Feature-Aging for Age-Invariant Face Recognition

Huiling Zhou, Kwok-Wai Wong, and Kin-Man Lam,
Centre for Signal Processing,
Department of Electronic and Information Engineering,
The Hong Kong Polytechnic University, Hong Kong
E-mail: hl.zhou@connect.polyu.hk, enkwwong@polyu.edu.hk, enkmlam@polyu.edu.hk

***Abstract*— Age-invariant face recognition has attracted some recent attention. In real applications, the age progression of those face images, stored in a face database for recognition and identification purposes, should also be considered, so as to achieve a higher accuracy level. In this paper, we propose a method to predict the aging of facial features so as to alleviate the effect of age progression on face recognition. The original facial feature and the aged facial feature of a face image should be correlated, so they are fused by using canonical correlation analysis to form a coherent feature for face recognition. The performance of our proposed approach is evaluated based on the FGNet database, and compared to some existing face recognition algorithms. Experiment results show that our proposed method can achieve a superior performance, when the query and probe face images have a large age difference.**

I. INTRODUCTION

A lot of research on human face recognition has been conducted in the past three decades [1][2]. Various face-recognition algorithms which can deal with faces under different facial expressions, lighting conditions, and poses have been proposed, and can achieve satisfactory performances. However, the changes in face appearance caused by age progression have received limited attention to date; this effect has a significant impact on the face-recognition algorithms.

There are two different approaches for age-invariant face recognition. The first is the generative approach [3][4]. In this approach, face images of other ages will be generated before face recognition is performed. For this approach, the age of a face image needs to be estimated, and sufficient training samples are necessary for learning the relationship between a face at two different ages. The second approach is based on discriminative models [5][6], which use facial features that are insensitive to age progression to achieve age-invariant face recognition. However, this approach cannot achieve a significant performance level if the two faces to be compared have a large age difference.

In this paper, we consider both the forward and backward prediction of aged features for face recognition. With a live query input, the age difference between the query face image and a gallery face in a face database can be computed, if the time, when the gallery face was captured, is known. Otherwise, age estimation [7] is needed to find out the age difference between the query and the gallery face images. For face recognition, different features can be used, such as Gabor wavelets [8], local binary pattern (LBP) [9], Locality Preserving Projections (LPP) [10], etc. In general, the age of the query input should be older than that of the gallery face. In our algorithm, we extract the Gabor features at a number of landmarks in the face images. Based on the estimated age difference between the two faces, backward prediction is performed for the Gabor features of the query face image, so that the Gabor features at a younger age are generated. Similarly, the Gabor features of the gallery face image are forward predicted to generate the corresponding features at an older age. In other words, to compare the query and gallery faces, we use both the features at the two different ages. As the facial features of a person at two different ages should be correlated with each other, canonical correlation analysis [11] (CCA) is employed to the two sets of features, at different ages, to generate more coherent features for face recognition.

The rest of this paper is organized as follows. In Section II, we will describe the Gabor features and CCA, and then the details of our proposed algorithm. We will present the learning of the forward and backward prediction, and the use of CCA to generate coherent features. In Section III, experimental setup and results will be described and discussed. Finally, conclusion and future work will be presented in Section IV.

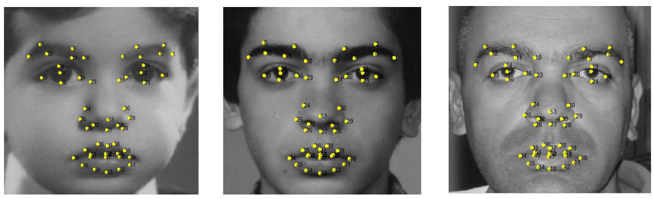

Fig. 1 Faces used for experiments are cropped and aligned with two eyes, with 53 facial feature points denoted as yellow dots.

II. DETAILS OF OUR APPROACH

In this section, we will first present the features used in our algorithm. Then, based on the corresponding features from two sets of training face images having a similar age, linear regression is employed to learn the forward prediction and the backward prediction of the facial features from a younger age to an older age and from an older age to a younger age, respectively. For the face images used in our experiments, 68 feature points have been labeled. As those feature points located on the face contour are not reliable for recognition, only 53 feature points are used in our algorithm, as shown in Fig. 1. The forward and backward prediction for each of these feature points will be learned. Having predicted the facial features of a face image at another age, CCA is then introduced to combine the features, at two different ages, to produce a coherent feature for face recognition.

A. Gabor wavelets

Gabor wavelets (GW) have been commonly used as local features for many pattern recognition and computer vision applications, such as texture retrieval, object detection, recognition, etc. It was found in [12, 13] that the Gabor functions are similar to the response of the two-dimensional receptive field profiles of the mammalian simple cortical cell. The Gabor features exhibit the desirable characteristics of capturing salient visual properties such as spatial localization, orientation selectivity, and spatial frequency selectivity. In the spatial domain, a GW is a complex exponential modulated by a Gaussian function, which is defined as follows [14]:

$$\Psi_{\omega,\theta}(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(xcos\theta+ysin\theta)^2+(-xsin\theta+ycos\theta)^2}{2\sigma^2}} \cdot \left[e^{i(\omega xcos\theta+\omega ysin\theta)} - e^{-\frac{\omega\sigma^2}{2}}\right], \quad (1)$$

where (x, y) represent the pixel position in the spatial domain, ω is the radial center frequency of the complex exponential, θ is the orientation of the GW, and s is the standard deviation of the Gaussian function. By using different center frequencies and orientations, a family of Gabor kernels can be produced, which can then be used to extract features from an image. In [15, 16], GWs have been employed for face recognition, and a promising performance can be achieved.

In the Gabor feature, only the Gabor magnitudes are used because the Gabor phases change linearly with small displacements. Gabor filters at five different scales and eight different orientations are adopted in our algorithm. Consequently, at each of the 53 facial feature points, its Gabor feature is formed by concatenating the outputs of the 40 (5×8) filters.

B. Forward and Backward Feature-Aging

In our algorithm, we predict the Gabor features of a face image from one age to another age. In our training set, we assign training face images into different groups according to their ages. Consider two different age groups, which are denoted as AG1 and AG2, respectively. Here, we assume that images in AG1 are younger than those in AG2. The training samples are denoted as $\mathbf{U}^1 = [\boldsymbol{u}_1^1, \ \ldots, \ \boldsymbol{u}_M^1]$ and $\mathbf{U}^2 = [\boldsymbol{u}_1^2, \ \ldots, \ \boldsymbol{u}_M^2]$ for AG1 and AG2, respectively, where M is the number of training pairs and $\boldsymbol{u}_j^i$ represents the Gabor features of the j^{th} training sample of the i^{th} age group at a particular landmark position in the face image. Linear regression is employed in the prediction. If the Gabor features of AG2 are predicted from AG1, "forward prediction" is performed and the Gabor features are under forward aging. Similarly, "backward prediction" is performed when we generate features from AG2 to age AG1.

For forward prediction, we assume that there is a linear mapping from the features $\mathbf{U}^1$ to $\mathbf{U}^2$, as follows:

$$f: \mathbf{U}^1 \rightarrow \mathbf{U}^2. \quad (2)$$

The mapping function should be complicated and nonlinear. However, to simplify the learning of the mapping function, linear mapping is adopted to obtain the approximated aged features. Therefore, the mapping can be written as follows:

$$\mathbf{U}^2 = \mathbf{A}_f \, \mathbf{U}^1, \quad (3)$$

where $\mathbf{A}_f$ is the forward mapping matrix, which can be computed by solving (3) as $\mathbf{U}^1$ and $\mathbf{U}^2$ are known. The dimension of $\mathbf{U}^1$ and $\mathbf{U}^2$ are 40×M, while the dimension of $\mathbf{A}_f$ is 40×40. $\mathbf{A}_f$ is computed as follows:

$$\mathbf{A}_f = \mathbf{U}^2(\mathbf{U}^1)^+, \quad (4)$$

where $(\mathbf{U}^1)^+ = (\mathbf{U}^1)^T(\mathbf{U}^1(\mathbf{U}^1)^T)^{-1}$ is the pseudo inverse of $\mathbf{U}^1$ and T represents the transpose operation. Having learned the linear mapping function $\mathbf{A}_f$ based on the training samples, a given Gabor feature can be aged from AG1 to AG2, as follows:

$$\boldsymbol{u}_j^2 = \mathbf{A}_f \, \boldsymbol{u}_j^1. \quad (5)$$

Similarly, when Garbor features at AG2 are available, we can predict the corresponding features at AG1 by learning the backward mapping function $\mathbf{A}_b$, as follows:

$$\mathbf{U}^1 = \mathbf{A}_b \, \mathbf{U}^2, \quad (6)$$

where $\mathbf{A}_b = \mathbf{U}^1(\mathbf{U}^2)^T(\mathbf{U}^2(\mathbf{U}^2)^T)^{-1}$.

C. Canonical Correlation Analysis

In our algorithm, features of a face subject at two different ages are available. As the features are extracted from the same person, it is expected that they are correlated. Canonical Correlation Analysis (CCA) is used to project the features at two different ages into a coherent subspace, where the correlation between the two projected features are maximized. The projected features are then combined to form a single coherent feature vector for age-invariant face recognition.

With the pairs of features, $\mathbf{U}^1$ and $\mathbf{U}^2$, for a particular position and at two ages, we use CCA to learn pairs of directions $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ that maximize the correlation between the projected features $\boldsymbol{v}_j^1 = \boldsymbol{\alpha}^T\boldsymbol{u}_j^1$ and $\boldsymbol{v}_j^2 = \boldsymbol{\beta}^T\boldsymbol{u}_j^2$. The projection matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ can be derived by maximizing the following criterion function:

$$J(\alpha, \beta) = \frac{\alpha^T \mathbf{C}_{12}\beta}{\sqrt{\alpha^T \mathbf{C}_{11}\alpha \cdot \beta^T \mathbf{C}_{22}\beta}}, \quad (7)$$

where $\mathbf{C}_{11} \in \mathrm{R}^{40\times40}$ and $\mathbf{C}_{22} \in \mathrm{R}^{40\times40}$ denote the covariance matrices of $\boldsymbol{u}_j^1$ and $\boldsymbol{u}_j^2$ (j =1, …, M), respectively, while $\mathbf{C}_{12} \in \mathrm{R}^{40\times40}$ is the covariance matrix of $\boldsymbol{u}_j^1$ and $\boldsymbol{u}_j^2$.

The original CCA does not include any class information when learning the projections $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. To improve the recognition performance, generalized CCA [17], which is an extension of CCA, is employed. The class information is incorporated in the covariance matrices, and they are defined as follows:

$$\mathbf{C}'_{11} = \sum_{i=1}^{c} P(\omega_i) \sum_{j=1}^{n_i} (\boldsymbol{u}_{ij}^1 - \overline{\boldsymbol{u}}_i^1)(\boldsymbol{u}_{ij}^1 - \overline{\boldsymbol{u}}_i^1)^T, \tag{8}$$

$$\mathbf{C}'_{22} = \sum_{i=1}^{c} P(\omega_i) \sum_{j=1}^{n_i} (\boldsymbol{u}_{ij}^2 - \overline{\boldsymbol{u}}_i^2)(\boldsymbol{u}_{ij}^2 - \overline{\boldsymbol{u}}_i^2)^T, \tag{9}$$

where $\mathbf{C}'_{11}$ and $\mathbf{C}'_{11}$ are the within-class scatter matrices of the training samples $\boldsymbol{u}_j^1$ and $\boldsymbol{u}_j^2$, respectively; $\boldsymbol{u}_{ij}^k$ denotes the j^{th} training sample in class i at age group k; $P(\omega_i)$ is the prior probability of class i; n_i is the number of training samples in class i; and $\overline{\boldsymbol{u}}_i^k$ represents the mean vector of the training samples $\boldsymbol{u}_{ij}^k$ (j = 1, …, n_i). The between-class covariance matrix of $\boldsymbol{u}_j^1$ and $\boldsymbol{u}_j^2$ is defined as follows:

$$\mathbf{L}_{12} = \sum_{i=1}^{M} (\boldsymbol{u}_i^1 - \overline{\boldsymbol{u}}^1)(\boldsymbol{u}_i^2 - \overline{\boldsymbol{u}}^2)^T, \tag{10}$$

where M is the total number of training samples, and $\overline{\boldsymbol{u}}^1$ and $\overline{\boldsymbol{u}}^2$ are the mean vectors of the training samples in $\mathbf{U}^1$ and $\mathbf{U}^2$, respectively. The pairs of vectors that maximize the correlation between the project vectors can be learned by maximizing the following generalized CCA criterion function:

$$J_g(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{\boldsymbol{\alpha}^T \mathbf{L}_{12} \boldsymbol{\beta}}{\sqrt{\boldsymbol{\alpha}^T \mathbf{C}'_{11} \boldsymbol{\alpha} \cdot \boldsymbol{\beta}^T \mathbf{C}'_{22} \boldsymbol{\beta}}}. \tag{11}$$

D. Age-Invariant Face Recognition

Two face images are matched based on the difference between their local Gabor features. Assume that p feature points are located in each face image, where the Gabor features are extracted. Therefore, p forward and backward prediction matrices are learned, as described in Section II.B, to change the ages of the Gabor features, according to the age difference between the query and gallery faces.

Assume that a gallery face is in AG1, while the query face is in AG2. Denote the Gabor feature of the gallery face and query face at the i^{th} feature point as $\boldsymbol{u}_i^{g1}$ and $\boldsymbol{u}_i^{q2}$, respectively. For the gallery face, its respective Gabor features are subject to the forward mapping so as to produce corresponding features of AG2, as follows:

$$\boldsymbol{u}_i^{g2} = \mathbf{A}_f^i \, \boldsymbol{u}_i^{g1}, \tag{12}$$

where $\mathbf{A}_f^i$ is the forward mapping function for the i^{th} feature point, and $\boldsymbol{u}_i^{g2}$ is the predicted feature at AG2. Similarly, for the Gabor features of the query input, backward mapping is applied, as follows:

$$\boldsymbol{u}_i^{q1} = \mathbf{A}_b^i \, \boldsymbol{u}_i^{q2}, \tag{13}$$

where $\mathbf{A}_b^i$ is the backward mapping function for the i^{th} feature point, and $\boldsymbol{u}_i^q$ is the predicted feature at AG1. Finally, these features at different ages are projected into corresponding subspaces to form coherent features, as follows:

$$\boldsymbol{v}_i^g = \begin{pmatrix} \boldsymbol{\alpha}^T \boldsymbol{u}_i^{g1} \\ \boldsymbol{\beta}^T \boldsymbol{u}_i^{g2} \end{pmatrix} \text{ and } \boldsymbol{v}_i^q = \begin{pmatrix} \boldsymbol{\alpha}^T \boldsymbol{u}_i^{q1} \\ \boldsymbol{\beta}^T \boldsymbol{u}_i^{q2} \end{pmatrix}, \tag{14}$$

where $\boldsymbol{v}_i^g$ and $\boldsymbol{v}_i^q$ are the features for the gallery face and the query face, respectively. Euclidean distance between the two feature vectors is computed, and the nearest neighbor rule is used for face recognition.

III. Experiments

A. Experimental setup

To compare our proposed age-invariant face recognition method with other standard methods, we conducted our experiment using the FGNet Aging Database [18]. It contains 1,002 face images of 82 people with ages ranging from 0 to 69, which is one of the most widely used datasets having the largest age range. However, some of the images in the database are of low quality, and suffer from blurring, camera artifacts, lighting variations, etc. We have further reduced the number of images, and have only retained 126 images of 42 people, which have relatively better quality and less variations. Each person has three images, each at a different age stage. All of these images are cropped to contain the face region only, with a size of 240×240, and are aligned based on both eyes, as shown in Fig. 1.

In order to learn the progression between different ages, we divide the images into three age groups, namely AG1 for ages from 0 to 6, AG2 from 7 to 18, and AG3 from 19 to 69. In the training phase, local Gabor features, at each of the 53 facial feature points in the training images, are extracted. Then, the linear mapping function and CCA projection matrices are learned for each pair of age groups. After training, the mapping function and the projection matrices are used to generate Gabor features from one age to another age and to form a coherent feature for face recognition. In our experiment, we randomly chose 12 people, with their images in three different age groups, for training each time, and the remaining people with different ages are used for testing. The experiments were repeated five times, so the performance is measured based on 150 images.

B. Experimental results and analysis

Firstly, we compare our proposed age-invariant face recognition method with the Gabor and the local Gabor methods so as to evaluate the improvement after combining aged feature prediction by linear mapping (LP) and generalized CCA (GCCA). The comparison results, in terms of recognition rates, are summarized in Table 1, where AgeInvLPG is our proposed method. We have also examined the improvement due to the use of aged feature prediction without performing fusion by GCCA, and the method is denoted as AgeInvLP. For the local Gabor method, only those Gabor features at the 53 feature points are considered. For the Gabor method, the Gabor features of the whole faces are used for representation. All the methods, compared in the experiment, have the same experimental setup, as described in Section III.A. Both the Gabor and local Gabor features are extracted using Gabor filters of five different scales and eight orientations.

Table 1 Face recognition rates (%) of the Gabor methods and Age-invariant methods with face images from different age groups.

Gallery vs. Query	Gabor	Local Gabor	AgeInvLP	AgeInvLPG
AG1 vs. AG2	4	8.7	10.5	**15.3**
AG1 vs. AG3	12.7	12	15.2	**19.3**
AG2 vs. AG1	16	13.3	33.3	**50.7**
AG2 vs. AG3	28.7	34	36.8	**58.7**
AG3 vs. AG1	12.7	10.7	14.3	**16.7**
AG3 vs. AG2	10	6	22.5	**32.7**

From Table 1, we can observe that the method based on local Gabor features, which only considers the features at the facial-feature points, can achieve a similar performance to the method based on all the Gabor features. The local Gabor method has a much smaller size than the Gabor method, so it is much more computationally efficient. However, we can see that both of these Gabor-based methods cannot recognize faces accurately due to the fact that the Gabor features have significant differences when faces have a great age difference. With our proposed age-invariant Gabor framework, a significant increase in recognition rates can be achieved, especially when the gallery and query face images come from two neighboring age groups. For the AgeInvLP method, i.e. aged features are predicted but no fusing by GCCA is performed, the recognition rate is increased by 1.8% to 20%, depending on the age difference between the face images. When compared to AgeInvLP, the AgeInvLPG method, i.e. the two features at different ages are fused by using GCCA, the improvement in terms of recognition rate is between 2.4% and 21.9%. This shows that the coherent feature, generated by projecting two features at different ages, is effective for age-invariant face recognition.

To give more comprehensive analysis, we have also compared our proposed age-invariant face recognition method, i.e. AgeInvLPG, with some conventional face recognition methods, including eigenface, LBP [9], and LPP [10]. For eigenface, all the principal components are selected for recognition, i.e. $M-1$ principal components. For LPP, the number of Laplacian faces used is also $M-1$. We also follow the same nearest-neighbor classifier approach for recognition, and set the number of nearest neighbors at 7. The comparison results, in terms of recognition rates, are summarized in Table 2.

From the results, we can see that our proposed algorithm can achieve better performance in terms of recognition rates. The eigenface and the LPP methods cannot work well when the face images, to be compared, have a large age difference, in particular, when either the query or gallery image involved belongs to the youngest age group, i.e. in the range of 0-6 years old. This may be due to the fact that the appearances of babies, and young children, change a lot in the first few years after birth. The LBP method can achieve better recognition performance than both PCA and LPP because it can capture more local texture information about faces. However, the performance of LBP still falls behind our proposed age-invariant Gabor-based method. When our method is compared to [6], which uses age-insensitive features for face recognition, the improvement in terms of the rank-1 recognition rate is similar. However, our approach can be applied to [6] to further improve its performance. This means that our proposed approach is effective and efficient for age-invariant face recognition.

Table 2 Face recognition rates (%) of different conventional face recognition methods and our proposed method with face images from different age groups.

Gallery vs. Query	PCA	LBP	LPP	AgeInvLPG
AG1 vs. AG2	6	8.5	9.9	**15.3**
AG1 vs. AG3	8	12.7	8.3	**19.3**
AG2 vs. AG1	18	40.4	24.6	**50.7**
AG2 vs. AG3	10.7	49.9	36.3	**58.7**
AG3 vs. AG1	3.3	11.8	5.5	**16.7**
AG3 vs. AG2	6.7	28.8	22.6	**32.7**

IV. CONCLUSIONS

In this paper, we have proposed a novel approach for age-invariant face recognition by predicting facial features at different ages. Since features at different ages should be correlated to each other, a coherent feature is formed by fusing the two features at different ages by the generalized CCA. By comparing it to those existing face recognition methods on face images with age differences, our algorithm can improve the recognition rate significantly.

A challenge of the research is that only a limited number of training samples are available. Therefore, in our experiments, we simply divide images into three age groups. Nevertheless, experiment results have proven the effectiveness of our proposed algorithm. In our future research, more training samples at different ages will be collected so that the learning of the mapping functions and the projection matrices can be more accurate. In addition, non-linear mapping will be investigated, and other types of features will also be employed.

ACKNOWLEDGEMENT

This work described in this paper was supported by an internal grant from The Hong Kong Polytechnic University, Hong Kong (Project No. G-YL21).

REFERENCES

[1] R. Chellappa, C.L. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," Proceedings of IEEE, 83 (5), pp. 705-740, 1995.
[2] W. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld, "Face Recognition: A Literature Survey," ACM Computing Survey, vol. 35, no. 4, pp. 399-458, 2003.
[3] Jinli Suo, Song-Chun Zhu, Shiguang Shan, and Xilin Chen, "A compositional and dynamic model for face aging," IEEE Trans. Pattern Anal. Machine Intell., vol. 32, no. 3, pp. 385-401, March 2010.
[4] X. Geng, Z. Zhoi, and K. Smith-Miles, "Automatic age estimation based on facial aging patterns," IEEE Trans. Pattern Anal. Machine Intell., vol. 29, no. 12, pp. 2234-2240, December 2007.
[5] Haibin Ling, Stefano Soatto, Narayanan Ramanathan, and David W. Jacobs, "Face verification across age progression using discriminative methods," IEEE Trans. on Information Forensics and Security, vol. 5, no. 1, pp. 82-91, March 2010.
[6] Zhifeng Li, Unsang Park, and Anil K. Jain, "A discriminative model for age invariant face recognition," IEEE Trans. on Information Forensics and Security, vol. 6, no. 3, pp. 1028-1037, September 2011.
[7] Li-Jun Hong; Di Wen; Chi Fang; Xiao-Qing Ding, "Face age estimation by using Bisection Search Tree," 2013 International Conference on Machine Learning and Cybernetics (ICMLC), pp. 370-374, 2013.
[8] T.S. Lee, Image Representation Using 2D Gabor Wavelets, IEEE Trans. Pattern Anal. Machine Intell., vol. 18, no. 10, pp. 959-971, 1996.
[9] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: application to face recognition," IEEE Trans. Pattern Anal. Machine Intell., vol. 28, no. 12, pp. 2037-2041, December 2006.
[10] X.F. He, S.C. Yan, Y.X. Hu, P. Niyogi, and H.J. Zhang. "Face recognition using Laplacianfaces." IEEE Trans. Pattern Anal. Machine Intell., vol 27, no. 3, pp. 328-340, 2005.
[11] J.R. Kettenring, Canonical analysis of several sets of variables, Biometrika 58 (3) (1971) 433-451.
[12] J.G. Daugman, Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters, Journal of the Optical Society of America A 2 (7) (1985) 1160-1169.
[13] J.G. Daugman, Two-dimensional spectral analysis of cortical receptive field profiles, Vision Res. 20 (10) (1980) 847-856.
[14] D. Liu, K.M. Lam, L.S. Shen, Optimal sampling of Gabor features for face recognition, Pattern Recognition Letters 25, (2) (2004) 267-276.
[15] Wiskott, J.M. Fellous, N. Krüger, C. Malsburg, Face recognition by elastic bunch graph matching, IEEE Trans. Pattern Analysis Machine Intelligence 19 (7) (1997) 775-779.
[16] C. Liu, Gabor-based kernel PCA with fractional power polynomial models for face recognition, IEEE Trans. Pattern Analysis Machine Intelligence 26 (5) (2004) 572-781.
[17] Q.S. Sun, Z.D. Liu, P.A. Heng, D.S. Xia, A Theorem on the Generalized Canonical Projective Vectors, Pattern Recognition 38 (3) (2005) 449-452.
[18] FGNet Aging Database. In http://www.fgnet.rsunit.com/.