2012

Measuring face familiarity and its application to 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

Publication Details

Zhan, C., Li, W. & Ogunbona, P. O. (2012). Measuring face familiarity and its application to face recognition. 2012 IEEE Workshop on the Applications of Computer Vision, WACV 2012 (pp. 177-183).

Measuring face familiarity and its application to face recognition

Abstract

The familiarity of faces is one of the key factors that come into play during human face analysis. However, there is very little research that studies face familiarity. In this paper, two methods are proposed to quantitatively measure the degree of familiarity of a face with respect to a known set. The methods are in accordance with the psychological study. In particular, non-negative matrix factorization (NMF) is extended to learn a localized non-overlapping subspace representation of commonly experienced facial patterns from known faces. The familiarity of a given face is then measured based on its reconstruction error after being projected into the learned extended NMF subspaces. A subjective study involving 50 subjects indicates the proposed familiarity measurement is in line with human judgments. Furthermore, the familiarity vector generated during the measuring process is employed for face recognition. Experiments based on the standard FERET evaluation protocol demonstrates the efficacy of the familiarity based representation for face recognition.

Keywords

measuring, application, face, recognition, familiarity, its

Disciplines

Engineering | Science and Technology Studies

Publication Details

Zhan, C., Li, W. & Ogunbona, P. O. (2012). Measuring face familiarity and its application to face recognition. 2012 IEEE Workshop on the Applications of Computer Vision, WACV 2012 (pp. 177-183).

Measuring Face Familiarity and Its Application to Face Recognition

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

Abstract

The familiarity of faces is one of the key factors that come into play during human face analysis. However, there is very little research that studies face familiarity. In this paper, two methods are proposed to quantitatively measure the degree of familiarity of a face with respect to a known set. The methods are in accordance with the psychological study. In particular, non-negative matrix factorization (NMF) is extended to learn a localized non-overlapping subspace representation of commonly experienced facial patterns from known faces. The familiarity of a given face is then measured based on its reconstruction error after being projected into the learned extended NMF subspaces. A subjective study involving 50 subjects indicates the proposed familiarity measurement is in line with human judgments. Furthermore, the familiarity vector generated during the measuring process is employed for face recognition. Experiments based on the standard FERET evaluation protocol demonstrates the efficacy of the familiarity based representation for face recognition.

1. Introduction

Automatic face analysis has been an active research topic over the last 20 years. To simulate the amazing ability of human face perception, computer vision and pattern recognition researchers have made great efforts to extract different kinds of information from facial images (videos), including identity, gender, pose and expression. However, there is very little research that studies face familiarity. In everyday life, we naturally categorize the faces we encounter as either familiar or unfamiliar. Psychological studies have found that familiarity is one of the key factors that affect human face processing, especially for face recognition. For example, although people are excellent at recognizing faces familiar to them, our ability to recognize unfamiliar face is rather poor [12].

The concept of face familiarity is not clearly defined in the literature. In most of the psychology studies, face familiarity is measured qualitatively based on previous exposures: Faces belonging to people we have seen a great deal and for long exposure durations are defined as familiar faces [12]; while unfamiliar faces refer to faces with which we have had none or only limited previous encounters [12]. With this definition, the degree of familiarity for "familiar" faces cannot be precisely measured. As a result, most experiments in the laboratory are conducted with "unfamiliar" faces. More importantly, the feeling of familiarity is not necessarily related to previous exposures. We often encounter a situation in which a particular face looks familiar, yet the person is totally unknown to us and only bear a semblance to several acquaintances. Psychologists refer to such a situation as the prototype effect [14]. Further studies based on experiments of old/new discrimination and familiarity ranking tasks found a higher level of prototype effect, where prototype faces combined based on previously viewed faces are regarded not only as familiar, but as more familiar than those that have been seen before. In such experiments, subjects are first asked to study some face images and remember them. Then in the test phase, the studied face images, unstudied face images and morphs between pairs of (or among multiple) studied images are added to the testing pool. Subjects are asked to conduct old/new judgments or familiarity ranking for all the images in the testing pool. The results show that many unstudied morphs are ranked more familiar or judged as "old" more often than their studied parents images [15].

The general concept of familiarity described by Mandler [8] may somehow account for the prototype effect. Mandler suggested two forms of familiarity: context-free familiarity and context-dependent familiarity. The first one is a sense of knowing we have encountered the target before; no specific source in the memory contributes to this feeling of knowing. The second one is a feeling of remembering, which is engendered when the target is matched to a previously encountered item in the memory. The prototype faces may only cause high feeling of context-free familiarity, however as Mandler suggested, we are not able to distinguish between these two forms of familiarity.

Wallis et al. [15] explain the prototype effect from representation point of view. They suggest that our internal representation of faces is based on combinations of reusable features that are abstracted from commonly experienced local facial patterns. Thus the local features of prototype faces would all have high possibility to match those in the memory, and therefore cause the prototype effect.

In this paper, we propose two methods to quantitatively measure the degree of familiarity of faces in accordance with Mandler's context-free and context-dependent cases. In particular, non-negative matrix factorization (NMF) is extended to learn a localized non-overlapping subspace representation of the "commonly experienced" facial patterns from the known faces. The familiarity of a given face is then measured based on its reconstruction error after being projected into the learned extended NMF subspaces. In the context-free based method, all the known faces are used to learn one subspace, and no specific known data directly contributes to the obtained familiarity. As for the context-dependent based method, one extended NMF subspace is learned for each of the known subjects from the person's own data, the familiarity is then measured based on specific known subjects. Furthermore, the familiarity vector generated during the process of context-dependent familiarity measurement is employed in face recognition.

The rest of the paper is organized as follows: In Section 2 a brief introduction is given on non-negative matrix factorization and its major extensions. Details of the proposed method are described in Section 3. In Section 4 the proposed method is evaluated based on simulation of psychological experiments. Section 5 presents the application of the proposed method to face recognition. Conclusions are drawn in Section 6.

2. Non-negative matrix factorization

Non-negative matrix factorization (NMF) [4] is a linear, non-negative approximate data representation. Given a non-negative data matrix $V = (v_{ij})_{m\times n}$, NMF finds non-negative matrices $W = (w_{ij})_{m\times r}$, and $H = (h_{ij})_{r\times n}$, 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 m pixel values of one of the n 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 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.

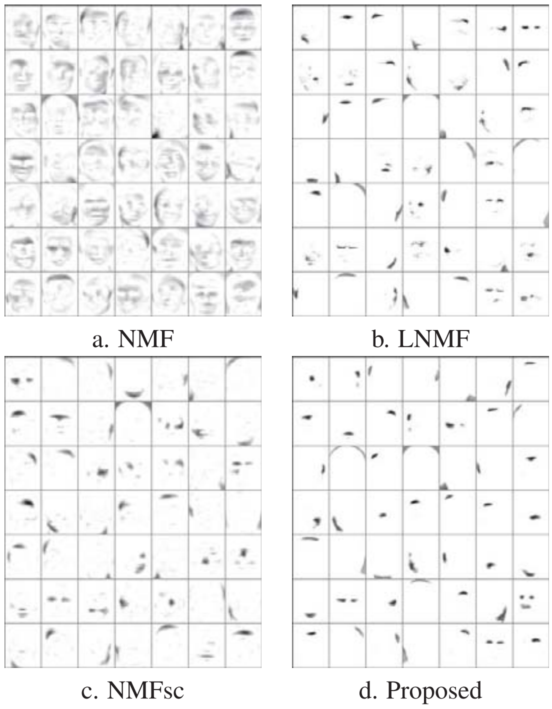

Figure 1. Basis images learned from ORL database using different methods ($r = 49$).

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 [10] proposed a gradient decent method for the optimization, Lee and Seung [5] devised a multiplicative algorithm to search a local optimum.

One of the issues with NMF is that it does not always give a part-based representation. As suggested by Li et al. [7], when NMF is applied on ORL face database [13], 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; we have reproduced the results). To improve the performance of NMF in learning part-based representation, Li et al. proposed a local NMF method (LNMF) [7], 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. Compared with NMF, we see that features gained by LNMF are more localized. However, some of the bases are still global. This is mainly due to the introduction of maximum sparsity constraint on coefficient matrix. Maximum sparsity in H makes sure that a basis component cannot be further decomposed into more components, thus the overlapping between basis images is reduced. However, a high sparseness in H forces each coefficient 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 become holistic prototypical faces.

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 because any given sample does not contain of all the available parts (bases). Hoyer [2] 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. We show the basis images learned from ORL database using NMFsc in Figure 1c, where the sparseness of W is set to 0.75 and the sparseness of H is unconstrained as the best result achieved in [2]. As can be seen from the figure, by only directly controlling the sparseness of the representation, NMFsc does not give a better part-based representation than LNMF.

3. The proposed method

3.1. Extended NMF

In the proposed method, we extend the NMF for producing localized, non-overlapping representation to mimic the abstract features suggested by Wallis et al.. Our extended NMF (ENMF) 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 $U = W^TW$, the orthogonality constraint can be imposed by minimizing $\sum_{i,j,i\neq j} U_{i,j}$. As introduced in Section 2, high sparsity in the coefficient matrix makes sure that a basis cannot be further decomposed into more components, while at the same time, leads basis images tend to be global. 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 ENMF is defined as:

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

where $U = W^TW$, α is a small positive constant. Then the ENMF is defined as following optimization problem:

$$\min_{W,H} E(W,H) \quad s.t.\ \ W,H \geq 0, \sum_i W_{ij} = 1 \ \ \forall j \quad (3)$$
$$sparseness(h_j) = S_h, \forall j$$

where h_j is the jth row of H; S_h are the desired sparsenesses of H; the sparseness is measured based on the relationship between the L_1 norm and the L_2 norm:

$$\text{sparseness}(\mathbf{x}) = \frac{\sqrt{n} - (\sum |x_i|)/\sqrt{\sum x_i^2}}{\sqrt{n}-1} \quad (4)$$

where n is the dimensionality of $\mathbf{x}$.

A local solution to the above minimization can be found by using the following two step update rules:

1. $$W_{ia} \leftarrow W_{ia}\frac{(VH^T)_{ia}}{(WHH^T)_{ia} + \alpha\sum_i W_{ia}} \quad (5)$$

2. $$H_{a\mu} \leftarrow H_{a\mu} - \eta_{a\mu}[W^T(WH-V)]_{a\mu} \quad (6)$$

   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. (For the projection method, please refer to [2].) $\eta_{a\mu}$ is the step size, and allowed to change at every iteration. we initially set $\eta_{a\mu}$ to 1, then multiply it by one-half at each subsequent iteration.

Figure 1d shows an example of the bases learned from ORL database using the proposed ENMF, S_h is set to 0.1 and α is set to 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.

3.2. Familiarity measure

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

$$L = W^{-1}S \quad (7)$$

where W^{-1} is the pseudo inverse matrix of W. Based on the obtained coefficient vector L, the sample S can be reconstructed by:

$$S' = WL \quad (8)$$

The reconstruction error between S and S' reflects the differences between S and the training set V. When V represents the faces we know, the reconstruction error indicates how similar is the sample face to the known faces, thus can be used as a measure of familiarity.

In the proposed methods, the reconstruction error between S and S' is calculated by mean square error (MSE):

$$MSE(S, S') = \frac{1}{n} E(S, S') = \frac{1}{n} \sum_{i,j} (S_{ij} - S'_{ij})^2 \quad (9)$$

where n is the number of pixel in the face image. Considering that the familiarity is actually a monotonic decreasing function of the reconstruction error, for consistency, as well as to mimic the human perception, we use the peak signal-to-noise ratio (PSNR) to measure the face familiarity. The PSNR is commonly used as an approximation to human perception of reconstruction quality and is calculated based on MSE:

$$PSNR(S, S') = 10log_{10}(\frac{MAX_I^2}{MSE(S, S')}) \quad (10)$$

where MAX_I is the maximum possible pixel value of the image.

3.2.1 Context-free based method

In the context-free based method, all the known faces are used to learn one ENMF subspace. A given face image is first projected into the learned subspace and reconstructed based on the projection coefficients. Then PSNR of the reconstruction is calculated as the measure of familiarity for the face. In this case, no specific known data in V directly contributes to the obtained familiarity, which is in line with the context-free familiarity suggested by Mandler.

3.2.2 Context-dependent based method

In the context-dependent based method, one personal ENMF subspace is learned for each of the known subjects from the person's own data. Then the familiarity of a given face S is measured according to following steps:

1. project S into each of the learned personal subspace
2. reconstruct S in each of the personal subspace
3. calculate PSNR for each of the reconstruction
4. sort all of the PSNR value in descending order
5. calculate the mean of the first k PSNR value as the measure of familiarity for S

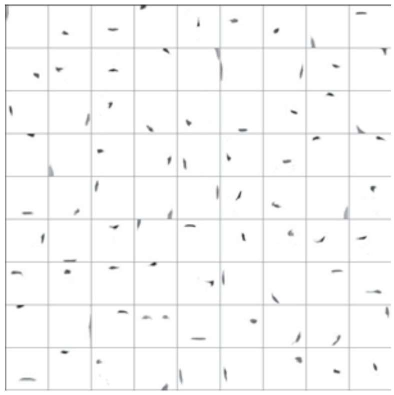

Figure 2. An example of the ENMF bases used in the experiments

Following the definition of context-dependent familiarity, the method measures the degree of familiarity based on specific known subject. According to the definition, k should always be one, as we match the target to the most similar source in the memory. However in this method, we leave k as a free parameter for further analysis.

4. Experimental results

4.1. Experimental setup

The face region of all images is first detected by the Viola-Jones face detection method. Then a modified version of the Viola-Jones face detection method [16] is employed to find the areas of mouth and eyes within the detected face. 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.

The ORL face database [13] is employed as the known set in the experiments. The dataset consists of 400 gray level face images from 40 subjects (10 images per subject). We applied the ENMF on ORL database with different parameters, while fixing the value of S_h, it is 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. Thus for the normalized 64×64 faces, we choose $r = 81$. With the number of bases fixed, best results are achieved by setting S_h to 0.1. The setting of r and S_h is then used for all the experiments in this paper. Figure 2 shows an example of the ENMF bases used in the experiments ($r = 81, S_h = 0.1$).

The BioId [1] database is employed as the unknown set in the experiments. The dataset consists of 1521 gray level face images from 23 subjects. All images are acquired in uncontrolled conditions and show roughly frontal view of faces.

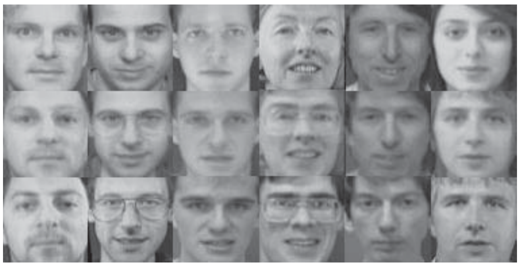

Figure 3. Examples of morph (the second row) and their parent faces

4.2. Simulation

In the experiment, we simulate the psychological experiment of familiarity ranking task. In particular, 25 pairs of face images are randomly selected from ORL database as known parent faces. For each pair of parent faces, a morph is generated, one parent face contribute 50% to the morph. All the morph images are generated nonlinearly using morphing software FantaMorph4 [3] based on manually labelled facial landmarks. Some morphs and their parent faces are shown in Figure 3. Then the familiarity degree of the morphs, the parent faces, and 50 unknown faces randomly selected from BioID database are measured by the proposed methods. To simulate different levels of previous exposures, ENMF subspaces are learned from different subsets of the ORL database. Each of the subsets contains z face images per subject. For the context-free based method, five ENMF subspaces are learned, with $z = 10, 8, 5, 2, 1$. For the context-dependent based method, four ENMF subspaces are learned for each subject, with $z = 10, 8, 5, 2$.

4.3. Subjective experiments

A subjective study of the familiarity ranking task is also conducted using exactly the same set of training and testing data as the simulation experiments. In particular, during the training phase, 50 participants are asked to study and remember face images from ORL database. The 50 participants are equally divided into 5 groups. For each group, z face images per subject from the ORL database are displayed sequentially to the participants and each image is displayed twice. A snapshot of the training interface ($z = 10$) is shown in Figure 4. In the testing phase, participants are asked to rate the familiarity of given test face images. Five points scale are used to measure the familiarity subjectively: very familiar (5), familiar (4), neutral (3), unfamiliar (2), totally unknown (1). A snapshot of the testing interface in subjective study is shown in Figure 5.

4.4. Results

Figure 6 shows the average familiarity of morphs, the known parent faces, and unknown faces measured by the context-free based method, and the results of the subjective

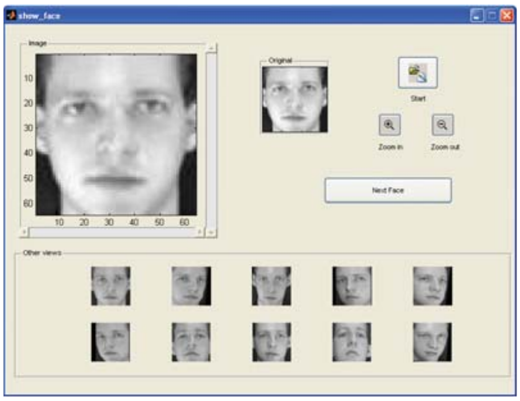

Figure 4. A snapshot of the training interface in subjective study

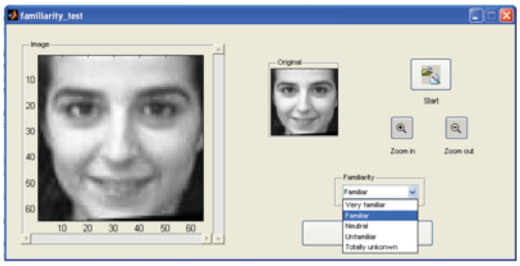

Figure 5. A snapshot of the testing interface in subjective study

ranking is shown in Figure 7. To compare the results, the objective measurements in PSNR and the subjective ranking are normalized to the range of 0 to 1. As can be seen from the figures, the resulting face familiarity measure of the proposed method has an overall consistent trend with the subjective ranking. Both results show a clear high level prototype effect that the unknown morphs built based on known faces are measured to be more familiar than their parent faces. As the number of face images per subject in the known set (the value of z) increases, higher degree of familiarity are obtained for all the testing faces. However, when the value of z exceeds 5, the increase of familiarity is limited especially for the unknown faces. This result is in line with our experience that the more we have seen someone, the more familiar he will become to us. As for a person we have never seen before, with more exposures of faces from other people, the probability of mismatching this person to the known ones would be increased. However this increase should be limited just as reflected in the experimental results.

The results of context-dependent based method is shown in Figure 8. Since similar effect for varying the value of z is observed as the above experiment, only the results for $z = 10$ is presented here so we can focus on the influence of k. We can see from the figure that unknown faces always obtain low level of familiarity for all the k. As expected, when $k = 1$, known faces are measured to be more famil-

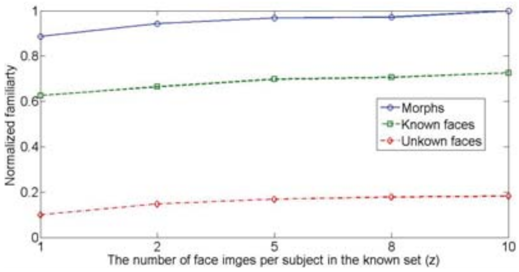

Figure 6. Results of context-free based method

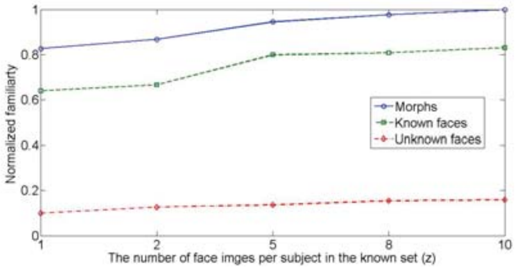

Figure 7. Results of human judgments

iar than the morphs. In this case the proposed method exactly model the context-dependent familiarity proposed by Mandler, a given face is matched to the most similar known subject. A known face would always find a good match (reconstruction), while the morph could only obtain a partial match. Thus no prototype effect is shown for $k = 1$. This case is corresponding to the situation when an accurate recognition of the known face occurs, where we are able to specifically recall when the particular encounter was. When $k = 2$, a morph could be partially matched to both its parent subjects, and that a known face still could find only one good match. Therefore, as can be seen from the figure at $k = 2$, the average familiarity for known faces and morphs are similar, the prototype effect begins to appear. As the value of k further increases, more known subjects contribute to the matching, until $k = 40$, the familiarity is measured based on all the known faces just like the context-free based method. So we can see for greater values of k, similar results are obtained as in the context-free based method, a higher level of prototype effect is shown.

5. Application to face recognition

Psychology studies [6] have suggested that context-free familiarity may introduce false alarms in face recognition, while context-dependent familiarity (referred to as specific familiarity in the literature) would lead to correct recognition of a target face. Following this suggestion, in this section, we apply the proposed context-dependent familiarity measure to face recognition.

As introduced in Section 3.2.2, to measure the context-

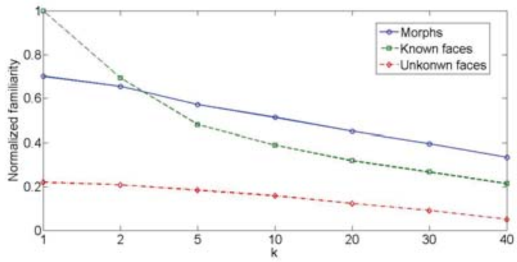

Figure 8. Results of context-dependent based method when $z = 10$

dependent familiarity, a given face is compared with all the known faces by calculating its reconstruction PSNR after being projected into each of the learned personal subspaces. Thus, each given face is associated with a familiarity vector that consists of the PSNR values. Formally, the familiarity vector of a given face S can be defined as

$$F_S = [PSNR(S, S'_1), PSNR(S, S'_2), \ldots, PSNR(S, S'_q)] \quad (11)$$

where q is the total number of known subjects, $S'_p(p = 1, 2, \ldots, q)$ is the reconstructed image in the personal subspace of the pth known subject. The familiarity vector reflects how similar is the given sample to each of the known person, it actually encodes the identity information of the sample indirectly with respect to the known set. To illustrate the discriminative nature of the familiarity vector for face recognition, familiarity vectors of three face images are plotted as curves in Figure 9. As can be seen in the figure, face1 and face2 are images from the same person. Although the face images are with different facial expressions, their familiarity vectors are very similar. As for the image from a different person, the familiarity vector of face3 is clearly different from the other two.

The discriminative power of the familiarity vector for face recognition is then tested according to the standard FERET evaluation protocol [11] 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). In the test, the familiarity vector is directly employed as feature vector to represent each face image. Nearest neighbor (NN) classifier is used for classification since there is only one training sample per subject, and Euclidean distance is employed as distance measure. Besides data from the ORL database, face images of 60 subjects from the AR database [9] are added to the known set. When varying the total number of known subjects used during the test, it is found that increasing the number of known subjects (the dimension of the familiarity vector) improves the recognition rate. For all the four test subsets, the best results are obtained when all the 100

Methods	fafb	fafc	dup.I	dup.II
PCA	0.85	0.65	0.44	0.22
LDA	0.94	0.73	0.55	0.31
LBP	0.97	0.79	0.66	0.64
EBGM	0.90	0.42	0.46	0.24
Familiarity vector (NMF)	0.87	0.71	0.61	0.55
Familiarity vector (LNMF)	0.89	0.73	0.62	0.58
Familiarity vector (NMFsc)	0.91	0.69	0.61	0.53
Familiarity vector (ENMF)	0.95	0.79	0.70	0.67

Table 1. The recognition rates of different methods based on the standard FERET evaluation protocol

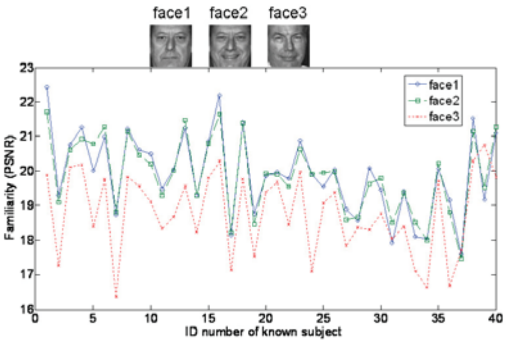

Figure 9. The familiarity vectors of three face images

available known subjects are used. We list the best results together with reported recognition rates of some popular face recognition methods in Table 1. To illustrate the efficiency of proposed ENMF representation, the familiarity vector is also generated based on traditional NMF, LNMF and NMFsc, the recognition results are included in Table 1 as well. It can be seen that the proposed method achieves competitive recognition rates on **fafb** and **fafc**, and outperforms all other methods on the **dup.I** and **dup.II** sets in which images are taken in different time. We believe by including more representative data into the known set, the discriminative ability of familiarity vector for face recognition would be further improved.

6. Conclusion

Based on the reconstruction error of a face image after being projected into learned ENMF subspaces, two methods are proposed in this paper to quantitatively measure the degree of familiarity of a face image with respect to a known set. Experiments on benchmark face database show that the proposed methods could effectively separate unknown faces from the known ones, and the results are also in line with subjective familiarity ranking. Furthermore, the familiarity vector generated during the measuring process is employed for face recognition. Preliminary results based on the standard FERET evaluation protocol demonstrates the efficacy of the familiarity based representation for face recognition.

References

[1] R. Frischholz and U. Dieckmann. Bioid: A multimodal biometric identification system. *Computer*, 33(2):64–68, 2000.

[2] P. O. Hoyer. Non-negative matrix factorization with sparseness constraints. *The Journal of Machine Learning Research*, 5(5):1457–1469, 2004.

[3] http://www.fantamorph.com/.

[4] D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401:788–791, 1999.

[5] D. D. Lee and H. S. Seung. Algorithms for non-negative matrix factorization. In *Proceedings of NIPS'2000*, pages 556–562, 2000.

[6] M. B. Lewis. Familiarity, target set and false positives in face recognition. *European Journal of Cognitive Psychology*, 9(4):437–459, 1997.

[7] S. Z. Li, X. Hou, H. Zhang, and Q. Cheng. Learning spatially localized, parts-based representation. In *Proceedings of CVPR'01*, volume 1, pages 207–212, 2001.

[8] G. Mandler. Recognizing: the judgement of previous occurrence. *Psychological Review*, 87:252–271, 1980.

[9] A. Martinez and R. Benavente. The AR face database. Technical Report 24, CVC, 1998.

[10] P. Paatero and U. Tapper. Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 5(2):1180–4009, 1994.

[11] P. Phillips, H. Wechsler, J. Huang, and P. Rauss. The FERET database and evaluation procedure for face recognition algorithms. *Image and Vision Computing*, 16:295–306, 1998.

[12] S. S. Rakover and B. Cahlon. *Face recognition: cognitive and computational processes*. John Benjamins Publishing Company, illustrated edition, 2001.

[13] F. Samaria and A. Harter. Parameterisation of a stochastic model for human face identification. *Proceedings of 2nd IEEE Workshop on Applications of Computer Vision*, pages 138–142, 1994.

[14] R. Solso and J. McCarthy. Prototype formation of faces: A case of pseudo-memory. *British Journal of Psychology*, 72:499–503, 1981.

[15] G. Wallis, U. E. Siebeck, K. Swann, V. B. Siegen, and H. H. Blthoff. The prototype effect revisited: Evidence for an abstract feature model of face recognition. *Journal of Vision*, 8:1–15, 2008.

[16] C. Zhan, W. Li, P. Ogunbona, and F. Safaei. Real-time facial feature point extraction. *Advances in Multimedia Information Processing - PCM2007, LNCS 4810,*, pages 88–97, 2007.