Systems and Computers in Japan, Vol. 21, No. 12, 1990
Translated from Denshi Joho Tsushin Gakkai Ronbunshi, Vol. 72-DII, No. 11, November 1989, pp. 1833-1838

A Method of Generating Template Patterns from a Few Sample Patterns in Character Recognition

Akihiro Suzuki, Shozo Makino, and Ken-iti Kido, Members
Research Center for Applied Information Sciences, Tohoku University, Sendai, Japan 980

Hiroshi Kanai and Yoshiyuki Kawazoe, Members
Education Center for Information Processing, Tohoku University, Sendai, Japan 980

SUMMARY

To recognize a character using a pattern matching method, it is necessary to prepare sufficiently many sample patterns for a template pattern. However, it is difficult to collect a sufficient number of samples of some characters because they do not appear often.

This paper proposes a new method of estimating a template pattern using the transfer function in a frequency range which is considered generally to be useful in such a case. The effectiveness of the method is demonstrated experimentally by using the printed Devanagari characters (in Buddhist literature) as an example.

1. Introduction

When an image enters into a computer, various noise and distortion components occur. Since such components influence the similarity and the calculation of distances, it is necessary to remove them in a preprocessing. For this purpose, the template pattern of a character is made by taking an average of many samples of each character. When an average pattern is made from N samples, the average amplitude of a noise component reduces $1/\sqrt{N}$ of that obtained from a single sample.

However, generally, it is difficult to obtain a sufficient number of samples for a character belonging to a new character set or a character only used occasionally. For such a case, conventionally a Gaussian filter [1] has been used to smooth the noisy components. The shape and size of such a filter has been selected by experience, and it has not been established to design an optimal filter for each character.

This paper proposes a method which designs an optimal filter for each character of interest using the transfer function in the frequency region so that a template pattern is estimated from a few samples. There have been a few works on automatic recognition of Devanagari characters used in Buddhist literature and Hindustani language [2 - 4]. Such a set of characters is relatively small in number but contains characters which are used very occasionally. Such characters are best suited for testing the proposed method. In this paper, Devanagari characters printed on the Hokkekyo (edited edition) [5] are used to test the effectiveness of the proposed method.

2. Principle

2.1. Principle of estimation of template pattern

The process of the method consists of: (1) estimation of a transfer function from a binary image in a frequency region; and (2) estimation of the template pattern from the transfer function.

(1) The transfer function is obtained by using characters which have many samples N, using the binary image sampled. Let the i-th sampled image be $s_i(x, y)$ ($i = 1, \ldots, N$), let the template pattern which is their average pattern be $t(x, y)$, and the two-dimensional Fourier transform $s_i(x, y)$ and

ISSN0882-1666/90/0012-0025$7.50/0

$t(x, y)$ be $S_i(x, y)$ and $T(x, y)$, respectively; x and y are the vertical and horizontal axes of the image plane, respectively, and let the size of the image be $N_x \times N_y$. Let the transfer function of the filter be $H(u, v)$. Using $H(u, v)$, the image pattern generated from $S_i(u, v)$ is defined by

$$T_i(u, v) = H(u, v) \cdot S_i(u, v) \cdot \begin{bmatrix} i = 1, \cdots, N \\ u = 1, \cdots, N_x \\ v = 1, \cdots, N_y \end{bmatrix} \tag{1}$$

The sum of the error powers of the sample pattern and the template pattern in the frequency region $\alpha(u, v)$ is given by

$$\begin{aligned} \alpha(u, v) &= \sum_{i=1}^{N} | T_i(u, v) - T(u, v) |^2 \\ &= \sum_{i=1}^{N} | H(u, v) \cdot S_i(u, v) - T(u, v) |^2 \end{aligned} \tag{2}$$

The filter function is determined so that $\alpha(u, v)$ becomes minimum. By applying a partial differentiation of $H(u, v)$ to $\alpha(u, v)$ in Eq. (2), and by making this value zero, the transfer function $H(u, v)$ is given by

$$H(u, v) = \frac{T(u, v) \sum_{i=1}^{N} S_i(u, v)^*}{\sum_{i=1}^{N} S_i(u, v) S_i(u, v)^*} \tag{3}$$

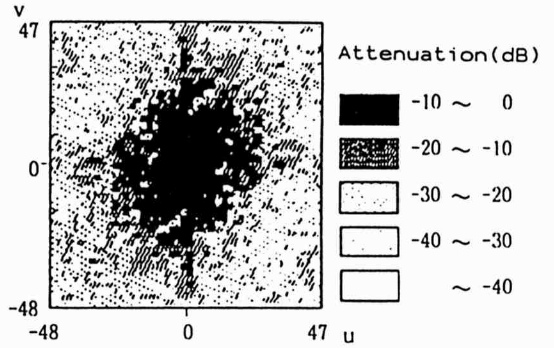
where * denotes a complex conjugate.

(2) By using the obtained transfer function $H(u, v)$, the template pattern of each character, including those for which few samples are obtained, is estimated.

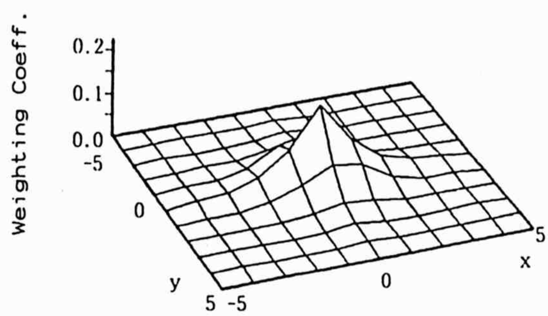
Although $H(u, v)$ has a size of $N_x \times N_y$ in the frequency region, an approximate mask $m(x, y)$ is designed on the image plane to reduce the computation. A two-dimensional Fourier transform is applied to the transfer function $H(u, v)$ so that the mask $m(x, y)$ is formed on the image plane. A part which has a very small value on $m(x, y)$ is approximated by zero so that the mask having $N_x \times N_y$ is reduced to a size of $(2w_x + 1) \times (2w_y + 1)$. The mask $m(x, y)$ is convoluted into a sample image $s_i(x, y)$ by using

$$t(x, y) = \sum_{x'=-w_x}^{w_x} \sum_{y'=-w_y}^{w_y} s_i(x + x', y + y') m(x', y') \tag{4}$$

so that the template pattern $t(x, y)$ is formed.



(a) Transfer function $H_1(u, v)$ of Devanagari charac. "अ" obtained from 10 samples



(b) Mask $m_1(x, y)$ obtained from reverse Fourier transformation of $H_1(u, v)$

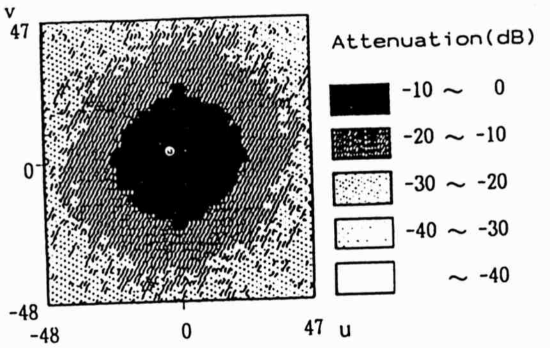
Fig. 1. Transfer function $H_1(u, v)$ and mask $m_1(x, y)$.

If a mask $m(x, y)$ is obtained by the foregoing process, a template pattern of a particular character usually obtained from an average pattern of many samples can be estimated from a single sample image. This method utilizes the fact that the characteristics of noise components are constant independently of each character.

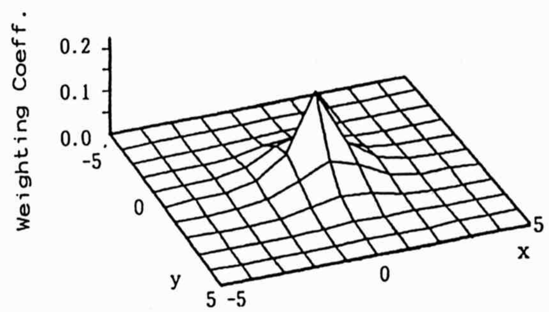
2.2. Stability of calculation of transfer function

Figure 1(a) shows the transfer function H_1 of a Devanagari character "अ" (an independent character for the vowel 'a') obtained by 10 samples using Eq. (3). The character pattern is an $N_x \times N_y = 96 \times 96$-dimensional binary image. This figure shows a trend of a low-pass filter as a whole, and some large values can be seen in the high-frequency region.

In a case where a character pattern has no power or very little power, the denominator of Eq. (3) becomes zero or a very small



(a) Transfer function $H_2(u, v)$ obtained from 633 samples of Devanagari character



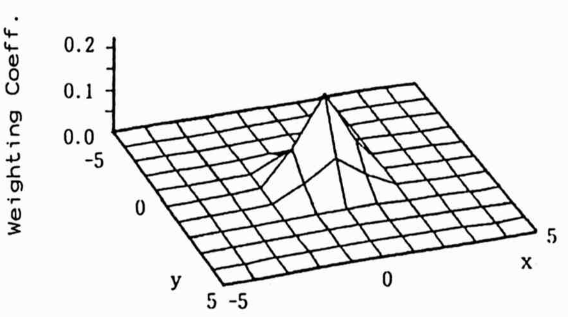
(b) Mask $m_2'(x, y)$ obtained from reverse Fourier transformation of $H_2'(u, v)$

Fig. 2. Transfer function $H_2(u, v)$ and mask $m_2(x, y)$.

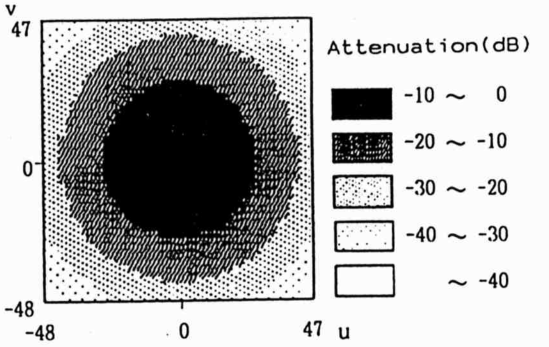
value at a frequency of (u, v), and the estimated value of $H(u, v)$ becomes very unstable. For a transfer function of a single sample of a character, it is possible that this instability affects fairly many frequency components.

To avoid this instability, it is better to use as many samples and characters as possible rather than a single sample. Therefore, the transfer function $H_2(u, v)$ was calculated from 663 Devanagari characters printed in [5]. This set of samples contains 89 different characters i.e., the sample sizes are different for each character, from only one or a few to 20 at the maximum. Comparing $H_2(u, v)$ and $H_1(u, v)$, the forementioned instability tends to reduce gradually from a low frequency to a high frequency.

Figure 1(a) shows a mask $m_1(x, y)$ obtained by applying a two-dimensional Fourier transform to the transfer function $H_1(u, v)$;



(a) Gaussian mask $m_2(x, y)$ obtained from $m_2(x, y)$



(b) Frequency characteristic $H_2'(u, v)$ of $m_2'(x, y)$

Fig. 3. Mask $m_2'(x, y)$ and its frequency property $H_2'(u, v)$.

Fig. 1(b) shows a mask $m_2(x, y)$ obtained by applying a two-dimensional Fourier transform to $H_2(u, v)$; $w_x = w_y = 5$ was used for the calculations.

2.3. Comparison with Gaussian filter

Let us approximate the mask $m_2(x, y)$ obtained by applying the inverse Fourier transform to the transfer function. Let us obtain the variances σ_x^2 and σ_y^2 of the weighting coefficients of the mask $m_2(x, y)$ in the x and y directions, respectively. A Gaussian mask $m_2'(x, y)$ using these variances is

$$m_2'(x, y) = A\exp\left[-\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2}\right)\right] \quad (5)$$

where A is the normalizing constant. Figure 3(a) shows this mask and Fig. 3(b) shows the frequency characteristic $H_2'(u, v)$ of $m_2'(x, y)$.

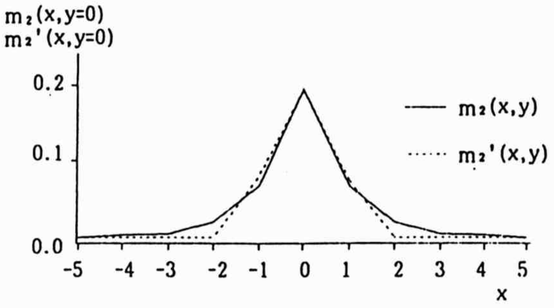


Fig. 4. Mask $m_2(x, y)$ and mask $m_2'(x, y)$.

Figure 4 shows the difference between the distributions of the weighing coefficients $m_2(x, y)$ and $m_2'(x, y)$ at $y = 0$. This figure illustrates well the difference between the transfer characteristics obtained from Eq. (3) and that of a conventional Gaussian filter. In other words, the mask designed by using the proposed method clearly contains information such as the characteristics of the Devanagari character and noises, although these are not represented in the Gaussian filter.

Figure 5 shows the Gaussian mask $m_1'(x, y)$ based on $m_1(x, y)$ and Eq. (5), and Fig. 5(b) shows its frequency property $H_1'(u, v)$. These figures also show a difference between the properties at $m_1(x, y)$ and $m_1'(x, y)$.

3. Recognition Experiments

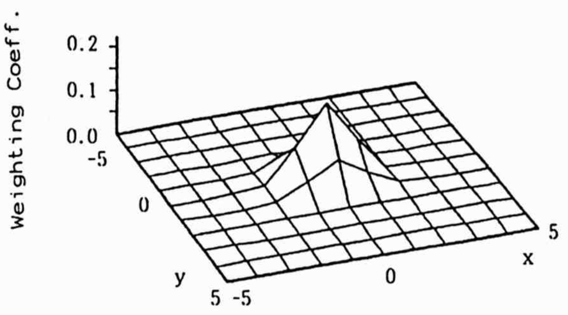
To confirm the effectiveness of the proposed method, an experiment was carried out using 7744 characters containing 89 kinds of Devanagari characters printed on the first 15 pages of [5].

(a) Experiment 1: Relationship between the formation of template pattern and recognition capability

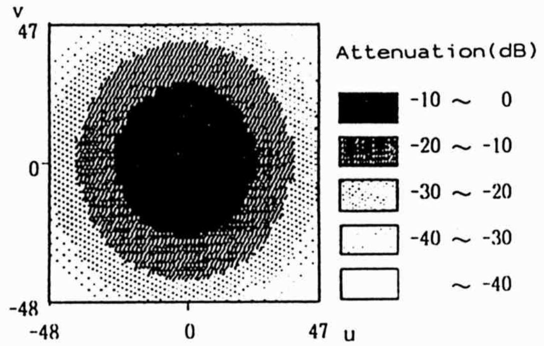
A template pattern was made using a single sample from each kind of character. The experiments were carried out using five different processes so that they can be compared:

(1) The sample is used directly;

(2) The sample is processed by using $m_1(x, y)$ (a mask designed using the proposed method for a single sample);



(a) Gaussian mask $m_1'(x, y)$ obtained from $m_2(x, y)$



(b) Frequency characteristic $H_1'(u, v)$ of $m_1'(x, y)$

Fig. 5. Mask $m_1'(x, y)$ and its frequency property $H_1'(u, v)$.

(3) The sample is processed by using $m_1'(x, y)$ (a Gaussian filter);

(4) The sample is processed by using $m_2'(x, y)$ (a mask designed by using the proposed method); and

(5) A single sample is processed by using $m_2'(x, y)$ (a Gaussian filter).

Similarly to (2) and (4) but with no preprocessing for unknown input patterns, experiments were conducted with (2a) and (4a). Again, similarly to (2) and (4) but with a preprocessing for the unknown input, experiments were performed on patterns using the mask. Table 1 shows the results.

To eliminate the influence of a peculiarity of the samples for making the template pattern, five characters for each of the 89 character types were prepared (except for a few which have less than five samples). Using the procedures (1) to (5), and using the whole set of 7744 characters, recognition experiments were carried out. The

Table 1. Experimental results of character recognition

	Method used for smoothing dictionary	Preprocessing for unknown input	Mean error rate and standard deviation (%)
(1)	none	none	3.0 ± 0.5
(2a)	m_1	none	2.0 ± 0.2
(2b)	m_1	used	1.5 ± 0.1
(3)	m_1'	none	2.1 ± 0.2
(4a)	m_2	none	2.0 ± 0.2
(4b)	m_2	used	1.4 ± 0.1
(5)	m_2'	none	2.2 ± 0.2

m_1: Mask designed by using the proposed method, from a single kind of character.
m_1': Gaussian filter.
m_2: Mask designed by using the proposed method, from many kinds of characters.
m_2': Gaussian filter.

A single sample was used for making the template pattern in all cases. The entire sample consists of 89 kinds of Devanagari characters, 7744 characters in all.

rightmost column in Table 1 shows the mean value and the standard deviation of the recognition rate obtained for each procedure (five times each).

The error rate in (4b) (with a mask designed by the proposed method) is less than that in (1) (a sample pattern alone). The standard deviation of the error rate in the former is smaller than that in the latter, and the result is very stable. The error rate in (4a) (with a mask designed by the proposed method) is less than in (5) (with a Gaussian mask). Hence, it is demonstrated that a mask using the proposed method is better than a conventional mask and that the proposed method produces a template pattern independently of samples used for making it.

(b) Experiment 2: Relationship between number of samples to make template pattern and recognition

To demonstrate the effectiveness of the proposed method for a small number of samples, the recognition experiments were carried out with different numbers of samples for the template pattern. The template patterns for these experiments are made by combining two sets of samples:

(1) a set of single samples for each kind of character (i.e., 89 characters in all); and

(2) a set of N characters sampled from the beginning part of [5].

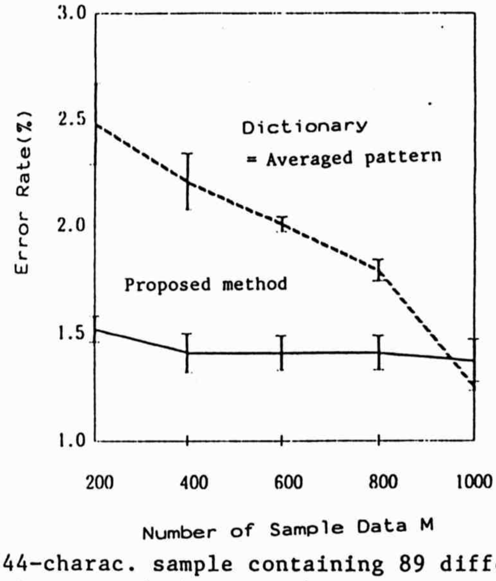
In other words, the template pattern was made by making a mask from (2) using the proposed method, and then by using this mask and each kind of character in (1).

For comparison,

(3) a set of $M(= N + 89)$ samples consisting of the samples in (1) and (2)

was used to make an average pattern for each kind of character, and this was another template pattern. Figure 6 shows the results with an average value and standard deviations. Note, five groups of samples were prepared as (1).

The results show that the proposed method can increase a recognition rate, even with a small number of samples for



(7744-charac. sample containing 89 different Devanagari characters)

Fig. 6. Relation between the number of sample data and error rates.

N in (2), which is comparable to a conventional method with a set of samples of about 1000 characters in (3). Therefore, the proposed method is useful to estimate a template pattern image which cannot collect a sufficient number of samples.

4. Conclusions

This paper proposes a new method to estimate a character with a small number of samples using a filter with a transfer function from a binary image to a template pattern in a frequency region. A set of printed Devanagari character (89 kinds, 7744 characters in all) were used for recognition experiments with a template pattern made by using the proposed method with a single sample, and a recognition error rate of about 1.4 percent was obtained. This rate is comparable with a result obtained by a conventional method using about 1000 samples. This shows that the proposed method is useful for recognizing characters for which only a few samples are available.

To further improve the recognition rate of the method, it is necessary to discriminate some very similar characters such as those appearing in Devanagari characters, since this is the largest single reason for error in the recognition. Taking this point into account, both the method of making a template pattern and the method of recognition will be examined in the future.

REFERENCES

1. Jun'ichiro Toriwaki. Digital image processing for image understanding (I). Shokodo Publ. (1988).
2. Akihiro Suzuki, Hiroshi Kanai, Shozo Makino, Yoshiyuki Kawazoe, and Ken'ichi Kido. Method of recognizing Devanagari characters with simultaneous segmentation. Papers of Technical Group on Pattern Recognition and Understanding, I.E.C.E., Japan, PRU88-81 (1988).
3. R. M. K. Sinha and H. N. Mahabala. Machine recognition of Devanagari script. IEEE Trans. Syst., Man. & Cybern., SMC-9, pp. 435-441 (1979).
4. Kurishnamachari Jayanty, Akihiro Suzuki, Hiroshi Kanai, Shozo Makino, Yoshiyuki Kawazoe, Masayuki Kimura, and Ken'iti Kido. An approach to Devanagari character recognition using outstanding features. Papers of Technical Group on Pattern Recognition and Understanding. I.E.C.E., Japan, PRU87-103 (1988).
5. H. Kern and B. Nanjo. Saddharmapundarika. Bibliotheca Buddhia X (1912).

AUTHORS (from left to right)




Akihiro Suzuki graduated in 1987 from the Dept. of Information Engineering, Faculty of Engineering, Tohoku University, where he obtained an M.S. in 1989. He is engaged in research on automatic recognition of characters. He is with NTT.

Hiroshi Kanai graduated in 1981 from the Faculty of Engineering, Tohoku University, and obtained a Ph.D. from there in 1986. He was appointed an Assistant at the Education Center for Information Processing at Tohoku University in 1986 and has been a Lecturer on the Faculty of Engineering there since 1990. He is engaged in research on digital processing of acoustic vibration signal and fault diagnosis, and pattern recognition. He was awarded the Ishikawa Prize in 1987. He is a member of Acoustical Society of Japan; I.E.I.C.E., Japan; the Institute of Mechanical Engineering, Japan; and IEEE.

Yoshiyuki Kawazoe graduated in 1970 from Second Science Department, Tohoku University, where he obtained a Dr. of Sci. degree in 1975. After appointment as an Assistant on the Faculty of General Education of Tohoku University and later to Associate Professor at the Education Center for Information Processing, he was appointed a Professor at the Institute for Materials Research in 1990. He was a research fellow at the Max Planck Research Institute (West Germany) in 1981 and a guest Professor at WACAE (Australia). He is engaged in research on information processing, parallel computations, character recognition, and theoretical physics. He is a member of the Institute of Physics, Japan; the Institute of Information Processing, Japan; and the Institute of India and Buddhism, Japan. He is the author of several books including Outline of Computers, Computers and Programming for Literary Fields, etc.

Shozo Makino graduated in 1969 from the Electronics Dept., Tohoku University, where he obtained a Dr. of Eng. degree in 1974. He was appointed an Assistant at the Research Institute of Electrical Communication of Tohoku University the same year, and later to an Assistant at the Research Center for Applied Information of Tohoku University, where presently he is an Associate Professor. He is engaged in research on automatic speech recognition using language information, acoustic digital-signal processing, image-signal processing, and character recognition. He is a member of the Acoustical Society of Japan and of the Institute of Information Processing, Japan.

AUTHOR

Ken'iti Kido graduated in 1948 from the Electrical Engineering Department, Tohoku University, where he was appointed an Assistant the same year. He has a Dr. of Eng. degree. After being appointed an Associate Professor in the same department, he was appointed a Professor at the Research Institute of Electric Communication of the University in 1963. Since 1976, he has been a Professor and Director of the Research Center for Applied Information Sciences, Tohoku University. He is engaged in research on acoustic instruments, building acoustics, noise control, and psychological acoustics. Presently, he is engaged in research on automatic speech recognition, digital-signal processing and its acoustic applications. He is the author of Acoustic Engineering (Corona Publ.); Introduction to Digital-Signal Processing (Maruzen Publ.); Basics of Electronic Computers, Parts 1 and 2 (Maruzen Publ.); Theory of Transient Phenomena (Asakura Shoten). He is a member of the Acoustical Society of Japan; IEE, Japan; the Society of Instrument and Control Engineers, Japan; the Institute of Information Processing, Japan; the Acoustical Society of Korea; IEEE, AES; and a Fellow of the Acoustical Society of America.