A Framework for 3D Object Recognition using the Kernel Constrained Mutual Subspace Method

Kazuhiro Fukui[*1], Bjorn Stenger[*2] and Osamu Yamaguchi[*2]

[*1] University of Tsukuba
[*2] Corporate Research & Development Center
Toshiba corporation

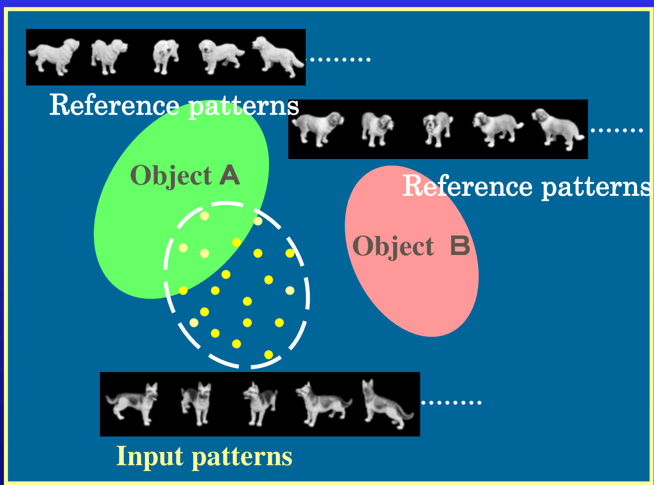
3D objection recognition

- Important problem in computer vision
- Many methods have been proposed
  - Geometry based method
  - View based method
    - Using single view pattern
    - Using multi-view patterns

View based method using multi-view patterns

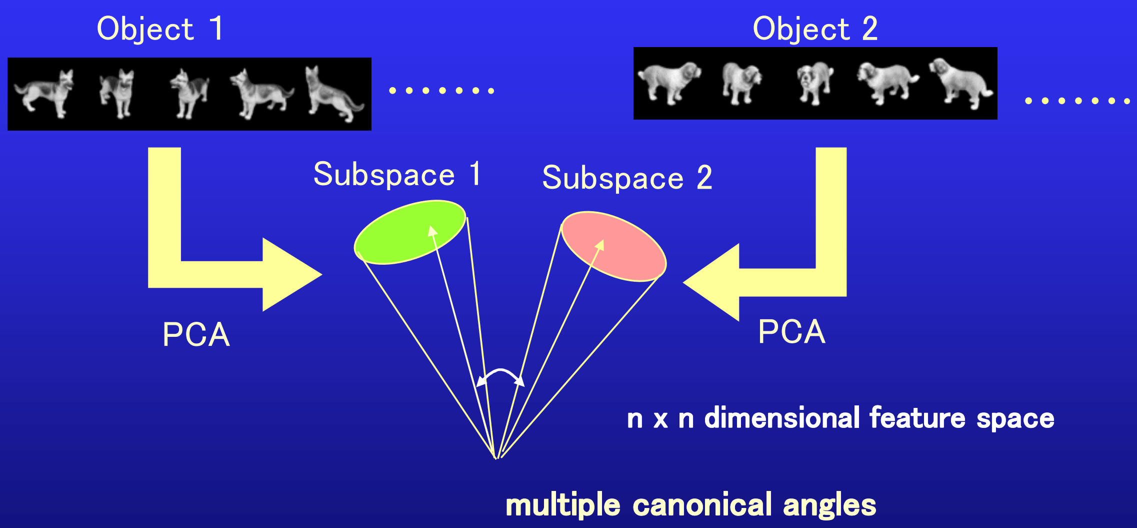
Let that an n x n pixel pattern is an n x n dimensional vector.

A object pattern is represented as a point in a n x n dimensional feature space



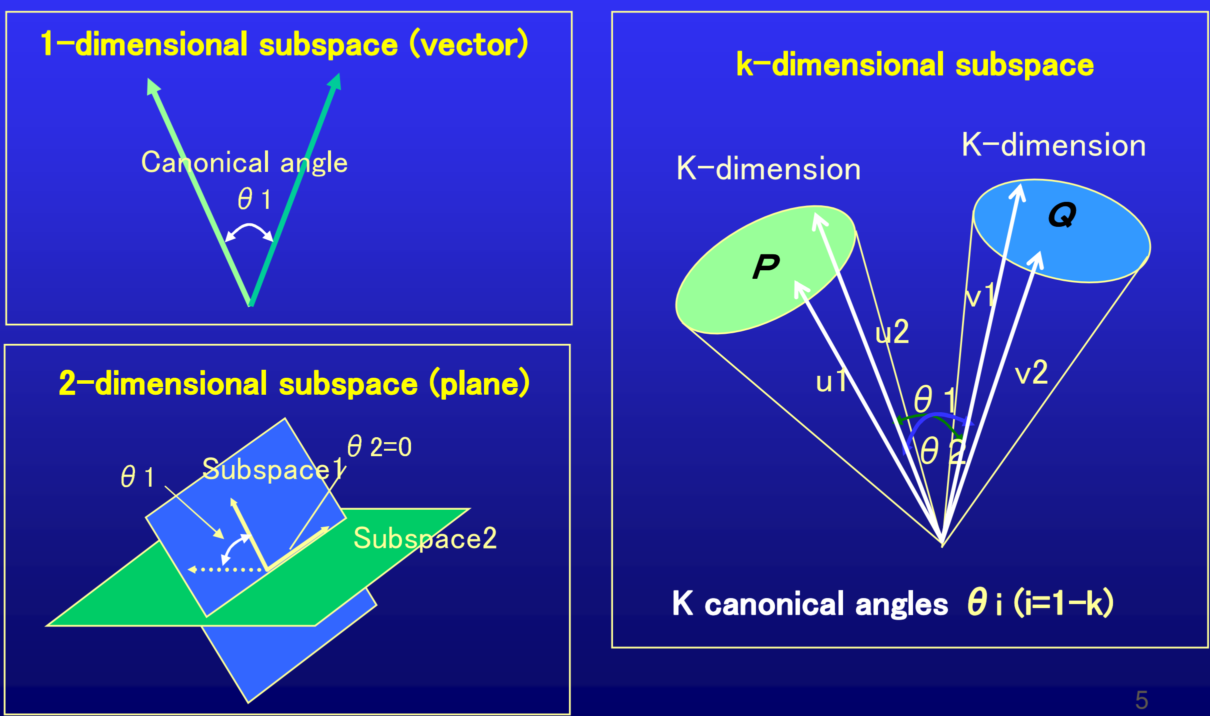
The distribution of view patterns reflects the 3D shape implicitly

How to measure the similarity between two distributions



Measure the similarity between two distributions with canonical angles

Canonical angles between two subspaces



Calculation of canonical angles by MSM

Given k dimensional subspace *P* and *Q* :

Canonical angles θi (i=1～k) are obtained

by Mutual Subspace Method (MSM)

$$\lambda_{\mathrm{i}} = \cos^2\theta_{\mathrm{i}} \quad (\mathrm{i} = 1 - k)$$

λ i: i-th largest eigenvalue of matrix **QPQ** or **PQP**

Projection matrix of subspace *P* $\mathbf{P} = \sum_{i=1}^{k} e_i e_i^t$

Projection matrix of subspace *Q* $\mathbf{Q} = \sum_{i=1}^{k} e_i e_i^t$

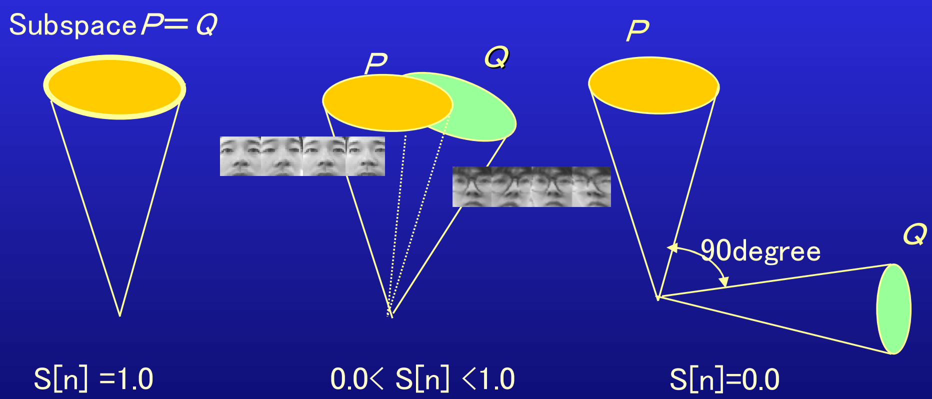
e i is the i-th basis vector of each subspace

Definition of similarity in MSM

$$\boxed{\text{Similarity } S[n] = \frac{1}{n}\sum_{i=1}^{n}\cos^2\theta_i}$$

λ : eigenvalue （$= \cos^2\theta_i$）

n : dimension



The structure of a subspace reflects the 3D shape of face

Limitation of linear MSM

MSM work well:

When the distribution of patterns can be represented by a linear subspace

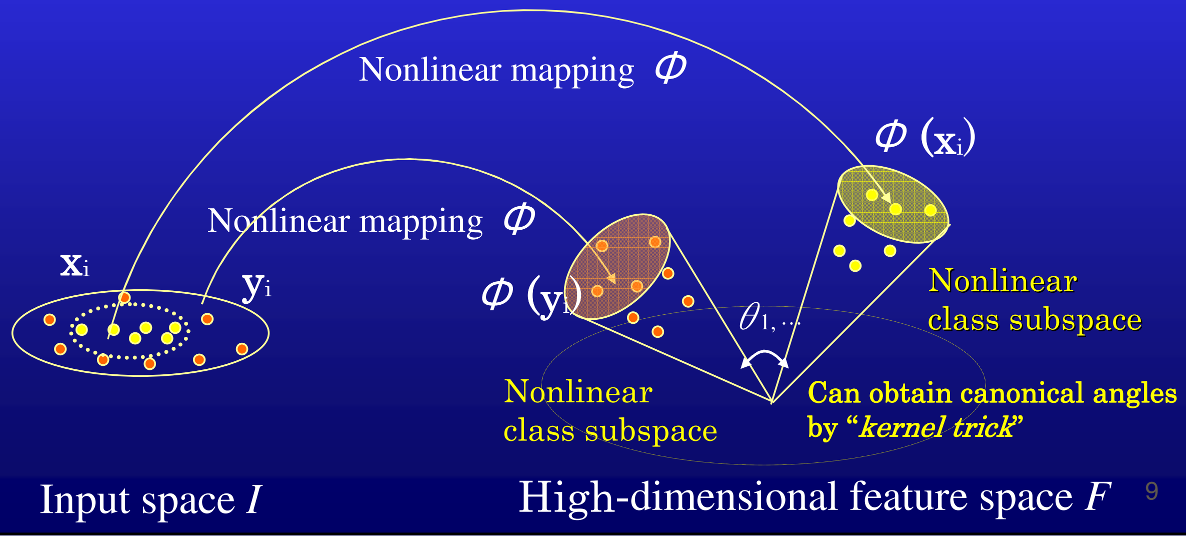
However,

When object, camera moves largely,

The distribution of view patterns *can not be* represented suitably by a linear subspace

Extension of MSM to Kernel MSM using nonlinear subspace

- MSM has been extended to the Kernel nonlinear MSM using the concept of nonlinear subspace (by H. Sakano, N. Mukawa, 2000, L. Wolf and A. Shashua, CVPR03)
- Nonlinear subspace is generated by applying kernel PCA to the nonlinear mapped view patterns
- Nonlinear subspace ＝ Linear subspace in feature space



Limitation of kernel MSM

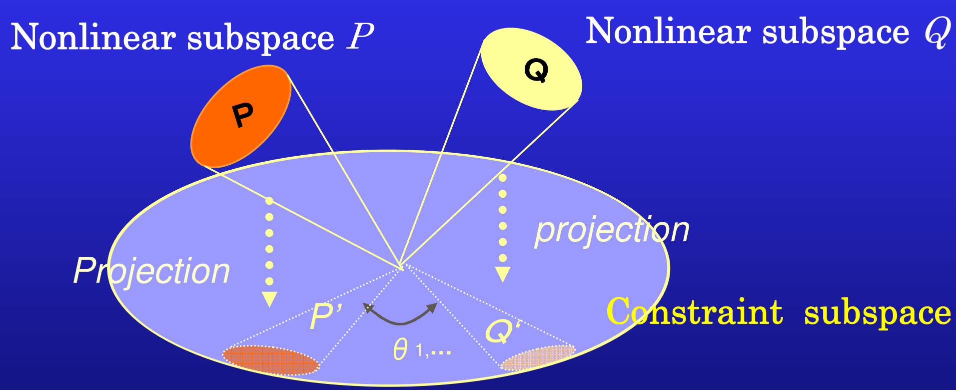
- Use of the kernel mutual subspace method can achieve a high recognition rate compared with the linear MSM

However,

- It's classification ability is still insufficient !

Improvement of classification ability of KMSM

Kernel Constrained Mutual Subspace Method（KCMSM）

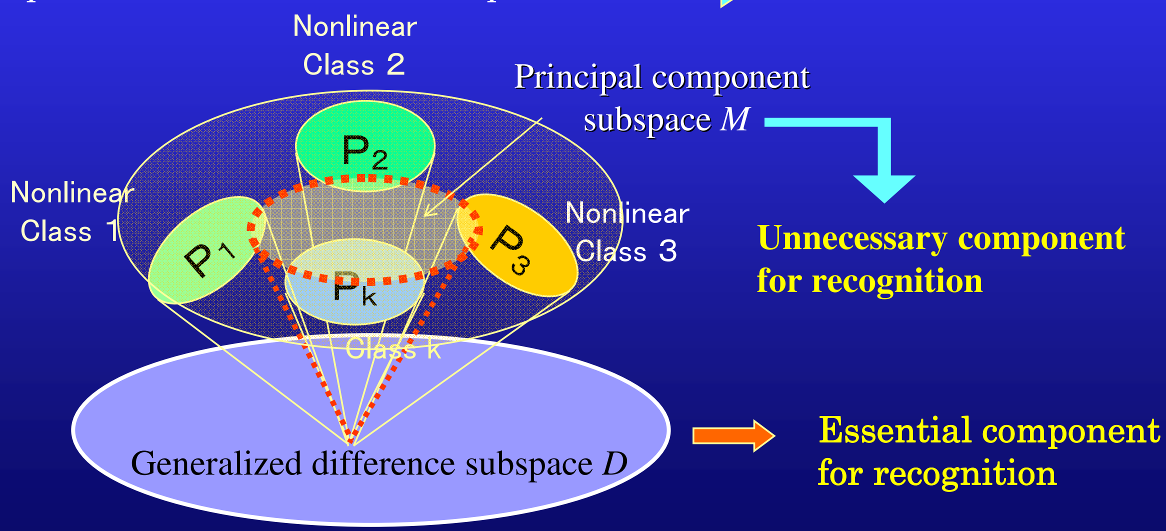


High-dimensional feature space F

Constrained subspace C： includes the effective component for recognition; the *different component* between nonlinear subspaces

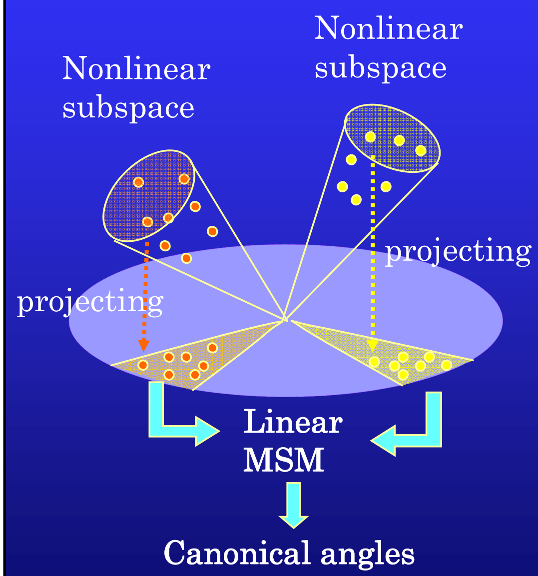
Introduction of Generalized difference subspace as the constraint subspace

Generalized difference subspace D is defined as the subspace which results by removing the principal component subspace M from sum space S of all nonlinear subspaces $\Rightarrow$ $M \perp D$



In high-dimensional feature space F

Calculation of canonical angles between projected nonlinear subspaces



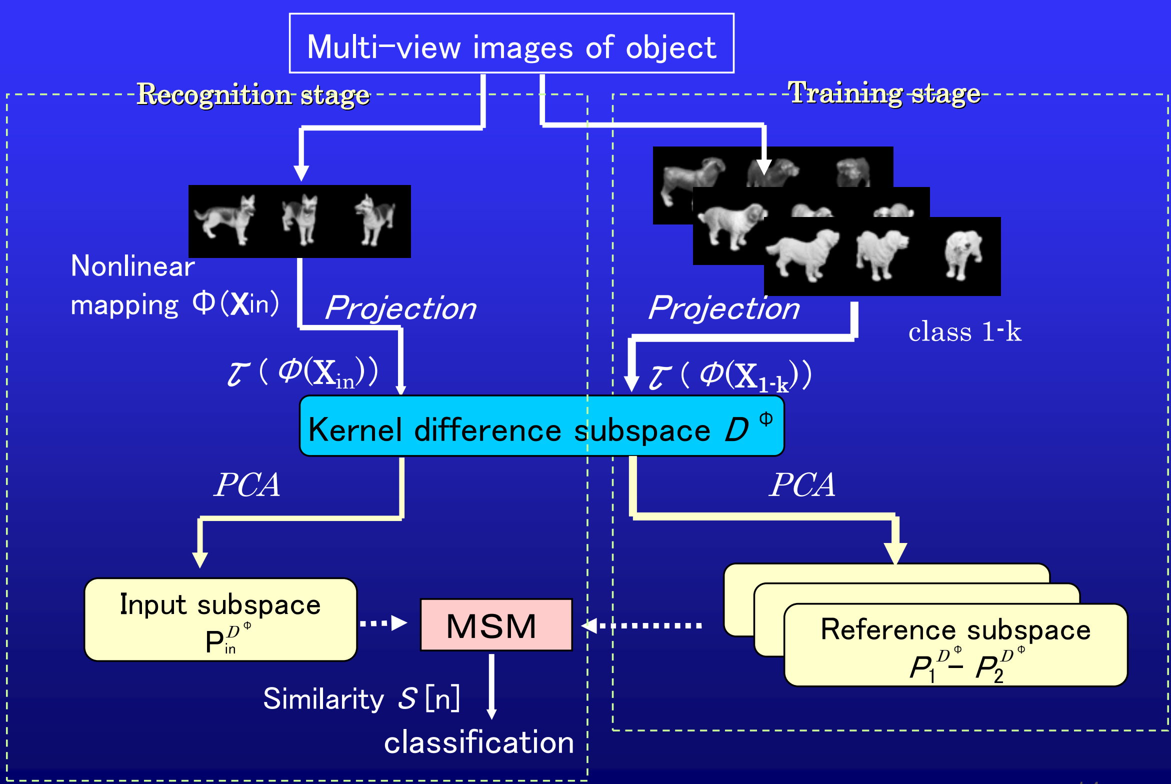
To calculate canonical angles between projected nonlinear subspaces in high-dimensional feature space

We need to calculate inner product of orthogonal basis vectors of projected nonlinear subspaces

This inner products can be calculated from all the nonlinear mapped view patterns by using "kernel trick"

In high-dimensional feature space F

Flow of object recognition using KCMSM

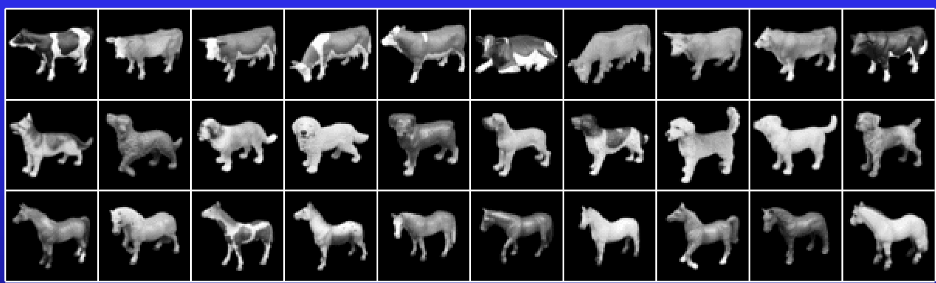


Experimental conditions

(1) Evaluation data

30 models selected from the dataset

ETH-80 : croppes-close128 (by B. Leibe and B. Schiele, CVPR03)



(2) Dimension of each subspace

Feature space: 225 (=15 x 15 pixel)

Input subspace: 7

Reference subspace: 7

Generalized difference subspace: 100-550

Kernel function:

$$k(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{\sigma^2}\right)$$

$\sigma = 0.05$

Training data and Test data

--- Example: all data of dog-1 ---

Training data

Reference nonlinear subspaces are generated from the training data.

Generalized difference subspace is generated from all the 20-dimensional reference nonlinear subspace.

Recognition rate (%)

		t=1	t=2	t=3	t=4
Linear	MSM	72.7	73.7	76.3	74.3
	CMSM-215	75.7	81.3	76.3	73.7
	CMSM-200	73.3	81.0	79.3	77.7
	CMSM-190	71.0	73.0	73.0	75.0
Nonlinear	KMSM	84.7	87.0	82.0	81.7
	KCMSM-550	83.0	85.3	85.7	86.3
	KCMSM-500	79.3	85.0	87.[illegible]	87.0
	KCMSM-450	82.0	88.0	89.3	89.7
	KCMSM-400	83.3	87.7	88.3	89.7
	KCMSM-300	81.0	87.7	88.7	89.0
	KCMSM-200	81.7	81.7	83.3	83.3
	KCMSM-100	57.7	62.7	68.0	65.3

t: number of canonical angles used for calculation of similarity

Classification ability (class separability)

		t=1	*t=2*	*t=3*	*t=4*
Linear	**MSM**	0.055	0.074	<u>0.082</u>	0.080
	CMSM-215	0.203	0.236	0.242	0.236
	CMSM-200	0.215	<u>0.257</u>	0.254	0.245
	CMSM-190	0.229	0.255	0.249	0.244
Nonlinear	**KMSM**	0.375	0.420	[illegible]420	<u>0.429</u>
	KCMSM-550	0.538	0.581	0.58[illegible]	0.5[illegible]8
	KCMSM-500	0.556	0.607	0.616	0.[illegible]2
	KCMSM-450	0.549	0.618	0.621	<u>0.621</u>
	KCMSM-400	0.529	0.601	0.607	0.609
	KCMSM-300	0.483	0.536	0.545	0.545
	KCMSM-200	0.340	0.385	0.403	0.408
	KCMSM-100	0.141	0.194	0.212	0.213

t: number of canonical angles used for calculation of similarity

Experimental result of face recognition

	CMSM	KMSM	**KCMSM**
Recog. rate(%)	88.89	89.75	**96.52**
Seperability	0.406	0.202	**0.643**

- Evaluation data: 25 person

  Person1-13 training data

  Person14-25 evaluation data

- Kernel function: Gaussian kernel with σ=0.1

Conclusion

- Proposed Kernel constrained mutual subspace method
  - Extension of Kernel MSM
  - Powerful feature extraction by projecting onto the generalized difference subspace
- Realized a new framework for 3D object recognition based on KCMSM

Feature works

- Reduce large computation time
- Experiment using more large data base

  for example, ALOI (Amsterdam library of image)
- Segmentation:

  Use of the feature invariant to position

  for example, HLAC: by otsu