Image-set based Face Recognition Using Boosted Global and Local Principal Angles

Xi Li[1], Kazuhiro Fukui[2], and Nanning Zheng[1]

[1] Xi'an Jiaotong University, China
lxaccv09@yahoo.com,   znn@xjtu.edu.cn
[2] University of Tsukuba, Japan
kf@cs.tsukuba.ac.jp

Abstract. Face recognition using image-set or video sequence as input tends to be more robust since image-set or video sequence provides much more information than single snapshot about the variation in the appearance of the target subject. Usually the distribution of such image-set approximately resides in a low dimensional linear subspace and the distance between image-set pairs can be defined based on the concept of principal angles between the corresponding subspace bases. Inspired by the work of[4, 14], this paper presents a robust framework for image-set based face recognition using boosted global and local principal angles. The original multi-class classification problem is firstly transformed into a binary classification task where the positive class is the principal angle based intra-class subspace "difference" and the negative one is the principal angle based inter-class subspace "difference". The principal angles are computed not only globally for the whole pattern space but also locally for a set of partitioned sub-patterns. The discriminative power of each principal angle for the global and each local sub-pattern is explicitly exploited by learning a strong classifier in a boosting manner. Extensive experiments on real life data sets show that the proposed method outperforms previous state-of-the-art algorithms in terms of classification accuracy.

1   Introduction

This paper presents a robust framework for image-set based face recognition using boosted global and local principal angles. Face recognition has been an active research field due to its various real life applications such as human-computer interface, non-intrusive public security et. al. Some traditional methods for face recognition include Principal Component Analysis(PCA), Independent Component Analysis(ICA), Linear Discriminative Analysis(LDA) and Bayesian method[1–4]. Recently face recognition using image-set or video sequence has attracted more and more attention within computer vision and pattern recognition community[5–12]. One reason is that the availability of modern cheap capture devices makes video sequence a more natural choice of input for face recognition tasks, for example in the surveillance scenario. More importantly, compared with single snapshot, a set or a sequence of images provides much more information about the variation in the appearance of the target subject. The variation always exists in the context of subject recognition or visual surveillance applications, where multiple shots of target subject under varying illumination and facial

expressions or surveillance system output over long periods of time are available. Previous studies show that more robust recognition performance can be achieved by fully exploiting these kind of information[6].

It is well known that the appearance distribution of image-set or sequence for a target subject captured under changing facial expressions and varying illumination conditions can be approximately represented by a low dimensional linear subspace. The principal angles between subspace pairs can be used as a distance measure of the corresponding image-set pairs. Many algorithms have been proposed for robust recognition using the concept of principal angle based subspace distances. A noteworthy work is the Mutual Subspace Method(MSM) presented by Yamaguchi et al [6]. In MSM, each image set is represented by the linear subspace spanned by the principal components of the data and the smallest principal angle between subspaces is exploited as distance measure. The original MSM can be further improved such as in the Constrained Mutual Subspace Method(CMSM)[8]. Instead of directly applying the classification using the principal angle based subspace distance, the underlying idea of CMSM is to learn a linear transformation while in the transformed space the corresponding inter-class subspace distances are larger than that in the original feature space. That is to say, the subspace bases for different classes in the transformed space are more orthogonal to each other. The above methods were further extended to their non-linear counterparts by using the kernel trick such as in [7, 9]. Kim et. al.[10] borrowed the idea of Linear Discriminant Analysis and presented an alternate method which tries to iteratively minimize the principal angles of within-class sets and maximize the principal angles of between-class sets.

Inspired by the work of[4, 14], this paper presents a robust framework for image-set based face recognition using boosted global and local principal angles. The original multi-class classification problem is firstly transformed into a binary classification task where the positive class is the principal angle based intra-class subspace "difference" and the negative one is the principal angle based inter-class subspace "difference". The principal angles are computed not only globally for the whole pattern space but also locally for a set of partitioned sub-patterns. This scheme is robust to local variants and can make the inherent linear principal angle based methods faithfully describe non-linear patterns to some extent. Furthermore, the discriminative power of each principal angle for the global pattern and each local sub-pattern is explicitly exploited and appropriately aggregated by learning a strong classifier in a boosting manner. Extensive experiments on real life data sets show that the proposed method outperforms the previous state-of-the-art algorithms in terms of classification accuracy.

The rest of this paper is organized as follows: we first overview the concept of the principal angles between subspace bases for the corresponding image-sets. We then discuss some previous principal angle based image-set classification methods and their drawbacks in section 2. Section 3 describes the proposed method in detail. Section 4 is the experimental results and section 5 draws the conclusion.

2 Image-set based face recognition using principal angles

We first define the image-set based face recognition problem as follows: Specifically, given C subjects n input image-sets $X_i \in R^{rc \times h_i}, i = 1, ..., n$ and corresponding subject identities $y_i \in \{1, 2, ..., C\}, i = 1, ..., n$, where rc is the image vector dimension and h_i is the number of images in the i-th image set or length of the i-th sequence. A face image pattern obtained from one view can be represented as a point in a high-dimensional feature space where an $r \times c$ pixel pattern is treated as a vector x in rc-dimensional space. Each input image set has the underlying k-dimensional subspace structure denoted as $U_i \in R^{rc \times k}, i = 1, ..., n$, which can approximately describe the variations in the appearance caused by different illuminations, varying poses and facial expression, et. al. For a test image-set X_{test}, the task is to predict its corresponding identity y_{test}.

2.1 Principal angles between subspace pairs

Recently in computer vision community, the concept of principal angles[13] is used as a distance measure for matching two image sets or sequences, where each of them can be approximated by a linear subspace[6]. If the principal angles between two subspaces derived from two image-sets are smaller enough, then the two sets are considered similar.

Generally, let U_A, U_B represent two k-dimensional linear subspaces. The principal angles $0 \leq \theta_1 \leq \theta_2 \leq ... \leq \theta_k \leq \pi/2$ between the two subspaces can be uniquely defined as[6, 13]:

$$cos^2(\theta_i) = \max_{\substack{u_i^A \perp u_j^A (j=1,2,...,i-1) \\ u_i^B \perp u_j^B (j=1,2,...,i-1)}} \frac{\|(u_i^A \cdot u_i^B)\|^2}{\|u_i^A\|^2 \|u_i^B\|^2} \tag{1}$$

where $u_i^A \in U_A, u_i^B \in U_B$. Denote $Dis(X_A, X_B)$ as the principal angle based distance between image-sets X_A, X_B with corresponding subspace bases U_A and U_B. It is a function of principal angles in the following form: $Dis(X_A, X_B) = f(\theta_1, \theta_2, ..., \theta_k)$. Different methods for image-set based recognition has different empirical form of the function f, which will be discussed in detail in the next subsection.

2.2 Image-set based recognition methods using principal angles

The original baseline Mutual Subspace Method(MSM)[6] performs the nearest neighbor search in the original feature space without feature extraction procedure. The distance between the subspace bases of the corresponding image-set pairs is defined as the minimum principal angle in the form of $Dis(X_A, X_B) = \theta_1$.

The Constrained Mutual Subspace Method(CMSM)[8] project the original feature onto the first d few eigenvectors of the $G = \sum_{i=1}^{C} P_i = \sum_{i=1}^{C} U_i U_i^T$, which is the sum of the projection matrix of all classes. The distance between the subspace bases of the

corresponding transformed image-set pairs is defined as the mean of the first t smallest principal angles as $Dis(X_A, U_B) = \frac{1}{t} \sum_{i=1}^{t} \theta_t$. Not only the dimension of the generalized difference subspace but also the number of t has great effect on the recognition rate and the selecting of appropriate choice is quite empirical and case-dependent. The Discriminative Analysis of Canonical Correlations(DCC)[10] is an alternate method which tries to iteratively minimize the principal angles of within-class sets and maximize the principal angles of between-class sets.

Generally, for a subspace base pairs U_A, U_B which have rank of k, there exist k principal angles. All above previous methods does not consider the different discriminative power of each principal angles. They either use the minimum principal angle or use the mean or weighted mean of a truncated form of all principal angles as the similarity measure. The main drawback of the these empirical schemes is that each principal angle based features has its own discriminative power to some extent, we should not only exploit as much as discriminative information of each of them but also the aggregation should be implemented in a more reasonable way, not just the mean or weighted mean of each principal angle based features. For example, Figure 1 is an example curve of the effect of each principal angle on the recognition rate.
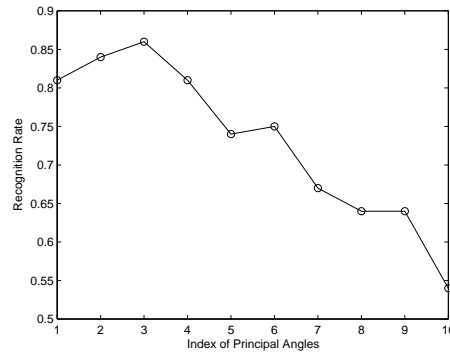
Fig. 1. The recognition performance of using each individual principal angle as the distance measure for a typical image-set based face recognition task.

Since the original principal angle is computed directly on the global pattern and its linear property in nature, it has the following inherent shortcomings: 1) it can not describe the nonlinear real life data faithfully. The real life high dimensional data, such as the vectorized image data, is often inherently nonlinear rather than simple normally distribution[18]. Wolf. L. et al [7] extended the MSM to its non-linear counterpart by using the kernel method. In their method, an input pattern is first mapped onto a higher dimensional feature space via a nonlinear map, then the MSM is applied to the linear subspaces generated from the mapped patterns. The performance is dependent on the form of the kernel function and the corresponding parameters to be tuned. K. Fukui et.

al. further extended the CMSM to the nonlinear case[9]. The Kernel-CMSM outper-forms the CMSM greatly but at the expense of prohibitively computation load which make the real life applications of Kernel-CMSM quite difficult. 2) the original principal angle based methods is performed on the whole global patterns and is less robust to the local variants. Each local region also contains different principal angle based discriminative power and should be aggregated in a reasonable way.
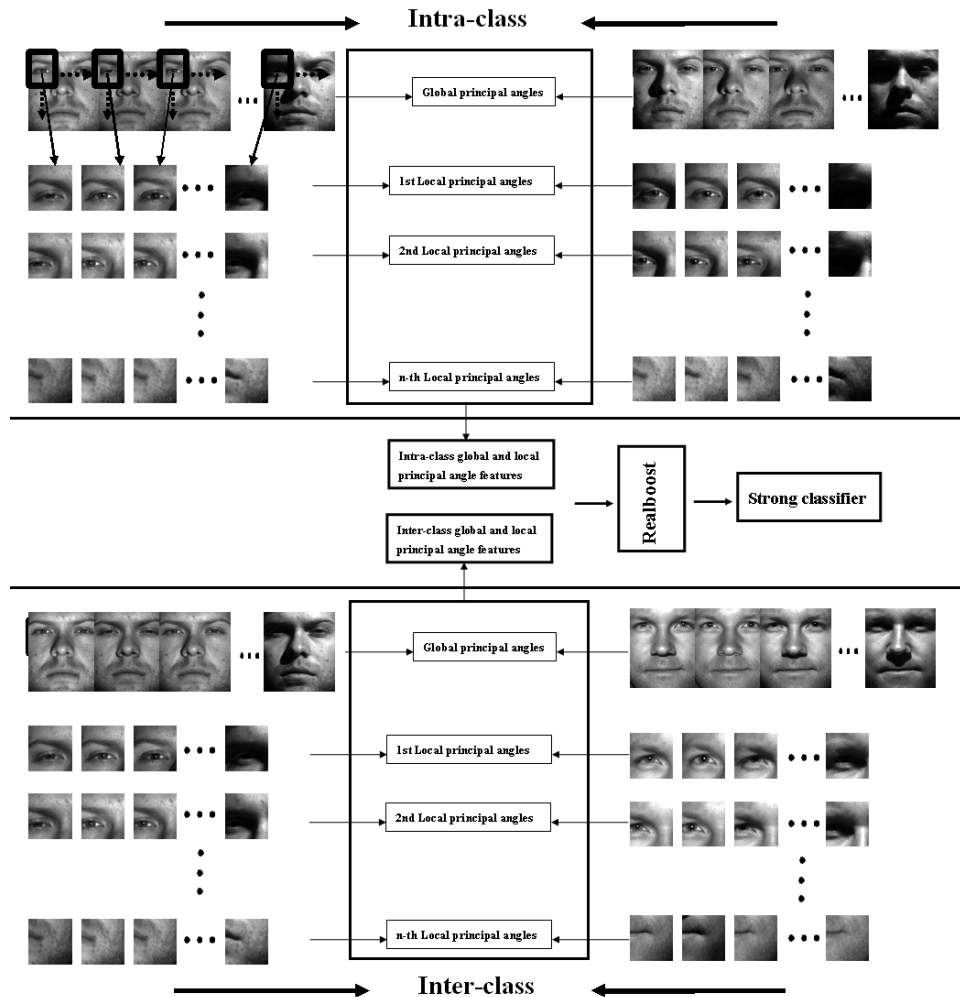
Fig. 2. The flowchart of the proposed method for image-set base face recognition using boosted global and local principal angles. the up-left part illustrates how to extract local sub-patterns using a sliding rectangle window as in[14].

3 Image-set based face recognition using boosted global and local principal angles

Based on the above analysis and inspired by the work of[4, 14], this paper presents a robust framework for image-set based face recognition using boosted global and local principal angles. The original multi-class classification problem is firstly transformed into a binary classification task where the positive class is the principal angle based intra-class subspace "difference" and the negative one is the principal angle based inter-class subspace "difference". Here we use the principal angles between subspace base pairs to resemble the facial image differences as in[4]. The principal angles are computed not only globally for the whole pattern space but also locally for a set of partitioned sub-patterns. The discriminative power of each principal angle for the global and each local subpattern is explicitly exploited and appropriately aggregated by learning a strong classifier in a boosting manner. Different contributions for the principal angle base classification made by different local sub-patterns of the whole global face pattern are emphasized also in the boosting manner.

3.1 The principal angle based intra-class and inter-class subspace difference

Besides the traditional PCA, ICA, LDA methods[1–3], the bayesian method proposed by Moghaddam et. al.[4] is one of the most successful algorithms for face recognition. In their method, the original multi-class face recognition problem is first converted into a two-class problem. Based on the Gaussian distribution assumption, the intra- and inter- personal subspace are learned to describe the variation in difference images of the same individual and different individuals respectively. The similarity score can be evaluated in terms of a posteriori probability using Bayesian rule. For more detail refer to literature[4].

Similar to the above approach, we denote Θ_I the principal angle based intra-class subspace "difference" , which describe the principal angle between image-set pairs corresponding to the same subject. The inter-class subspace "difference" Θ_E can be defined in a similar way which describe the principal angle between image-set pairs corresponding to the different subject. More specific, given a subspace pairs U_i, U_j with principal angles $PA(U_i, U_j) = [\theta_1, \theta_2, ..., \theta_k]$

$$\begin{cases} PA(U_i, U_j) \in \Theta_I, & if \quad C_i = C_j \\ PA(U_i, U_j) \in \Theta_E, & if \quad C_i \neq C_j \end{cases}$$

Due to its inherently linear property, the principal angles computed on the global image patterns can not describe the nonlinear real life face patterns faithfully and is not robust to local variants. We partition the original face image to a set of equally sized sub-images as in[14]. Then all those sub-pattern sharing the same original components are respectively collected from the training set to compose a corresponding sub-pattern's training set. For each sub-pattern's training set, the intra-class and inter-class subspace "difference" are computed in a similar way to that of global counterpart. The global and local principal angles are concatenated to form the principal angle features for intra-class and inter-class respectively, which are provided as inputs to the boosting

procedure described in the next subsection. For extracting each sub-pattern, we slide a rectangle window on the whole image pattern in left-to-right and up-to-down order as in[14]. Figure 2 illustrate the flowchart of the proposed methods.

3.2 Principal angle boosting

Boosting is a classifier ensembling method and has success applications such as in the face/object detection. In this paper, we treat each principal angle as the input features for the inter-class(Θ_E) and intra-class(Θ_I) classification problem and learn the strong classifier using boosting to combine the discriminative power of each principal angle. We use the Realboost algorithm proposed in [15], which is an extension of the discrete adaboost algorithm. Realboost deals with confidence-rated weak classifier which has a real-valued output and the confidence of the strong classifier can be easily evaluated. Here the confidence resembles the intra-personal and inter-personal likelihood as in[4]. The procedure of the boost learning algorithm can be described in detail as follows:

Training:
Input: Given the training data $\{\Omega_i \in \Gamma\}, i = 1, ..., n$ and the corresponding labels $y_i = \{-1, +1\}, i = 1, ..., n$. $\Omega = \{\theta^g, \theta^{l_1}, \theta^{l_2}, ..., \theta^{l_m}\}$ where θ^g is the principal angle features for global patterns and $\theta^{l_j}, j = 1, ..., m$ is the principal angle features for the j-th local sub-patterns.
　　　　The number of the maximum boosting step T.
Procedures:
1) Initialize the sample weights $W_1(i) = 1/n, i = 1, ..., n$,
2) Repeat for $t = 1, ..., T$,
　　2.1) For each principal angle feature based weak classifier:
　　　　a) Partition the space Γ into several disjoint blocks as $\Gamma_1, \Gamma_2, ..., \Gamma_z$
　　　　b) Under the current weighting W_t, calculate
　　　　$W_l^j = Pr(\Omega_i \in \Gamma_j, y_i = l) = \sum_{i:\Omega_i \in \Gamma_j, y_i = l} W_t(i)$
where $l = \{-1, +1\}$
　　　　c) $\forall \Omega \in \Gamma_j, h(\Omega) = \frac{1}{2} ln(\frac{W_{+1}^j + \epsilon}{W_{-1}^j + \epsilon})$
where ϵ is a small positive constant.
　　　　d) calculate the normalization factor:
　　　　$Z = 2 \sum_j \sqrt{(W_{+1}^j W_{-1}^j)}$
　　2.2) Select the weak learner for the minimum Z and set the corresponding function h as h_t
　　2.3) Update the sample weight as
　　　　$W_{t+1}(i) = W_t(i) exp(-y_i h_t(x_i))$
and re-normalize so that the sum of elements of W equals to 1.
3) The final strong classifier is
　　　　$H(\Omega) = sign[\sum_{t=1}^T h_t(\Omega)]$
And the confidence of the output of H can be defined as
　　　　$Confidence(\Omega) = |\sum_{t=1}^T h_t(\Omega)|$

Testing:

For a given test image-set, its principal angle based "difference" with each training image-set is computed first and classified using the learned strong classifier. The label of the training image-set with the highest intra-class confidence score is selected as the output label.
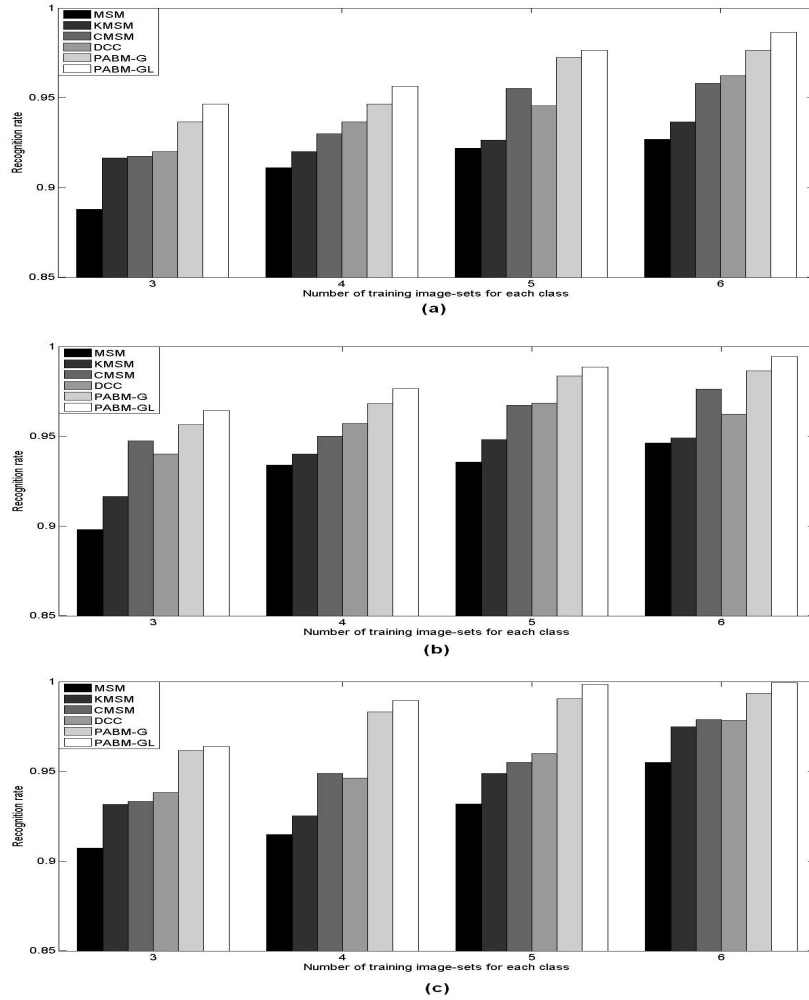
Fig. 3. Classification rate comparison between MSM,KMSM,CMSM,DCC,the proposed method with global principal angles boosting only(PABM-G) and the proposed method with both global and local principal angles boosting(PABM-GL) for a)CMU-PIE dataset 2) YaleB dataset and c) a self-collected face video dataset.

4 Experimental result

In this section we test the proposed method using real life facial image databases, which include the CMU-PIE database[16], the YaleB face database[17] and a self-collected face video database. We compare the performance of the following algorithms: 1) The baseline Mutual Subspace Method(MSM)[6]; 2) The Kernel Mutual Subspace Method(KMSM)[7]; 3) The Constrained Mutual Subspace Method(CMSM)[8]; 4) The Discriminative Analysis of Canonical Correlations(DCC)[10]; 5) The proposed Principal Angle Boosting Method using only Global patterns(PABM-G); 6) The proposed Principal Angle Boosting Method using both Global and Local patterns(PABM-GL);

The linear subspace of each image-set was learned using principal component analysis and the corresponding dimension was chosen to represent 98% data energy. For CMSM, the dimension of the generalization difference subspace was empirically set to be the 95% of the full image dimension. For KMSM, a six-degree polynomial kernel was used as in[7]. For PABM-GL, we extract the local sub-patterns using a sliding rectangle window with size 8×8 in pixels and steps in vertical and horizontal direction are set to be 4 pixels.

• For the YaleB face database we used images of 38 subjects. Firstly 80 near-frontal images under different illuminations per subject were selected. Then the face regions are cropped and resize to 24×21 in pixels. The 80 images are divided into 10 image-sets with each image set has 8 images.

• For the CMU-PIE face database we used images of 45 subjects and for each subject 160 near-frontal images were selected which cover variations in facial expression and illumination. Then the face regions are cropped and resize to 24×24 in pixels. The 160 images are divided into 16 image-sets with each image set has 10 images with different illumination conditions.

• To further illustrate the performance of the proposed method, we self-collected a face video database which has 20 subjects and for each subject 10 separate sequences were recorded with large variation in illumination and facial express. Each specific sequence contains about 180 frames and the face region was extracted automatically using a face detector. The face regions were histogram equalized and resized to 25×25 in pixels.

For all the above three datasets each face vector was normalized to the unit. Then we randomly selected 3,4,5,6 images-sets from each class into training sets and the remaining into testing set, respectively. For each number of training sets, the random partition procedure repeated for 10 times and the average classification results are computed.

The recognition rate comparison for the above experiments is demonstrated in Figure 3. It can be seen that the proposed method of recognition using boosted global principal angles outperforms the MSM,KMSM,CMSM,and DCC consistently. The performance can be further improved by integrating local principal angle boosting.

5 Conclusion

This paper presents a robust framework for image-set based face recognition using boosted global and local principal angles. The original multi-class classification problem is firstly transformed into a binary classification task where the positive class is the

principal angle based intra-class subspace "difference" and the negative one is the principal angle based inter-class subspace "difference". The principal angles are computed not only globally for the whole pattern space but also locally for a set of partitioned sub-patterns. The discriminative power of each principal angle for the global and each local subpattern is explicitly exploited by learning a strong classifier in a boosting manner. Experiments on real life data sets demonstrate the superior performance of the proposed method to previous state-of-the-art algorithms in terms of classification accuracy..

References

1. M. Turk, A. Pentland, Eigenfaces for Recognition, Journal of Cognitive Neurosicence, Vol. 3, No. 1, pp. 71-86, 1991
2. K. Etemad, R. Chellappa, Discriminant Analysis for Recognition of Human Face Images, Journal of the Optical Society of America A, Vol. 14, No. 8, pp. 1724-1733, 1997
3. M.S. Bartlett, J.R. Movellan, T.J. Sejnowski, Face Recognition by Independent Component Analysis, IEEE Trans. on Neural Networks, Vol. 13, No. 6, pp. 1450-1464, 2002
4. B. Moghaddam, T. Jebara, A. Pentland, Bayesian Face Recognition, Pattern Recognition, Vol. 33, Issue 11, pp. 1771-1782, 2000
5. G. Shakhnarovich, J. W. Fisher and T. Darrell. Face recognition from long-term observations. In Proceedings of European Conference on Computer Vision, pp.851-865, 2002
6. O. Yamaguchi, K. Fukui and K. Maeda. Face recognition using temporal image sequence. International Conference on Automatic Face and Gesture Recognition, pp.318-323, 1998
7. L. Wolf and A. Shashua. Learning over Sets using Kernel Principal Angles. Journal of Machine Learning Research,4(10),pp.913-931, 2003
8. K. Fukui. and O. Yamaguchi. Face recognition using multi-viewpoint patterns for robot vision. 11th International Symposium of Robotics Research, pp.192-201, 2003
9. K. Fukui, B. Stenger, O. Yamaguchi, "A framework for 3D object recognition using the kernel constrained mutual subspace method", Asian Conference on Computer Vision, part-I, pp.315-324, 2006.
10. T. K. Kim, J. Kittler and R. Cipolla, Discriminative Learning and Recognition of Image Set Classes Using Canonical Correlations, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.29, No.6, pp. 1005-1018, 2007
11. W. Fan, D.Y. Yeung. Locally linear models on face appearance manifolds with application to dual-subspace based classification. IEEE Conference on Computer Vision and Pattern Recognition, 2006.
12. R.P.Wang, Sh. G. Shan, X.L. Chen, W. Gao. Manifold-Manifold Distance with Application to Face Recognition based on Image Set. IEEE Conference on Computer Vision and Pattern Recognition, 2008.
13. H. Hotelling, Relations between two sets of variates, Biometrika, 28, pp. 321-372, 1936
14. S.C. Chen and Y. L Zhu, Subpattern-based principal component analysis, Pattern Recognition, 37(1): 1081-1083, 2004
15. R. E. Schapire and Y. Singer, Improved boosting algorithms using confidence-rated predictions, Machine Learning, Vol. 37, pp.297-336, 1999
16. T. Sim, S. Baker, and M. Bsat, The CMU Pose, Illumination, and Expression Database IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 12, pp. 1615 - 1618, 2003
17. A. S. Georghiades, P. N. Belhumeur and D. J. Kriegman. From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.23, No.6, pp:643-660, 2001
18. B. Schlkopf and A.J. Smola: Learning with Kernels. MIT Press, Cambridge, MA, USA, 2002