AN APPROACH TO REAL-TIME PATTERN RECOGNITION OF IMAGES BY AISS

Tsutomu Shibata Shigesumi Kuwashima
OKK INC.
Oh-yoh Keisoku Kenkyusho

22-3 Kitasenzoku 3-Chome, Ohta-Ku,
Tokyo, Japan

ABSTRACT

AISS (Adaptive Intelligent Sensing System), a real-time pattern recognition system with learning capability has been developed by OKK INC.
AISS was designed on the concept that indispensable keys for intelligent sensing are the following three.

a) Real-time processing
b) Adaptivity
c) Real-time monitoring capability

AISS processes images according the following 4 steps.

STEP 1) Feature extraction by using primitive masks based on Nth-order auto-correlation. ($N \leq 2$)
STEP 2) Learning by supervised teaching
STEP 3) Analyzing
STEP 4) Discriminating

The new hardware gives a real-time feature extraction and a real-time decision based on discriminant-algorithm at the rate 30Hz. And new parameters are presented. They are useful to monitor the condition of the systems.
We are going to apply AISS to various fields particularly pattern recognition of movement in the near future.
In this paper an outline of hardware and software of AISS are described.

INTRODUCTION

Neuro Computers, Fuzzy systems and Expert Systems have been under focus these years. However it is difficult to evaluate them.
It is to be said that they seem to have an endless potential of capability in the future. Is it true? We throw doubt on their future in the following points they have.

1) Lack of time(=cost) scale
2) Vagueness of evaluation criterion
3) Lack of generality of applications

Because Pattern Recognition is one of the intelligent sensing systems and we (OKK INC., an R & D company for sensing systems and sensors.) think that indispensable keys for intelligent sensing are the following three.

a) Real-time processing
b) Adaptivity
c) Real-time monitoring capability

AISS was developed on this concept and the system is an approach to real-time pattern recognition.
In this paper we introduce principles, hardware and software of AISS and emphasize the capability for application with moving images.

SYSTEM

This system (Fig. 1) consists of AISS, an analogue RGB monitor, an industrial television camera (CCD) and a host computer. A videodisc recorder is optionally available to store up to a lot of images and to read them automatically.

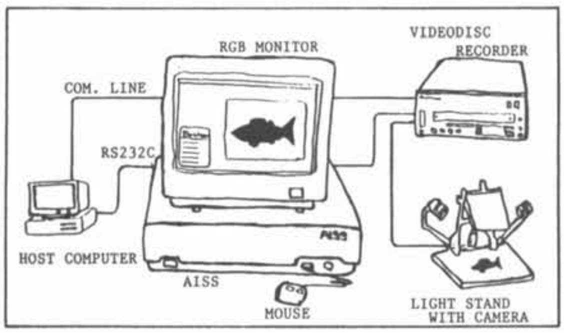

FIG1. SYSTEM CONFIGURATION

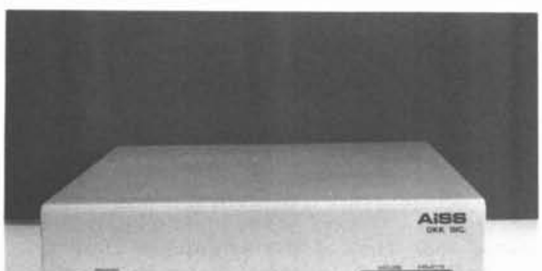

Fig.2 PHOTOGRAPH OF AISS

This work was carried out under the cooperation with ETL and finantially supported in part by the grant of the City of Tokyo.

PRINCIPLES

The hardware of AISS scans the area within a window of a digitized sample picture by 25 different mask patterns and counts the number of positions where the picture matched with the mask patterns. This process gives 25 parameter values for one picture. In other words one 25-dimension vector per picture. So we assign m different pieces of data each having n-parameters(n=25) to classes(categories) of data in the learning step. We express one data as n-dimension vector x,m-pieces of data set as m x n matrix A and the class number set as m-dimension vector m. For example xi belongs to No.mi-class. The problem is when A and m are given to infer to which class an unclassified data vector x of a digitized picture belongs. The important point with this kind of problem is to recognize the contents and limit of the given data when classifying new image data without letting it exceed those limits. In the following explanations we refer to the analyzing and knowing process of the limits of data A and m as the "analyzing step" and the inferring process of a class of unknown data x as the "discriminant step".

In the analyzing step we limit the space to that being spanned by all data vectors by applying PCA(principal component analysis) to A (without using class information m). Thus we can find the least dimensional space that can describe data vectors.
PCA is executed as follows. (Ref.2,3)

1. Produce deviation data matrix

$$A_1 = \begin{pmatrix} x1^t - \bar{x}^t \\ \cdots\cdots \\ xn^t - \bar{x}^t \end{pmatrix} \qquad x^t = \text{transpose of } x \qquad \bar{x} = \text{mean vector}$$

2. Get principal component vector (axis) 11,....,lp (p<n) satisfying following condition

$$Z^t = \begin{pmatrix} z1^t \\ \cdots \\ zn^t \end{pmatrix} = LA_1{}^t = \begin{pmatrix} 11^t \\ \cdots \\ lp^t \end{pmatrix} A_1{}^t$$

$Z^tZ=R$ (p x p diagonal matrix)

$LL^t=Ip$ (p x p unit matrix)

The scattergram of first and second components makes the data-distribution clearer than that of the original.
We define Mahalanobis distance matrix M and class separation Sp as follows:

$$M = \{a_{ij}\},\ a_{ij} = (\bar{x}i - \bar{x}j)^t Sj (\bar{x}i - \bar{x}j)$$

Sj is covariance matrix of class j

$\bar{x}i$ is mean vector of class i

$\bar{x}j$ is mean vector of class j

$$Sp = \frac{1}{2 \times kC_2} \Sigma a_{ij}$$

k: number of classes

a_{ij} is invariant after any kind of linear transformation C
y=Cx (C is full-rank matrix)
Therefore Sp is constant too. The only thing we can do is preventing it from being reduced when we reduce the data-dimension.
We use the following steps to find linear transformation. The purpose is to get both as little separation within classes as possible and as much separation between different classes as possible. (Ref.1,5)

1. Apply Mahalanobis transformation to Z

   $Z2^t = C^tZ^t$

   C is determined so that Euclidean distance in Z2-space equals Mahalanobis distance in Z.

2. Apply PCA to Z2 so that we get more separation between different class-mean vectors.

   $Y^t = B^tZ2^t$ (B: p x q, $B^tB=Iq$)

   Y is not orthogonal transformation because C is not.

In the discriminant step we
1. judge whether or not unknown vector x belongs to space spanned by all vectors given in the learning step.
2. search the class whose mean vector is nearest to x.
3. answer the inferred class with confidence rate Blg.

At the beginning of the analyzing step, we use PCA.
The purpose is to divide space into on-stage space and off-stage space. On-stage space is spanned by A whose all vectors are given in the learning step. Off-stage space is complement space to on-stage space.
Unknown vector x is represented by

$$x = xon + xoff$$

(xon : on-stage component
xoff : off-stage component)

We introduce the confidence rate

$Blg = ||xon||^2/||x||^2$ .... formula 1

PCA constructs the axis so that $\Sigma\ ||xioff||^2$ is minimal. Therefore most of the learning data x's can be thought to belong to

P_8	P_7	P_6
P_5	P_0	P_4
P_1	P_2	P_3

Fig.3 3x3 MASK

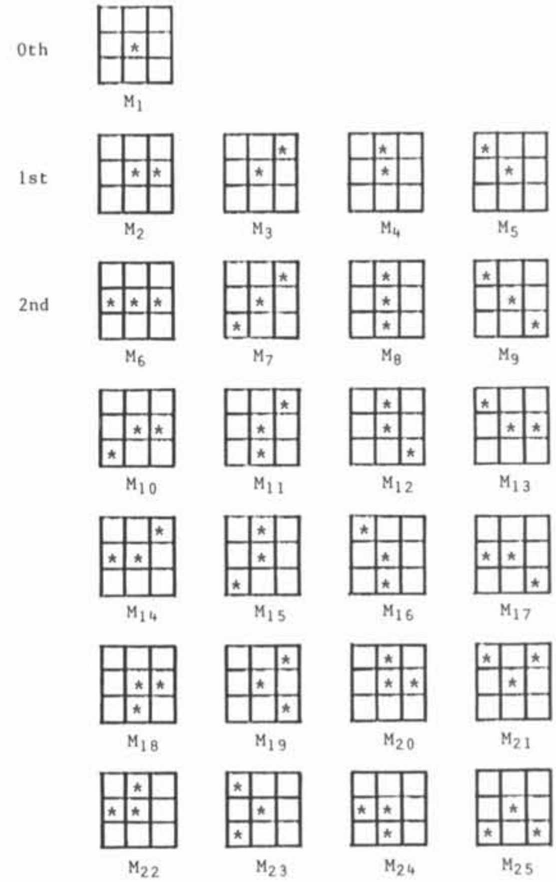

Fig.4 Nth-ORDER AUTOCORRELATION MASKS (N=0,1,2)

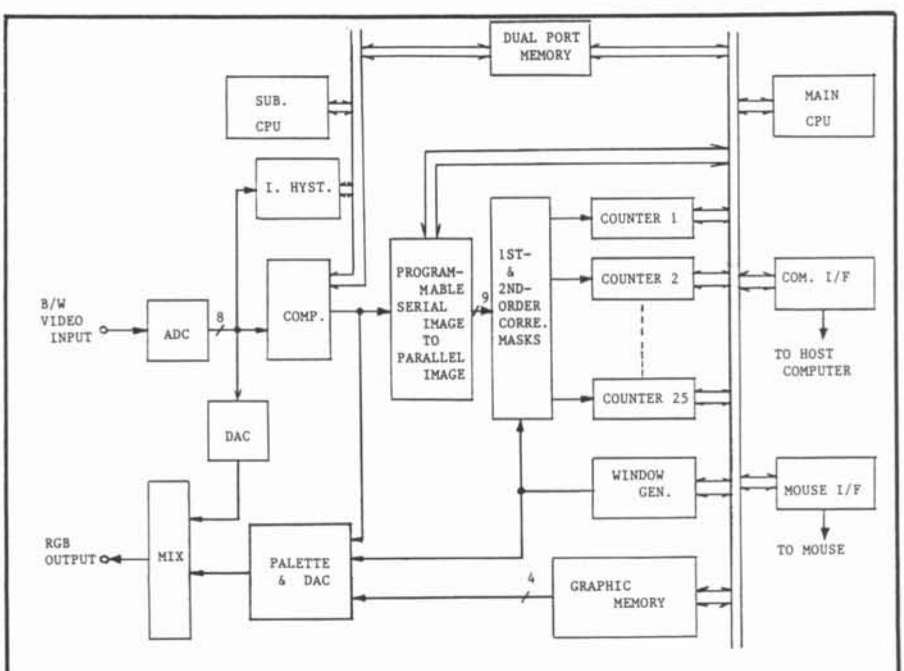

Fig.5 AISS BLOCK DIAGRAM

on-stage space. In the rest of our discussion we infer using xon and ignore xoff, so the check of formula 1 is inevitable.
Next by z=Lx, we get z, the coordinate of xon on on-stage space and by $y=B^tC^tz$ we get transformation of z to Y-space.

We prepared yi in the transformation of class mean xi. Therefore we can infer the class j so that ||y-yi|| is the minimum and the confidence rate is Blg.

HARDWARE

SPECIFICATION

Video Input	:	RS170
Image Resolution	:	320x240 total graphic coordinates 256x192 effective area
Number of Masks	:	25
Mask Size	:	3x3, 5x5, 7x7
Pattern of Masks	:	0th-, 1th-, 2nd-order Autocorrelation masks
Processing Rates	:	30Hz
Monitor Output	:	R,G,B,Sync.
Communication Port	:	RS232C
Threshold Selection	:	Manually adjustable, Adaptive (to be implemented.)

BLOCK DIAGRAM

Fig.5 shows the block diagram of AISS.

SERIAL TO PARALLEL UNIT (Fig.6)

Local images of 7x7, totally forty-nine pixels, are cut out from the image one after another according to the system clock, and then nine pixels of them are selected to be outputted parallel. The selection from three sizes (3x3, 5x5, 7x7) with similar squares is possible. (Fig.3)

MATCHING UNITS

The nine parallel pixels are matched with the masks (Fig.9) respectively and successively. The results of matching are calculated to the counters only when the centor of the pixels is within the window.

MASKS (Fig.4)

We adopted a simple mask based method which relies on Nth-order Autocorrelation for a feature extraction. The masks are orthogonal to each other and invariant against parallel displacement. And the result of the extraction method is an efficient reduction of the image information. The discussion about the masks is presented in detail in Ref.4.

WINDOW GENERATOR

Images are processed within the window. The window generator generates a window of any rectangular shape located anywhere inside the effective picture area. And it is possible to scan a window of the sequence of the video frame rate.

SOFTWARE IMPLEMENTED IN AISS

AISS is a multi-window system with a pointing device (mouse) and menus. Fig.8 shows the main menu. On choosing an item from the main menu, a sub menu is activated for initial setting, sampling, execution, testing or communication. These operations can also be controled by commands from a host computer. Therefore interactive systems with a complex algorithm can be realized, combining an AISS and a computer.

SOFTWARE

Fig.7 shows the algorithm of the software.

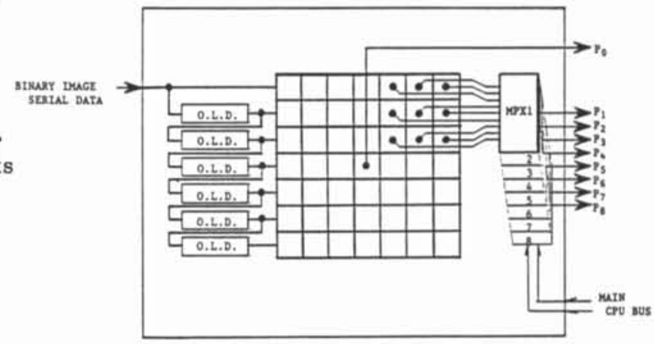

Fig.6 PROGRAMMABLE SERIAL IMAGE TO PARALLEL IMAGE UNIT

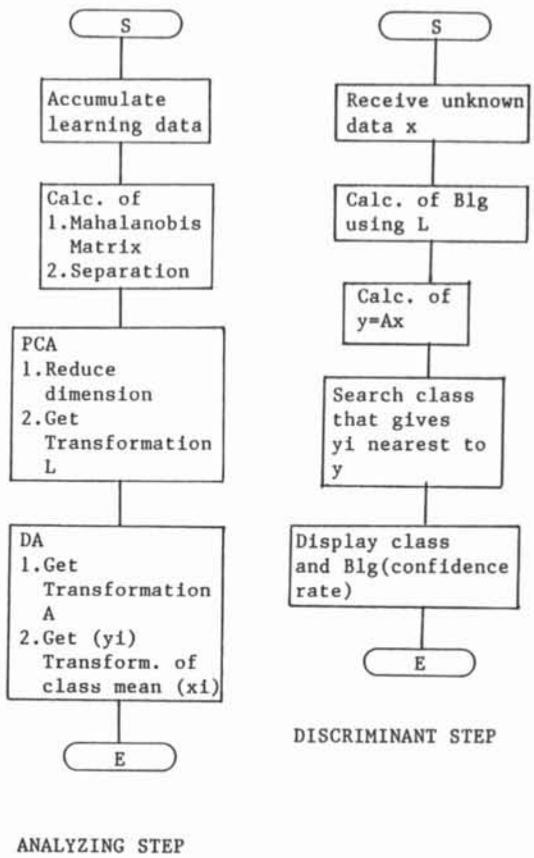

Fig.7 ALGORITHM

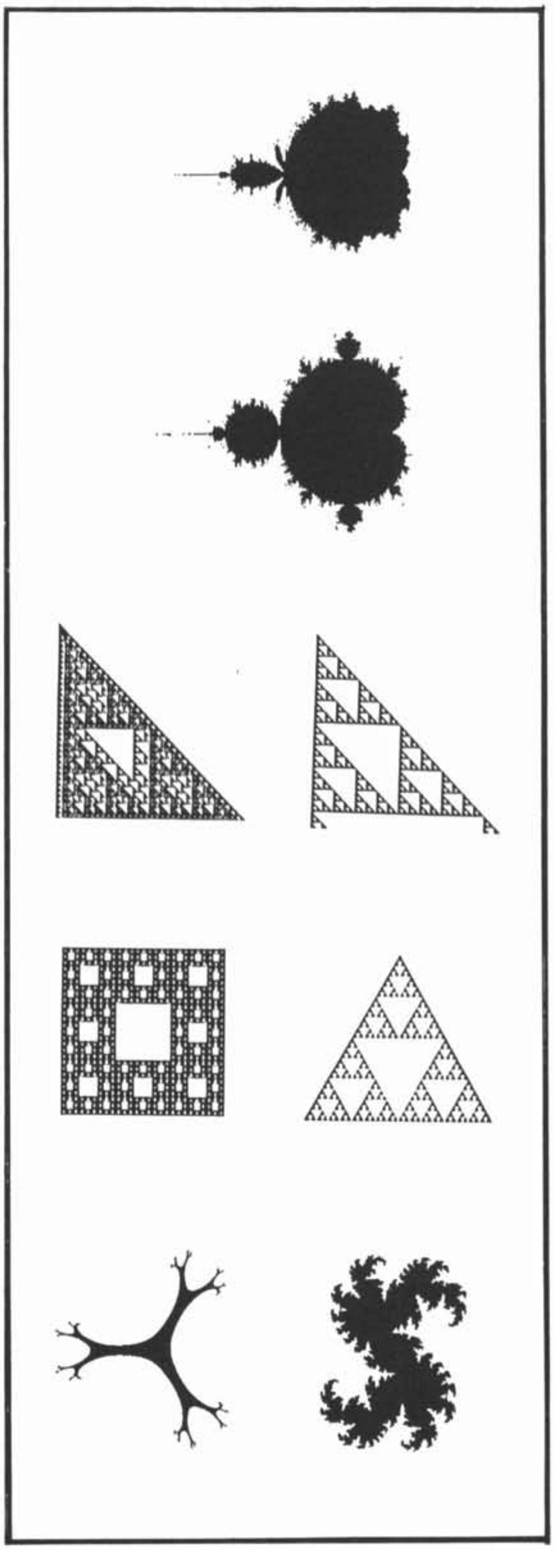

Fig.9 EXAMPLE 1,FRACTAL PICTURES
(drawn by programs in Ref.6)

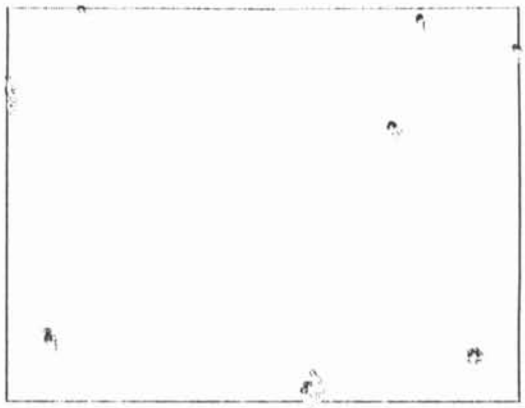

Fig.8 RESULT OF PCA(FRACTAL PICTURES)

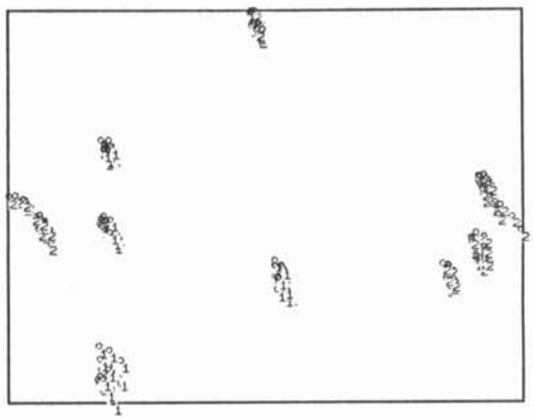

Fig.10 RESULT OF PCA(FACES)

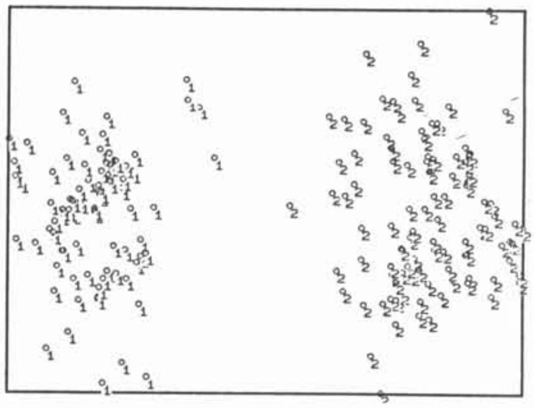

Fig.11 RESULT OF DA(FACES)

EXPERIMENT AND RESULTS

We applied our algorithm to two examples in order to confirm the effectiveness of the algorithm.

In general good separation can be achieved with applying only PCA with i picture = 1 class as shown in **Fig.8**. We treat 8 fractal pictures, inputting each 20 times, slightly rotating them (a few degrees) each time. (Sp=2826742239.7 min(a_{ij})=7439) Even if Con < 0.99999 (Con=accumulated contribution rate), the number of dimensions can be reduced to 4. This makes DA(discriminant analysis) simple and easy. In the discriminant step, Blg rejects data efficiently which is far from the learning vectors due to reduced dimensions.

When one class contains plural pictures, PCA alone does not give good separation. We use many axes by letting Con(the accumulated contribution rate)=1 and apply DA to get good separation. The result is good separation but the Blg system lacks trustworthiness due to the Blg being close to 1 most of the time. However the Sp value in the analysing step(Sp=157.9, min(a_{ij})=86) has warned us that many dimensions are required even though good separation could be achived. **Fig.10** and **Fig.11** show the scattergram of the pictures of faces, **10** is the result of PCA and **11** the result of DA.

Class 1 represents faces with glasses.(Fig.12)
Class 2 represents faces without glasses.(Fig.13)

As you can see, in the first example the Blg-system performs well but not in the second. However the irreliable performance is predictable in the analyzing step by monitoring the Sp value or the min(a_{ij}) value. This means the Sp-Blg system is honest to learning data information. Therefore Sp and Blg can be considered useful monitoring criterion in real-time pattern recognition.

CONCLUSION

We developed AISS. The new hardware gives a real-time feature extraction and a real-time decision based on discriminant-algorithm at the rate 30Hz. According to this capability AISS can process moving images.
And we presented new parameters, a criterion value (Sp) and a confidence rate (Blg) in this paper. The parameters are useful for pattern recognition systems to monitor the condition of the systems. Therefore they help users to judge whether AISS is good for an application the users are being faced with.

At the view point described in Introduction, we are going to apply AISS to various fields particularlly pattern recognition of movement in the near future.

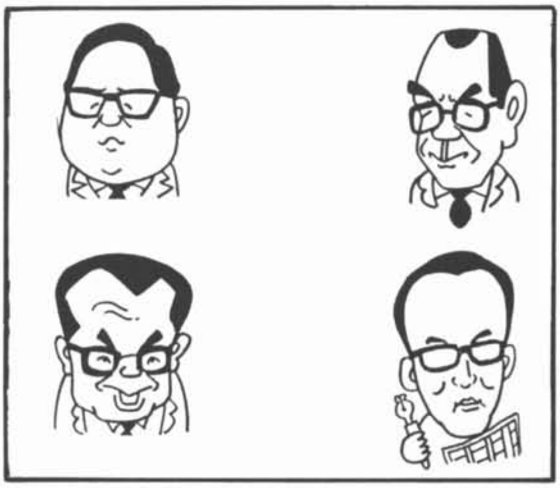

Fig.12 EXAMPLE 2, CLASS 1(FACES WITH GLASSES)

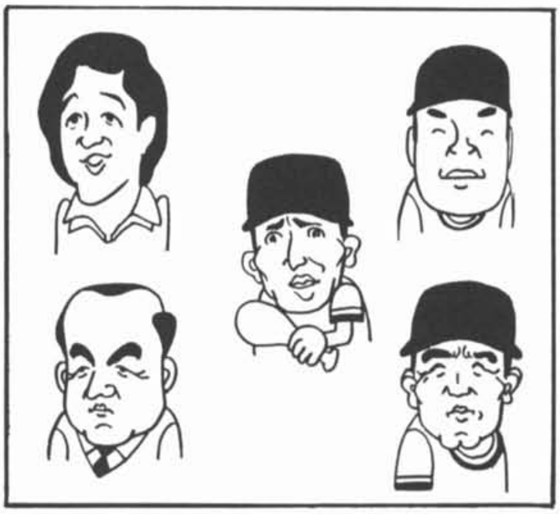

Fig.13 EXAMPLE 2, CLASS 2(FACES WITHOUT GLASSES)

REFERENCE

1. Otsu.N,Mathematical Studies on Feature Extraction in Pattern Recognition, Researches of the Electrotechnical Laboratory, No.818, July 1981 (in Japanese)
2. Oja.E, Subspace Methods of Pattern Recognition, John Wily & Sons, 1983 (Japanese translation)
3. Enkawa.T, Multivariate Data Analysis, Asakura Shoten, 1988 (in Japanese)
4. Otsu.N et al., A Method of Pictorial Feature Extraction Based on Nth-Order Autocorrelation Masks, PRL78-31, Society of Electronics and Communication (in Japanese)
5. Otsu.N et al., A New Scheme for Practical Flexible and Intelligent Vision Systems, IAPR Workshop on Computer Vision, 1988 (to be presented)
6. Tsukuda.T, World of Fractal, Sankaido, 1987