2010

Finding distinctive facial areas for face recognition

Ce Zhan
University of Wollongong, czhan@uow.edu.au

Wanqing Li
University of Wollongong, wanqing@uow.edu.au

Philip O. Ogunbona
University of Wollongong, philipo@uow.edu.au

Finding distinctive facial areas for face recognition

Abstract

One of the key issues for local appearance based face recognition methods is that how to find the most discriminative facial areas. Most of the existing methods take the assumption that anatomical facial components, such as the eyes, nose, and mouth, are the most useful areas for recognition. Other more elaborate methods locate the most salient parts within the face according to a pre-specified criterion. In this paper, a novel method is proposed to identify the discriminative facial areas for face recognition. Unlike the existing methods that only analyze the given face, the proposed method identifies the distinctive areas of each individual's face by its comparison to the general population. In particular, non-negative matrix factorization (NMF) is extended to learn a localized non-overlapping subspace representation of the facial patterns from a generic face image database. In the learned subspace, the degree of distinctiveness for any facial area is measured depends on the probability of this area is belong to a general face. For evaluation, the proposed method is tested on exaggerated face images and applied in exiting face recognition systems. Experimental results demonstrate the efficiency of the proposed method.

Keywords

Finding, distinctive, facial, areas, for, face, recognition

Disciplines

Physical Sciences and Mathematics

Publication Details

Zhan, C., Li, W. & Ogunbona, P. (2010). Finding distinctive facial areas for face recognition. 2010 11th International Conference on Control, Automation, Robotics and Vision (ICARCV) (pp. 1848-1853). Piscataway, New Jersey, USA: IEEE.

2010 11th Int. Conf. Control, Automation, Robotics and Vision
Singapore, 7-10th December 2010

Finding Distinctive Facial Areas for Face Recognition

Ce Zhan, Wanqing Li, and Philip Ogunbona
School of Computer Science and Software Engineering
University of Wollongong, Australia
Email: {cz847, wanqing, philipo}@uow.edu.au

***Abstract*—One of the key issues for local appearance based face recognition methods is that how to find the most discriminative facial areas. Most of the existing methods take the assumption that anatomical facial components, such as the eyes, nose, and mouth, are the most useful areas for recognition. Other more elaborate methods locate the most salient parts within the face according to a pre-specified criterion. In this paper, a novel method is proposed to identify the discriminative facial areas for face recognition. Unlike the existing methods that only analyze the given face, the proposed method identifies the distinctive areas of each individual's face by its comparison to the general population. In particular, non-negative matrix factorization (NMF) is extended to learn a localized non-overlapping subspace representation of the facial patterns from a generic face image database. In the learned subspace, the degree of distinctiveness for any facial area is measured depends on the probability of this area is belong to a general face. For evaluation, the proposed method is tested on exaggerated face images and applied in exiting face recognition systems. Experimental results demonstrate the efficiency of the proposed method.**

***Index Terms*—Face Recognition, Feature Extraction, Saliency Detection, NMF**

I. Introduction

Face recognition has been one of the most active research topics in computer vision for more than 20 years. Although significant progress have been achieved [1], reliable face recognition under unconstrained conditions still remains a difficult problem far from being solved. This is mainly due to the fact that facial appearances are easily affected by variations of pose, illumination, expression, occlusion and other factors. At the same time, sample images per subject are often limited in realistic application scenarios. Insufficient samples (comparing with the dimensionality of the feature space) makes it even more difficult to train a robust face recognition system, since intra-person variations could not be well estimated.

Among numerous methods recently proposed to solve the face recognition problem in real life environment, local appearance based methods have received more and more attention. Unlike holistic methods that make use of the global information from the entire face and treat all facial areas equally important (e.g. eigenface [2], LDA [3]), local methods rely on features extracted from different facial parts. Comparing with global features, local features are generally more robust to the above mentioned variations, since most of the variations in appearance affect only part of the face. Further more, local based approaches also seem more suitable for handling the small sample problem, for they provide additional flexibility to recognize a face based on its parts. However, an additional challenging issue is brought to local appearance based methods, that is how to identify the most discriminative facial areas. The amount of useful information for recognition in one face is not uniformly distributed within the face image, thus different facial areas where the local features are extracted from should not be treat equally important in local methods. For those methods that extract features only on selected regions (e.g. [4], [5]), features extracted from non-discriminative image areas increases the required processing resources, on the other side, the non-discriminative features may drift or bias the classifier's responses. For local methods that simply partition the face image into sub-regions, and directly extract local features on each of the sub-regions (e.g. [6], [7]), a weight is always set for each region based on the importance of the information it contains. Still, bias is introduced when high weights are set for non-discriminative facial regions.

Rather than identifying the discriminative facial areas adaptively, some of the local based face recognition methods make an assumption that anatomical facial components, such as the eyes, nose, and mouth, are the most discriminative areas. For example, in [6], predefined weights are set to equally partitioned square facial regions. Eyes and mouth areas are given higher weights while cheek regions are given zero weight. In the famous Elastic Bunch Graph Matching (EBGM) method [4], Gabor responses are extracted at predefined fiducial points (facial landmarks), which are the key points on facial components such as eye corners, mouth corners, eyebrow corners and nostril corners. Unfortunately, the assumption is not always true, characteristic facial features can not be reduced to just eyes, nose and mouth. For some of the faces, other facial features such as philtrums, chins and the gaps around the eyes, nose and mouth could be more discriminative for recognition, especially when they contain scars, spots, dimples and lines. Other more elaborate methods often employ interesting points detectors to locate salient facial regions, so that features extracted from the region around the detected points could be invariant to factors like noise, illumination and viewpoint. Recently, scale invariant feature transform (SIFT) [8] has been successfully used in face recognition [5] [9] to extract local features. The SIFT descriptor detect key points in an image by means of a local optimization process applied to the difference of Gaussians image, filtered at different scales and orientations. Unlike the SIFT descriptor tends to look for blob-like features, the Harris-Laplace detector [10] looks for points in the image whose value of cornerness is locally maximal. Thus, it finds corner-like or junction-like features in images, and is employed to select local regions for face recognition in [11]. Kepenekci et al. [12] detect interesting points based on Gabor features, they believe that points with high-energized Gabor wavelet response contain more information on a face image, and only extract Gabor features at these points to improve EBGM.

In this paper, we propose a novel method to identify the discriminative facial areas for face recognition. Unlike the above mentioned "interesting regions" based methods that only analyze the given face, the proposed method find the characteristic areas of each individual's face by its comparison to the general population. The selected regions are the most distinctive facial areas that make the given face different from other faces "we know". The idea of comparative analysis is in accordance with the findings of human face perception. Psychologists have suggested that individual faces are represented in relative terms: human observers store a model of general face, against which all other faces are compared [13]. This is why we often use the words "long nose" or "small eyes" to describe a given face. These judgments are made based on the comparison to the general known faces, that act like references and give us the knowledge about how long a nose should be or how big the eyes should be. Particularly, in the proposed method, non-negative matrix factorization (NMF) is extended to learn a localized non-overlapping subspace representation of the facial patterns from a generic face image database. In the learned subspace, the degree of distinctiveness for any facial area is measured depends on the probability of this area is belong to a general face. Each local area of the general face is modeled using Gaussian Mixture Model (GMM) based on the generic face database.

The rest of the paper is organized as follows: In Section II a brief introduction is given on non-negative matrix factorization and its major extensions. Detail of the proposed method is described in Section III. Section IV presents the implementation of the proposed method and the experimental results. Conclusions are drawn in Section V.

II. Non-negative matrix factorization

Non-negative matrix factorization (NMF) [14] is a linear, non-negative approximate data representation. Given a non-negative data matrix $V = (v_{ij})_{m\times n}$, NMF finds the non-negative matrix $W = (w_{ij})_{m\times r}$, and the non-negative matrix $H = (h_{ij})_{r\times n}$, such that $V \approx WH$. The rank r of the factorization is generally chosen to satisfy $(n+m)r < mn$, so that the product WH can be regarded as a compressed form of the data in V. Let V represents a face database, each column of V contains n pixel values of one of the m face images in the database. Then, each face in V can be represented by a linear combination of r columns of W, the columns are called basis vectors (images). Each column of H is called a coefficient vector, that is in one-to-one correspondence with a face in V and describes how strongly each basis is present in the face. Since entries in W and H are all non-negative, only

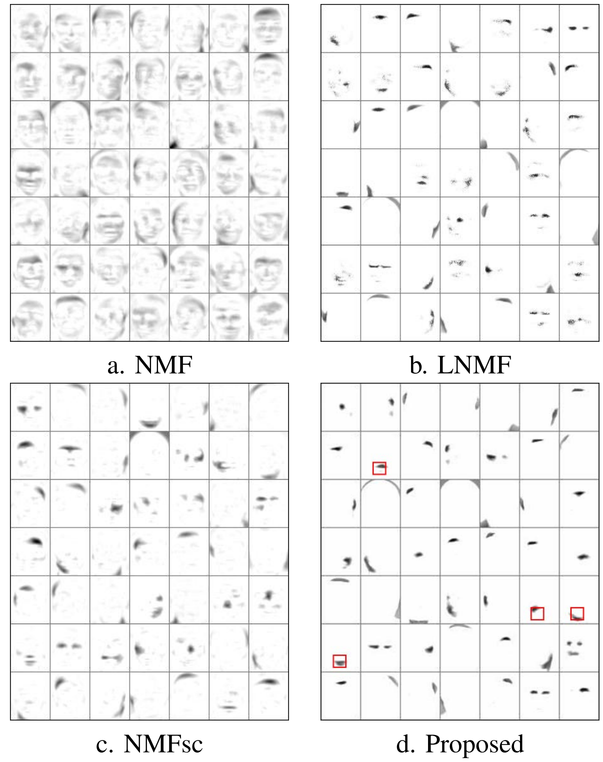

Fig. 1. Basis images learned from ORL database using different methods ($r = 49$), d also shows a group of extended NMF bases related to the "mouth" area (marked with red squares).

additive combinations of the basis vectors are allowed. Thus, NMF naturally leads to a part-based representation, the learned basis images tend to match intuitive facial features like mouth, nose and eyes.

NMF can be taken as an optimization problem, where W and H are chosen to minimize the reconstruction error between V and WH. Various error functions (objective functions) have been proposed, one widely used is the Euclidean distance function:

$$E(W,H) = \|V - WH\|^2 = \sum_{i,j}(V_{ij} - (WH_{ij}))^2 \quad (1)$$

Although the minimization problem is convex in W and H separately, it is not convex in both simultaneously. Paatero and Tapper [15] proposed a gradient decent method for the optimization, Lee and Seung [16] devised a multiplicative algorithm to reach a local optimum.

One of the issues of NMF is that it does not always give a part-based representation. As suggested by Li et al. [17], [18], when NMF is applied on ORL face database [19], in which faces are not well aligned, the learned basis images are holistic rather than local part-based (as can be seen in Figure 1a, the results are reproduced by us). To improve the performance of NMF in learning part-based representation, Li et al. proposed a local NMF method (LNMF) [17], [18], that adds three additional constraints on NMF: Maximum Sparsity in H, Maximum Expressiveness of W, Maximum

Orthogonality of W. Figure 1b shows the basis images learned from ORL database using LNMF. Comparing with NMF, we see that features gained by LNMF are more localized. However, some of the bases are still global and overlapped with each other. Furthermore, since more constraints are imposed, the convergence of LNMF is time consuming.

As an effect of part-based decomposition, NMF usually produces sparse representation. W is sparse since the learned bases tend to be non-global. H is often sparse due to that any given sample does not consist of all the available parts (bases). Hoyer [20] proposed a method called NMF with sparseness constraints (NMFsc), and suggested that by explicitly controlling the sparseness of W and H, NMF could give a more meaningful part-based representation. In NMFsc, the level of sparseness is measured based on the relationship between the L_1 norm and the L_2 norm:

$$sparseness(\mathbf{x}) = \frac{\sqrt{n} - (\sum |x_i|)/\sqrt{\sum x_i^2}}{\sqrt{n} - 1} \tag{2}$$

where n is the dimensionality of **x**. Then NMFsc is defined as the following optimization problem:

$$\begin{aligned} \min_{W,H} E(W,H) \quad s.t.\ \ & W, H \geq 0, \sum_i W_{ij} = 1 \ \ \forall j \\ & sparseness(w_j) = S_w, \forall j, \\ & sparseness(h_j) = S_h, \forall j \end{aligned}$$

Where w_j is the jth column of W and h_j is the jth row of H; S_w and S_h are the desired sparsenesses of W and H respectively. We show the basis images learned from ORL database using NMFsc in Figure 1c, where S_w is set to 0.75 and S_h is unconstrained as the best result achieved in [20]. As can be seen from the figure, NMFsc does not give a better part-based representation than LNMF. However, directly control the sparseness of the representation is very useful for many applications.

III. The Proposed Method

A. Local representation via extended NMF

One of the advantages for NMF based face representation is that when the basis images are localized and not overlapped with each other, only limited bases contribute to one local facial area, at the same time each basis related to limited facial area. Thus, one facial area can be represented by the corresponding basis coefficients as a low dimension feature vector. Further more, each row of the coefficient matrix H describes how strongly the corresponding basis is present in all the faces from the database. For localized non-overlapping bases, the properties of one specific facial area from all faces in the database can be represented by only few rows of H. This representation is ideal for us to learn the general model of each facial area based on the database.

In the proposed method, we extend the NMF for producing the desired localized, non-overlapping representation. Inspired by LNMF and NMFsc, our extended NMF impose orthogonality constraint on basis matrix W while controlling the sparseness of coefficient matrix H. To reduce the overlapping between basis images, different bases should be as orthogonal as possible so as to minimize the redundancy. Denote $U = W^T W$, the orthogonality constraint can be imposed by minimizing $\sum_{i,j,i\neq j} U_{i,j}$. As introduced in Section II, for learning localized bases, LNMF adds two more constraints to maximize the sparsity in H. Maximum sparsity in the coefficient matrix makes sure that a basis component cannot be further decomposed into more components, thus the overlapping between basis images is further reduced. However, a high sparseness in H forces each coefficient try to represent more of the image, and then the basis images tend to be global. Consider the extreme case when only one element in each column of H is allowed to be nonzero, then the NMF reduces to vector quantization (VQ), and all the basis images turn to holistic prototypical faces. Therefore, we chose to explicitly control the sparseness level of H, so that a compromise can be made between localization and overlapping and the value of the sparseness could be set based on different application scenarios.

The objective function of the extended NMF is defined as:

$$E(W,H) = \frac{1}{2} \sum_{i,j} (V_{ij} - (WH)_{ij})^2 + \beta \sum_{i,j,i\neq j} U_{i,j} \tag{3}$$

Where $U = W^T W$, β is a small positive constant. Then the extended NMF is defined as following optimization problem:

$$\begin{aligned} \min_{W,H} E(W,H) \quad s.t.\ \ & W, H \geq 0, \sum_i W_{ij} = 1 \ \ \forall j \\ & sparseness(h_j) = S_h, \forall j \end{aligned} \tag{4}$$

Where h_j is the jth row of H; S_h are the desired sparsenesses of H; the sparseness is measured based on formula (2). A local solution to the above minimization can be found by using the following two step update rules:

1)
$$W_{i\alpha} \leftarrow W_{i\alpha} \frac{(VH^T)_{i\alpha}}{(WHH^T)_{i\alpha} + \beta \sum_i W_{i\alpha}} \tag{5}$$

2)
$$H_{\alpha\mu} \leftarrow H_{\alpha\mu} - \mu_H [W^T(WH - V)]_{\alpha\mu} \tag{6}$$

Then project each row of H to be non-negative, have unit L_2 norm, and L_1 norm set to achieve desired sparseness S_h. (μ_H is a small positive constant. For the projection method, please refer to [20].)

Given a new sample face S (same size as the face images in the database V), its coefficient vector in the learned subspace W can be obtained by

$$L = W^{-1} S \tag{7}$$

Where W^{-1} is the pseudo inverse matrix of W. Since the extended NMF learns localized non-overlapping basis images, as discussed above, we are able to represent each local facial area by corresponding basis coefficients, that is only a small subset of L. Figure 1d shows an example of the bases used to represent a "mouth" area. The basis images here are learned from the ORL database using the proposed extended NMF, S_h is set to 0.1. As can be seen from the figure, more localized, less overlapped basis images are obtained, and limited bases contribute to each specific local facial area.

B. Measuring the distinctiveness

As introduced in Section I, we measure the distinctiveness of each facial area by its comparison to the general face. Although it is definitely present in the human visual system and the concept is clear in our mind, the appearance of a general face is not easy to be defined in computer vision. An alternative way is to train a general model for each of the facial area based on a generic database which consists of images from subjects other than those under consideration. The generic database can be considered as a collection of known faces as in the human visual system. By analyzing these "known faces", the knowledge of how each facial area generally looks like can be obtained.

In the proposed method, the extended NMF is applied on the generic database to learn a "general face" subspace and all the analyses are conducted in the subspace. Denote g_k as a local area k of a general face, l_k as the local area k sampled from a face image. Then, the probability of l_k belong to g_k can be computed with the Bayes rule

$$P(g_k|l_k) = \frac{p(l_k|g_k)P(g_k)}{p(l_k)} \tag{8}$$

Where $p(l_k|g_k)$ is the probability density function (pdf) of g_k, $P(g_k)$ is the priori probability and $p(l_k)$ is merely a scaling factor. We approximate the unknow pdf $p(l_k|g_k)$ by using Gaussian Mixture Model (GMM) [21]. The GMM is defined as

$$p(l_k|g_k;\Theta_k) = \sum_{c=1}^{C} \alpha_c N(l_k;\mu_c,\Sigma_c) \tag{9}$$

where $N(l_k;\mu_c,\Sigma_c)$ is the Gaussian pdf with mean value μ_c and covariance matrix Σ_c, α_c are positive weights of the component c and $\sum_{c=1}^{C}\alpha_c = 1$. Then, the Gaussian mixture probability density function can be completely defined by the parameter list

$$\Theta_k = \{\alpha_1,\mu_1,\Sigma_1,\ldots,\alpha_C,\mu_C,\Sigma_C\} \tag{10}$$

In the learned "general face" subspace, by employing the proposed representation described in Section III-A, l_k is a feature vector that consists of basis coefficients related to facial area k; the parameters of Gaussian mixture pdf for each g_k can be estimated based on corresponding rows in the coefficient matrix H. The distinctiveness of a given facial area l_k can be measured based on

$$d_k = \frac{1}{P(g_k|l_k)} \tag{11}$$

Fig. 2. The normalization process

IV. Experimental results

A. The implementation

The face region of all images is first detected by the Viola-Jones face detection method [22]. Then a modified version of the Viola-Jones face detection method is employed to find the areas of mouth and eyes within the detected face. For details on the method, readers are referred to our previous work [23]. Once the eyes and mouth have been localized, using the differences between the x and y coordinates, the original image is rotated so that the centers of eyes and mouth are at the same pixel coordinates in all images. Then the face area is cropped and resized to a final 64×64 face image. Figure 2 shows the whole normalization process.

The BioID face database [24] is used as the generic database in the implementation. The dataset consists of 1521 gray level face images from different subjects. All images are taken in uncontrolled conditions and show roughly frontal view of faces. Some normalized examples can be seen in Figure 4. We applied the extended NMF on BioID database with different parameters, and found that as the number of basis (rank r) increases, the obtained basis images become more localized. However too localized bases are meaningless and just reduced to pixel level. Considering the dimension of feature vector in the subspace, we choose $r = 100$ for the normalized 64×64 faces. With the number of bases fixed, best results are archived by setting S_h to 0.1. Figure 3 shows the bases used in the implementation ($r = 100, S_h = 0.1$). Example of original face images and their corresponding reconstructed images in the learned subspaces is showed in Figure 5.

The face image is then partitioned into local facial areas. The size, shape and overlaps of the facial areas can be determined by different application scenarios. For each of the facial area k, a Gaussian mixture model for general face is trained. The parameters of the Gaussian mixture pdfs are estimated with Expectation Maximization (EM) algorithm. For details on the method, please referred to [21], [25] .

B. Exaggerated samples

One direct way to evaluate the efficiency of the proposed method is to conduct a perceptual experiment that ask human observers to point out the most distinctive areas on face images and then compare the results with ours. However, such a experiment is too expensive to carry out. Here, alternatively, we apply some random transforms on one local area of face images so that the facial area is exaggerated and obviously different from regular ones. Then the proposed method is applied on the transformed images to see if the exaggerated

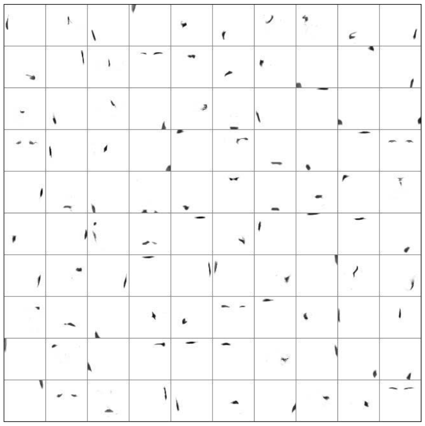

Fig. 3. The bases used in the implementation

Fig. 4. Examples of normalized faces from BioID database

facial area is detected. In the experiment, face images are divided by a 16×16 sliding-window, the sliding step is set to 4 pixels. Some typical examples are showed in Figure 6, the 10 most distinctive facial areas measured by the proposed method are marked in red squares. We can see that most detected distinctive areas converge on the exaggerated facial parts.

C. Applications in face recognition

In Section I, we have grouped the local appearance based face recognition approaches into two categories: one kind of the approaches extract features only on selected areas, and the other simply partition the face image into sub-regions, and directly extract local features on each of them. In this section, for each kind of local approaches, one representative work is modified by employing the proposed method for finding the discriminative facial areas.

Ahonen et al. [6] extract local binary pattern (LBP) histograms from equally partitioned 7×7 facial regions, and employ the nearest neighbor classifier for recognition. In the computed feature space, a weighted Chi square is used as the dissimilarity measure for the nearest neighbor classifier. Due to the stability and simplicity of the LBP based face representation, this method has attracted plenty of attention in face recognition domain. In the experiment, we apply the proposed method in Ahonen et al.'s work to measure the distinctiveness of each non-overlapped 7×7 facial region for the sample face. Then, in the classification stage, rather than using the predefined weights (reviewed in Section I), each of the facial region is weighted based on the obtained distinctiveness.

Fig. 5. Examples of original face image and the corresponding reconstructed image in the learned subspace

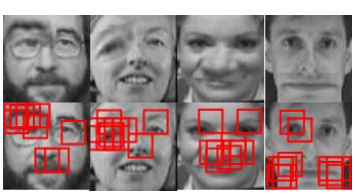

Fig. 6. Results for exaggerated faces

In Kepenekci et al.'s work [12] (referred as selected Gabor method in the rest of paper), 40 Gabor wavelets are convolved with the face, then interesting points with high-energized Gabor wavelet response are found on each of the 40 Gabor filtered images. Instead of using predefined facial points as in EBGM, Gabor features are extracted at theses interesting points. Since the number of the points and their locations vary for different face images, the correspondences of extracted Gabor jets between two facial images are unknown, thus in the recognition stage, only Gabor jets with similarity above a preset threshold are taken into consideration. The image similarity of two facial images is calculated as the mean of the similarities of the selected jets. Finally, the overall similarity of a test image and a sample face is a weighted sum of the image similarity and the number of similar jets. Still, in the experiment, the selected Gabor method is modified by using the proposed method to detect the interesting points. The face image is partitioned into 8×8 sub-regions with 50% overlaps, and the distinctiveness of each sub-region is measured. Then we select the 100 most distinctive regions, and extract gabor features at the center of each selected region.

Both the LBP and the selected Gabor methods reported good results based on the FERET database with only one training sample per subject. To show the efficiency of the proposed method, the two modified systems are also tested according to the standard FERET evaluation protocol [26] with the gallery set including 1196 frontal images of 1196 persons and four probe sets: **fafb** (1195 images with expression variations); **fafc** (194 images with illumination variations); **dup.I** (722 images taken in less than 18 months); **dup.II** (234 images taken about 18 months later). The testing results (in average recognition rate) as well as the performance of original methods are showed in Table I. For comparison, reported results of two SIFT feature based methods, SIFT_Grid [5] and SIFT_Person-specific [9] are also included in the Table. It can be seen that by employing the proposed method to select and emphasize the distinctive facial areas, the modified approaches achieved better recognition rate on **dup.I** and **dup.II** sets, in which images are taken in different time.

V. Conclusion

Locating the discriminative facial areas is crucial for local appearance based face recognition methods, no matter it

Methods	fafb	fafc	dup.I	dup.II
Selected Gabor [12]	0.96	0.70	0.58	0.47
Modified selected Gabor	0.97	0.67	0.64	0.59
LBP [6]	0.97	0.79	0.66	0.64
Modified LBP	0.97	0.77	0.68	0.67
SIFT_GRID [5]	0.94	0.35	0.53	0.36
SIFT_Person-specific [9]	0.97	0.47	0.61	0.53

TABLE I
THE RECOGNITION RATES OF DIFFERENT METHODS ON FERET PROBE DATA SETS

extracts features only on selected areas, or simply partitions the face image into regions. Most of the existing methods take the assumption that anatomical facial components, such as the eyes, nose, and mouth, are the most useful areas for recognition. Other more elaborate methods locate the most salient parts within the face according to a pre-specified criterion. In this paper, a different method is proposed to find the characteristic areas of each individual's face by its comparison to the general population. The selected regions are the most distinctive facial areas that make the given face different from other faces "we know". In particular, we extend NMF to learn a localized non-overlapping subspace representation of the facial patterns from a generic face image database. In the learned subspace, the degree of distinctiveness for any facial area is measured depends on the probability of this area is belong to a general face. Each local area of the general face is modeled using Gaussian Mixture Model (GMM) based on the generic face database. When apply the proposed method in existing face recognition systems, the experiment results show that it do improve the original approaches, especially for face images that are taken in different time. Although the method is proposed for face recognition, it also can be used in other related applications such as caricature generation and facial image retrieval.

REFERENCES

[1] W. Zhao, R. Chellappa, A. Rosenfeld, and P. Phillips, "Face recognition: A literature survey," *ACM Computing Surveys*, vol. 35, no. 4, pp. 399–458, 2003.

[2] Matthew Turk and Alex Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–76, 1991.

[3] J. Cohn, A. Zlochower, J. Lien, and T. Kanade, "Automated face analysis by feature point tracking has high concurrent validity with manual facs coding," *Psychophysiology*, vol. 36, pp. 35–43, 1999.

[4] L. Wiskott, J.M. Fellous, N. Kuiger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 775–779, 1997.

[5] M. Bicego, A. Lagorio, E. Grosso, and M. Tistarelli, "On the use of sift features for face authentication," *Proceedings of CVPR Workshop 2006*, vol. 17, no. 22, pp. 35–35, 2006.

[6] T. Ahonen, A. Hadid, and M. Pietikinen, "Face recognition with local binary patterns," in *Proceedinds of ECCV'04*, 2004, pp. 469–481.

[7] X. Tan, S.C. Chen, Z.-H. Zhou, and F. Zhang, "Recognizing partially occluded, expression variant faces from single training image per person with som and soft knn ensemble," *IEEE Transactions on Neural Networks*, vol. 16, no. 4, pp. 875–886, 2005.

[8] David G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[9] Jun Luo, Yong Ma, Erina Takikawa, Shihong Lao, Masato Kawade, and Bao-Liang Lu, "Person-specific sift features for face recognition," in *Proceedings of ICASSP'07*, 2007.

[10] Krystian Mikolajczyk and Cordelia Schmid, "Scale & affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 60, no. 1, pp. 63–86, 2004.

[11] C. Fernandez and M.A. Vicente, "Face recognition using multiple interest point detectors and sift descriptors," in *Proceeding of FG'08*, 2008, pp. 1–7.

[12] B. Kepenekci, F. Boray Tek, and G. Bozdagi Akar, "Occluded face recognition based on gabor wavelets," in *Proceedings of ICIP'02*, 2002, vol. 1, pp. 293–296.

[13] T. Valentine, "A unified account of the effects of distinctiveness, inversion, and race in face recognition," *Quarterly Journal of Experimental Psychology*, vol. 43, no. A, pp. 161–204, 1991.

[14] Daniel D. Lee and H. Sebastian Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, pp. 788–791, 1999.

[15] P. Paatero and U. Tapper, "Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values," *Environmetrics*, vol. 5, no. 2, pp. 1180–4009, 1994.

[16] Daniel D. Lee and H. Sebastian Seung, "Algorithms for non-negative matrix factorization," in *Proceedings of NIPS'2000*, 2000, pp. 556–562.

[17] Stan Z. Li, Xinwen Hou, Hongjiang Zhang, and QianSheng Cheng, "Learning spatially localized, parts-based representation," in *Proceedings of CVPR'01*, 2001, vol. 1, pp. 207–212.

[18] Tao Feng, Stan Z. Li, Heung-Yeung Shum, and Hongjiang Zhang, "Local non-negative matrix factorization as a visual representation," in *Proceedings of ICDL'02*, 2002, pp. 178–183.

[19] Ferdinando Samaria and Andy Harter, "Parameterisation of a stochastic model for human face identification," *Proceedings of 2nd IEEE Workshop on Applications of Computer Vision*, pp. 138–142, 1994.

[20] Patrik O. Hoyer, "Non-negative matrix factorization with sparseness constraints," *The Journal of Machine Learning Research*, vol. 5, no. 5, pp. 1457–1469, 2004.

[21] Pekka Paalanen, Joni-Kristian Kamarainen, Jarmo Ilonen, and Heikki Klviinen, "Feature representation and discrimination based on gaussian mixture model probability densitiespractices and algorithms," *Pattern Recognition*, vol. 39, no. 7, pp. 1346–1358, 2006.

[22] P. Viola and M. Jones, "Robust real-time object detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.

[23] Ce Zhan, Wanqing Li, Philip Ogunbona, and Farzad Safaei, "Real-time facial feature point extraction," *Advances in Multimedia Information Processing - PCM2007, LNCS 4810,*, pp. 88–97, 2007.

[24] R. Frischholz and U. Dieckmann, "Bioid: A multimodal biometric identification system," *Computer*, vol. 33, no. 2, pp. 64–68, 2000.

[25] J.J. Verbeek, N. Vlassis, and B. Krse, "Efficient greedy learning of gaussian mixture models," *Neural Computation*, vol. 15, no. 2, pp. 469–485, 2003.

[26] P.J. Phillips, H. Wechsler, J. Huang, and P. Rauss, "The feret database and evaluation procedure for face recognition algorithms," *Image and Vision Computing*, vol. 16, pp. 295–306, 1998.