Handprinted Kanji Character Recognition by the Background Feature Pattern Matching Method using Non-linear Normalization

Atsushi YOSHIDA and Yasuo HONGO
FUJIFACOM Corporation, Tokyo, JAPAN

Abstract

Handprinted character recognition technique gives a human-friendly man-machine interface. In Japan, many kinds of characters, Kanji, Hiragana, Katakana, alphabet, numerals, are commonly used. The most important subject in handprinted character recognition is the ability to compensate for shape variations of handprinted characters. To solve this problem, several normalization methods have been proposed. Linear normalization methods have been generally applied, but also some non-linear normalization methods have been proposed for handprinted Kanji character recognition and their effectivenesses were shown.

In this paper, a novel non-linear normalization method is proposed. Two types of extended image planes and three types of line density are newly defined. Experiments for handprinted Kanji character recognition by the background feature pattern matching method are performed in order to demonstrate its ability for six combinations and results are compared each other. A high recognition rate of 97.8% was achieved on handprinted character recognition experiments for more than 19,000 character images included in the handprinted Kanji character database, ETL8B.

1 Introduction

Character recognition methods are roughly classified into the pattern matching method (correlation method) and the structural analysis method. In the pattern matching method, it is important to normalize the size and position of the unknown pattern with those of the dictionary pattern prior to the matching process. As a technique for normalization of handprinted characters, linear normalization is insufficient because the balance of the character shape varies each time even if they are printed by the same person. For this reason, several non-linear normalization methods have been proposed.

Two non-linear normalization methods[1] are proposed by Yamada et al.. One is the following way. The number of lines in both the X-axis and Y-axis directions and the cumulative function with the number of lines is calculated, and then the image is shaped by sampling pitch where the line density is a certain constant in both the X-axis and Y-axis directions. Accordingly, the portion where lines are dense is expanded and the portion where lines are sparse is contracted. In a case of the character "道 (michi)", a human has a tendency to lengthen the last stroke (lowest part) rightward. In this method, the right-lower line is naturally shrunk. But in a case of the character "言 (gen)", the part "口 (kuchi)" is constracted vertically so that the upper part may become small. This result looks unnatural to human eyes. Therefore, another is proposed. The line interval in both the X-axis and Y-axis directions is measured and the cumulative function with maximum value of the inverse values of the line intervals at each point is calculated. In a case of the character "言" by this method, interval of vertical lines is almost equalized, and transformation becomes more natural. Here, the important point is that the feature is defined by the value integrated X-axis direction and Y-axis direction.

With the non-linear normalization method[2] proposed by Tsukumo, line interval is measured in both the X-axis and Y-axis directions. The cumulative function is calculated using the inverse of the measured line interval. In a case of the character "田 (ta)", the sizes of four inner-spaces are equalized. However, in this method, the X-axis direction and the Y-axis direction are independent.

In these methods using line interval, however, because of the definition of line interval, the line interval is a certain constant on the outside of character. In this paper, two types of extended image planes are defined so that the line interval in the character's peripheral area are defined. Furthermore, three types of line density are defined as functions of line interval in the X-axis and Y-axis directions as the feature of each point, and a normalization is made so that the sum of the feature in the sampling pitch is equalized. The comparison of conventional methods is shown in Table 1.

Table 1 Comparison of conventional methods

Method	Feature	X,Y dependence
reference[1]-1	Number of lines	Independent
reference[2]	Inverse number of line interval	Independent
reference[1]-2	Inverse number of line interval	Integrated
This method	Function of line interval	Integrated

Experiments are performed for six combinations and results are compared each other. The details of these experiments are reported below.

2 Definition of line density

2.1 Definition of extended image plane

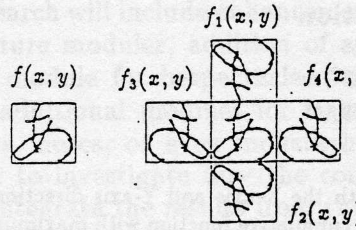
Let us define the two types of extended image planes, R_M and R_C, on which input character patterns $f(x,y)$ are configured as adjacent to each other.

[Definition 1] Mirror image plane R_M :

On plane R_M , the input character pattern $f(x,y)$ is in a line symmetrical relation with the upper, lower, left and right patterns $f_1(x,y) \sim f_4(x,y)$. This relation is shown in Fig. 1.

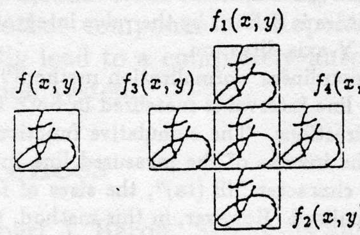
[Definition 2] Cyclic plane R_C :

On plane R_C, the input character pattern $f(x,y)$ is in a parallel positional relation with the upper, lower, left and right patterns $f_1(x,y) \sim f_4(x,y)$. This relation is shown in Fig. 2.



(a) original image (b) extended image

Fig. 1 Mirror image plane R_M



(a) original image (b) extended image

Fig. 2 Cyclic plane R_C

2.2 Definition of line interval

Assuming that input character patterns $f(x,y)$ are configured on plane R as shown in Fig. 1 or Fig. 2, the line interval at each point in the circumscribed rectangle of the character pattern $f(x,y)$ is calculated. The width of the circumscribed rectangle is W_x, and the height of that is W_y.

First, the rising edge positions, $X_+^L(x,y)$ and $X_+^R(x,y)$, where the image changes from white "0" to black "1", and falling edge positions, $X_-^L(x,y)$ and $X_-^R(x,y)$, where the image changes from black "1" to white "0" as shown in Fig. 3 are calculated for four directions by eqs. (1) $\sim$ (4). Hatched areas represent black areas, the white areas represent background areas.

$$X_+^L(x,y) = \max\{x' | x' \le x, \bar{f}(x'-1,y) * f(x',y) = 1\} \quad (1)$$
$$X_+^R(x,y) = \min\{x' | x' > x, \bar{f}(x'-1,y) * f(x',y) = 1\} \quad (2)$$
$$X_-^L(x,y) = \max\{x' | x' \le x, f(x'-1,y) * \bar{f}(x',y) = 1\} \quad (3)$$
$$X_-^R(x,y) = \min\{x' | x' > x, f(x'-1,y) * \bar{f}(x',y) = 1\} \quad (4)$$

where; $\bar{f}(x,y)$ is logical NOT of $f(x,y)$.

Next, line interval $L_x(x,y)$ in the X-direction is calculated by the following equation.

$$L_x(x,y) = (X_+^R - X_+^L + X_-^R - X_-^L)/2 \quad (5)$$

The calculation can be made by eq. (5) in the case of Fig. 3 (a) and (b). In the case where four values, i.e., $X_+^R(x,y)$, $X_+^L(x,y)$, $X_-^R(x,y)$ and $X_-^L(x,y)$, are uncertain as shown in Fig. 3 (c), however, $L_x(x,y)$ is changed to $4W_x$.

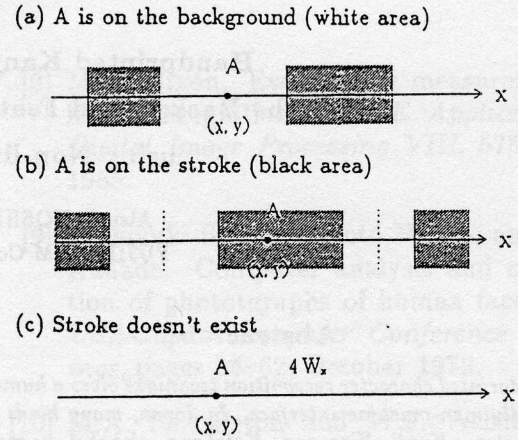


Fig. 3 Definition of line interval L_x

The line interval, $L_x(x,y)$ is calculated as described above. The Y-direction, the line interval, $L_y(x,y)$ is identically calculated by the following equations.

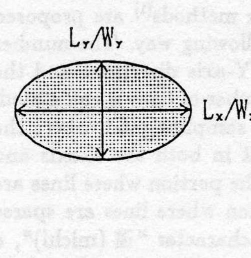
$$Y_+^L(x,y) = \max\{y' | y' \le y, \bar{f}(x'-1,y) * f(x',y) = 1\} \quad (6)$$
$$Y_+^R(x,y) = \min\{y' | y' > y, \bar{f}(x'-1,y) * f(x',y) = 1\} \quad (7)$$
$$Y_-^L(x,y) = \max\{y' | y' \le y, f(x'-1,y) * \bar{f}(x',y) = 1\} \quad (8)$$
$$Y_-^R(x,y) = \min\{y' | y' > y, f(x'-1,y) * \bar{f}(x',y) = 1\} \quad (9)$$

$$L_y(x,y) = (Y_+^R - Y_+^L + Y_-^R - Y_-^L)/2 \quad (10)$$

In the case where four values, i.e., $Y_+^R(x,y)$, $Y_+^L(x,y)$, $Y_-^R(x,y)$ and $Y_-^L(x,y)$, are uncertain, $L_y(x,y)$ is changed to $4W_y$.

2.3 Definition of line density

Function λ which describes the size of ellipse associated from the background area of the character pattern is defined in three ways as shown in Fig. 4. The inverse number of this descriptive function is defined as line density ρ. Three types of line density, ρ_E, ρ_P and ρ_A, are conducted from corresponding λ_E, λ_P and λ_A. Line density $\rho(x,y)$ is defined as a function of line intervals $L_x(x,y)$ and $L_y(x,y)$.



(a) $\lambda_E = \pi \cdot \dfrac{L_X}{W_X} \cdot \dfrac{L_Y}{W_Y}$

(b) $\lambda_P = \pi \cdot \left(\dfrac{L_X}{W_X} + \dfrac{L_Y}{W_Y} \right)$

(c) $\lambda_A = \lambda_E / \lambda_P$

$$= \frac{L_X}{W_X} \cdot \frac{L_Y}{W_Y} \Big/ \left(\frac{L_X}{W_X} + \frac{L_Y}{W_Y} \right)$$

Fig. 4 Definition of function λ

[Definition (a)] Area of ellipse λ_E :

Line density $\rho_E(x,y)$ is calculated as the product of the inverse numbers of line intervals in the X- and Y-directions.

$$\rho_E(x,y) = \frac{1}{\lambda_E} = \left(\frac{W_x}{L_x} \cdot \frac{W_y}{L_y}\right) \qquad (11)$$

[Definition (b)] Peripheral length of ellipse λ_P :

Line density $\rho_P(x,y)$ is calculated as the inverse number of the sum of line intervals in the X- and Y-directions.

$$\rho_P(x,y) = \frac{1}{\lambda_P} = \left(\frac{L_x}{W_x} + \frac{L_y}{W_y}\right) \qquad (12)$$

[Definition (c)] Mean depth of ellipse λ_A :

Line density $\rho_A(x,y)$ is calculated as the sum of the inverse numbers of line intervals in the X- and Y-directions.

$$\rho_A(x,y) = \frac{1}{\lambda_A} = \left(\frac{W_x}{L_x} + \frac{W_y}{L_y}\right) \qquad (13)$$

3 Non-linear normalization

Projection functions $h_x(x)$ and $h_y(y)$ in the X-direction and Y-direction of $\rho_i(x,y)(i = A,P,E)$ are calculated by the following expressions, respectively.

$$h_x(x) = \sum_{y=1}^{W_y} \rho_i(x,y) \quad (i = A,P,E) \qquad (14)$$

$$h_y(y) = \sum_{x=1}^{W_x} \rho_i(x,y) \quad (i = A,P,E) \qquad (15)$$

Next, the cumulative density functions $C_x(X)$ and $C_y(Y)$ in intervals $[1, X]$ and $[1, Y]$, respectively, are calculated.

$$C_x(X) = \sum_{x=1}^{X} h_x(x) \qquad (16)$$

$$C_y(Y) = \sum_{y=1}^{Y} h_y(y) \qquad (17)$$

Then the following equation is realized.

$$S = C_x(W_x) = C_y(W_y) \qquad (18)$$

Mapping function $X(i)$ and $Y(j)$ are defined as follows in order to obtain the normalizing pattern $g(i,j)$ $(i = 1 \sim M, j = 1 \sim N)$.

$$X(i) = \min\{x | C_x(x) \geq i \cdot \frac{S}{M}\} \qquad (19)$$

$$Y(j) = \min\{y | C_y(y) \geq j \cdot \frac{S}{N}\} \qquad (20)$$

The normalized pattern is obtained by the eq. (21). The Process of normalization is shown in Fig. 5.

$$g(i,j) = f(X(i), Y(j)) \qquad (21)$$

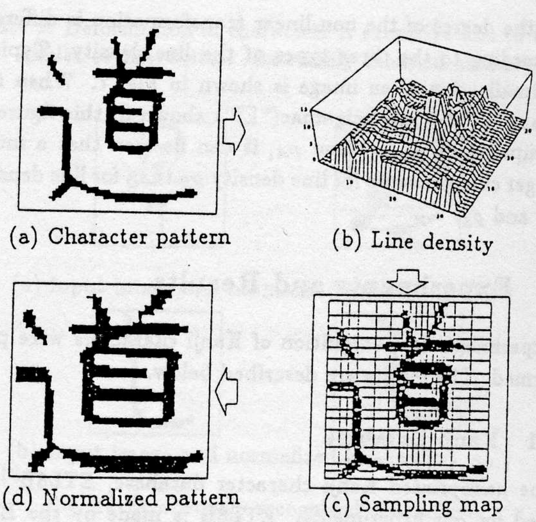


Fig. 5 Process of normalization

3.1 Comparison of two types of extended image plane

With R_M, the line interval on the right-hand side of the character is determined by the shape on the same side. On the other hand, with R_C , the line interval on the right-hand side of the character is related to the shape on the left-hand side. Therefore, the balance of the shape of the character normalized with R_M is different from that of the character normalized with R_C. Typical normalization of an image is shown in Fig. 6. It can be seen in this figure that the position of the vertical stroke of the character "土" is different between R_M and R_C after normalization.

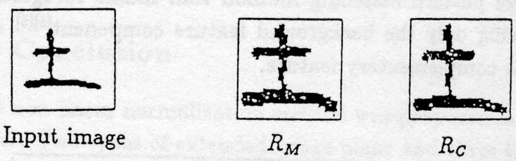


Fig. 6 Typical normalized images

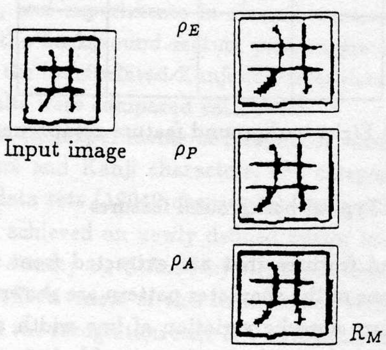


Fig. 7 Typical normalized images

3.2 Comparison of three types of line density

When the size of background portion varies, the variation of line density ρ_E is larger than those of ρ_P and ρ_A. That is, the degree of the non-linear transformation is different according to the three types of the line density. Typical normalization of an image is shown in Fig. 7. When the shape of "井" in kunigamae("□") shown in this figure is compared by ρ_E, ρ_P and ρ_A, it can be seen that a much larger change occurs for line density ρ_E than for line density ρ_P and ρ_A.

4 Experiments and Results

Experiments for recognition of Kanji characters were performed. The results are described below.

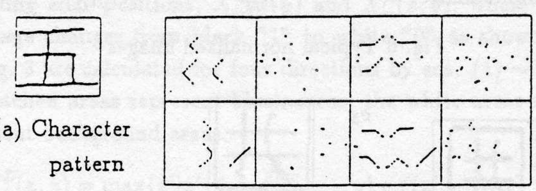
4.1 Kanji Database

The handprinted Kanji character database, ETL8B[3], is used for our experiments. ETL8B is made by the Electrotechnical Laboratory of MITI. It consists of Hiragana characters of 75 categories and Kanji characters of 881 categories, 956 characters in total.

A dictionary pattern was prepared as the learning sample using odd data sets of 80 persons ($952 \times 80 = 76160$ characters). In addition, even data sets of 20 persons ($952 \times 20 = 19040$ characters) were used as unknown patterns, and recognition and comparison were made.

4.2 Background feature pattern matching method

Yasuda, et al. contrive the complementary feature and good results are obtained for recognition of handprinted Kanji characters[3][5]. In this paper we use the background feature pattern matching method that makes recognition by using only the background feature component[3][5] out of the complementary feature.



(a) Character pattern

(b) Eight background feature compornents

Fig. 8 Typical background features

The background features that are extracted from the background portions of the character pattern are shown in Fig. 8. This feature absorbs variation of line width and variation of position, and is obtained as a stable feature.

Experiments are performed by the method shown in the flow of Fig. 9. Line density measurement and feature extraction from the original image are executed, and the feature image is normalized by the line density obtained from the original image.

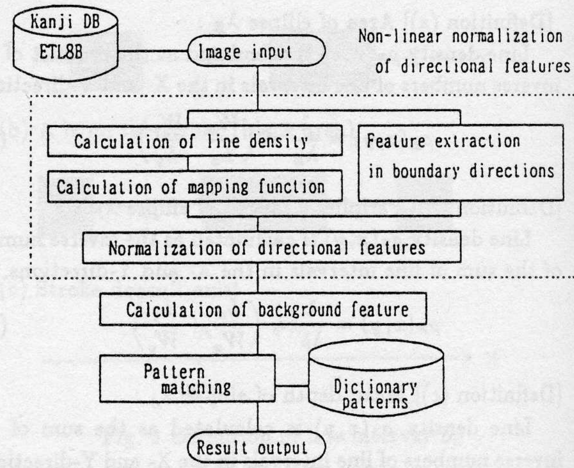


Fig. 9 Flow of processing

Pattern matching is executed using components in eight directions of background feature ($16 \times 16 \times 8$). Furthermore, the shift similarity method[4] is used. The shift similarity is the highest value of nine correlations which are obtained when the feature pattern is shifted in the range of ± 1 mesh in X and Y direction.

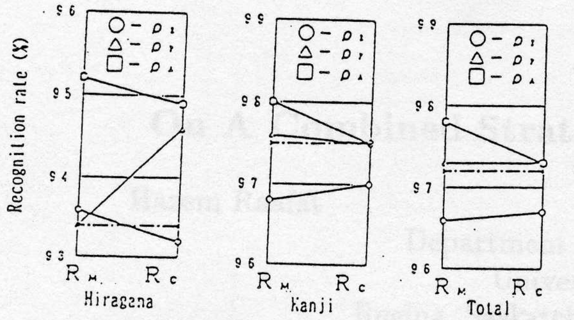
4.3 Results

The results in the definition of each image plane and line density are shown in Table 2 and Fig. 10. A comparison of recognition rate between datasets No. 2 and No. 34 in particular is shown in Table 3. The character quality in dataset No. 2 is high, while the character quality in dataset No. 34 are low. For comparison, the results by non-linear normalization method of reference [1]-2, which had the best recognition rate of reference [1], [2], are shown in Table. 2, 3, and Fig. 10.

Table 2 Recognition rate by category

Parameter condition		Recognition rate (%) "Hiragana"	"Kanji"	Total
R_M	ρ_E	93.6	96.8	96.6
	ρ_P	93.4	97.6	97.3
	ρ_A	95.2	98.0	97.8
R_C	ρ_E	93.2	97.0	96.7
	ρ_P	94.6	97.5	97.3
	ρ_A	94.9	97.5	97.3
Reference[1]-2		93.4	97.5	97.2

The highest value of total recognition rate was obtained with plane R_M and line density ρ_A. It was learned that both R_M and R_C are better than the method of reference [1]-2 when the definition of line density was ρ_A and ρ_P respectively. In the case where the definition of line density was ρ_E, no uplift was observed regardless of the definition



(One-dot chain lines indicate what was obtained by the method of reference [1]-2.)

Fig. 10 Recoginition rate by category

of the plane. Furthermore, a drop in the recognition rate was observed particularly for Kanji characters. The highest value was obtained with both dataset No. 2 and No. 34 when the definition of plane was R_M and the definition of line density was ρ_A. With dataset No. 2, uplift was observed with both R_M and R_C as compared to the method of reference [1]-2 regardless of the definition of line density. With data set No. 34, on the contrary, the recognition rate dropped for both R_M and R_C when the definition of line density was ρ_E.

Table 3 Recognition rate by each data set

Parameter condition		Recognition rate (%)	
		Dataset No.2	Dataset No.34
R_M	ρ_E	98.7	89.6
	ρ_P	98.9	91.1
	ρ_A	99.5	92.5
R_C	ρ_E	98.8	90.2
	ρ_P	99.5	91.2
	ρ_A	99.3	91.6
Reference[1]-2		98.5	90.5

4.4 Reading errors

Japanese characters have so many kinds that many groups of similar characters exist. For this reason, handprinted character recognition becomes more difficult. The reason why a character is erroneously recognized is listed below.

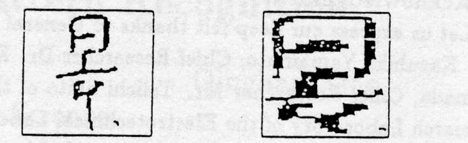
Case 1: Vocal variation in Hiragana such as "ば (ba)" and "ぱ (pa)"; and "ぼ (bo)" and "ぽ (po)".

Case 2: Difference by a short stroke such as "方 (hou)" and "万 (man)".

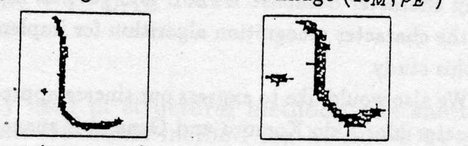
Case 3: Difference in the stroke length such as "土 (tuti)" and "士 (si)".

Case 4: Reading error dependent on the quality of the character due to batter, blur and noise. They are shown in Fig. 12.

Case 5: Deformation of characters is excessive due to the habits of the writer. Extremely inclined characters, etc.



(a) Input image and normalized image (R_M, ρ_E)



(b) Input image and normalized image (R_M, ρ_E)

Fig. 12 Erroneously recognized characters

In this non-linear normalization method, the balance of the shape of character is normalized. In the cases of 1, 2, therefore, its effctiveness cannot be expected. The processing of details of a part of character pattern is needed. In the case of 3, this non-linear normalization has a bad influence. Therefore, measuring the stroke length on the input character image is needed. In the cace of 4, in this non-linear normalization method the balance of the shape of normalized pattern is sensitive to noises. Removing noises or smoothing image is needed before character normalization. In the cace of 5, its effectiveness can be expected according to the degree of the shape variation. However, the improvement of normalization is needed.

5 Conclusion

New non-linear normalization method was proposed in this paper. Two types of extended image plane and three types of line density was defined. This non-linear normalization was applied to background features extracted from the image, and experiments in six combinations was conducted by the background feature pattern matching method using the handprinted Kanji character database ETL8B and results were compared each other.

In the experiments of recognition using Hiragana characters and Kanji characters, 952 categories in total, and 20 data sets (19040 characters), recognition rate of 97.8% was achieved on newly defined mirror image plane R_M in line density ρ_A defined by the mean depth of ellipse, and the effectiveness of this method was verified as compared with the recognition rate of 97.2% achieved with the background feature pattern matching method using non-linear normalization method[1]-2 of Yamada, et al..

The following items can be considered as subjects for future study. (1) Expansion of categories to be read, (2) analysis of reading error, (3) improvement of non-linear normalization method, (4) improvement of recognition method, etc. We intend to lay emphasis on these subjects in order to enhance the recognition rate.

Acknowledgement

Let us express our deep felt thanks to General Manager Dr. Kazuhiko Yamamoto, Chief Researcher Dr. Hiromitsu Yamada, Chief Researcher Mr. Taiichi Saito of the Image Research Laboratory of the Electrotechnical Laboratory of MITI and Dr. Michio Yasuda, who provided handprinted kanji character database ETL8B and gave us instructions on the character recognition algorithm for implementation of this study.

We also would like to express our sincere appreciation to Director Mr. Akio Komuro and General Manager Mr. Hiroshi Adachi of No. 5 Engineering Division of FUJIFACOM Corporation, to chief researcher Dr. Kunihiko Edamatsu of the Fuji Electric Research & Development Laboratory and also to relevant persons who made significant contributions to study.

References

[1] H.Yamada, et al.:"A Nonlinear Normalization Method for Handprinted Kanji Character Recognition - Line Density Equalization -" Pattern Recognition. Vol 23. No. 9, pp.1023-1029, 1990 (The Journal of the Pattern Recognition Society)

[2] J.Tsukumo, H.Tanaka:"Classification of Handprinted Chinese characters using nonlinear normalization methods" 9th ICPR, pp.168-171, 1988

[3] S.Mori, et al.:"Research on Machine Recognition of Handprinted characters" IEEE Trans. Pattern Anal. Mach. Intell. PAMI-6, pp.386-405, 1984

[4] H.Yamada, et al.:"An Improved Correlation Method - Shift similarity -" paper of IEICE Japan, J64-D, 10, pp.970-976, 1981 (in Japanese)

[5] M.Yasuda, et al.:"An Improved Correlation Method for Character Recognition - Handprinted Chinese Character Recognition in a Reciprocal Feature Field -" paper of IEICE Japan, J68-D, 3, pp.353-360, 1985 (in Japanese)