SICE Annual Conference in Fukui, August 4-6, 2003
Fukui University, Japan

A Gait Recognition Method using HMM

Kota Iwamoto, Kazuyuki Sonobe, Naohisa Komatsu
Waseda University, Okubo 3-4-1, Shinjyuku-ku, Tokyo, Japan

Abstract: New biometrics identification methods, including person identification using a walk characteristic (Gait Recognition) is attracting attention. In this paper, an identification method on Gait Recognition is proposed. In the proposed method, outlines of person are represented by P-style Fourier Descriptions, and these time-serial walking characteristics are using HMM. It is shown that the proposed method brings the high rate of identification and is effective in indoor environment with simulation results.

Keywords: Gait Recogniton, Person Identification, HMM(Hidden Markov Model), PCA(Principal Component Analysis), Biometrics

1 Introduction

It follows on progress of a rapid information society in recent years that a close-up of a security issue comes to be taken greatly. Then, the necessity for the authentication of an individual in access control, a machine interface, etc. has been increasing. As the technique of individual authentication, biometrics authentication is performed actively.

Among them, the individual identification using a person's walking characteristics are observed as completely new biometrics (below, Gait Recognition)[1, 2]. Gait Recognition is performed a discrimination by catching the person's walking characteristic from a person's walking image from the camera. It has the advantages which are not in other biometrics mentioned below.

- It is uncontacting, and a user is not forced special operation for identification.
- A user is not conscious of being recognized, since identification is performed at the time of the usual walking operation.
- Identification can be performed from a long distance.
- It is easy to acquire the data since man usually needs to walk at the time of movement.

As mentioned above, application of Gait Recognition such as access control being friendly to a user, camera surveillance system, and criminal investigation can be considered.

However, the conventional Gait Recognition method has the problem such as, (1) The unit of identification is not defined. (2) The feature vector is not effective enough. (3) The method of identification doesn't reflect time-series nature of the walking characteristic. In this paper, Gait Recognition which has improved the issues mentioned above is proposed.

2 The summary of the proposal method

2.1 The feature and the flow of the proposal method

The feature of the proposal method is shown below.

The unit of identification: In order to clarify the starting point and the ending point of the object for identification, let one Gait cycle (1GC) be the unit of identification focusing on the periodicity of a walking.
The feature vector: The outline information on a silhouette image is used as a feature vector so that a person's form can be expressed efficiently.
The identification method: Focusing on time-serial transition of the feature parameter, we use HMM (Hidden Markov Model) which is robust to recognition of a time-series pattern.

The flow of the proposal method is shown in Fig.1. There is a training phase and a identification phase, and they process to every 1 gait cycle (1GC) in both phases. In a training phase, 1 GC is extracted from a walking image (GC extraction). And then the feature vector (outline information) is extracted from 1 GC, and used for generation of HMM. In a identification phase, in a similar way 1 GC is extracted from a walking image (GC extraction), and after transforming into the feature vector it is checked to the HMM which was computed in the training phase. Then, the result of identification for the walking cycle is obtained.

2.2 Gait Cycle (GC)

Murray *et al.* expressed Gait as "a total walking cycle"[4]. Namely, there is periodicity in a person's walking pattern (Gait), and it can be considered that walking operation is a periodic signal. As shown in Fig.2, period of walking (Gait Cycle : GC) can be defined as the time of heel strike between the same leg. Heel strike means the time of the heel's grounding on the ground.

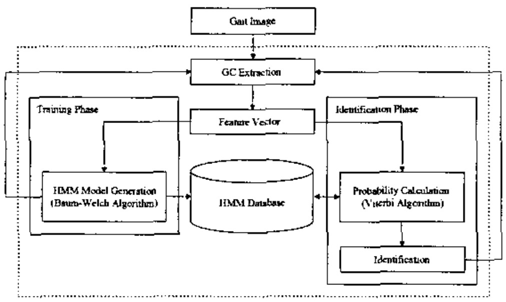

Fig.1: The flow of the proposal method

Then, 1 Gait Cycle (1GC) will be used as a unit to perform identification in this research. The advantage obtained by making 1 Gait Cycle the unit of identification is that

- A clear starting point and ending point of identification can be set, and exact identification is anticipated.
- A completely equal walking part can be extracted as a candidate for identification to every person.
- It can be always used as common identification unit without depending on the conditions such as time, place, and camera.

It is necessary to cut out every 1GC from each person's walking image (GC extraction) in order to use Gate Cycle. In a heel strike point, the interval d_{feet} between each leg will generally show the local maximum. In this research, we consider the local maximum of this d_{feet} as a heel strike point, and GC extraction is performed by this. The method of obtaining d_{feet} for an extraction person is shown in Fig.3. It is considered that the domain of the bottom one sixth of an extracted person's height is the leg region, and the horizontal distance between the left end and the right end in the leg region is defined as d_{feet}.

Fig.2: Gait Cycle(GC)

3 The feature vector

In this chapter, we propose the feature vector acquired in each frame of a walking image.

It cannot necessarily be said that the features which are used by the conventional researches about image

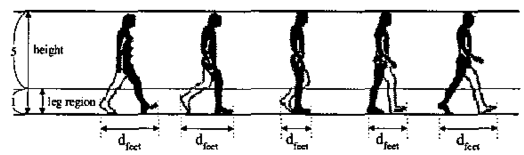

Fig.3: Method of obtaining d_{feet}

recognition are parameters suitable for Gait Recognition because the acquisition accuracy and the validity of the feature parameter is not satisfactory.

In the research in which an image itself is used as a feature, the luminosity information of the pixel in the square window centering on subject is used as the feature vector[2, 5]. All the information of the image such as the shade, form, and ups and downs are reflected in the feature vector since the image itself is defined as a feature parameter.

On the other hand, the optical flow data is used widely in fields such as gesture recognition as the feature parameter[1, 6]. Since the move direction and the speed of a subject are reflected in the feature vector, it is considered as the parameter which is effective in recognition of sequential images. But, as for optical flow data, there is a problem in accuracy of acquisition. And application to a real system is not easy since it takes a lot of time to operate.

There are also some researches using the mesh feature[7]. Although the mesh feature reflects the rough form of a subject, it is considered that the validity as a feature parameter is small.

Furthermore, for the three features mentioned above, processing such as setting the position of a subject in the window (centering), expansion, and reduction are required. Then they affect the acquisition accuracy of the feature parameter greatly.

3.1 Person's outline information

In this reserch, person's outline in each frame is used as the feature parameter. A person's outline shows the features, such as the man's form, on the whole. There is a fact that man can discriminate a person who is at a distance by watching his walking. So it is considered that a person's outline is sufficient information as the feature parameter of Gait Recognition. The advantages which make person's outline into the feature parameter in Gait Recognition are shown below.

- A person's outline is easily acquired by subtracting the background. Therefore, the accuracy and the stability of acquisition for the feature parameter are high.
- Redundant information is not included in the process of identification since we only consider the

shape information as a feature parameter.

- By the technique mentioned below, the setup of a window domain, processing such as position setting, expansion, and reduction are unnecessary.

A technique of describing closed curvilinear information including an outline of a subject efficiently, namely the P-style Fourier Description[3]. P-style Fourier Description is the method of describing the curve on the plane in the frequency domain. This can be obtained by Fourier deployment of the complex function which possesses the whole curved rate function of the curvilinear in the index part of an exponential function. The reproduction curve by the low-pass portion of Fourier Description is a visual good approximation to the original curve. And the pattern information for each class is gathered in the low-pass portion.

In this research, the expression of the feature vector using the transformation whose rate of compression is higher is proposed based on P-expression which is used for expressing the curvilinear of P-style Fourier Description. And it is used as a feature parameter in Gait Recognition. P-style Fourier Description is suitable for expressing arbitrary form. But in this research Karhunen-Loeve (KL) Expansion, which is the optimal rectangular for the compression to the given arbitrary data sets, is used since application is limited to a person's outline form.

The flow of the feature vector acquisition is shown in Fig.4. An outline curve is first calculated from a person's silhouette image. Next the vector represented with P expression is generated. In the training stage, the eigen vector is computed by Principle Component Analysis. In the identification stage, the feature vector is obtained by the mapping the vector expressed with P expression to the partial space using the eigen vector obtained by training.

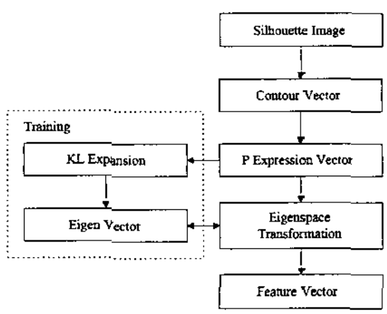

Fig.4: The flow of the feature vector acquisition

3.2 Acquisition of an outline information

In this section, the technique of acquiring a person's outline is explained.

A person's silhouette image is first acquired by subtracting the background. Next, the outline curve C is obtained by pursuing the outline of a silhouette image. Let the pixel of the topmost part of a silhouette image be the starting point of the pursuit of outline. The pixel of a person's topmost part in a walking image can be regarded as a starting point which is common to any person. It is because the topmost part is certain to correspond to the top of a head for every person. The pursuit of outline is started letting the coordinates of the starting point for pursuit on complex plane be $Z(0)$. The outline curve C is expressed as m arbitrary complex coordinates points

$$Z(0), Z(1), Z(2), ..., Z(m-1), Z(m) \quad (1)$$

where $Z(m) = Z(0)$.

Next, let δ denote the positive real number which is small enough and let n denote the positive integer which is large enough. The coordinate points of outline mentioned above are approximately converted into n coordinate points

$$z(0), z(1), z(2), ..., z(n-1), z(n) \quad (2)$$

where the length between contiguity coordinates serves as δ. This is for unifying the number of coordinates of each frame into n to every person. Thus,

$$\delta = |z(j) - z(j-1)| (j = 1, ..., n) \quad (3)$$

Here, the angle between each vector $z(j) - z(j-1)$ and the real axis is defined as $\theta(j)$ with counterclockwise direction positive as shown in Fig. 5. By using θ, complex function ω is defined as

$$\begin{aligned} \omega(j) &= \exp(i\theta(j)) & (4) \\ &= \frac{z(j+1) - z(k)}{\delta} & (5) \end{aligned}$$

and it is called "P-expression" of the outline curve C.

Thus, complex value by "P expression" $\omega(j)$ can be obtained from person's outline curve. By putting the real part and imaginary part of omega (j) in order one by one, $2n$-dimensional P expression vector $\boldsymbol{x}$

$$\begin{aligned} \boldsymbol{x} = \; & [Re(\omega(0)), Im(\omega(0)), Re(\omega(1)), Im(\omega(1)), \\ & ...Re(\omega(n)), Im(\omega(n))] \quad (6) \end{aligned}$$

is obtained.

3.3 Principle Component Analysis

It is difficult to use P expression vector x itself as a feature vector since its number of dimension is large and it has many redundancies. Therefore, P expression vector is compressed by Principle Component Analysis (PCA), and the feature vector $\boldsymbol{o}$ is obtained.

It is first necessary to calculate the eigen vector $\boldsymbol{v}$ in the training stage in order to perform the compression of

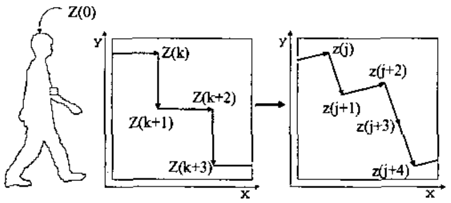

Fig.5: The acquisition method of a coordinates point

the datas by PCA. Now, as data for training, the datas (P expression vector x) of J frames per each person are prepared to I persons. Let xij denote the P expression vector of the jth frame for person i, then the following matrix X is formed.

$$\boldsymbol{X} = [\boldsymbol{x}_{1,1}, \boldsymbol{x}_{1,2}, ...\boldsymbol{x}_{1,J}, ..., \boldsymbol{x}_{i,j}, ..., \boldsymbol{x}_{I,J}] \quad (7)$$

This X contains a total of $I \times J$ elements of a P-expression vector $\boldsymbol{x}$. Let $\boldsymbol{m}$ denote an average vector, then $\boldsymbol{\Sigma}$ which is covariance matrix of $\boldsymbol{X}$ is calculated as

$$\boldsymbol{\Sigma} = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} (\boldsymbol{x}_{i,j} - \boldsymbol{m})(\boldsymbol{x}_{i,j} - \boldsymbol{m})^T \quad (8)$$

An eigen equation given by

$$\boldsymbol{\Sigma} \boldsymbol{v}_k = \lambda_k \boldsymbol{v}_k \quad (9)$$

is solved, and eigen vectors $(\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_k)$ corresponding to k large eigen value $(\lambda_1 \geq \lambda_2 \geq ..., \geq \lambda_k)$ are calculated. Then, the feature vector $\boldsymbol{o}$ is calculated by the formula below from P-expression vector $\boldsymbol{x}$.

$$\boldsymbol{o} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, ..., \boldsymbol{v}_k]^T (\boldsymbol{x} - \boldsymbol{m}) \quad (10)$$

4 GC-HMM

In this chapter, the technique of performing identification of person with the observed feature vector series $\boldsymbol{O}$ is described.

It is considered that the individual characteristics of the man's outside such as his form is also contained in the feature vector $\boldsymbol{o}$ which shows outline information. However, if we consider the walking characteristic (Gait) as the time-serial operation, it hopes that individual nature appears more greatly in the situation of time-serial transition of the feature vector.

Then, in this research, the individual nature in the walking characteristic is regarded as time-serial transition of person's outline. By doing so, the probable change on the time-axis of the feature parameter can be reflected on identification together with the individual nature of information of person's outline itself. Therefore, it can be expected that individual characteristics can be extracted clearly.

In this research, the method of identification using Hidden Markov Model (HMM) which is robust to time-series pattern recognition will be proposed. HMM is mainly used for recognition of voice or signature. However, it is also used for sequential image recognition such as gesture recognition[6] and recognition of a tennis stroke[7], and they have performed the good results.

First, a silhouette image is obtained from a camera by person's walking operation, next the feature vector $\boldsymbol{o}$ is computed. Then, the observed feature vector series $\boldsymbol{O}$ is acquired. Thus, the walking motion can be assumed as a HMM model which outputs the observed feature vector. It is considered that every person owns HMM which outputs the feature vector $\boldsymbol{o}$ with original probability. It is called GC-HMM, since this HMM corresponds to 1GC. GC-HMM is generated to every person, and a person is discriminated by using it.

In addition, in this research, Left-to-Right model is used as HMM because of irreversibility of time. Left-to-Right model is generally used for speech recognition and so on since there is a characteristic that the state of a system goes to the right from the left in a Left-to-Right model.

Training: In the training process, the probability for the change of state A and the probability for the output of symbol B of GC-HMM are presumed for every person, and then GC-HMM is generated. Here, let GC-HMM be set to λ and let a person be set to x. If the GC-HMM of every person is expressed as λ^x, GC-HMM corresponding to every person is described as

$$\lambda^x = (A^x, B^x) \quad (11)$$

It is necessary to estimate the probability for the change of state A and the probability for the output of symbol B which are the internal parameters. This is the problem of the model presumption for HMM. When the feature vector series for training $\boldsymbol{O}$ and the model λ are given, the model parameters A and B are calculated so that the probability $P(\boldsymbol{O}|\lambda)$ becomes the maximum. $P(\boldsymbol{O}|\lambda)$ shows the probability that the feature vector series $\boldsymbol{O}$ will be outputted from the model λ.

The Baum-Welch algorithm (the method of re-presuming the parameter) is used as a method of calculating the local optimal value of the parameter of the model λ which makes the probability for generation of an observed series the maximum. In the Baum-Welch algorithm, the value of the model parameter converges by calculating recursively with the re-presumption algorithm after giving the proper initial value to the probability for the change of state A and the probability for the output of symbol B which are the model parameters.

Identification: In the identification process, the result of identification (the optimal model) is delived by regarding the feature vector series $\boldsymbol{O}$ cut out by every 1GC as input. This is the problem of evaluation for HMM. It is necessary to calculate the probability that the feature vector series $\boldsymbol{O}$ is generated from model λ when the feature vector series $\boldsymbol{O}$ and model λ are given. And the person who has the model which shows the maximum similarity, i.e., for whom the probability which outputs the series is the highest (the optimal model) in regard to the input feature vector series $\boldsymbol{O}$, becomes the identification result $\hat{x}$.

$$\hat{x} = \arg\max_{x} P(O|\lambda^{x}) \tag{12}$$

Viterbi algorithm is used for calculation of probability $P(O|\lambda)$. It is the algorithm which calculates the optimal state series (the optimal course) $S = s_1, s_2, ..., s_T$ and the probability on the course.

5 Evaluation experiment

The experiments mentioned below were conducted based on the proposal method. In addition, the terms in experiments is shown in Table 1.

Experiment 1: The clothes of test data are the same as that of training data.

Experiment 2A: The clothes of test data are different from that of training data (indoor clothes).

Experiment 2B: The clothes of test data are different from that of training data (coat).

Table 1: The terms in experiments

HMM		Left-to-Right model
probability density function		mixed Gauss-distribution
P expression vector		4096 dimensions
feature vector		10,20,30,50,100 (dimensions)
the angle of walking for camera		0 degree (just right)
training data		21 persons(total 400GC)
test data	exp.1	21 persons(total 160GC)
	exp.2A	4 persons(total 36GC)
	exp.2B	2 persons(total 12GC)

5.1 Experiment 1

In the case where the clothes of test data are the same as that of training data, the effect of identification by HMM were investigated changing the value of the HMM parameter. The date and hour of shooting is also the same date as training data. The rate of identification was computed by changing the number of dimensions of the feature vector, the number of states of HMM, and the number of mixtures of Gaussian distribution (the number of mixed distributions).

The result when the feature vector is 20 dimensions and 50 dimensions is shown in Fig.7 and Fig.9 as example result. A person's outline curve reproduced from the feature vector $\boldsymbol{o}$ is shown in Fig.6 (when the feature vector is 20 dimensions) and Fig.8 (50 dimensions) together.

The rate of identification was low when the number of mixed distributions was 1 in every number of states. But the rate of identification became high by more than two states. Moreover, the comparatively high rate of identification could be obtained by choosing the suitable number of states and mixed distributions by the case of every dimension of vector. Especially, with the seven states, 21 persons were able to be discriminated among 21 persons by two or more mixed distribution numbers. As the general trend, it can be said that it is in the tendency which the rate of identification also goes up respectively as the number of vector dimensions, the number of states, and the number of mixed distributions go up. This is because every parameter can express the more detailed walking characteristic as a number goes up.

As mentioned above, it can be said that the validity of using HMM for identification of person by the walking characteristic was able to be confirmed.

Fig.6: Reproduced line by 20 dimensions

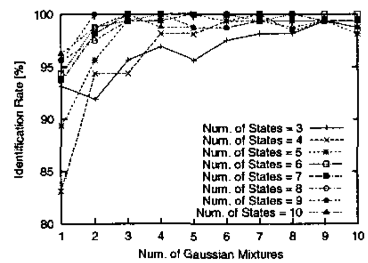

Fig.7: The rate of identification by 20 dimensions in experiment 1

Fig.8: Reproduced line by 50 dimensions

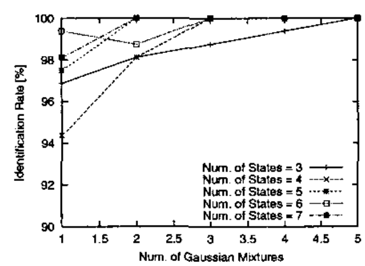

Fig.9: The rate of identification by 50 dimensions in experiment 1

5.2 Experiment 2

Next, in experiment 2, the rate of identification at the time of wearing different clothes from the data for training was computed in order to investigate the influence on identification by change of clothes. It was investigated about when wearing the interior clothes (T-shirt, sweater, or sweat shirt) considered that person's outline does not change much (exp.2A) and when wearing the coat considered that an outside changes sharply (exp.2B). For 10, 20, 30, 50, and 100 dimensions of the feature vector, estimation was performed with the number of states; 5 and the number of mixed distributions; 10 by which the especially good rate of identification was obtained in the experiment 1. As a result, the number of GC discriminated within each ranking among 36 GCs [2A] and 12 GCs [2B] in total for 21 candidates was counted. The result of experiments 2A and 2B respectively is shown in Table 2 and Table 3.

As it turns out when Table 2 is seen, the probability which is discriminated by 1 grade when the clothes differs from the data for training (the interior clothes) is low compared to experiment 1 (the same clothes as the data for training). However, it turns out that it is discriminated in general within 2 - 3 grade. This result means that the object can be extracted to quite high ranking by the case of wearing the interior clothes even when it differs from clothes of training data. Therefore the learned HMM can be said to be effective. This is considered to be because the fine changes of clothes have a tendency to be absorbed owing to using the terms of low dimensions of PCA as a feature vector. In addition, the difference in the tendency by the number of vector dimensions was not especially seen.

On the other hand, as it turns out when Table 3 is seen, in the case of having a coat on, the rate of identification was very low. It is considered that the value of the feature vector changes a lot and the effect of this technique using HMM did not appear since a person's form may change a lot clearly at the time of wearing a coat. Moreover, when a heavy clothes like a coat is worn, it is considered to be very difficult to perform Gait Recognition since the way of waving hands, how to walk, a step, etc. change a lot.

From the result mentioned above, it is considered that this proposal method is fully applicable in the indoor environment where a man does not have on clothes such as a coat by which his outline changes a lot.

Table 2: The result of expriment 2A [GC] (in 36 GCs)

a	b	c	top	within 2nd	within 3rd	within 5th
10	5	10	24	31	33	35
10	8	5	21	31	33	34
20	5	10	26	33	33	35
20	8	5	19	32	34	35
30	3	10	18	30	33	34
30	5	10	24	31	34	34
50	3	10	22	31	34	34
50	5	10	18	30	33	34
100	3	10	23	32	33	34
100	5	10	20	31	34	36

Table 3: The result of expriment 2B [GC] (in 12 GCs)

a	b	c	top	within 2nd	within 3rd	within 5th
10	5	10	1	4	5	7
10	8	5	0	0	2	6
20	5	10	0	6	7	8
20	8	5	0	0	0	5
30	3	10	0	1	5	8
30	5	10	0	3	3	8
50	3	10	0	2	6	8
50	5	10	1	2	6	8
100	3	10	0	0	0	8
100	5	10	0	2	5	9

(a: dimensions
b: states
c: mixed distributions)

6 Conclusion

In this paper, a new Gait Recognition method has been proposed. In particular, definition of 1GC as the unit of identification, the use of a person's outline as the feature parameter, and the technique of identification by using HMM to the person identification by the walking characteristic (Gait) have been carried out. And verification by evaluation experiment was performed, and as a result the validity of proposed method has been shown.

The future subjects are mentioned below. First, in this examination only the case when every person walks at the angle of 0 degree to a camera was examined. But in fact a person walks at all angles to a camera, so it is necessary to carry out correspondence to two or more walking angles. Next, the identification result was obtained to every 1GC and the rate of identification was computed in this examination. However, the information on two or more GCs can be acquired while a camera is catching and pursuing a walking person. Moreover, the acquisition technique of the feature vector which is not easily influenced of clothes or dimension selection of the feature vector are also needed since the rate of identification fell sharply when a coat was worn in this examination.

Acknowledgment

This research was supported (in part) by a Grant-in Aid for the WABOT-HOUSE Project by Gifu Prefecture.

References

[1] P.S.Huang, C.J.Harris, and M.S.Nixon, "Human gait recognition in canonical space using temporal templates" IEE Proc.-Vis. Image Signal Process, vol.146, no.2, pp.93–100, April 1999.

[2] Hiroshi Murase and Rie Sakai, "Moving object recognition in eigenspace representation: gait analysis and lip reading" Pattern Recognition Letters, vol.17, pp155–162, 1996.

[3] Yoshinori Uesaka, "A new Fourier Descriptor Applicable to Open Curves" IEICE, vol.J67-A, no.3, pp.166–173, March 1984.

[4] M.P.Murray, A.B.Drought, and R.C.Kory. "Walking patterns of normal men" Journal of Bone and Joint Surgery, vol.46-A,N

[5] Hiroshi Murase, Shree K.Nayar, "3D Object Recognition from Appearance - Parametric Eigenspace Method-" IEICE, vol.J77-D-II, pp.2179-2187, November 1994.

[6] Tadashi Hata, Yoshio Iwai, Masahiko Yachida, "Robust Gesture Recogniton by Using Image Motion and Data Compression" IEICE, vol.J81-D-II, No.9, pp.1983-1992, September 1998.

[7] Junji Yamato, Jun Ohya, Kenichiro Ishii, "Recognizing Human Action in Time-Sequential Images Using Hidden Markov Models" IEICE, vol.J76-D-II, No.12, pp.2556-2563, December 1993.