The Kernel Orthogonal Mutual Subspace Method and its Application to 3D Object Recognition

Kazuhiro Fukui[1] and Osamu Yamaguchi[2]

[1] Department of Computer Science, University of Tsukuba, Japan
[2] Corporate Research and Development Center, Toshiba corporation, Japan

Abstract. This paper proposes the kernel orthogonal mutual subspace method (KOMSM) for 3D object recognition. KOMSM is a kernel-based method for classifying sets of patterns such as video frames or multi-view images. It classifies objects based on the canonical angles between the nonlinear subspaces, which are generated from the image patterns of each object class by kernel PCA. This methodology has been introduced in the kernel mutual subspace method (KMSM). However, KOMSM is different from KMSM in that nonlinear class subspaces are orthogonalized based on the framework proposed by Fukunaga and Koontz before calculating the canonical angles. This orthogonalization provides a powerful feature extraction method for improving the performance of KMSM. The validity of KOMSM is demonstrated through experiments using face images and images from a public database.

1 Introduction

This paper introduces the kernel orthogonal mutual subspace method (KOMSM) for 3D object recognition. KOMSM is an appearance-based method for classifying a set of patterns such as a video frames or images obtained from a multi-camera system. As the set of such patterns generally has highly nonlinear structure, we have to tackle a nonlinear classification problem of multiple sets of patterns. The kernel mutual subspace method (KMSM) is one suitable method for this task. KMSM is a nonlinear extension of the mutual subspace method (MSM)[3] by using the kernel trick. MSM classifies sets of patterns based on the canonical angles θ between linear class subspaces, which represent the distribution of the training set of each class respectively as shown in Fig.1. In this method an w×h image pattern is represented as a vector in w×h-dimensional space (called input space $\mathcal{I}$). Although MSM has the ability to handle the variability of patterns to achieve higher performance compared to other methods[1], its performance drops significantly when the pattern distributions have highly nonlinear structure. In such cases class distributions cannot be represented by a linear subspace without overlapping each other. The kernel mutual subspace method (KMSM)[4, 5] has been proposed in order to solve this problem. In this method an input pattern $\mathbf{x}$ is mapped into a high (in some cases infinite) dimensional feature space $\mathcal{F}$ via a nonlinear map ϕ. Consequently, KMSM carries out

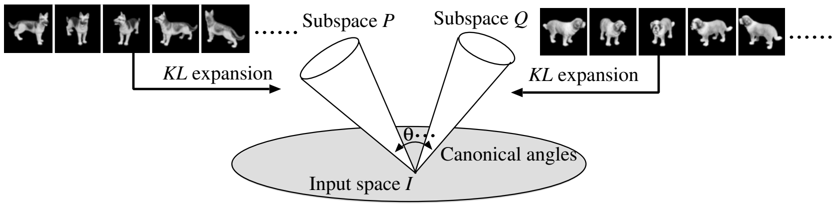

Fig. 1. Similarity between two distributions of view patterns.

the MSM on the linear subspaces [3]generated from the mapped patterns $\{\phi(\mathbf{x})\}$ using the Karhunen-Loève (KL) expansion, also known as principal component analysis (PCA).

KMSM works well since each subspace can be generated without overlapping with another subspace in the feature space $\mathcal{F}$. However its classification performance is still insufficient for many applications in practice, because the nonlinear class subspaces are generated independently of each other [1]. There is no reason to assume a priori that it is the optimal nonlinear class subspace in terms of classification performance, while each nonlinear class subspace represents the distribution of the training patterns well in terms of a least-mean-square approximation. This suggests that there is room for improving the performance of KMSM.

In order to improve the performance of KMSM the kernel constrained mutual subspace method (KCMSM)[11] has been proposed. In this method, each nonlinear class subspace is projected onto a discrimination space called the constraint subspace. This projection extracts a common subspace of all the nonlinear class subspaces from each nonlinear class subspace, so that the canonical angles between nonlinear class subspaces are enlarged to approach orthogonal relation. KCMSM could significantly improve the performance of KMSM in many applications, such as face recognition and 3D object recognition[11]. This implies that the concept of orthogonalization is essential for improving the performance of KMSM. Therefore, in this paper we apply the framework proposed by Fukunaga and Koontz[2], which achieves perfect orthogonalization, while the orthogonalization achieved in KCMSM is only approximate. We orthogonalize the nonlinear class subspaces using this framework and then apply KMSM to the orthogonalized nonlinear class subspaces. Fukunaga and Koontz's method has been applied to related linear subspace methods including MSM, improving their performance [1, 7–9].

In Fukunaga and Koontz's method orthogonalization is achieved by applying a whitening transformation matrix to the training patterns or orthogonal basis vectors of each class subspace. Thus, the main task we need to solve is

[3] Note that a linear subspace in the feature space $\mathcal{F}$ is an nonlinear subspace in the input space $\mathcal{I}$.

to construct the whitening transformation matrix for orthogonalizing the linear subspaces in the feature space $\mathcal{F}$.

This paper is organized as follows. Section 2 describes the calculation of the canonical angles. In Section 3, we compute the whitening transformation. In Section 4, we define the kernel whitening transformation and orthogonalize nonlinear class subspaces. Then we construct the KOMSM by applying the KMSM to the orthogonalized nonlinear class subspaces. Section 5 demonstrates the effectiveness of KOMSM through experiments. We conclude in section 6.

2 Canonical angles between subspaces

Given an m_p-dimensional linear input subspace $\mathcal{P}$ and an m_q-dimensional linear reference subspace $\mathcal{Q}$ in the n-dimensional feature space, the canonical angles $\{0 \leq \theta_1, \ldots, \theta_{n_p} \leq \frac{\pi}{2}\}$ between $\mathcal{P}$ and $\mathcal{Q}$ (for convenience $m_p \leq m_q$) are uniquely defined as [12]:

$$\cos(\theta_k) = \max_{\mathbf{u}\in\mathcal{P}} \max_{\mathbf{v}\in\mathcal{Q}} \mathbf{u}^\top \mathbf{v} \tag{1}$$

$$\text{subject to :}$$

$$\mathbf{u}_i^\top \mathbf{u}_i = \mathbf{v}_i^\top \mathbf{v}_i = 1;\ \mathbf{u}_i{}^\top \mathbf{u}_j = 0; \mathbf{v}_i{}^\top \mathbf{v}_j = 0;\ i \neq j, i = 1{\sim}m_p,\ i = j{\sim}m_q,$$

Let $\mathbf{\Phi}_i$ and $\mathbf{\Psi}_i$ denote the i-th n-dimensional orthonormal basis vectors of the subspaces $\mathcal{P}$ and $\mathcal{Q}$, respectively. These orthogonal basis vectors can be obtained as the eigenvectors of the correlation matrix $\sum_{i=1}^{l} \mathbf{x}_i\mathbf{x}_i^\top$ calculated from the l learning patterns $\{\mathbf{x}\}$ of each class.

A practical method of finding the canonical angles is by computing the matrix $\mathbf{X}$=$\mathbf{A}^\top\mathbf{B}$, where $\mathbf{A} = [\mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_{m_p}]$ and $\mathbf{B}$=$[\mathbf{\Psi}_1, \ldots, \mathbf{\Psi}_{m_q}]$. Let $\{\kappa_1, \ldots, \kappa_{m_p}\}$ be the singular values of the matrix $\mathbf{X}$. The canonical angles between the two subspaces can be obtained as $\{cos^{-1}(\kappa_1), \ldots, cos^{-1}(\kappa_{m_p})\}$. In practice, we consider the value of the mean of the canonical angles, $S[t] = \frac{1}{t}\sum_{i=1}^{t} \cos^2\theta_i$, as the similarity between two subspaces. The value S reflects the structural similarity between two subspaces.

3 Orthogonalization by the whitening transformation

In this section, we will describe how to calculate the whitening matrix $\mathbf{O}$ for orthogonalizing r m-dimensional linear class subspaces with the orthogonal basis vectors $\mathbf{e}_i (i = 1{\sim}m)$ in the n-dimensional input space $\mathcal{I}$.

At first, we define the projection matrix corresponding to the projection onto the class i subspace $\mathcal{P}_i$ by $\mathbf{P}_i$= $\sum_{j=1}^{m} \mathbf{e}_j\mathbf{e}_j^\top$ where $\mathbf{e}_j$ is the j-th orthogonal basis vector of $\mathcal{P}_j$. Then we define the total projection matrix $\mathbf{G}$=$\sum_{i=1}^{r} \mathbf{P}_i$.

Using the eigenvectors and the eigenvalues of the total projection matrix $\mathbf{G}$, the $v{\times}n$ whitening matrix $\mathbf{O}$ is defined by the following equation:

$$\mathbf{O} = \mathbf{\Lambda}^{-1/2}\mathbf{H}^\top \ , \tag{2}$$

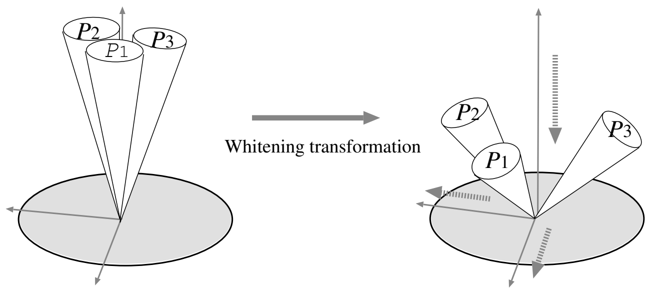

Fig. 2. Concept of orthogonalization of subspaces by the whitening transformation.

where $v = r \times m$, $(v = n,\ if\ v > n)$, $\mathbf{\Lambda}$ is the $v \times v$ diagonal matrix with the i-th highest eigenvalue of the matrix $\mathbf{G}$ as the i-th diagonal component, and $\mathbf{H}$ is the $n \times v$ matrix whose i-th column vector is the eigenvector of the matrix $\mathbf{G}$ corresponding to the i-th highest eigenvalue. We can confirm that the matrix $\mathbf{O}$ can whiten the matrix $\mathbf{G}$ so that r class subspaces are orthogonalized as proved in [8]. Fig.2 shows the concept of orthogonalization by the whitening transformation. We can see that the orthogonalization is achieved by whitening, i.e. homogenizing the variances in all directions:

All transformed basis vectors are orthogonal to each other when the multiple of the number r of classes and the dimension m of each class is smaller than the dimension f of the input space[8]. On the other hand, since the transformed basis vectors of an input subspace are no longer orthogonal, all transformed basis vectors need to be orthogonalized again by Gram-Schmidt orthogonalization.

4 The proposed method

In this section, we will first describe the concept of nonlinear subspaces, and then construct the kernel whitening matrix in the feature space $\mathcal{F}$.

4.1 Nonlinear subspace by kernel PCA

The nonlinear function ϕ maps the patterns $\mathbf{x} = (x_1, \ldots, x_n)^\top$ of an n-dimensional input space $\mathcal{I}$ onto an f-dimensional feature space $\mathcal{F}$: $\phi : \Re^n \rightarrow \Re^f,\ \mathbf{x} \rightarrow \phi(\mathbf{x})$. To perform PCA on the mapped patterns, we need to calculate the inner product $(\phi(\mathbf{x}) \cdot \phi(\mathbf{y}))$ between the function values. However, this calculation is difficult, because the dimension f of the feature space $\mathcal{F}$ can be very high, possibly infinite. However, if the nonlinear map ϕ is defined through a kernel function $k(\mathbf{x}, \mathbf{y})$

which satisfies Mercer's conditions, the inner products $(\phi(\mathbf{x}) \cdot \phi(\mathbf{y}))$ can be calculated from the inner products $(\mathbf{x} \cdot \mathbf{y})$. This technique is known as the "kernel trick".

A common choice is to use an exponential function: $k(\mathbf{x}, \mathbf{y})=\exp\left(-\frac{||\mathbf{x}-\mathbf{y}||^2}{\sigma^2}\right)$. The function ϕ with this kernel function maps an input pattern onto an infinite feature space $\mathcal{F}$. The PCA of the mapped patterns is called kernel PCA[10]. Note that the linear subspace in the feature space $\mathcal{F}$ is a nonlinear subspace in the input space $\mathcal{I}$.

4.2 Kernel whitening transformation

In the following, we will explain how to generate the kernel whitening matrix $\mathbf{O}_\phi$ from all the basis vectors of r d-dimensional nonlinear class subspaces $\mathcal{V}_k (k = 1 \sim r)$, that is, $r \times d$ basis vectors. This calculation corresponds to the kernel PCA for the basis vectors of all the classes.

Assume that a class k nonlinear subspace $\mathcal{V}_k$ is generated from l learning patterns $\mathbf{x}_i^k (i = 1 \sim l)$. The d basis vectors $\mathbf{e}_i^k (i = 1 \sim d)$, which expand the subspace $\mathcal{V}_k$, are defined by the following equation:

$$\mathbf{e}_i^k = \sum_{j=1}^{l} \mathrm{a}_{ij}^k \phi(\mathbf{x}_j^k) \ , \tag{3}$$

where the coefficient a_{ij}^k is the j-th component of the eigenvector $\mathbf{a}_i$ corresponding to the i-th largest eigenvalue λ_i of the m×m kernel matrix $\mathbf{K}$ defined by $\mathrm{k}_{ij}=(\phi(\mathbf{x}_i^k), \phi(\mathbf{x}_j^k))$. $\mathbf{a}_i$ is normalized to satisfy $\lambda_i(\mathbf{a}_i \cdot \mathbf{a}_i)=1$.

Next, assume that $\mathbf{E}$ is the matrix where all basis vectors are arranged as the column component:

$$\mathbf{E} = [\mathbf{e}_1^1, \ldots, \mathbf{e}_d^1, \ldots, \mathbf{e}_1^r, \ldots, \mathbf{e}_d^r] \ . \tag{4}$$

Then, we solve the eigenvalue problem of the matrix $\mathbf{Q}$ defined by the following equation:

$$\begin{aligned} &\mathbf{Qb} = \beta \mathbf{b} \\ &\mathrm{Q}_{ij} = (\mathbf{E}_i \cdot \mathbf{E}_j), \quad (i, j = 1 \sim r \times d) \ , \end{aligned} \tag{5}$$

where $\mathbf{E}_i$ means the i-th column component of the matrix $\mathbf{E}$. In the above equation, the inner product between i-th basis vector $\mathbf{e}_i^k$ of class k and j-th basis vector $\mathbf{e}_j^{k^*}$ of class k^* can be actually calculated as the linear combination of a kernel function value $k(\mathbf{x}^k, \mathbf{x}^{k^*})$ of $\mathbf{x}^k$ and $\mathbf{x}^{k^*}$.

$$(\mathbf{e}_i^k \cdot \mathbf{e}_j^{k^*}) = (\sum_{s=1}^{l} \mathrm{a}_{is}^k \phi(\mathbf{x}_s^k) \cdot \sum_{t=1}^{l} \mathrm{a}_{jt}^{k^*} \phi(\mathbf{x}_t^{k^*})) \tag{6}$$

$$= \sum_{s=1}^{l} \sum_{t=1}^{l} \mathrm{a}_{is}^k \mathrm{a}_{jt}^{k^*} (\phi(\mathbf{x}_s^k) \cdot \phi(\mathbf{x}_t^{k^*})) \tag{7}$$

$$= \sum_{s=1}^{l} \sum_{t=1}^{l} \mathrm{a}_{is}^k \mathrm{a}_{jt}^{k^*} k(\mathbf{x}_s^k, \mathbf{x}_t^{k^*}) \ . \tag{8}$$

The i-th row vector $\mathbf{O}_{\phi_i}$ of the kernel whitening matrix $\mathbf{O}_\phi$ can be represented as the linear combination of the basis vectors $\mathbf{E}_j$ ($j = 1 \sim r{\times}d$) using the eigenvector $\mathbf{b}_i$ corresponding to the eigenvalue β_i as the combination coefficient.

$$\mathbf{O}_{\phi_i} = \sum_{j=1}^{r\times d} \frac{\mathrm{b}_{ij}}{\sqrt{\beta_i}} \mathbf{E}_j \ , \tag{9}$$

where the vector $\mathbf{b}_i$ is normalized to satisfy that $\beta_i(\mathbf{b}_i \cdot \mathbf{b}_i)$ is equal to 1. The row vectors of $\mathbf{O}_\phi$ with eigenvalues β lower than a threshold value are discarded, since their reliability is low. Moreover, assume that $\mathbf{E}[j]$ is the $\eta(j)$-th basis vector of the class $\zeta(j)$. Then the above equation can be changed as follows:

$$\mathbf{O}_{\phi_i} = \sum_{j=1}^{r\times d} \frac{\mathrm{b}_{ij}}{\sqrt{\beta_i}} \sum_{s=1}^{l} \mathrm{a}_{\eta(j)s}^{\zeta(j)} \phi(\mathbf{x}_s^{\zeta(j)}) \tag{10}$$

$$= \sum_{j=1}^{r\times d} \sum_{s=1}^{l} \frac{\mathrm{b}_{ij}}{\sqrt{\beta_i}} a_{\eta(j)s}^{\zeta(j)} \phi(\mathbf{x}_s^{\zeta(j)}) \ . \tag{11}$$

Although we cannot calculate this vector $\mathbf{O}_{\phi_i}$, the inner product with the mapped vector $\phi(\mathbf{x})$ can be calculated.

4.3 Whitening transformation of the mapped patterns

The mapped vector $\phi(\mathbf{x})$ is transformed by the kernel whitening matrix. This can be calculated from an input vector $\mathbf{x}$ and all $r{\times}l$ learning vectors $\mathbf{x}_s^k (s = 1{\sim}l, k = 1{\sim}r)$ using the following equation:

$$(\phi(\mathbf{x}) \cdot \mathbf{O}_{\phi_i}) = \sum_{j=1}^{r\times d} \sum_{s=1}^{l} \frac{\mathrm{b}_{ij}}{\sqrt{\beta_i}} \mathrm{a}_{\eta(j)s}^{\zeta(j)} (\phi(\mathbf{x}) \cdot \phi(\mathbf{x}_s^{\zeta(j)})) \tag{12}$$

$$= \sum_{j=1}^{r\times d} \sum_{s=1}^{l} \frac{\mathrm{b}_{ij}}{\sqrt{\beta_i}} \mathrm{a}_{\eta(j)s}^{\zeta(j)} k(\mathbf{x}, \mathbf{x}_s^{\zeta(j)}) \ . \tag{13}$$

Finally, the whitening transformed vector $\chi(\phi(\mathbf{x}))$ of the mapped vector $\phi(\mathbf{x})$ is represented as $(z_1,\ z_2, \ldots,\ z_{n_o})^\top,\ \ z_i = (\phi(\mathbf{x}) \cdot \mathbf{O}_{\phi_i}),\ \ 1{\leq}i{\leq}n_o{\leq}r{\times}d$, where n_o is the row number of $\mathbf{O}_\phi$ mentioned above.

4.4 The KOMSM Algorithm

We construct the KOMSM by applying the linear MSM to the linear subspaces generated from the whitening transformed vector $\chi(\phi(\mathbf{x}))$ in the feature space $\mathcal{F}$ as follows.

In the learning stage:

1. The nonlinear mapped $\phi(\mathbf{x}_i^k)$ of all the patterns $\mathbf{x}_i^k (i = 1 \sim l)$ belonging to class k are transformed by the kernel whitening matrix $\mathbf{O}_\phi$.

2. The basis vectors of the n-dimensional linear orthogonal reference k subspace $\mathcal{P}_k^{\mathbf{O}_\phi}$ are obtained as the eigenvectors of the correlation matrix generated from the whitening transformed pattern set $\{\chi(\phi(\mathbf{x}_1^k)), \ldots, \chi(\phi(\mathbf{x}_l^k))\}$, corresponding to the n highest values.
3. Similarly the other linear reference orthogonal subspaces are generated on the nonlinear feature space $\mathcal{F}$.

In the recognition stage:

1. The linear input orthogonal subspace $\mathcal{P}_{in}^{\mathbf{O}_\phi}$ is also generated from the whitening transformed pattern set $\{\chi(\phi(\mathbf{x}_1^{in})), \ldots, \chi(\phi(\mathbf{x}_l^{in}))\}$.
2. The canonical angles between the linear orthogonal input subspace $\mathcal{P}_{in}^{\mathbf{O}_\phi}$ and the linear orthogonal reference subspaces $\mathcal{P}_k^{\mathbf{O}_\phi}$ are calculated as the similarity.
3. Finally the object class is determined as the linear orthogonal reference subspace with the highest similarity S, given that S is above a threshold value.

In the above process, it is possible to replace the generation process of the nonlinear orthogonal subspaces with the following processes. Firstly the input subspace and the reference subspaces are generated from the set of the nonlinear mapped patterns. Next the basis vectors of the generated subspaces are transformed by the kernel whitening matrix, and then the whitening transformed basis vectors are orthogonalized by the Gram-Schmidt method.

5 Evaluation experiments

We compared the performances of KOMSM with previous methods (MSM[3], CMSM[6], OMSM[8], KMSM[4], KCMSM[11]) using the public data base of multi-view images (Cropped-close128 of ETH-80)[13] and the data set of front face images collected by ourselves.

5.1 3D object recognition (Experiment-I)

Thirty similar models were selected as the evaluation data from the ETH-80 database as shown in Fig.3. The images of each model were captured from 41 views as shown in Fig.3. All images are cropped, so that they contain only the object without any border area.

The odd numbered images (21 frames) and the even numbered images (20 frames) were used for training and evaluation, respectively. We prepared 10 datasets for each model by making the start frame number i changes from 1 to 10 where 10 frames from i-th frame to $i + 9$-th is one set. The total number of the evaluation trials is 9000(=10×30×30). The evaluation was performed in terms of the recognition rate and the equal error rate (EER), which represents the error rate at the point where the false accept rate (FAR) is equal to the false reject rate (FRR).

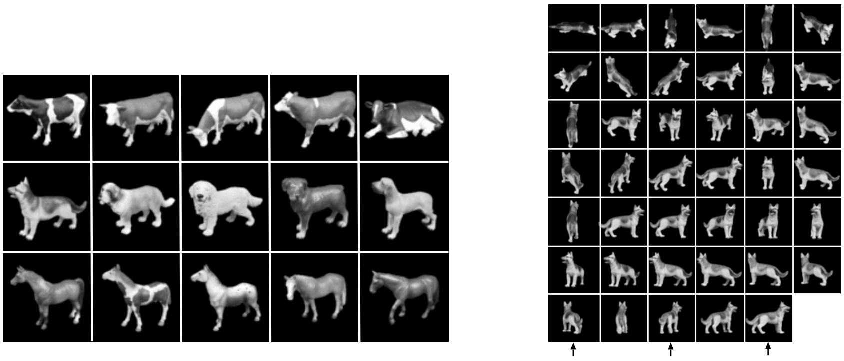

Fig. 3. Left: Evaluation data, Bottom: cow, Middle: dog, Bottom: horse. This figure shows five of ten models. Right: All view-patterns of dog1, the rows indicated by arrows are used as training data.

Table 1. Recognition rate and EER of each method (%) (Experiment-I).

Method	MSM	CMSM-200	OMSM-100	KMSM	KCMSM-400	**KOMSM-550**
Rate	78.6	92.33	89.6	96.33	99.67	99.67
EER	16.6	4.7	7.7	5.2	1.0	1.0

The experimental conditions are as follows. We converted the 180×180 pixel color images to 15×15 pixels monochrome images and use them as the evaluation data. Thus, the dimension n of a pattern is 225(=15×15). The dimensions of the input subspace and the reference subspaces were set to 7 for all methods. The whitening matrix $\mathbf{O}$ and the kernel whitening matrix $\mathbf{O}_\phi$ were generated from thirty 20-dimensional linear class subspaces and thirty 20-dimensional nonlinear class subspaces respectively. The row numbers of $\mathbf{O}$ and $\mathbf{O}_\phi$ were set to 100 and 550, respectively, on the basis of the eigenvalues β in Eq.(6). The dimensions of constraint subspaces of the CMSM and KCMSM were set to 200 and 400, respectively, as used in [11]. The length of input vectors was not normalized for all the methods. A Gaussian kernel with $\sigma^2 = 1e + 6$ was used for all the kernel methods.

Table 1 shows the recognition rate and EER of each method. The recognition of multiple view images is typically a nonlinear problem. This is clearly shown by the experimental results that the performance of the nonlinear methods (KMSM, KCMSM and KOMSM) is superior to that of the linear methods (MSM, CMSM and OMSM). The performance of MSM was improved by the nonlinear extension of MSM to KMSM where the recognition rate increased from 78.6% to 96.3% and the EER decreased from 16.6% to 5.2%. KOMSM further improved the performance of KMSM where the recognition rate increased from 96.3% to 99.6% and the EER decreased from 5.7% to 1.0%. This confirms the effectiveness of the

Fig. 4. Face images: From left, Lighting1∼Lighting10.

Table 2. Recognition rate and EER of each method (%)(Experiment-II)

Method	MSM	CMSM-200	OMSM	KMSM	KCMSM-1050	**KOMSM**
Recognition rate	91.74	91.30	97.09	91.15	97.40	97.42
EER	12.0	7.5	6.3	11.0	4.3	3.5

orthogonalization of the nonlinear subspaces, which serves as a feature extraction step in the feature space $\mathcal{F}$.

5.2 Recognition of face image (Experiment-II)

We conducted the evaluation experiment of all the methods using the face images of 50 persons captured under 10 kinds of lighting. We cropped the 15 ×15 pixel face images from the 320 × 240 pixel input images based on the positions of pupils and nostrils.

The normalized face patterns of subjects 1-25 in lighting conditions L1-L10 were used for generating the difference subspace $\mathcal{D}$, the kernel difference subspace $\mathcal{D}_\phi$, the whitening matrix $\mathbf{O}$ and the kernel whitening matrix $\mathbf{O}_\phi$. The face patterns extracted from the images of the other subjects, 26-50, in lighting conditions L1-L10 were used for evaluation. The number of the data of each person is 150∼180 frames for each lighting condition. The data was divided into 15∼18 sub datasets by every 10 frames. The input subspaces were generated from these sub datasets. The dimension of the input subspace and reference subspaces were set to 7 for all the methods. The difference subspace $\mathcal{D}$ and the whitening matrix $\mathbf{O}$ were generated from 25 60-dimensional linear subspaces of 1∼25 persons. The kernel difference subspace $\mathcal{D}_\phi$ and the kernel whitening matrix $\mathbf{O}_\phi$ were generated from 25 60-dimensional nonlinear class subspaces. The row numbers of of $\mathbf{O}$ and $\mathbf{O}_\phi$ were set to the full dimensions, 225 and 1500, respectively. The dimensions of the generalized difference subspace and the kernel generalized difference subspace were set to 200 and 1050, respectively. We used a Gaussian kernel with $\sigma^2 = 1.0$ for all nonlinear methods.

Table 2 shows the recognition rate and the EER of each method. The difference between the recognition rates of OMSM and KOMSM was small, while the EER decreased from 6.3% to 3.5%. This implies that the data sets used in this task do not exhibit highly nonlinear structure. The good performance of KOMSM and KCMSM are also demonstrated in this experiment. In particular the EER of KOMSM is very low.

Although the performance of KOMSM and KCMSM are at the same level regardless of their different principals of orthogonalization, KOMSM has an advantage in selecting parameters compared to KCMSM. KOMSM does not have any parameters to be tuned. In contrast, the dimension of the constraint subspace used in KCMSM needs to be carefully selected in prior experiments, since the performance of KCMSM strongly depends on this value.

6 Conclusion

In this paper we have introduced the kernel orthogonal mutual subspace method (KOMSM) and applied it to 3D object recognition. The essence of the KOMSM is to orthogonalize nonlinear subspaces based on Fukunaga and Koontz's framework before applying the KMSM. We have confirmed that this orthogonalization provides a strong feature extraction method for the KMSM and that the performance of the KMSM is improved significantly. This was shown by evaluation on multi-view image sets of 3D objects as well as frontal face images. In feature work, we will attempt to find the efficient computation of the eigen problems of the matrixes $\mathbf{K}$ and $\mathbf{Q}$ with the sizes which are proportional to the numbers of the classes and the training patterns.

References

1. E. Oja: Subspace methods of pattern recognition. Research Studies Press (1983)
2. K. Fukunaga, W. C. G. Koontz: Application of the Karhunen-Loeve expansion to feature selection and ordering. IEEE Trans. Computers Vol.C-19 (1970) 311–318
3. K. Maeda, S. Watanabe: A pattern matching method with local structure: Trans. IEICE Vol.J68-D No.3 (1985) 345–352 (in Japanese)
4. H. Sakano, N. Mukawa, T. Nakamura: Kernel Mutual Subspace Method and its Application for Object Recognition. Electronics and Communications in Japan, Vol.88 (2005) 45–53
5. L. Wolf, A. Shashua: Kernel principal angles for classification machines with applications to image sequence interpretation. CVPR03 (2003) 635–642
6. K. Fukui, O. Yamaguchi: Face recognition using multi-viewpoint patterns for robot vision. 11th International Symposium of Robotics Research (ISRR) (2003) 192–201
7. K. Nagao, M. Sohma: Weak orthogonalization of face and perturbation for recognition. CVPR98 (1998) 845–852
8. T. Kawahara, M. Nishiyama, O. Yamaguchi: Face recognition by orthogonal mutual subspace method. CVIM-151 (2005) 17–24 (in Japanese)
9. Tae-Kyun Kim, J. Kittler, R. Cipolla: Incremental Learning of Locally Orthogonal Subspaces for Set-based Object Recognition. BMVC06 (2006) 559–568
10. B. Schölkopf, A. Smola, K.-R. Müller: Nonlinear principal component analysis as a kernel eigenvalue problem. Neural Computation Vol. 10 (1998) 1299–1319
11. K. Fukui, B. Stenger, O. Yamaguchi: A framework for 3D object recognition using the kernel constrained mutual subspace method. ACCV06 (2006) 315–324
12. F. Chatelin: Eigenvalues of matrices. John Wiley & Sons, Chichester (1993)
13. B. Leibe, B. Schiele: Analyzing appearance and contour based methods for object categorization. CVPR03 (2003) 409–415