using Class Specific Linear Projection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, 1997, pp. 711-720

[12] L. Wiskott, J.Fellous, N. Krüger and C. von der Malsburg, Face Recognition by Elastic Bunch Graph Matching, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, 1997, pp. 775-779

[13] T. F. Cootes, G. J. Edwards and C. J. Taylor, Active Appearance Models, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, 2001, pp. 681-685

[14] T. F. Cootes, C. J. Taylor, D. H. Cooper and J Graham, Active Shape Models–Their Training and Application, *Comput. Vis. Image Underst.*, vol. 61, 1995, pp. 38-59

[15] S. Biswas, G. Aggarwal and R. Chellappa, Robust Estimation of Albedo for Illumination-Invariant Matching and Shape Recovery, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, 2009, pp. 884-899

[16] A. Hadid, M. Pietikainen and T. Ahonen, Face Description with Local Binary Patterns: Application to Face Recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, 2006, pp. 2037-2041

[17] S. Zhou and R. Chellappa, Multiple-Exemplar Discriminan Analysis for Face Recognition, *In Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04)*, vol. 4, 2004, pp. 191-194

[18] G. Aggarwal, S. Biswas and R. Chellappa, UMD Experiments with FRGC Data, *In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 3, 2005, pp. 172

[19] P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min and W. Worek, Overview of the Face Recognition Grand Challenge, *In IEEE Conference on Computer Vision and Pattern Recognition*, 2005, pp. 947-954

[20] J. Max, Quantizing for Minimum Distortion, *IRE Transaction on Information Theory*, vol. 6, 1960, pp. 7-12

[21] G. Borgefors, Hierarchial Chamfer Matching: A Parametric Edge Matching Algorithm, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, 1988, pp. 849-865

[22] R.M. Sakia, The Box-Cox Transformation Technique: A Review, *The Statistician*, vol. 41, 1992, pp. 169-178

[23] J.R. Beveridge, D. Bolme1, B.A. Draper and M. Teixeira, The CSU Face Identification Evaluation System, *Machine Vision and Applications*, vol. 16, 2005, pp. 128-138