ACCV2007

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

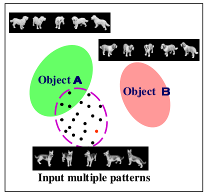
Kazuhiro Fukui, Osamu Yamaguchi
University of Tsukuba, Toshiba Corporation

University of Tsukuba

Motivation

Task of classifying sets of patterns such as video frames or multi-view images is essential in computer vision.

- 3D object recognition with multi-camera system
- Gesture recognition using video image



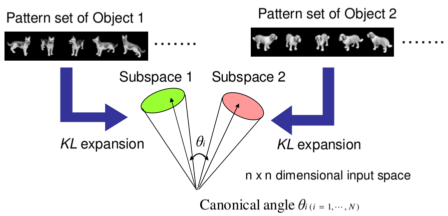
Using multiple patterns

Let that an n x n pixel pattern is an n x n dimensional vector.
A object pattern is represented as a point in a n x n dimensional feature space

Our contribution

- We propose the Kernel Orthogonal Mutual Subspace Method (KOMSM).
  (1)Each class set of patterns is represented by an nonlinear subspace.
  (2)Nonlinear class subspaces are orthogonalized by Fukunaga and Koontz's framework.
  (3) The canonical angles between the orthogonalized nonlinear class subspaces are calculated by the kernel Mutual Subspace Method (KMSM).
- The orthogonalization provides a powerful feature extraction method for improving the performance of KMSM.

How to measure the similarity between distributions ?



Measure the similarity between two distributions with canonical angles

Calculation of canonical angles

The canonical angles θ_k between M-dimensional subspace P and N-dimensional subspace Q (for convenience N ≤ M) are uniquely defined as follows:

$$\cos\theta_k = \max_{\mathbf{u}\in P}\max_{\mathbf{v}\in Q}\mathbf{u}^T\mathbf{v} \quad (k=1,\cdots,N)$$

subject to:

$$\mathbf{u}_i^T\mathbf{u}_i=\mathbf{v}_i^T\mathbf{v}_i=1;\quad \mathbf{u}_i^T\mathbf{u}_j=\mathbf{v}_i^T\mathbf{v}_j=0;$$
$$i\neq j, i=1,\cdots,N,\ j=1,\cdots,N$$

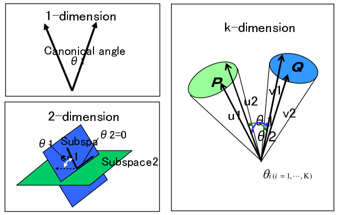
$\cos\ \theta_i = \lambda_i\ (i=1\sim N)$

θ_i : i-th largest canonical angles

λ_i : i-th largest singular value of matrix $\mathbf{A}^T\mathbf{B}$

where $\mathbf{A}=[\boldsymbol{\varphi}_1,\boldsymbol{\varphi}_2,\cdots,\boldsymbol{\varphi}_M]$, $\mathbf{B}=[\boldsymbol{\psi}_1,\boldsymbol{\psi}_2,\cdots,\boldsymbol{\psi}_N]$,

$\boldsymbol{\varphi}_{i\,(i=1-M)}$, $\boldsymbol{\psi}_{i\,(i=1-N)}$ are orthonomal basis of subspace P and Q, respectively.



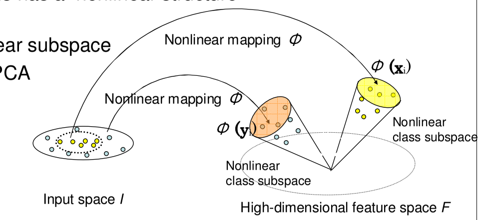
Canonical angles between two subspaces

Kernel Mutual Subspace Method (KMSM)

Distribution of patterns has a nonlinear structure — Sakano, H.; Mukawa, N. 2000

⇩

Introduction of nonlinear subspace generated by kernel PCA



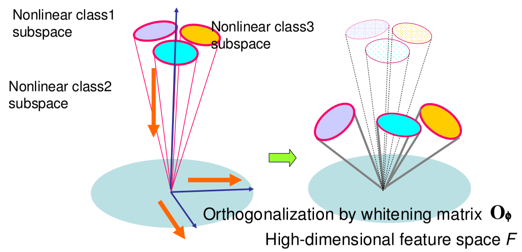
KMSM is a method of classifying sets of pattern by using the canonical angles between input nonlinear subspace and nonlinear reference subspaces as the similarity.

Improvement of the performance of KMSM

The classification ability of KMSM is still insufficient, because each class subspace are not optimal in terms of classification performance, while it represents the distribution of the training patterns well in terms of a least-mean-square approximation.

⇩

We orthogonalize nonlinear class subspaces each other by whitening transformation, before calculating the canonical angles between nonlinear subspaces



Equalization of distribution in each principal component direction

Algorithm of KOMSM

In learning stage:

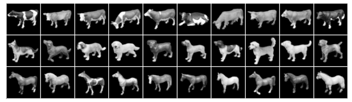
1. Patterns {X^k} belonging to class k are mapped to ϕ(X^k) by nonlinear function ϕ.
2. The nonlinear mapped ϕ(X^k) are transformed by the kernel whitening matrix $\mathbf{O}_\phi$.
3. The basis vectors of the n-dimensional linear orthogonal reference k subspace are obtained as the eigenvectors of the correlation matrix generated from the whitening transformed pattern set , corresponding to the n highest values.

In recognition stage:

1. A linear input orthogonal subspace is also generated from the whitening transformed pattern set.
2. The canonical angles between the linear orthogonal input subspace and the linear orthogonal reference subspaces 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.

Experiments

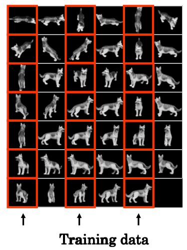
Experiment-I (3D object recognition)



30 models selected from the open dataset ETH-80 : croppes-close128

B. Leibe and B. Schiele, CVPR03

Example: all data of dog-1



Experimental conditions

Dimension of Input space: 225 (=15 x 15 pixels)
Input subspace: 7
Reference subspace: 7
Kernel function: Gaussian kernel with δ=0.05

Experimental results

	MSM	CMSM	OMSM	KMSM	KCMSM	KOMSM
Recognition 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

* EER: Error Equal Rate

Experiment-II (Face recognition)

50 subjects, 10 Illumination conditions
Subjects 1-25 are for generating whitening matrix $\mathbf{O}_\phi$,
Subjects 26-50 are for testing
Dimension of input space: 225 =15x15pixels

Experimental results

	MSM	CMSM	OMSM	KMSM	KCMSM	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

MSM: Linear MSM, CMSM:Constrained MSM, OMSM: Whitening MSM (Orthogonal MSM), KMSM: Kernel MSM, KCMSM: Kernel CMSM, KOMSM: Kernel OMSM (Proposed method)