

Classifiers under Continuous Observations

Hitoshi Sakano and Takashi Suenaga

NTT Data Corp.
Kayabacho Tower, 1-21-2, Shinkawa, Chuo-ku
Tokyo, 104-0033, Japan
{sakano,suenaga}@rd.nttdata.co.jp

Abstract. Many researchers have reported that recognition accuracy improves when several images are continuously input into a recognition system. We call this recognition scheme a *continuous observation-based scheme* (COBS). The COBS is not only a useful and robust object recognition technique, it also offers a new direction in statistical pattern classification research. The main problem in statistical pattern recognition for the COBS is how to define the measure of similarity between two distributions. In this paper, we introduce some classifiers for use with continuous observations. We also experimentally demonstrate the effectiveness of continuous observation by comparing various classifiers.

1 Introduction

Our research opens new directions in the field of statistical pattern recognition for highly accurate 3D object recognition.

Continuous observations may improve the robustness of object recognition systems. They enable a model to be extracted that accounts for changes in input when the pose or lighting conditions change and that also reduces the noise in an object image recognition problem (see Fig. 1.). In this paper, we discuss the problem for statistical pattern recognition when continuously input data are assumed. We call the method a *continuous observation-based scheme* (COBS).

It is reasonable to assume continuous observations in experimental studies of object recognition because we ordinarily use video cameras to capture object images. When we are working on image streams, statistical processing of the images in the stream may improve the accuracy of object recognition systems. Statistical processing reduces noise and enables invariant features to be extracted from the images .

We considered the following research issues from the viewpoint of statistical pattern recognition:

1. What type of statistical techniques can be effectively applied to input images?
2. How can we define the similarities between training images or templates and input images?

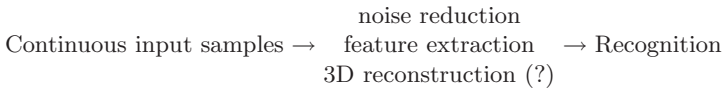


Fig. 1. Continuous observation-based scheme (COBS)

Some procedures have already been proposed under the COBS. Maeda proposed the *mutual subspace method* (MSM) [1], which uses principal component analysis to extract invariant features from input images. Yamaguchi demonstrated the effectiveness of the MSM for facial image recognition experimentally [2]. We proposed the *kernel mutual subspace method* (KMS) [3,4]. The KMS, which is one of the most powerful algorithms for object recognition, is based on kernel principal component analysis (KPCA) [5].

In this paper, we propose some classifiers for use when continuous observations are assumed. We also describe experiments to compare the classifiers experimentally.

In the next section, we describe [a problem in the COBS?] and in section 3 we introduce some classifiers used in the COBS. In section 4, we describe experiments carried out to clarify the properties of the classifiers.

2 Problem Setting

When COBS is assumed, we must consider the two issues described above. Generally, an object recognition algorithm consists of templates $T(\cdot)$ and the calculation of the similarity measure $S(T; \mathbf{x}')$ between the template and the input image \mathbf{x}' . The template is calculated from training data $\mathbf{x}_i, i = 1, \dots, m$, where m is the number of training images. PCA, a multilayer perceptron, sample means, and other statistical techniques are used to calculate the templates T .

Under the COBS, we must define a new measure of similarity between the templates and object images $s(T; T')$ where T' is calculated from features extracted from the input images $\mathbf{x}'_i, i = 1, \dots, m'$, where m' is the number of input images.

We then have to decide what kind of statistical analysis to apply? We can create many classifiers under the COBS.

3 Classifiers in COBS

We describe some of the classifiers used in the COBS. These classifiers are based on CLAFIC [6] (except for the multiple potential function classifier) because the subspace method is regarded as the standard method for object image recognition [7,8].

3.1 Sample Mean

First, we use a sample mean as the simplest statistical technique under CobS. It also has the advantage of reducing noise. When several images have been input into the recognition system, the system uses the sample mean

$$\boldsymbol{\mu}' = \frac{1}{m'} \sum_{i=1}^{m'} \boldsymbol{x}'_i \quad (1)$$

as the input image, where m' is the number of input images. For example, the similarity measure in CLAFIC is defined as

$$S_{CLAFIC}(V_1, \dots, V_l; \boldsymbol{x}') = \sum_{i=1}^l (V_i \cdot \boldsymbol{x}')^2, \quad (2)$$

when a single input is assumed, where V_i is the i th eigenvector of a covariance matrix and l is the number of dimensions of the subspace. However, when multiple data input is assumed, the similarity measure is defined as

$$s_{mean}(V_1, \dots, V_l; \boldsymbol{\mu}') = \sum_{i=1}^l (V_i \cdot \boldsymbol{\mu}')^2, \quad (3)$$

using the sample mean.

3.2 Multiple Potential Function Classifier

The definition of a conventional *potential function classifier* (PFC) [9] is

$$S_{PFC}(T; \boldsymbol{x}') = \sum_{i=1}^m k(\boldsymbol{x}_i, \boldsymbol{x}'), \quad (4)$$

where $k(\cdot, \cdot)$ is a bell-like nonlinear function called the potential function or kernel, \boldsymbol{x}' is an input image, and m is the number of training samples. The templates of the potential function classifiers are themselves training samples. The classifier is simply extended under the COBS as

$$s_{MPFC}(T; \boldsymbol{x}'_1, \boldsymbol{x}'_2, \dots, \boldsymbol{x}'_{m'}) = \sum_{j=1}^{m'} \sum_{i=1}^m k(\boldsymbol{x}_i, \boldsymbol{x}'_j). \quad (5)$$

In this paper, the extended form of the potential function classifier is called a *multiple potential function classifier* (MPFC) in the COBS.

3.3 Mutual Subspace Method

Yamaguchi demonstrated the effectiveness of the MSM for facial image recognition experimentally [2]. In the MSM, PCA is performed on images obtained by continuous observations and "MSM similarity" is measured. MSM similarity is defined by the angle between two subspaces of the input images and training images. The two subspaces used for the MSM similarity measurement are computed from continuous observation data or preregistered training data.

In the learning phase of the MSM, the [basis'OK?] obtained from the training data is registered as a template. (This is the same as in the conventional subspace method.) In the recognition phase,

1. the PCA basis is calculated from the input images.
2. the following matrix is calculated.

$$Z = (Z_{ij}) = \sum_{k=1}^M (\mathbf{V}_i \cdot \mathbf{V}'_k)(\mathbf{V}'_k \cdot \mathbf{V}_j), \quad (6)$$

where \mathbf{V} is the basis of the training data subspace and \mathbf{V}' is the basis of the subspace obtained from the input data.

3. The maximum eigenvalue of Z , which is the angle between two subspaces, is obtained [10].

The maximum eigenvalue of Z is regarded as the similarity measure $s_{MSM}(V_1, \dots, V_l; V'_1, \dots, V'_{l'})$, where l is the number of dimensions of the subspace of the template and l' is the number of extracted features. The MSM therefore consists of continuous observation samples. This ensures the method has robust capability for object image recognition because it is easy to obtain several images by observing the object in the form of a motion image sequence.

3.4 Kernel Mutual Subspace Method

In this section, we describe the *kernel mutual subspace method*. The effectiveness of the method was demonstrated in facial image recognition experiments, which were reported previously [3].

First, we must briefly review kernel principal component analysis (KPCA). KPCA is performed by carrying out singular value decomposition in functional space \mathcal{F} for a given set of data $\mathbf{x}_i, i = 1, \dots, m$ in the n dimensional feature space R^n . We can define the functional space \mathcal{F} , which is related to the feature space, by a possibly non-linear map

$$\Psi : \mathcal{R}^N \rightarrow \mathcal{F}, \mathbf{x} \rightarrow \mathbf{X}. \quad (7)$$

Note that the functional space \mathcal{F} may have large, possibly infinite dimensionality.

In functional space \mathcal{F} , the covariance matrix takes the form

$$\bar{C} = \frac{1}{m} \sum_{j=1}^m (\Psi(\mathbf{x}_j)\Psi(\mathbf{x}_j)^T). \quad (8)$$

The basis of KPCA is given by diagonalizing the matrix. However, the matrix is too large (sometimes infinitely so) to compute. We use the $m \times m$ kernel matrix

$$K_{ij} = \Psi(\mathbf{x}_i) \cdot \Psi(\mathbf{x}_j), \quad (9)$$

to solve the eigenvalue problem,

$$m\lambda\alpha = \alpha K \quad (10)$$

for non-zero eigenvalues. The α_i denotes the column vector with entries $\alpha_1, \dots, \alpha_m$.

To extract the principal component, we need to compute the projection onto eigenvectors V in \mathcal{F} . Let \mathbf{x}' be the input feature vector, with an image $\Psi(\mathbf{x}')$ in \mathcal{F} . Then

$$V \cdot \Psi(\mathbf{x}') = \frac{1}{\lambda} \sum_{i=1}^m \alpha_i \Psi(\mathbf{x}_i) \cdot \Psi(\mathbf{x}') \quad (11)$$

may be called the nonlinear principle components corresponding to Ψ . Since map $\Psi(\cdot)$ makes a very large, possibly infinite, functional space, the cost of computing the dot product is extremely large (or infinite). To avoid this problem, Schölkopf introduced the Mercer kernel, which satisfies

$$k(\mathbf{x}, \mathbf{y}) = \Psi(\mathbf{x}) \cdot \Psi(\mathbf{y}). \quad (12)$$

When we use this kernel function, the computation of the dot product $\Psi(\mathbf{x}) \cdot \Psi(\mathbf{y})$ replaces the computation of function $k(\mathbf{x}, \mathbf{y})$. That is,

$$V \cdot \Psi(\mathbf{x}) = \frac{1}{\lambda} \sum_{i=1}^m \alpha_i k(\mathbf{x}_i, \mathbf{x}). \quad (13)$$

This result shows that we can calculate a projection for the nonlinear principal components in finite time without an explicit form of the nonlinear basis V .

Now we can describe our proposed kernel mutual subspace method (KMS), which combines MSM and KPCA. We first define a similarity measure for the KMS in functional space \mathcal{F} . Practical applications demand lower computational costs. Therefore, we must prove that the proposed method takes a finite time to compute the angle in functional space \mathcal{F} .

Let W be an eigenvector calculated from continuous images input into the recognition system. Then, we can describe V and W by

$$\begin{aligned} V &= \sum_{i=1}^m \alpha_i \Psi(\mathbf{x}_i) \\ W &= \sum_{j=1}^{m'} \alpha_j \Psi(\mathbf{x}'_j), \end{aligned} \quad (14)$$

where m and m' are the number of samples for training and test images. The similarity measure is then computed by dot product $V \cdot W$:

$$V \cdot W = \sum_{i=1}^m \alpha_i \Psi(\mathbf{x}_i) \cdot \sum_{j=1}^{m'} \alpha'_j \Psi(\mathbf{x}'_j) \quad (15)$$

$$= \sum_{i=1}^m \sum_{j=1}^{m'} \alpha_i \alpha'_j \Psi(\mathbf{x}_i) \cdot \Psi(\mathbf{x}'_j). \quad (16)$$

If we substitute (12), the equations can be written as,

$$V \cdot W = \sum_{i=1}^m \sum_{j=1}^{m'} \alpha_i \alpha'_j k(\mathbf{x}_i, \mathbf{x}'_j). \quad (17)$$

Because the numbers of m and m' are limited, this form shows that this method takes a finite time to compute the dot product of the basis of two subspaces. The method for obtaining the angles between two subspaces can be derived by substituting (17) into (6).

4 Experiment

We compared the classifiers described above in facial image recognition experiments. We employed CLAFIC as a reference recognition method when single image input was assumed because the subspace method is regarded as the standard object recognition method [7,8].

We used 15 individuals from the UMIST data set [11]. The data had a nonlinear structure, as shown in Fig. 2. The facial images were manually segmented and normalized to 15×15 pixel images (225 dimensionality vector). The number of training images was 10 per person, and 599 test images were created from the remaining data. We determined $m' = 5$ in this experiment.

We used a Gaussian radial basis function as a nonlinear kernel function for the KMS and MPFC. The kernel parameter and the number of dimensions of subspaces were selected based on the results of a preliminary experiment. Experiments were performed to compare the effectiveness of using CLAFIC, the sample mean, the MPFC, the MSM, and the KMS. The results are shown in Table 1. These results showed the effectiveness of continuous observations. They also showed that:

- in this case, redundant information was apparently needed to improve the accuracy of the linear methods (sample mean, MSM).
- unexpectedly, the accuracy of the sample mean method was high. This result is inconsistent with Yamaguchi's results [2]. We think the inconsistency was caused by the non-linearity of the distribution. Linear PCA fails when the distribution has a nonlinear structure.
- the MPFC was the least accurate method.

Table 2 lists the properties of the methods.

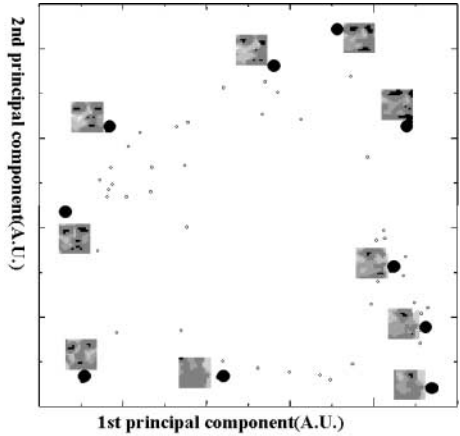


Fig. 2. Scatter plot of facial images in the UMIST data using PCA

Table 1. Recognition rate for each method

method	CLAFIC	Mean	MPFC	MSM	KMS
accuracy(%)	95.7	99.8	94.0	99.6	100.0
# of dim(train.)	5	5	-	6	3
# of dim(input)	-	0	-	2	2

5 Conclusion

We have described a *continuous observation-based system* for object image recognition and some classifiers used in the system. We also clarified the properties of the classifiers in an experimental comparison. The experimental results show the effectiveness of the classifiers in the COBS in terms of recognition accuracy. We believe the COBS offers a new research direction for statistical pattern recognition.

In future work, we will introduce other classifiers under the COBS and clarify their properties.

Acknowledgment

We are grateful to Prof. N.S. Allinson of the University of Manchester Institute of Science and Technology for permitting us to use the UMIST face database.

Table 2. Properties of each classifier

Method	Accuracy cal. cost.		comments
CLAFIC	low	low	single observation
Sample Mean	high	low	most simple
MPFC	low	high	easy training
MSM	high	mid	redundant expression
KMS	high	high	compact expression

References

1. Ken-ichi Maeda and Sadaichi Watanabe, "Pattern Matching Method with Local Structure", Trans. IEICE(D), Vol. 68-D. No. 3, pp. 345-352(1985) (in Japanese) [799](#)
2. O. Yamaguchi, K. Fukui and K. Maeda, "Face Recognition using Temporal Image Sequence", In Proc. IEEE 4thtl. Conf. on Face and Gesture Recognition, pp. 318-323 (1998) [799](#), [801](#), [803](#)
3. H. Sakano, et. al., "Kernel mutual subspace method for robust facial image recognition", in Proc. IEEE Intl. Conf. of Knowledge Engineering System, pp.245-248, (2000) [799](#), [801](#)
4. H. Sakano, "Kernel Mutual Subspace Method for Object Recognition", Trans. IEICE(D-II), Vol. J84-D-II, No. 8, pp. 1549-1556, (2001) (in Japanese) [799](#)
5. B. Schölkopf, et al., "Nonlinear component analysis as a kernel eigenvalue problem", Neural Computation, Vol. 10, pp. 1299-1319 (1998) [799](#)
6. S. Watanabe and N. Pakvasa, "Subspace method of pattern recognition", Proc. 1st IJ CPR, pp. 25- 32 (1973) [799](#)
7. M. Turk and A. Pentland, "Recognition Using Eigenface", Proc. CVPR, pp. 568-591 (1991) [799](#), [803](#)
8. H. Murase and S. K. Nayer, "Visual learning and recognition of 3-D object from appearance", International Journal of Computer Vision, Vol. 14, pp. 5-24, (1995) [799](#), [803](#)
9. M. A. Aizerman, et. al., "Theoretical foundations of the potential function method in pattern recognition learning", Automation and Remote Control, Vol. 25, pp.821-837, (1964) [800](#)
10. F. Chatelin, "Valeurs propres de matrices", Masson, Paris (1988) [801](#)
11. D. B. Graham and N. S. Allinson, "Characterizing Virtual Eigensignatures for General Purpose Face Recognition", in H. Wechsler, et al. ed. "Face Recognition From Theory to Applications", Springer Verlag, (1998) [803](#)