Compound Mutual Subspace Method for 3D object recognition: A theoretical extension of Mutual Subspace Method

Naoki Akihiro, and Kazuhiro Fukui

Graduate School of Systems and Information Engineering,
University of Tsukuba, Japan

Abstract. In this paper, we propose the Compound Mutual Subspace Method (CPMSM) as a theoretical extension of the Mutual Subspace Method, which can efficiently handle multiple sets of patterns by representing them as subspaces. The proposed method is based on the observation that there are two types of subspace perturbations. One type is due to variations within a class and is therefore defined as "within-class subspace". The other type, named "between-class subspace", is characterized by differences between two classes. Our key idea for CPMSM is to suppress within-class subspace perturbations while emphasizing between-class subspace perturbations in measuring the similarity between two subspaces. The validity of CPMSM is demonstrated through an evaluation experiment using face images from the public database VidTIMIT.

1 Introduction

In this paper, we propose the Compound Mutual Subspace Method (CPMSM), which has the ability to classify similar sets of patterns accurately. Then we apply it in a face recognition experiment based on multiple images.

Subspace-based methods have recently attracted attention from many researchers who are interested in recognition of 3D objects, such as faces. The mutual subspace method (MSM)[1] is one of the most effective and efficient methods for object recognition, as it can efficiently handle multiple images[2][3][4]. In subspace-based methods, including MSM, a pattern composed of $n \times n$ pixels is usually regarded as a vector $\mathbf{x}$ in n^2-dimensional space. MSM represents a set of patterns $\{\mathbf{x}\}$ from each class through a low-dimensional linear subspace generated from the set by using the Karhunen-Loève (KL) expansion, which is also known as principal component analysis (PCA). Finally, the similarity between two sets of patterns can be readily measured by using canonical angles θ_i between two subspaces, as shown in Fig.1.

Even though MSM is capable of absorbing differences in appearance caused by changes in view point or illumination, compared with conventional methods using a single input image, such as the subspace method[5], the classification performance of MSM is still not sufficiently high. One reason for this is that a

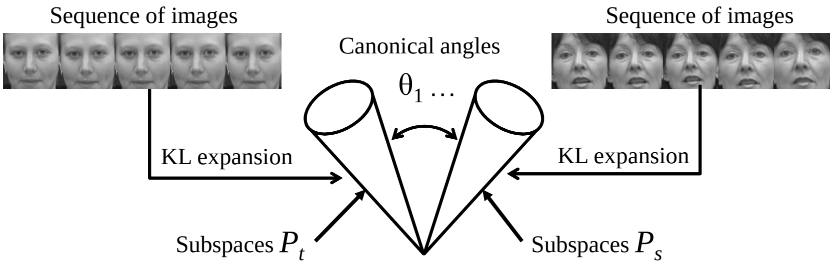


Fig. 1. Concept of Mutual Subspace Method.

subspace which provides a satisfactory representation of the distribution of the training patterns in terms of a least-mean-square approximation is not always optimal in terms of classification performance. Many extended methods have been proposed[6][7][8] for improving the classification performance of MSM, including the nonlinear extensions[9][10][11][12] using a kernel trick. In this paper, we focus on the Constrained MSM (CMSM) and the Orthogonal MSM (OMSM)[13] since they have been used in the development of the recognition engine of the state-of-the-art face recognition system "FacePass" and have achieved extremely high scores in the Face Recognition Vendor Test (FRVT) 2006[14].

The essence of these methods is to apply MSM to sets of class subspaces which have been orthogonalized with respect to each other in advance. The implementation of orthogonalization is different in the two methods. In OMSM, all the class subspaces are orthogonalized by using the Fukunaga-Koontz framework[15]. The kernel OMSM executes this operation in extremely high-dimensional feature space in order to ensure complete orthogonalization. CMSM achieves approximate orthogonalization of all the class subspaces by projecting them onto the generalized difference subspace $\mathcal{D}$. The kernel CMSM executes the projection in a high-dimensional feature space.

In this paper, we also aim to improve the performance of MSM by introducing the concept of "difference subspace" between two subspaces. This approach is notably different from the orthogonalization operation used in CMSM and OMSM. Our approach is based on the observation that there are two types of subspace perturbations. One type occurs due to differences within a class, while the other is due to differences between separate classes. In this paper, we refer to the former as "within-class subspace $\mathcal{D}_W$" and the latter as "between-class subspace $\mathcal{D}_B$".

It should be noted that MSM does not distinguish within-class subspace perturbations from between-class subspace perturbations. Thus, MSM cannot distinguish an input subspace between a subspace of a rival class and a subspace of the same class when they have the same canonical angles as a similarity to the input subspace.

This leads us to develop a proper strategy for suppressing within-class subspace perturbations while emphasizing between-class subspace perturbations.

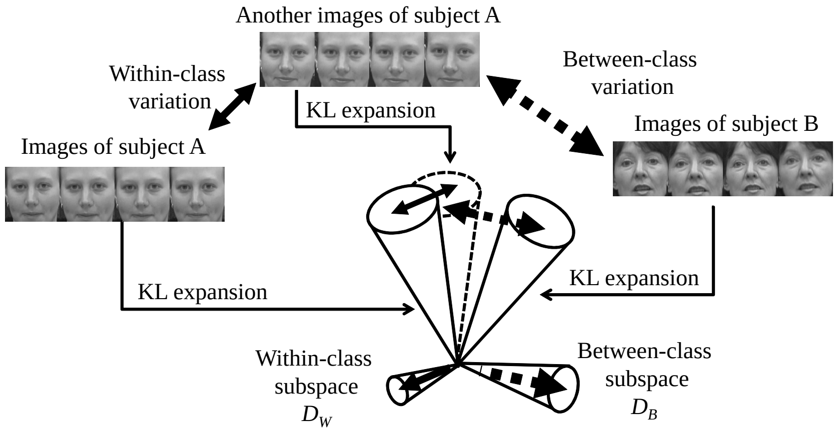


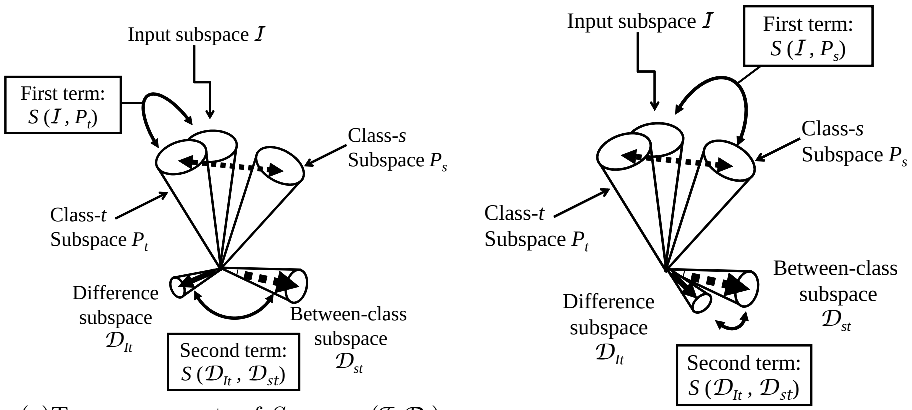
Fig. 2. Two types of difference subspaces.

To realize such a strategy, we introduce the concept of "difference subspace" between two subspaces. The concept of difference subspace is a natural extension of the difference vector between two vectors. We can obtain a within-class subspace $\mathcal{D}_W$ as the difference subspace between two subspaces of the same class, as shown in Fig.2. On the other hand, we can obtain a between-class subspace $\mathcal{D}_B$ as the difference subspace between subspaces belonging to different classes.

The essence of the proposed method is to classify a difference subspace $\mathcal{D}_I$ between an unknown input subspace $\mathcal{I}$ and a class-t subspace $\mathcal{P}_t$ into one of two types of subspaces $\mathcal{D}_W$ and $\mathcal{D}_B$ by using canonical angles. The similarity obtained through this classification is used to correct the similarity obtained with MSM. We refer to the MSM which takes into account $\mathcal{D}_W$ and $\mathcal{D}_B$ utilizing difference subspaces as the "Compound Mutual Subspace Method" (CPMSM).

The advantage of the proposed method is that it can be applied only to limited pairs of class subspaces which are too close and can be easily misclassified. This restriction can reduce the computation time as compared to both CMSM, which projects all class subspaces onto the constraint subspace, and OMSM, which performs orthogonalization of all class subspaces. In addition, the proposed method can be used as post-processing for existing methods, such as MSM, CMSM, and OMSM. Here, we evaluate CPMSM by applying it to a face recognition experiment using a public database containing face images (VidTIMIT audio-video database) [17].

This paper is organized as follows. In Section 2, we explain the concept behind the proposed method and describe the algorithm of CPMSM. In Section 3, the effectiveness of our method is demonstrated through evaluation experiments using a public database containing face images. Finally, Section 4 concludes the paper.



(a)Two components of $S_{CPMSM}(\mathcal{I}, \mathcal{P}_t)$ (b)Two components of $S_{CPMSM}(\mathcal{I}, \mathcal{P}_s)$

Fig. 3. Similarity of the input subspace $\mathcal{I}$ to each class subspace. This figure shows the case that the input subspace belongs to class t, (a) the terms of the similarity to $\mathcal{P}_t$, (b) the terms of the similarity to $\mathcal{P}_s$.

2 Compound Mutual Subspace Method (CPMSM)

In this section, we first explain the basic principle of CPMSM. Then, we define a new similarity for CPMSM based on the concept of difference subspace.

2.1 The basic principle of CPMSM

The basic principle of CPMSM can be explained as follows. When the difference subspace $\mathcal{D}_I$ between $\mathcal{I}$ and the class-t subspace $\mathcal{P}_t$ is similar to the between-class subspace $\mathcal{D}_B$ and dissimilar to the within-class subspace $\mathcal{D}_W$, the input subspace $\mathcal{I}$ should be classified into class t. On the other hand, when $\mathcal{D}_I$ is similar to the within-class subspace $\mathcal{D}_W$ and dissimilar to $\mathcal{D}_B$, the input subspace $\mathcal{I}$ can be considered to belong to some similar rival class rather than to the class-t subspace $\mathcal{P}_t$. The similarity between difference subspaces can be measured by using canonical angles since a difference subspace is a linear subspace, as will be mentioned later.

In practical calculation of the similarity, it is only necessary to measure the similarity between the subspaces $\mathcal{D}_I$ and $\mathcal{D}_B$ since $\mathcal{D}_I$ is projected onto an orthogonal complement of $\mathcal{P}_t$ in such a way that the projected $\mathcal{D}_I$ has no components belonging to the within-class subspace $\mathcal{D}_W$.

2.2 Calculation of similarity in CPMSM

The similarity S_{CPMSM} consists of two terms, as follows:

$$S_{CPMSM}(\mathcal{I}, \mathcal{P}_t) = (1 - \mu)S(\mathcal{I}, \mathcal{P}_t) - \mu S(\mathcal{D}_{It}, \mathcal{D}_{st}) \ , \tag{1}$$

where μ is a weighting parameter which should be determined experimentally.

In the above equation, the first term $S(\mathcal{I}, \mathcal{P}_t)$ indicates the similarity between the input subspace $\mathcal{I}$ and the class-t subspace $\mathcal{P}_t$. This similarity is obtained by using MSM. The second term $S(\mathcal{D}_{It}, \mathcal{D}_{st})$ is the regulation term, which can be obtained as the similarity between two difference subspaces $\mathcal{D}_{It}$ and $\mathcal{D}_{st}$, where $\mathcal{D}_{st}$ is the difference subspace between the subspace of class t and that of its similar rival class s.

In the following paragraphs, we will explain how to apply the above similarity to the task of classifying an input subspace into one of two similar classes, subspace $\mathcal{P}_t$ and $\mathcal{P}_s$, by using Fig. 3. In this case, we can obtain the following two similarities for the input subspace.

$$S_{CPMSM}(\mathcal{I}, \mathcal{P}_t) = (1-\mu)S(\mathcal{I}, \mathcal{P}_t) - \mu S(\mathcal{D}_{It}, \mathcal{D}_{st}) \ , \tag{2}$$

$$S_{CPMSM}(\mathcal{I}, \mathcal{P}_s) = (1-\mu)S(\mathcal{I}, \mathcal{P}_s) - \mu S(\mathcal{D}_{Is}, \mathcal{D}_{st}) \ , \tag{3}$$

where the former is the similarity for class t and the latter is that for class s. The input subspace is classified into the class with higher similarity.

The proposed idea of similarity shares common features with the method used in Bayesian face recognition[18] in that it is based on the analysis of image differences, that is, a difference vector between two image pattern vectors. However, that method can not handle complex situations, such as the relation between two sets of image pattern vectors. In addition, a single image is used as an input in the Bayesian method.

2.3 Measure of similarity between two subspaces

The measure of similarity between two subspaces is defined through canonical angles. Assume that we have an N-dimensional subspace $\mathcal{P}_t$ and an M -dimensional subspace $\mathcal{P}_s$ (assume $N \leq M$ for convenience). In this case, we can obtain N canonical angles θ_i, $(i = 1{\sim}N)$ between $\mathcal{P}_t$ and $\mathcal{P}_s$ by solving the eigenvalue equation of the following matrix $\mathbf{S}$ [1]:

$$\mathbf{S}\mathbf{a} = \lambda\mathbf{a} \ \ . \tag{4}$$

$$\mathrm{S}_{ij} = \sum_{l=1}^{N}(\mathbf{\Phi}_i \cdot \mathbf{\Psi}_l)(\mathbf{\Psi}_l \cdot \mathbf{\Phi}_j) \ \ , \tag{5}$$

where $\mathbf{\Phi}_i$ and $\mathbf{\Psi}_i$ are the i-th orthonormal basis vectors that span subspace $\mathcal{P}_t$ and $\mathcal{P}_s$, respectively. The value of $cos^2\theta_i$ for the i-th smallest canonical angle θ_i is obtained as the i-th largest eigenvalue of the matrix $\mathbf{S}$. Finally, the measure of similarity between two subspaces is defined with n canonical angles as the following equation (this measure of similarity is used for MSM):

$$S[n] = \frac{1}{n}\sum_{i=1}^{n} cos^2\theta_i \ \ . \tag{6}$$

2.4 Definition of difference subspace

The difference subspace is considered a natural generalization of the difference vector between two vectors[16]. A difference subspace is spanned by a set of difference vectors $\mathbf{d}_i$ between canonical vectors, $\mathbf{u}_i$ and $\mathbf{v}_i$, which form the i-th canonical angle. The canonical vectors are calculated from the following equations:

$$\mathbf{u}_i = \sum_{l=1}^{N} \mathrm{a}_{kl} \mathbf{\Psi}_l \quad . \tag{7}$$

$$\mathbf{v}_i = \sum_{l=1}^{N} \mathrm{a}'_{kl} \mathbf{\Phi}_l \quad . \tag{8}$$

In the above equations, the coefficient a_{kl} is the l-th element of the k-th eigenvector $\mathbf{a}_k$, corresponding to the k-th smallest eigenvalue of matrix $\mathbf{S}$ in Eq.(4). Furthermore, the coefficient $\mathrm{a}^{'}_{kl}$ is the l-th element of the k-th eigenvector $\mathbf{a}^{'}_k$ of matrix $\mathbf{S}^{'}$, where $\mathrm{S}^{'}{}_{ij} = \sum_{l=1}^{M} (\mathbf{\Psi}_i \cdot \mathbf{\Phi}_l)(\mathbf{\Phi}_l \cdot \mathbf{\Psi}_j)$.

2.5 Flow of the classification process using similarity in CPMSM

The process of classifying an input image set by using CPMSM is given as follows.

- Learning
  - Apply KL expansion on classes s and t of training image sets to obtain the reference subspaces $\mathcal{P}_t$ and $\mathcal{P}_s$.
  - Obtain the difference subspace $\mathcal{D}_{st}$ by using Eqs. 7 and 8.
- Testing

  step 1
  - Apply KL expansion on input image set to obtain the input subspace $\mathcal{I}$.
  - Calculate the similarities $S(\mathcal{I}, \mathcal{P}_t)$ and $S(\mathcal{I}, \mathcal{P}_s)$ by using Eq. 6.

  step 2
  - Obtain the difference subspaces $\mathcal{D}_{Is}$ and $\mathcal{D}_{It}$ by using Eqs. 7 and 8.

  step 3
  - Calculate the similarities $S(\mathcal{D}_{Is}, \mathcal{D}_{st})$ and $S(\mathcal{D}_{It}, \mathcal{D}_{st})$ by using Eq. 6.

  step 4
  - Combine $S(\mathcal{I}, \mathcal{P}_t)$ with $S(\mathcal{D}_{It}, \mathcal{D}_{st})$ to obtain $S_{CPMSM}(\mathcal{I}, \mathcal{P}_t)$ in Eq. 2.
  - Combine $S(\mathcal{I}, \mathcal{P}_s)$ with $S(\mathcal{D}_{Is}, \mathcal{D}_{st})$ to obtain $S_{CPMSM}(\mathcal{I}, \mathcal{P}_s)$ in Eq. 3.
- Identification
  - Compare the obtained similarity $S_{CPMSM}(\mathcal{I}, \mathcal{P}_t)$ with $S_{CPMSM}(\mathcal{I}, \mathcal{P}_t)$. The input subspace is classified into the class which has higher similarity.

3 Validation of the proposed method by using a database containing face images

The proposed method was designed to distinguish classes that are difficult to distinguish with MSM. To demonstrate the validity of the proposed method, it is necessary to find such pairs in the data set in advance. For this purpose, we carried out a face recognition experiment using MSM and selected pairs that were frequently misclassified, after which we applied the proposed method to those pairs.

3.1 Setup of experiment for face recognition

We used face images from the VidTIMIT audio-video database [17]. This database contains face data for 43 subjects. Three sequences of images are available for each subject. In order to conduct a face recognition experiment, the face region was extracted from each of these images by using the face detection function distributed with OpenCV ver. 1.0. We carefully removed false positives and obtained 140 images for each sequence of images. These cropped face images were converted into 15×15 pixels grayscale images, and 225 dimensional vectors were obtained.

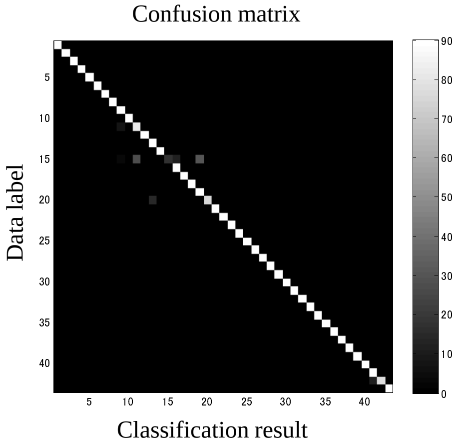
For the classification experiment in this paper, one sequence of images was used to prepare test data sets, and the others were used to prepare training data sets. Every third frame of the image sequence was used as a starting image of the test data set.

The parameters for the experiments of one class were set as follows. Number of training images used to generate reference subspace is 280. Number of testing images used to generate testing subspace is 30. Dimension of the reference subspace is 20. Dimension of the testing subspace is 7. Dimension of the between-class subspace is 20. Dimension of the difference subspace between the input subspace and either reference subspace is 7. Number of trials is 90.

3.2 Extraction of frequently misclassified pairs.

We conducted a classification experiment for all subjects contained in the database. To examine which input data is classified into which class, we constructed a confusion matrix. The confusion matrix is a table with a horizontal axis representing the results from the classifier and a vertical axis representing the labeled class. The classification frequency was plotted on this table.

The results from this experiment are plotted in Fig. 5. The color codes for the frequency are given in the legend on the right. From this confusion matrix, we can see that misclassification occurs only in certain specific similar pairs, namely, the six pairs that involve subjects No.11, No.15, No.20 and No.42, as shown in Fig.5. The total recognition rate for all 43 subjects was 97.2%, as shown in Fig.4. By contrast, the recognition rate of all subjects except the mentioned four subjects was 100%.



Summery of recognition rate using MSM

Classification Rate For All subjects	97.2 %
No.11 (fedw0)	91.1%
No.15 (fgjd0)	20.0%
No.20 (fjem0)	83.3%
No.42 (mstk0)	86.7%

* These 4 subjects are classified into not collect subject.

* The symbols in parentheses correspond to the subject in the database [17].

Fig. 4. Summary of classification with Mutual Subspace Method for 43 subjects

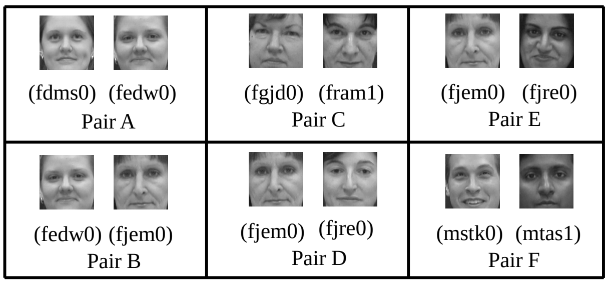


Fig. 5. Frequently misclassified pairs.

3.3 Classification results for frequently misclassified pairs

To evaluate the validity of CPMSM, we compared the performance of CPMSM with that of MSM and CMSM. These methods were applied in distinguishing between pairs as obtained in the previous section. To compare the performance of these methods, we used recognition rate and EER. EER is the error rate at the threshold value where the false accept rate (FAR) is equal to the false reject rate (FRR).

The performance of CPMSM depends on the weighting parameter μ in Eq.(3). We select the optimal value experimentally for each pair, as shown in Fig. 6. Note that when the weighting parameter μ is 0, CPMSM is equivalent to MSM.

From Tables 1 and 2, it can be seen that the recognition rate and EER in CPMSM have been improved in comparison to those in MSM for all pairs. The average recognition rate for all pairs increased from 0.950 to 0.959, and the

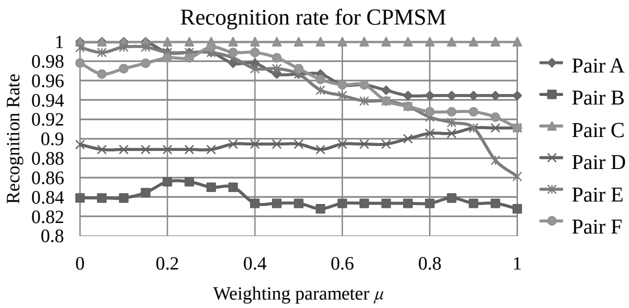


Fig. 6. Relation between recognition rate and μ. In the case of $\mu = 0$, CPMSM is equivalent to MSM.

Table 1. Recognition rate

Confused Pairs	CPMSM	CMSM	MSM
Pair A	1.0	0.961	1.0
Pair B	0.856	0.850	0.839
Pair C	1.0	0.961	1.0
Pair D	0.911	0.911	0.894
Pair E	0.994	0.989	0.994
Pair F	0.994	0.961	0.978
Average	0.959	0.939	0.950

Table 2. Equal Error Rate

Confused Pairs	CPMSM	CMSM	MSM
Pair A	0.078	0.103	0.217
Pair B	0.150	0.217	0.286
Pair C	0.006	0.067	0.156
Pair D	0.094	0.139	0.222
Pair E	0.106	0.072	0.211
Pair F	0.144	0.139	0.217
Average	0.096	0.123	0.218

average EER decreased from 0.218 to 0.096. From these results, we can confirm the validity of CPMSM and its ability to improve the performance of MSM.

4 Conclusions

In this paper, we have proposed the Compound Mutual Subspace Method (CPMSM) for face recognition. The advantage of CPMSM is its strong ability to distinguish between specific highly similar pairs among a large number of combinations of subjects. This characteristics can reduce the computation time and can improve the overall recognition rate by improving the performance with respect to a small number of pairs. The strong ability to distinguish between similar pairs was achieved by introducing a regulation term into the measure of similarity in MSM. The validity of the proposed method has been demonstrated through evaluation experiments with face images taken from the VidTIMIT public database.

References

1. Yamaguchi, O., Fukui, K., Maeda, K.: Face recognition using temporal image sequence. Proc. 3rd. Int. Conf. on Face & Gesture Recognition (1998) 318-323

2. Sakano, H., Suenaga, T.: Classifiers under continuous observation. Structural, Syntactic, and Statistical Pattern Recognition. **2396** (2009) 631-663
3. Faggian, N., Paplinski, A., Chin, T.-J.: Face Recognition From Video using Active Appearance Model Segmentation. 18th International Conference on Pattern Recognition **1** (2006) 287-290
4. Beveridge, J. R., Draper, B. A., Chang, J.-M., Kirby, M., Kley, H., Peterson, C.: Principal Angles Separate Subject Illumination Spaces in YDB and CMU-PIE. IEEE Trans. on Pattern Analysis and Machine Intelligence **31** (2009) 351-363
5. Watanabe, S., Pakvasa, N.: Subspace method of pattern recognition. Proc. 1st. Int. J. Conf. on Pattern Recognition (1973)
6. Li, X., Fukui, K., Zheng, N.: Boosting constrained mutual subspace method for robust image-set based object recognition. Proc. 21st. Int. Joint Conference on Artificial Intelligence (2009) 1132-1137
7. Kim, T.-K., Kittler, J., Cipolla, R.: Learning discriminative canonical correlations for object recognition with image sets. Computer Vision - ECCV 2006, **3953** (2006) 251-262
8. Chin, T.-J., Suter, D.:A new distance criterion for face recognition using image sets. Computer Vision - ACCV 2006 **3851** (2006) 549-558
9. Sakano, H., Mukawa, N.: Kernel Mutual Subspace Method for Robust Facial Image Recognition, Proc. 4th. Int. Conf. Knowledge-Based Intelligent Engineering Systems and Allied Technologies **1** (2000) 245-248
10. Wolf, L., Shashua, A.: Learning over Sets using Kernel Principal Angles. Journal of Machine Learning Research, **4** (2003) 913-931
11. Fukui, K., Yamaguchi, O.: The Kernel Orthogonal Mutual Subspace Method and its Application to 3D Object Recognition. Computer Vision - ACCV 2007 **4844** (2007) 467-476
12. Fukui, K., Stenger, B., Yamaguchi, O.: A Framework for 3D Object Recognition Using the Kernel Constrained Mutual Subspace Method. Computer Vision - ACCV 2006 **3852** (2006) 315-324
13. Kawahara, T., Nishiyama, M., Kozakaya, T., Yamaguchi, O.: Face Recognition based on Whitening Transformation of Distribution of Subspaces. Workshop on ACCV2007, Subspace2007 (2007) 97-103
14. Phillips, P., Scruggs, W., O 'Tools, A., Flynn, P., Bowyer, K., Schott, C., Sharpe, M.: FRVT 2006 and ICE 2006 Large-Scale Results. Technical Report NISTIR 7408, NIST, (2007)
15. Fukunaga, K., Koontz, W.L.G.: Applications of the Karhunen-Loève Expansion to Feature Selection and Ordering. IEEE Trans. Computers, **C-19** (1970) 311-318
16. Fukui, K., Yamaguchi, O.: Face Recognition Using Multi-viewpoint Patterns for Robot Vision. Robotics Research **15** (2005) 192-201
17. Sanderson, C.: Biometric Person Recognition: Face, Speech and Fusion. VDM-Verlag (2008)
18. Moghaddam, B., Jebara, T., Pentland, A.: Bayesian Face Recognition. J. of Pattern Recognition, **33** (2000) 1771-1782
19. Chatelin, F.: Eigenvalues of matrices. John Willey and Sons, Chichester (1993)
20. Basri, R., Jacobs, D.: Lambertian Reflectance and Linear Subspaces. IEEE Trans. on Pattern Analysis and Machine Intelligence (2003) 218-233