Boosting Constrained Mutual Subspace Method for Robust Image-Set Based Object Recognition

Xi Li
Xi'an Jiaotong University, China
lxconfxian@gmail.com

Kazuhiro Fukui
Tsukuba University, Japan
kf@cs.tsukuba.ac.jp

Nanning Zheng
Xi'an Jiaotong University, China
znn@xjtu.edu.cn

Abstract

Object 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 snap-shot about the variability in the appearance of the target subject. Constrained Mutual Subspace Method (CMSM) is one of the state-of-the-art algorithms for image-set based object recognition by first projecting the image-set patterns onto the so-called generalized difference subspace then classifying based on the principal angle based mutual subspace distance. By treating the subspace bases for each image-set patterns as basic elements in the grassmann manifold, this paper presents a framework for robust image-set based recognition by CMSM-based ensemble learning in a boosting way. The proposed Boosting Constrained Mutual Subspace Method(BCMSM) improves the original CMSM in the following ways: a) The proposed BCMSM algorithm is insensitive to the dimension of the generalized differnce subspace while the performance of the original CMSM algorithm is quite dependent on the dimension and the selecting of optimum choice is quite empirical and case-dependent; b) By taking advantage of both boosting and CMSM techniques, the generalization ability is improved and much higher classification performance can be achieved. Extensive experiments on real-life data sets (two face recognition tasks and one 3D object category classification task) show that the proposed method outperforms the previous state-of-the-art algorithms greatly in terms of classification accuracy.

1 Introduction

Recently object/face recognition with image-set has attracted more and more attention within computer vision and pattern recognition communities. Compared with single snapshot, a set or a sequence of images provides much more information about the variability in the appearance of the target subject. The variability always exist in the context of object category classification, visual surveillance or face recognition applications, where multiple-viewpoint shots of target object or surveillance system output over a period of time with varying illumination conditions or changing facial expressions are available. Previous studies show that more robust object/face recognition performance can be achieved by fully exploiting these kind of information[Shakhnarovich *et al.*, 2002; Yamaguchi *et al.*, 1998; Wolf and Shashua, 2003; Fukui and Yamaguchi, 2003; Fukui *et al.*, 2006]. Fig1 shows the superiority of image-set based recognition to the traditional method using a single snap-shot of image pattern.

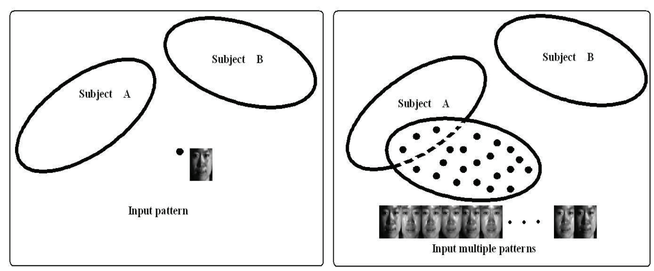

Figure 1: Left: The traditional recognition method using single input pattern; Right: Image-set based recognition using multiple patterns

It is well known that the appearance distribution of image-set or sequence for a target subject captured under complex conditions, for example changing facial expressions, varying illuminations or multiple viewpoints, can be approximately represented by a low dimensional linear subspace. The principal angles[Hotelling, 1936] between subspace pairs can be used as a distance measure of the corresponding image-set pairs. Many algorithms have been proposed for image-set based object recognition using the concept of principal angle based subspace distances. A noteworthy work is the Mutual Subspace Method(MSM) presented by Yamaguchi et al. [Yamaguchi *et al.*, 1998]. 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. K. Fukui et al. proposed the Constrained Mutual Subspace Method(CMSM)[Fukui and Yamaguchi, 2003] which outperforms the original MSM greatly. Instead of directly applying the classification using the principal angle based subspace distances, 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 thus a higher classification rate can be achieved. The above methods were further extended to their non-linear counterpart by using the kernel trick such as in [Wolf and Shashua, 2003; Fukui *et al.*, 2006]. These kernel based extensions improve the recognition performance at the expense of prohibitively computational burden and the difficulty of choosing optimum model and kernel parameters.

Recently ensemble-based learning with boosting has received more and more attention within pattern recognition and machine learning communities. Boosting is a classifier ensembling method and has successful applications such as in the face/object detection. The main idea of boosting is to sequentially learn a base classifier on a weighted version of the training sample set. The weight distribution of each sample is updated at each iteration in such a way that more emphasis are put on those samples that are misclassified in the preceding iteration. It has been proved theoretically and demonstrated experimentally that in case that the learned base weak learners perform slightly better that random guess, the final ensembled classifier can achieve very accurate result. The boosting algorithm is appealing for its ability of overfitting preventing and generalization error reducing. The boosting algorithm can use not only weak learners such as stump function but also some rather strong learners for example LDA(Linear Discriminant Analysis)-like algorithms[Masip and Vitria, 2006; Lu *et al.*, 2003; Dai and Yeung, 2007].

By treating the subspace bases for each image-set patterns as basic elements in the grassmann manifold, this paper presents a framework for robust image-set based recognition by CMSM-based ensemble learning in a boosting way. In this paper we will show that procedures of CMSM computation can be incorporated into the proposed boosting framework seamlessly. The proposed Boosting Constrained Mutual Subspace Method(BCMSM) improves the original CMSM in the following ways: a) The proposed BCMSM algorithm is insensitive to the dimension of the generalized difference subspace while the performance of the original CMSM algorithm is quite dependent on the dimension and the selecting of optimum choice is quite empirical and case-dependent; b) By taking advantage of both boosting and CMSM techniques, the generalization ability is improved and much higher classification performance can be achieved. Extensive experiments on real-life data sets show that the proposed method outperforms the previous state-of-the-art algorithms greatly in terms of classification accuracy.

The rest of this paper is organized as follows: In section 2 we first review the concept of the principal angles based distance metric between subspace bases for the corresponding image-sets. Then some previous principal angles based image-set classification methods and their drawbacks are discussed, especially the Constrained Mutual Subspace Method(CMSM), which is one of the state-of-the-art image-set based recognition algorithms and also the basic learner in the proposed boosting framework. Section 3 describes the proposed method in detail and we will show how to boost the performance of the CMSM in an ensemble learning way. Section 4 is the experimental results and section 5 draws the conclusion.

2 Image-set based recognition using principal angles

We first define the problem of image-set based recognition as follows: Given C classes n input image-sets $X_j^i \in R^{rc\times h_j^i}, i = 1,...,C, j = 1,...,n_i, \sum_{i=1}^{C} n_i = n$ and the X_j^i denotes the j-th image-set instance for class i. r and c represent the number of rows and columns respectively for the original image in image-set and rc is the dimension of the vectorized image. h_j^i is the number of images in the j-th image-set or length of the j-th sequence for the i-th class. The collection of image-set instances for each class approximately reside in a K_i-dimensional subspace $\Omega_i \in R^{rc\times K_i}$ while each input image-set has its own underlying k_j^i-dimensional subspace structure denoted as $U_j^i \in R^{rc\times k_j^i}, i = 1,...,C, j = 1,...,n_i$, which can approximately describe the variations in the appearance caused by different illuminations, varying poses and changing facial expressions et al.[1] For a test image-set $X_{test} \in R^{rc\times h}$, the task is to predict its corresponding identity $y_{test} \in \{1,...,C\}$.

2.1 Principal angles between subspace pairs

The concept of principal angles between linear subspaces was first presented in the 19-th century and then translated into the statistics as canonical correlations[Hotelling, 1936]. Its successful applications include data analysis, random process, stochastic realization and system identification. Recently in computer vision community, the concept of principal angles is used as a distance measure for matching two image-set pairs or sequence pairs, where each of them can be approximated by a linear subspace[Yamaguchi *et al.*, 1998]. If the principal angles between two subspace pairs are small enough, then the corresponding image-set pairs from which subspace pairs derived are considered similar. Generally, let U_A,U_B represent two k-dimensional linear subspaces. The principal angles $0 \le \theta_1 \le \theta_2 \le ... \le \theta_k \le \pi/2$ between the two subspaces can be uniquely defined as[Yamaguchi *et al.*, 1998]:

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

$u_i^A \in U_A, u_i^B \in U_B$.

Denote $d(X_A, X_B) = d(U_A, U_B) = f(\theta_1, \theta_2, ..., \theta_k)$ 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 and different methods for image-set based recognition has different empirical form of the function f.

[1] In the following discussion, for the simplicity we assume that each image-set instance has the same number of frames which is represented as h. The subspace for each image-set pattern has the same dimension of k and the approximate subspace for each class has the same dimension of K.

2.2 Image-set based recognition using principal angles

The original baseline Mutual Subspace Method(MSM) 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 $d(\cdot,\cdot) = \theta_1$[Yamaguchi *et al.*, 1998].

K. Fukui et al. proposed the Constrained Mutual Subspace Method(CMSM) to further improve the performance of the original MSM by feature extraction. Specifically, the Constrained Mutual Subspace Method(CMSM) projects the original feature onto the eigenvectors of $G = \sum_{i=1}^{C} P_i = \sum_{i=1}^{C} \Omega_i {\Omega_i}^T$ corresponding to the d smallest eigenvalues, where G is the sum of the projection matrix of all classes. The space spanned by those eigenvectors is defined as the generalized difference subspace[Fukui and Yamaguchi, 2003]. Projecting the image pattern on the generalized difference subspace has the effect of expanding the principal angle based distances between each class and making the subspace bases for each class as orthogonal as possible. Thus higher classification rate can be achieved. The generalized difference subspace includes only the essential component for recognition task since it is orthogonal to the principal component subspace, which represents the intersections of reference subspaces of each class. The distance between the subspace bases of the corresponding transformed image-set pairs is defined as the mean of the square of the first t smallest principal angles as $d(\cdot,\cdot) = \frac{1}{t}\sum_{i=1}^{t} \theta_t^2$. More specifically, the CMSM algorithm can be summarized as follows:

Algorithm 1: **Constrained Mutual Subspace Method(CMSM)**:

Training:

1) Compute sum of the projection matrix of all classes as $G = \sum_{i=1}^{C} P_i = \sum_{i=1}^{C} \Omega_i {\Omega_i}^T$ and eigendecompose G;

2) Output projection matrix W as the eigenvectors of G corresponding to the first d smallest eigenvalues;

3) Project training image-sets using projection matrix W as $\hat{X}_j^i = W^T X_j^i, i = 1, ..., C, j = 1, ..., n_i$

Testing:

1) Project testing image-set using projection matrix W as $\hat{X}_{test} = W^T X_{test}$;

2) Classify using the projected training and testing image-set patterns using baseline Mutual Subspace Method(MSM).

Although CMSM algorithm outperforms the original Mutual Subspace Method(MSM) greatly and is one of the state-of-the-art image-set based recognition methods, it still suffers some drawbacks:

a) CMSM performs the principal angle based classification after projecting the input patterns onto the generalized difference subspace which represents the subspace orthogonal to the common subspace of the subspace bases for all input classes. But there does not exist a theoretical way of choosing the dimension of the generalized difference subspace. In fact, the dimension of the generalized difference subspace has great effect on the classification performance in terms of recognition rate and the selecting of optimum choice of the dimension is quite empirical and case-dependent. For example, Fig2 demonstrates a typical curve of recognition rate vs. the dimension of the generalized difference subspace.

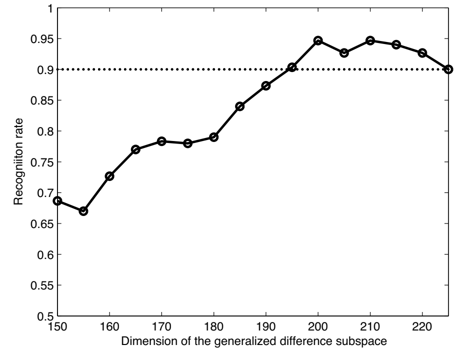

Figure 2: A typical curve of recognition rate vs. the dimension of the generalized difference subspace. The dimension of the original feature is 225 so when the dimension of the generalized difference subspace equals to 225 the Constrained Mutual Subspace Method(CMSM)(Solid line) degenerates to the original Mutual Subspace Method(MSM)(Dashed line)

b) More severely, CMSM assumes that the distribution of each class can be represented using a simple linear subspace, while in the real life applications the actual distributions are usually highly complex and reside in a nonlinear manifold. Thus CMSM's linear subspace assumption limits its generalization ability and has severe overfitting problem under certain circumstances. Instead, in this paper we treat the subspace bases for each image-set patterns in a specific class as points in the grassmann manifold, and compute the representative prototype as weighted Karcher mean[Boothby, 2002; Begelfor and Werman, 2006]. The constructed representative prototype emphasizes different region of the nonlinear manifold according to the weight distribution. This will be discussed in detail in the next section.

3 Boosting Constrained Mutual Subspace Method(BCMSM)

Just as discussed above, the CMSM algorithm has its own drawbacks and in this section, by treating the subspace base for each image-set instance as a point in the grassmann manifold space, a framework for robust image-set based object recognition by CMSM ensemble learning in a boosting way is presented.

Since our proposed method is based on the algorithm of adaboost.M2, we first review the adaboost.M2 as follows. Adaboost.M2 is a multiclass extension of the original binary adaboost.M1 algorithm and aims to extend the communication between the boosting algorithm and the corresponding basic learner, which is the projection matrix learned by CMSM in the proposed method. The adaboost.M2 allows the basic learner to generate more expressive hypotheses indicating a degree of belief. For example, the hypotheses output a continuous confidence of each label rather than a single label. More specifically, a hypothesis h takes a sample y and a class label c as the inputs and produces a plausibility score $h(y, c) \in [0, 1]$ as the output. Furthermore, a more sophisticated error measure pseudo-loss ϵ_t is introduced with respect to the so-called mislabel distribution $D_t(n, c)$, where t is the boosting step index. (n, c) is a mislabel pair where n is the index of a training sample and c is an incorrect label associated with the sampling n. let B denotes the set of all mislabels as follows:

$$B = (n, c) : n = 1, ..., N, c \neq c_n \quad (1)$$

A mislabel distribution $D_t(n, c)$ is a distribution defined over the set B of all mislabels.

At each boosting step t, the pseudo-loss ϵ_t of the hypothesis h_t with respect to $D_t(n, c)$ can be computed as follows:

$$\epsilon_t = \frac{1}{2} \sum_{(n,c)\in B} D_t(n, c)(1 - h_t(y_n, c_n) + h_t(y_n, c)) \quad (2)$$

By manipulating the mislabel distribution the resulting adaboost.M2 boosting algorithm can focus not only on the hard-to-classify examples, but more specifically also on the incorrect labels that are hardest to discriminate. For more detail of adaboost.M2 algorithm refer to literature[Freund and R.E.Schapire, 1996].

From the mislabel distribution $D_t(n, c)$ we can also derive the pseudo sample weight and pseudo class weight respectively as follows:

$$w_t^{sample}(n) = \sum_{c\neq c_n} D_t(n, c) \quad (3)$$

$$w_t^{class}(c) = \sum_{n\neq c_n} D_t(n, c) \quad (4)$$

Pseudo sample weight and pseudo class weight describe the difficult extent of correctly classifying for a specific sample or for a specific class, respectively. Using the above definitions, we next discuss how to construct the Constrained Mutual Subspace Method(CMSM) based basic learner under the corresponding weights on each round of boosting. And how to incorporate the CMSM into the framework of ensemble learning based boosting seamlessly. Intuitively, The pseudo weights can effect the CMSM computation in the following two ways:

1)As discussed above, a direct linear subspace representation for each class can not describe the pattern distribution faithfully. On each boosting round, we should focus more on those hard-to-classify image-set samples patterns, that is to say, those hard-to-classify samples should have more contribution to the construction of the prototype subspace base for the corresponding class, which will be fed into the CMSM procedure as input of computing the generalized difference subspace. The k-dimensional real vector subspace bases of the rc-dimensional vector space R^{rc} can be represented as points of grassmann manifold denoted as $G(k, rc)$[Boothby, 2002]. Similar to the Euclidean spaces, the Karcher mean of points on the grassmannian manifold can be defined as:

$$\tilde{\Omega} = argmin_U \sum_{n=1}^{N} d^2(U_n, U) \quad (5)$$

which minimizes the sum of squared grassmann distance, where $d(\cdot, \cdot)$is the principal angle based geodesic distance metric between two subspaces. It has been shown that unique Karcher mean exists if distributions are limited to a sufficiently small region of grassmann manifold, which is usually satisfied in the object recognition scenarios. In this paper we compute the Karcher mean by adopting the method proposed in[Begelfor and Werman, 2006], which is described briefly as follows:

Algorithm 2: Compute Karcher Mean on Grassmann Manifold:
Input: Points $U_1, U2, ..., U_N \in G(k, rc)$
Output: Karcher mean $\tilde{\Omega}$
Procedure:
1) Initialize $\tilde{\Omega} = U_1$
2) Compute $W = \frac{1}{N} \sum_{n=1}^{N} Log_{\tilde{\Omega}}(U_n)$
where Log denotes log map on grassmann manifold[Begelfor and Werman, 2006].
3) if $||W||$ is small enough, return $\tilde{\Omega}$, else
4) Singular Value Decomposition W as $W = USV^T$
and update $\tilde{\Omega} = \tilde{\Omega}V \cos(S) + U \sin(S)$
5) iterate till convergence reached.

Here we extend the concept of Karcher mean to its weighting version. The extension is straightforward and for each class we define its weighted Karcher mean as follows:

$$\tilde{\Omega}_i = argmin_U \frac{1}{\sum_{j=1}^{N_i} w_t^{sample}(n)} \sum_{j=1}^{N_i} w_t^{sample}(n) * d^2(U_j^i, U) \quad (6)$$

$n = \sum_{l=1}^{i-1} N_l + j$ denotes the index for the j-th image-set instance in the i-th class. The weighted Karcher mean can be computed easily based on **Algorithm 2** by changing step 2 to its weighted counterpart.

2)We should focus not only on the hard-to-classify samples, but also on the incorrect labels that are hardest to discriminate. The pseudo class weight can be incorporated into the CMSM as follows. When computing the the generalized difference subspace we put more weights on those classes which is harder to be correctly classified during the preceding boosting steps

$$G = \sum_{i=1}^{C} d_t^{class}(i) \tilde{\Omega}_i * \tilde{\Omega}_i^T \quad (7)$$

Next we eigen-decompose the matrix G and the eigenvectors corresponding to the d smallest eigenvaules are chosen as the generalized difference subspace, which forms the projection matrix of the CMSM and is used as the base learner to generate hypothesis.

The propose Boosting Constrained Mutual Subspace Method(BCMSM) can be summarized in detail as follows:

Algorithm 3 : Boosting Constrained Mutual Subspace Method(BCMSM)
Input: C classes n input training image-sets $X_j^i \in R^{rc\times h}, i = 1, ..., C, j = 1, ..., n_i, \sum_{i=1}^{C} n_i = n$. X_j^i denotes the j-th image-set instance for class i. r and c represent the number of rows and columns respectively for the original image in image-set and rc is the dimension of the vectorized image; The CMSM-based base learner described in **Algorithm**

Table 1: Recognition rate comparison of different methods on PIE, FaceVD and ETH-80 dataset.

Recognition rate	PIE				FaceVD				ETH80			
Training set number	3	4	5	6	3	4	5	6	3	4	5	6
MSM[Yamaguchi *et al.*, 1998]	87.6	89.6	90.7	91.2	80.7	81.4	85.1	85.5	80.5	81.5	83.5	83.5
KMSM[Wolf and Shashua, 2003]	91.7	90.3	92.9	92.5	83.1	85.2	84.9	86.3	86.5	86.3	88.5	91.3
CMSM[Fukui and Yamaguchi, 2003]	91.2	92.1	92.5	93.3	84.9	87.2	88.6	88.7	84.5	87.0	89.5	90.3
The proposed BCMSM	**95.0**	**95.3**	**97.6**	**98.3**	**98.5**	**99.1**	**99.7**	**99.9**	**92.5**	**94.0**	**95.0**	**95.5**

1; The number of maximum iterations T.
Initialization:
$D_1(n,c) = \frac{1}{N(C-1)}, D_1(n,c_n) = 0$, Compute pseudo sample weight d_1^{sample} and pseudo class weight d_1^{class} using equation 3 and 4 respectively. For each image-set instance X_j^i compute its corresponding subspace base U_j^i using principal component analysis.
Procedures:
For $t = 1, ..., T$,

1)For each class, compute the weighted Karcher mean using **Algorithm 2** based on the pseudo sample weight d_t^{sample};

2)Compute the weighted sum of the projection matrix G for each class using equation 7 based on the pseudo class weight d_t^{class};

3)CMSM based learning using **Algorithm 1**. Specifically, eigen-decompose the matrix G and the eigenvectors corresponding to the d smallest eigenvalues are chosen as the generalized difference subspace and projecting all the image-sets X_j^i onto the generalized difference subspace as $\tilde{X}_j^i$;

4)Get hypothesis $\{h_t(y_n, c) \in [0,1]\}$ by applying the nearest mean classifier, where $y_{\sum_{l=1}^{i-1} n_l + j} = \tilde{X}_j^i$;

5)Compute the pseuo-loss of the current hypothesis h_t using equation 2;

6)Update the mislabel distribution as follows:

$$D_{t+1}(n,c) = D_t(n,c)\beta_t^{\frac{1}{2}(1+h_t(y_n,c_n)-h_t(y_n,c))}$$

where $\beta_t = \epsilon_t/(1-\epsilon_t)$ and then normalize it to a distribution function as: $D_{t+1}(n,c) = \frac{D_{t+1}(n,c)}{\sum_n \sum_c D_{t+1}(n,c)}$

7) Update the pseudo sample weight and pseudo class weight using equation 3 and 4 respectively;

Output: The final strong hypothesis:

$$H(y) = argmax_c \sum_{t=1}^{T} (log\frac{1}{\beta_t}) h_t(y,c)$$

4 Experimental results

In this section we test the proposed image-set based object recognition method using real life datasets, which include the PIE face dataset[Sim *et al.*, 2003], a self-collected face video database(FaceVD) and the ETH-80 multiview 3D object database[Geusebroek *et al.*, 2005]. We compare the performance of the following algorithms: 1) The baseline Mutual Subspace Method(MSM)[Yamaguchi *et al.*, 1998]; 2) The Kernel Mutual Subspace Method(KMSM)[Wolf and Shashua, 2003]; 3) The Constrained Mutual Subspace Method(CMSM)[Fukui and Yamaguchi, 2003]; 4) The proposed Boosting Constrained Mutual Subspace Method(BCMSM);

The linear subspace of each image-set is 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 set corresponding to the best recognition result for each training and testing round. For KMSM, a six-degree polynomial kernel was used as in[Wolf and Shashua, 2003]. For the proposed boosting constrained mutual subspace method(BCMSM), the maximum iteration number was set to be 25.

• 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 15×15 in pixels. The 160 images are divided into 10 image-sets with each image-set has 16 images with different illumination and pose conditions.

• To further illustrate the performance of the proposed method, we self-collected a face video database(FaceVD) which has 20 subjects and for each subject 10 separate sequences were recorded with large variation in illumination and facial expression. 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 16×16 in pixels.

• For the category classification task we use the public multiview 3D object database—ETH-80 dataset to evaluate the proposed method. This data set has 8 categories (e.g. Cow,Horse,Dog,Cup,Apple,Pearl,Peach and Car) and each category has 10 objects. For each object images taken from 41 different view points was collected and forms the image-set pattern.

For each dataset, we randomly partition 10 image-sets/sequences of each subject/category into training set which include $3, 4, 5, 6$ objects and the remaining for the test set. The random partition procedure repeated for 10 times and the average classification results are computed. The comparison of recognition performance for the above experiments are demonstrated in Table1. It can be seen that the proposed method of recognition using Boosting Constrained Mutual Subspace Method(BCMSM) outperforms the MSM,KMSM and CMSM greatly and consistently. The proposed BCMSM algorithm ensembles a set of CMSM-based basic learner in a boosting way while each learner is constructed with emphasis on certain region of the nonlinear distribution. The boosting procedure is illustrated in Fig 3. As expected, the training error and testing error decrease which consistent with the boosting theory. And the selection of the dimension of generalized difference subspace for the basic CMSM based learner has little effect on the recognition performance. Refer the caption for detail.

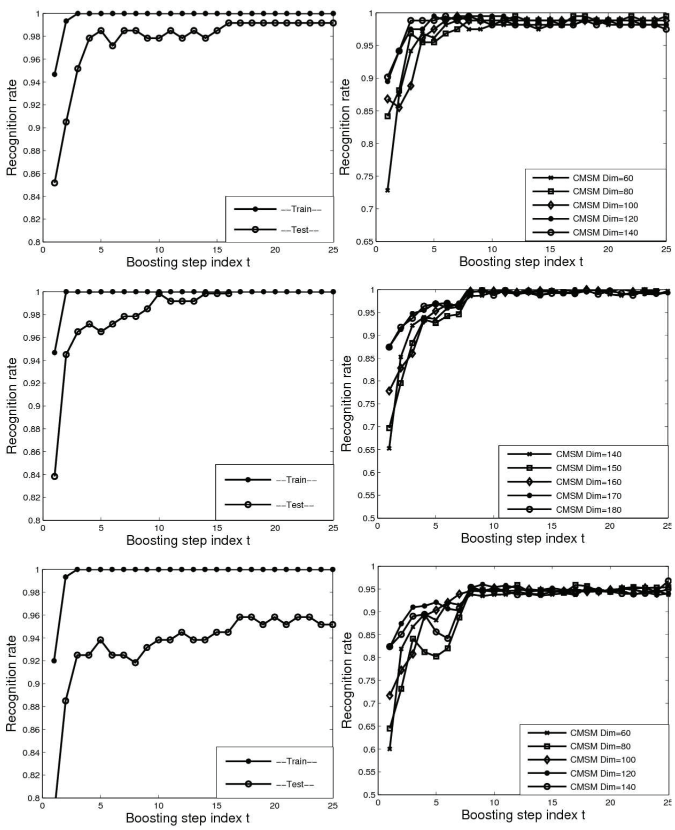

Figure 3: Left column: the training and testing recognition rate vs. boosting step index, for PIE, FaceVD and ETH-80 datasets respectively. Right column: The testing recognition rate with different dimension of generalized difference subspace(CMSM Dim) vs. boosting step index. It can be clearly seen that the proposed BCMSM is insensitive to the dimension of of generalized difference subspace for the basic learner.

5 Conclusion

By treating the subspace base for each image-set instance as a point in the grassmann manifold space, this paper presents a framework for robust image-set based object recognition by CMSM based ensemble learning in a boosting way. The proposed Boosting Constrained Mutual Subspace Method(BCMCM) is insensitive to the dimension of the generalized difference subspace, and more importantly, by taking advantage of both boosting and CMSM techniques, the generalization ability is improved and higher classification performance can be achieved. Extensive experiments on real life data sets demonstrate the superiority of the proposed method.

References

[Begelfor and Werman, 2006] E. Begelfor and M. Werman. Affine invariance revisited. In *Proceedings of International Conference on Computer Vision and Pattern Recognition*, pages 2087–2094, 2006.

[Boothby, 2002] W.M. Boothby. *An introduction to differentiable Manifolds and Riemannian Geometry*. Academic Press, 2002.

[Dai and Yeung, 2007] G. Dai and Dit-Yan Yeung. Boosting kernel discriminant analysis and its application on tissue classification of gene expression data. In *Proceedings of International Joint Conference on Artificial Intelligence*, pages 744–749, 2007.

[Freund and R.E.Schapire, 1996] Y. Freund and R.E.Schapire. Experiments with a new boosting algorithm. In *Proceedings of International Confernce on Machine Learning*, pages 148–156, 1996.

[Fukui and Yamaguchi, 2003] K. Fukui and O. Yamaguchi. Face recognition using multi-viewpoint patterns for robot vision. In *Proceedings of 11th International Symposium of Robotics Research*, pages 192–201, 2003.

[Fukui *et al.*, 2006] K. Fukui, B. Stenger, and O. Yamaguchi. A framework for 3d object recognition using the kernel constrained mutual subspace method. In *Proceedings of Asain Conference on Computer Vision*, pages 315–324, 2006.

[Geusebroek *et al.*, 2005] J.M. Geusebroek, G.J. Burghouts, , and A.W.M. Smeulders. The amsterdam library of object images. *International Journal of Computer Vision*, 61(1):103–112, 2005.

[Hotelling, 1936] H. Hotelling. Relations between two sets of variates. In *Biometrika*, pages 321–372, 1936.

[Lu *et al.*, 2003] J.W. Lu, K.N. Plataniotis, and A.N. Venetsanopoulos. Boosting linear discriminant analysis for face recognition. In *Proceedings of International Conference on Image Processing*, pages 657–660, 2003.

[Masip and Vitria, 2006] D. Masip and J. Vitria. Boosted discriminant projections for nearest neighbour classification. *Pattern Recognition*, 39(2):174–170, 2006.

[Shakhnarovich *et al.*, 2002] Gregory Shakhnarovich, John W. Fisher, and Trevor Darrell. Face recognition from long-term observations. In *Proceedings of the 7th European Conference on Computer Vision*, pages 851–865, September 2002.

[Sim *et al.*, 2003] T. Sim, S. Baker, and M. Bsat. The cmu pose, illumination, and expression database. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(12):1615–1618, 2003.

[Wolf and Shashua, 2003] L. Wolf and A. Shashua. Learning over sets using kernel principal angles. *Journal of Machine Learning Research*, 4(10):913–931, 2003.

[Yamaguchi *et al.*, 1998] O. Yamaguchi, K. Fukui, and K. Maeda. Face recognition using temporal image sequence. In *Proceedings of the 3rd International Conference on Automatic Face and Gesture Recognition*, pages 318–323, 1998.