Available online at www.sciencedirect.com
SCIENCE DIRECT
NEUROCOMPUTING
ELSEVIER
Neurocomputing 69 (2006) 1796–1801
www.elsevier.com/locate/neucom

Letters

Sub-intrapersonal space analysis for face recognition

Xiaoyang Tan, Jun Liu, Songcan Chen*

Department of Computer Science & Engineering, Nanjing University of Aeronautics & Astronautics, 29 Yudao Street, Nanjing, Jiangsu 210016, China

Received 7 August 2005; received in revised form 17 September 2005; accepted 19 September 2005
Available online 21 February 2006
Communicated by R.W. Newcomb

Abstract

Bayesian subspace analysis has been successfully applied in face recognition. However, it suffers from its operating on a whole face difference and using one global linear subspace to represent the similarity model. We develop a novel approach to address these problems. The proposed method operates directly on a set of partitioned local regions of the global face differences, and a separate Gaussian distribution is used to model each sub-intrapersonal space, accordingly. By combining all the local models, we can represent the complex intrapersonal variations more accurately. We further improve the system performance by reducing the contribution of local subspaces containing large variations using a smoothing method. The experiments on several standard face sets show that the proposed method is competitive.

Keywords: Bayesian analysis; Principal component analysis (PCA); Face recognition; Sub-intrapersonal space analysis

1. Introduction

Subspace analysis has attracted much attention in face recognition over the last decade. The essence of subspace analysis is to find the intrinsic face manifold in a low-dimensional space [4]. Although the face image is usually represented in a high-dimensional pixel space, the low-dimensional manifold pursuit is reasonable considering the regularity of facial configuration (e.g., the positions of nose and eyes in a face image). Eigenface [1], Fisherface [2], Lapalacianface [4] and the Bayesian method [3] are four representative subspace methods in the field. Among them, Eigenface does not consider the class information; Fisherface uses the class information but its decision boundaries are both crisp and simple (linear) in nature; Lapalacianface seeks to extract more discriminating information using a local information preserving embedding technique, provided that sufficient training samples are given. The Bayesian method also uses supervised information, but in a way different from the aforementioned methods, i.e., it tries to construct the similarity model (i.e., intrapersonal space) of the same individual in a soft (probabilistic) way. This makes it easier to adapt to unknown samples.

More specifically, in a Bayesian method, the intrapersonal space is constructed by collecting all the difference images (denoted as Δ, $\Delta \in R^n$, where n is the dimension of face image vectors) between any two image pairs belonging to the same individual. From these face differences, the typical intrapersonal variations (denoted by Ω_I) in the same individual will be learned and represented as a likelihood function, i.e., $P(\Delta|\Omega_I)$. By assuming Ω_I to be a high-dimensional Gaussian distribution, the intrapersonal likelihood (called ML measure) is estimated as $P(\Delta|\Omega_I) = (2\pi)^{-n/2}|\Sigma_I|^{-1/2}\exp\{-(1/2)\Delta^{\mathrm{T}}\Sigma_I^{-1}\Delta\}$, where Σ_I is the covariance matrix on the intrapersonal difference set $\{\Delta|\Delta \in \Omega_I\}$. It can be shown the ML estimation of $P(\Delta|\Omega_I)$ is mathematically equalvalent to PCA if the prior knowledge is not considered [3], and the subsequent recognition is reduced to the estimation of the Mahalanobis distance between a probe face and a face in the gallery set, i.e., $d_{\mathrm{F}}^2(\Delta, \Sigma_I) = \Delta^{\mathrm{T}}\Sigma_I^{-1}\Delta$ in the principal subspace. By solving the eigenvalue problem on Σ_I, we can calculate the distance using only the first p principal components, that is, $d_{\mathrm{F}}^2(\Delta, \Sigma_I) = \sum_{i=1}^{p} y_i^2/\lambda_i$, where y_i is the

*Corresponding author. Tel: +86 25 84892452; fax: +86 25 84498069.
E-mail addresses: x.tan@nuaa.edu.cn (X. Tan), j.liu@nuaa.edu.cn (J. Liu), s.chen@nuaa.edu.cn (S. Chen).

0925-2312/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2005.09.018

ith principal component of Δ and λ_i is the corresponding eigenvalue.

Despite the advantages of the Bayesian method, it may be unable to handle the complex situation, where a dataset contains significant transformation difference caused by large lighting, pose and expression variations. Based on the previous analysis, we know that the intrapersonal difference set $\{\Delta | \Delta \in \Omega_I\}$ plays a critical role in the method, and the PCA approach is used to learn the so-needed intrapersonal space. However, the number of training samples per class used to construct the intrapersonal difference set is usually small. Moreover, the standard PCA working on the global face pattern is inadequate to faithfully learn the nonlinear intrapersonal variation. This is due to PCA's preference to mapping directions with maximal variations, while those directions with minimal variations may be unfortunately ignored as noise. In other words, PCA is prone to be fooled by large variations. The Bayesian method tries to circumvent this problem with the Mahalanobis distance, which gives more weight to projectors with small variations. However, due to the fact that the Bayesian method still operates on the global patterns, this distance measure cannot guarantee to effectively reduce the intrapersonal variation.

In this paper, we present a novel method to improve the Bayesian method. Our strategy is to learn a set of local intrapersonal subspaces rather than a global one to capture the complex intrapersonal variations. More specifically, the proposed method operates directly on a set of partitioned local regions of the global face differences, and then constructs corresponding sub-intrapersonal spaces separately using a simple Gaussian distribution. Finally, all the sub-intrapersonal spaces are combined within a probabilistic framework for subsequent recognition.

The significance of the proposed method is threefolds. First, since the whole complex intrapersonal variation is represented by a set of local low-dimensional Gaussian distributions rather than one single high dimensional Gaussian distribution, it is expected that the variation can be modeled more faithfully. Second, experimental results show that most of local intrapersonal variations in a dataset are relatively small and only a small portion of them are large but contrarily dominate the whole intrapersonal variations. In this paper, we use a smoothing method to effectively control the contribution of local models with large variation, thus effectively reducing the intrapersonal variation. Third, due to the low dimensionality of local regions, the learning procedure of the method is very efficient, making it suitable for large datasets with very high dimensionality.

In Section 2 the proposed algorithm is described in detail. Experiments are carried out in Section 3, where the comparing results between the proposed method and several state of the art subspace algorithms are presented on several standard face databases. We conclude in Section 4.

2. Local intrapersonal space analysis

In the proposed method, we decompose the global intrapersonal variation manifold into M local spaces, and use a simple Gaussian distribution to represent each of them. More specifically, in the training stage, the intrapersonal difference sample set is first constructed by computing all the difference images between any two image pairs belonging to the same individual. Then the obtained difference images are partitioned into local regions. In this paper for simplicity, we adopt the equally sized partition scheme [5], that is, each difference image is partitioned into M equally sized local regions (sub-patterns) in a non-overlapping manner. Each local region is further concatenated into corresponding column vectors with dimensionality of l. Then we collect these vectors at the same position of all difference face images to form a training set, in this way, the M separate local difference sets are formed (i.e., $\Delta_k|_{k=1}^{M}$) and the corresponding local intrapersonal variation is denoted as $\Omega_{I,k}|_{k=1}^{M}$. Under the independent assumption among local regions, the global intrapersonal likelihood $P(\Delta|\Omega_I)$ can then be expressed as the product of M local intrapersonal likelihoods $P(\Delta_k|\Omega_{I,k})$, i.e.:

$$P(\Delta|\Omega_I) = \prod_{k=1}^{M} P(\Delta_k|\Omega_{I,k}). \tag{1}$$

If a Gaussian distribution is assumed on each local subspace, it follows that

$$P(\Delta_k|\Omega_{I,k}) = (2\pi)^{-l/2}|\Sigma_{I,k}|^{-1/2} \exp\left\{-\frac{1}{2}\Delta_k^{\mathrm{T}}\Sigma_{I,k}^{-1}\Delta_k\right\}, \tag{2}$$

where $\Sigma_{I,k}$ is the covariance matrix on the kth intrapersonal difference set $\{\Delta_k|\Delta_k \in \Omega_{I,k}\}$, i.e., $\Sigma_{I,k} = \Delta_k\Delta_k^{\mathrm{T}}$. As mentioned before, when no prior knowledge is available, the ML estimation of $P(\Delta_k|\Omega_{I,k})$ is PCA [3], which in turn reduces to solve an eigenvalue problems of the covariance matrix $\Sigma_{I,k}$, and the resulting local intrapersonal subspace is spanned by the first q eigenvector of the $\Sigma_{I,k}$. In each subspace, we can either adopt the squared Euclidean distance or the squared Mahalanobis distance as the "distance" measure for recognition. The square of the Euclidean distance is defined to be

$$(d_{\mathrm{E}}^k)^2 = \sum_{i=1}^{q}(y_i^k)^2 \tag{3}$$

and that of the Mahalanobis distance:

$$(d_{\mathrm{F}}^k)^2 = \sum_{i=1}^{q}(y_i^k)^2/\lambda_i^k, \tag{4}$$

where y_i^k is the ith principal component of the kth local region and λ_i^k the corresponding eigenvalue. Obviously, d_{E}^k is a special case of d_{F}^k if λ_i^k is not considered. In practice, the choice of Eqs. (3) or (4) is dataset-dependent. Empirically, if the local regions contain large-scale variations as in the ORL dataset, the squared Mahalanobis distance (Eq. (4))

can be used, otherwise, the squared Euclidean distance (Eq. (3)) is preferred. In the following experiments, the squared Euclidean distance (Eq. (3)) is adopted as the default setting. This is mainly due to the observation that the local variations contained in each face regions are generally very small (see experimental section), and the Euclidean-based distance could be helpful to reduce the risk of amplifying irrelevant dimensions with small variance.

Now, by applying a logarithmic transformation on both sides of Eq. (1) and combining it with Eq. (2), we obtain the total squared distance D between any two images I_1, I_2:

$$D(I_1, I_2) = \sum_{k=1}^{M} \Delta_k^{\mathrm{T}} \Sigma_{I,k}^{-1} \Delta_k \triangleq \sum_{k=1}^{M} (d^k)^2, \tag{5}$$

Local Intrapersonal Subspace Analysis Algorithm

Input: A set of training samples $T = \{x_{c,i}; c = 1,\ldots,C; i = 1,\ldots,J_c\}$,where $x_{c,i}$ is the i-th image belonging to the c-th class, and J_c is the total number of faces in the c-th class.

Training:

Step1: Construct global intrapersonal difference set for each person $\{\Delta_{c,k_c} \mid \Delta_{c,k_c} = x_{c,i} - x_{c,j}; i \neq j\}$, and the overall set is $\Delta = \{\Delta_{c,k_c} \mid c = 1,\ldots,C; k_c = 1,\ldots,K_c\}$,where K_c is the total number of difference images in the c-th class.

Step2: Partition each difference image Δ_{c,k_c} into M local regions, that is, $\Delta_{c,k_c} = \{\Delta_{c,k_c,i} \mid i = 1,\ldots,M\}$.

Step3: Group all the local regions at the same position over the whole difference set. Each group is denoted by $g_i = \{\Delta_{c,k_c,i} \mid c = 1,\ldots,C; k_c = 1,\ldots,K_c\}$, and the total group is $G = \{g_i \mid i = 1,\ldots,M\}$.

Step4: Apply PCA on each group g_i to obtain its corresponding local intrapersonal space s_i, which is spanned by the first p eigenvetors of the covariance matrix on group g_i.

Recognition:

Step1: Compute a probe face image t's difference with each image in the gallery set, $\{\Delta^t_{c,i} \mid \Delta^t_{c,i} = t - x_{c,i}; i = 1,\ldots,J_c; c = 1,\ldots,C\}$.

Step2: Partition each difference image into M local regions as the same way above, that is, $\Delta^t_{c,i} = \{\Delta^t_{c,i,k} \mid k = 1,\ldots,M\}$

Step3: Project each local region $\Delta^t_{c,i,k}$ into its corresponding local intrapersonal space s_k, and compute the square of distances in the eigenspace $(d^k_{c,i})^2$ using either e.q.(3) or e.q.(4).

Step4: Smooth each local squared distance $(d^k_{c,i})^2 \mid_{i=1}^{M}$ using e.q.(6).

Step5: Sum all the smoothed squared distances with e.q.(7) to obtain the total similarity $s(t, x_{c,i})$ between the probe t and a gallery image $x_{c,i}$ in T.

Step 6: Recognize the face class with the maximal similarity to the probe face image (e.q(8)).

Fig. 1. Algorithm of local intrapersonal space analysis.

where $(d^k)^2$ is the kth local square-of-distances, which can be evaluated using either Eqs. (3) or (4).

Another key issue for the proposed method is how to reduce the overall intrapersonal variation. We address the problem using a traditional smoothing method. Here the exponential function is chosen for that purpose, that is,

$$d_s^k = \exp(-(d^k)^2/2), \quad (6)$$

where d_s^k is the smoothed version of kth local squared distance defined in (3) or (4). Such a transformation not only controls the contribution of each local model, but also changes the squared distance into a similarity measure, with the property that the larger the squared distance, the smaller the similarity. Consequently, the influence of large local variations on the total similarity is smoothed and the overall intrapersonal variation is effectively reduced.

To this end, simply by replacing $(d^k)^2$ with d_s^k in Eq. (5), we obtain the total similarity of a probe t to the prototype image I_c of each class ($c = 1,\ldots,C$), that is,

$$S(t, I_c) = \sum_{k=1}^{M} d_s^k(\Delta_{t,c}^k), \quad (7)$$

where $\Delta_{t,c}^k$ is the kth local intrapersonal difference between t and I_c. And we recognize the face class with the maximal similarity to the probe face image, i.e.,

$$\text{label}(x) = \arg \max_{c=1,\ldots,C} (S(t, I_c)). \quad (8)$$

The detailed description of the above-proposed algorithms is shown in Fig. 1.

3. Experimental results

In the first experiment, we demonstrate how the global pattern may cover the truth of intrapersonal variation. The faces are from the ORL database (http://www.uk.research.att.com/facedatabase.html). It contains 40 persons with 10 images for each person. All the images are cropped to the size of 56×46 pixels. The face images contain significant intrapersonal variations caused by rotation, expression and sample size. We partitioned the difference images into 8×23 local regions (sub-patterns), and computed the intrapersonal pairwise squared Euclidean distance for each region. The histogram of the obtained distances is depicted in Fig. 2. Fig. 2 indicates that 83.6% local distances are less than 0.4×10^6 (small compared to the largest one, 1.9×10^6), however, their contribution to the overall intrapersonal distances is only 61.5%. In other words, *nearly 40% total variation is due to about 16.0% local regions with large variation*. This very insight, which seems to be ignored by most previous researches, provides one of the major justifications of our algorithm. It actually suggests an efficient and effective way to reduce the overall intrapersonal variation, i.e., by punishing those local models with large variation.

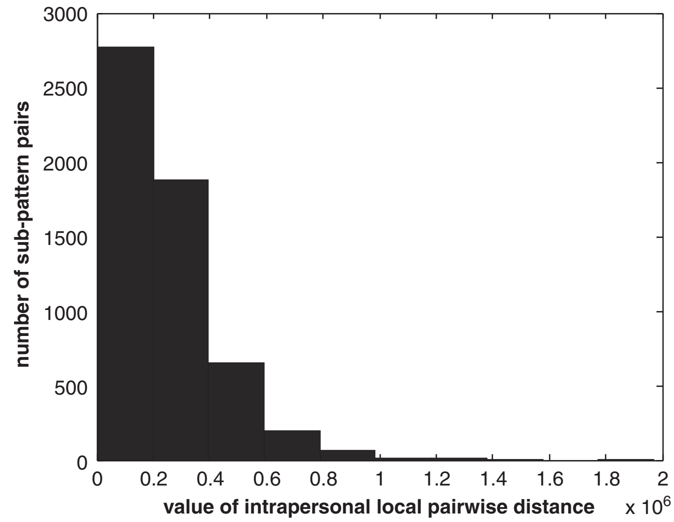

Fig. 2. Histograms of intrapersonal local pairwise distances on the ORL dataset.

Next, we want to compare the performance of the local intrapersonal space analysis to the global intrapersonal space analysis (i.e., the standard Bayesian method) as well as other state of the art subspace analysis methods, including eigenface, fisherface, Lapalacianface and a local PCA method named subpattern-based PCA (SpPCA [5]). In the experiments, 98% information in the sense of reconstruction is kept in the PCA subspaces for all the compared methods. For fisherface and Lapalacianfaces, ($C-1$) projectors are extracted, where C is the number of total classes. The experiments are conducted on four well-known databases, i.e., AR [6], Yale [2], ORL and FERET [7]. The AR dataset contains 100 subjects and each subject has 26 face images taken in two sessions. For each session, there are 13 face images. Here the first 7 faces from the first session of each person are used for training (700 faces in total), and the first 7 faces from the second session of each person (700 faces in total) for testing. The 1400 images are all cropped into the same size of 66×48 pixels. Face images in this dataset have very significant intrapersonal variations including large expression and lighting changes. The Yale set contains 165 face images of 15 persons, with each person having 11 images. The first 6 faces of each person are used for training and the latter 5 for testing. All the images are cropped into 50×50 pixels. This dataset is used to examine the system performance when both facial expressions and illumination are varied. On the ORL dataset, the first 5 images of each person are selected for training and the latter 5 for testing. This dataset is challenging for its variations in pose, expression and sample size. On this dataset, the squared Mahalanobis distance (Eq. (4)) is used. Finally, a larger dataset from FERET is used. This dataset contains 1195 subjects with two faces for each person. Images of 195 persons are randomly selected for training, and the remaining 1000 persons are used for testing. So, there are a total of 390

Table 1
Classification accuracy (%) comparison of our method with other methods on four datasets

Dataset	Our method	Bayesian	Eigenface	Fisherface	Lapalacianface	SpPCA
AR	89.6(22×6*)	84.3	74.1	85.3	85.1	77.7
ORL	91.5(8×2)	82.5	88.5	85.5	85.0	90.0
YALE	84.0(10×10)	76.0	77.3	82.7	82.7	81.3
FERET	91.3(12×12)	89.3	76.8	85.2	87.7	79.7

*The size of local region obtaining the corresponding performance.

Table 2
Sensitivity of the proposed method to the size of local regions

	AR(66×48[a])			ORL(56×46)			YALE(50×50)			FERET(60×60)		
	6×8[b]	11×8	22×6	8×2	28×2	8×23	5×5	10×10	25×10	10×5	12×12	15×15
Accuracy (%)	89.3	89.1	89.6	91.5	91.0	91.0	84.0	84.0	84.0	89.2	91.3	89.9
Mean (%)		89.3			91.2			84.0			90.1	
Std		0.06			0.08			0.0			1.14	

[a]Size of the original image.
[b]Size of the local region.

images in the training set, 1000 face images in the gallery, and 1000 face images for probe. All the images are cropped into 60×60 pixels. The experimental results are given in Table 1. The proposed local intrapersonal space method outperforms all other methods consistently of all the face datasets.

Finally, we study the sensitivity of the proposed method to the size of the local regions. The top 1 matching rates under different local region size in the four datasets are shown in Table 2. The results reveal that our algorithm is insensitive to the size of local region.

4. Conclusions

We proposed a novel method to model the intrapersonal variation. The proposed method decomposes the complex intrapersonal manifold into a set of local models, and uses a separate simple Gaussian distribution to represent each of them. In addition, we effectively reduce overall intrapersonal variation by reducing the contribution of those local models with large local variation with a smoothing method. Experimental results on four well-known face datasets reveal that the proposed local-pattern-based method achieves a higher accuracy than the traditional global-pattern-based Bayesian algorithm as well as Eigenface, Fisherface and the recently proposed Lapalacianface when only a few images are available. In the future work, we will focus on addressing the partial occlusion problems using the proposed method.

Acknowledgements

This work was supported by the National Natural Science Foundation of China under the Grant No. 60473035 and the Natural Science Foundation of Jiangsu Province under the Grant No. BK2005122.

References

[1] M. Turk, A. Pentland, Eigenfaces for recognition, J. Cognitive Neurosci. 3 (1) (1991) 71–86.
[2] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs. Fisherfaces : Recognition using class specific linear projection, IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 711–720.
[3] B. Moghaddam, Bayesian face recognition, Pattern Recogn. 13 (11) (2000) 1771–1782.
[4] X. He, X. Yan, et al., Face recognition using Laplacianfaces, IEEE Trans. Pattern Anal. Mach. Intell. 27 (3) (2005) 328–340.
[5] S. Chen, Y. Zhu, Subpattern-based principal component analysis, Pattern Recogn. 37 (5) (2004) 1081–1083.
[6] A. M. Martínez, R. Benavente, The AR Face Database, Computer Vision Center (CVC) Technical Report, #24, June 1998, available at http://sampl.ece.ohio-state.edu/~aleix/.
[7] P.J. Phillips, H. Moon, S. Rizvi, P. Rauss, The FERET evaluation methodology for face recognition algorithms, IEEE Trans. Pattern Anal. Mach. Intell. 22 (10) (2000) 1090–1103.

Xiaoyang Tan received the B.Sc. and M.Sc. degrees in computer science from Nanjing University of Aeronautics & Astronautics (NUAA), China, in 1993 and 1996, respectively. Then he joined the Department of Computer Science & Engineering of NUAA as an assistant lecturer in 1996. He received a Ph.D. degree in computer science at Nanjing University in 2005, and is now an associate professor at NUAA. His research interests include pattern recognition, machine learning and neural computing.

Jun Liu received a B.Sc. from Nantong Institute, Jiangsu, PR China. Currently he is a Ph.D. student at the Department of Computer Science & Engineering, NUAA. His research interests focus on Neural Computing and Pattern Recognition.

Songcan Chen received the B.Sc. degree in mathematics from Hangzhou University (now merged into Zhejiang University) in 1983. In December 1985, he completed the M.Sc. degree in computer applications at Shanghai Jiaotong University and then worked at Nanjing university of Aeronautics & Astronautics (NUAA) in January 1986 as an assistant lecturer. There, he received a Ph.D. degree in communication and information systems in 1997. Since 1998, as a full professor, he has been with the Department of Computer Science and Engineering at NUAA. His research interests include pattern recognition, machine learning and neural computing. In these fields, he has authored or coauthored over 70 scientific journal papers.