Recognition of Handprinted Japanese Characters by Relaxation Matching with the Integrated Dictionary

Kenji Iwata Yuuichi Hoshino Masayoshi Okamoto
Kunihito Kato Kazuhiko Yamamoto

Department of Information Science, Faculty of Engineering
Gifu University
1-1 Yanagido, Gifu 501-1193, Japan
e-mail: iwata@yam.info.gifu-u.ac.jp

Abstract

In the character recognition, the multi-template method is effective, but it has problems of dictionary size and process time. Therefore, we propose integrated dictionary for the recognition system of Japanese handprinted characters by relaxation matching. This system is modified from the previous relaxation method. The shapes of the inner holes are extracted in feature extraction. High quality data which is a printed character is used for the initial data of the dictionary. No suitable case is neglected in the association of plural segments. In the other, the allowance for the position of long segments is expanded by path points. The feature extraction by polygonal approximation is unstable by the deformation of the handprinted character. To solve this problem, integrated dictionary is important. This system was tested on common database ETL9 with the recognition result of 95.81% for the unknown data.

1 Introduction

Many methods have been researched for character recognition[1], for example, the pattern matching method[2] and the structural analysis method based on the Newton's equations of motion[3].

Relaxation is a technique for using contextual information to reduce local ambiguities. Relaxation operation was introduced by Rosenfeld[4] to understand the picture processing field from mathematical field. We had applied a relaxation method for recognition of handprinted Kanji characters and we had good results[5]. It was extended to Dempster-Shafer theory to solve the problem of noisy information[6].

In this paper, we had several improvements for the previous recognition system[7] as follows:

1. The shapes of inner holes are extracted in the feature extraction.
2. No suitable case is neglected in the association of plural segments.
3. In the other, allowance for the position of long segments is expanded by path points.
4. High quality data which is the printed character is used for the initial data of the dictionary.

In the character recognition, the multi-template method was known effective[2]. In a recognition system by relaxation matching, it is effective, too. This system has a problem in the feature extraction by polygonal approximation. It has the case of extracting the features which are not applied by relaxation matching, because some Kanji characters have a complicated structure and a tendency for the lines to touch. The multi-template for one category is one solution of this problem. If the number of templates increases, perhaps, the performance of recognition will rise, but the size of the dictionary and process time become huge. In this paper, we propose integrated dictionary which is integrate the plural. template to one template. It can express the variety of the input pattern, but increases of the dictionary size and process time are little.

Recognition system by relaxation matching with some improvements is described in **2**. We propose integrating the dictionary which is described in **3**. There is one template for one category, but its performance is near that of the multi template for one category. Effectiveness of these method is shown through the experiments which is described in **4**. Our system is tested on handprinted character database ETL9[8] which was composed by the first level of JIS Kanji characters.

2 Recognition system

Fig.1 shows a recognition system by relaxation matching. Tracing the edge of the character, approximating the input pattern by plural segments, and extracting the segment feature are described in **2.1**. Learning about the dictionary using by relaxation matching is described in **2.2**. The initial probability between the input segment and the dictionary segment is described in **2.3**. Relaxation matching by spring connection of neighbors is described in **2.4**. Choosing

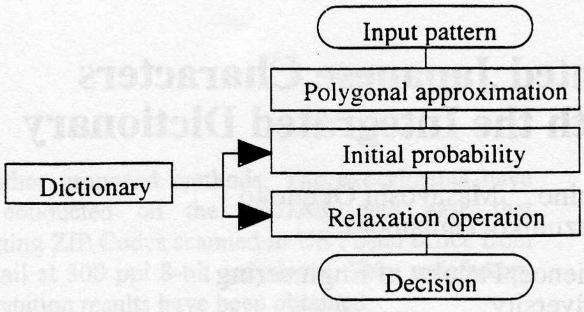


Fig.1 Recognition System.

the most suitable correspondence and calculating the distance are described in **2.5**.

2.1 Feature Extraction

The shape of the input pattern is represented by polygonal approximation. It obtained some closely point lists from tracing the edge of the input pattern. If some points have a the sharp curve, it segments the curves at these points[5]. (Segmentation method 1)

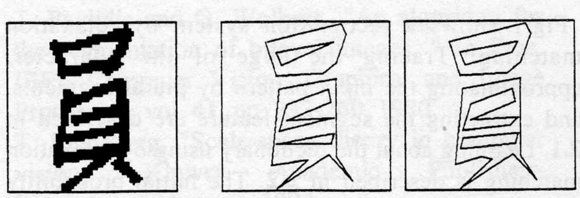
It segments the curves for extracting the shape of the inner holes as in Fig.2(a). If an angle which is made by approximate line segments in front and behind a segmentation point i is less than 30° and difference of the length of these lines is less than 5 pixels, point i has the possibility of failure segmentation and is investigated. It assumes the case is segmented at two points which is the periphery of point i and the case is segmented at only one point i. An assumption which has the smallest sum of the distance between each point on an edge curve and the approximate point which is corresponded to the point is decided. (Segmentation method 2)

Fig.2(b) was applied segmentation method 1. Fig.2(c) was applied segmentation method 2.

The feature of each line segment $\mathbf{O}_k$ is represented as:

$$\begin{aligned}\mathbf{O} &= [\mathbf{O}_k]_{k=1}^{\varphi} \\ &= [\ x_{sk}, y_{sk}, x_{ek}, y_{ek}, \theta_k, L_k, \\ &\quad x_{p1k}, y_{p1k}, x_{p2k}, y_{p2k}, x_{p3k}, y_{p3k}\ ]_{k=1}^{\varphi}\end{aligned} \tag{1}$$

Where φ is the number of the line segment of the polygon; (x_{sk}, y_{sk}) and (x_{ek}, y_{ek}) are the position of the starting point and the ending point of segment k; θ_k is the direction; L_k is the length; (x_{pnk}, y_{pnk}) is the position of the n quarter path point.



(a) Input pattern (b) Result of method 1 (c) Result of method 2

Fig.2 Polygonal approximation.

2.2 Learning about the dictionary

The dictionary feature of each line segment $\mathbf{T}_i$ is added some parameters to $\mathbf{O}_k$.

$$\begin{aligned}\mathbf{T} &= [\mathbf{T}_i]_{i=1}^{\varphi} \\ &= [\ x_{si}, y_{si}, x_{ei}, y_{ei}, \theta_i, L_i, \\ &\quad x_{p1i}, y_{p1i}, x_{p2i}, y_{p2i}, x_{p3i}, y_{p3i}, \\ &\quad [E_j]_{j=1}^{n_{si}}, [E'_j]_{j'=1}^{n_{ei}}, P_{di}\ ]_{i=1}^{\varphi}\end{aligned} \tag{2}$$

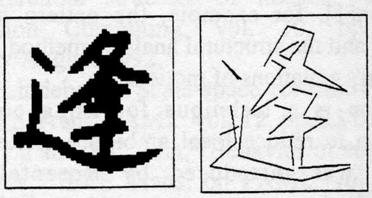
Where E_j and E'_j are the name of neighboring segments for the starting point and the ending point of segment i; n_{si} and n_{ei} are their number; P_{di} is the probability of appearance.

The process of learning is the following. It is initialized by substituting a feature of the first learning data for a mask feature. The other learning data are chosen which is the most suitable correspondence of each segment by relaxation matching described in **2.3**, **2.4** and **2.5**. It substitutes the average of each feature for the mask features.

The influence of initial data is strong in this learning method. Therefore it is necessary to use high quality data for the initial data. We prepared high quality printed character data for the initial data.

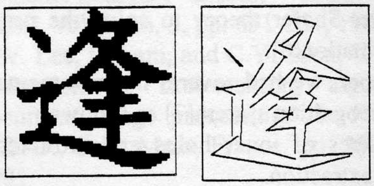
Fig.3 shows a learning result using the handprinted character for the initial data. This data isn't suitable for expressing the property of this category, because the number of segments isn't enough. Fig.4 shows a learning result using the high quality printed character for the initial data. This character has enough segments.

2.3 Initial probability



(a) Initial data. (b) Learning result.

Fig.3 Example of learning result using handprinted character for initial data.



(a) Initial data. (b) Learning result.

Fig.4 Example of learning result using high quality printed character for initial data.

It can consider one mask segment i as two mask segments i_s and i_e, if it satisfies the condition as:

$$\begin{cases} l_{es} \le c_1 \\ l_{es} < l_{ss} \quad , \quad l_{es} < l_{ee} \\ l_{se} > l_{ss} \quad , \quad l_{se} > l_{ee} \end{cases} \tag{3}$$

Where:

$$c_1 = 15$$

$$l_{ss} = \sqrt{(x_{sis} - x_{sie})^2 + (y_{sis} - y_{sie})^2}$$

$$l_{se} = \sqrt{(x_{sis} - x_{eie})^2 + (y_{sis} - y_{eie})^2} \tag{4}$$

$$l_{es} = \sqrt{(x_{eis} - x_{sie})^2 + (y_{eis} - y_{sie})^2}$$

$$l_{ee} = \sqrt{(x_{eis} - x_{eie})^2 + (y_{eis} - y_{eie})^2}$$

This condition is able to neglect the no suitable case as two segments are side by side, or one segment has the opposite direction of other segment. The case of considering one input segment k as two input segments k_s and k_e, is the same.

The association function between mask segment i and input segment j is evaluated, if the following conditions are satisfied:

$$\begin{cases} ds_{ik} \le c_1 \\ de_{ik} \le c_1 \\ |\theta_i - \theta_k| \le c_2 \\ |L_i - L_k| \le c_3 \end{cases} \tag{5}$$

Where:

$$ds_{ik} = \sqrt{dx_{sik}^2 + dy_{sik}^2}$$

$$de_{ik} = \sqrt{dx_{eik}^2 + dy_{eik}^2}$$

$$dx_{sik} = x_{si} - x_{sk} \quad , \quad dy_{sik} = y_{si} - y_{sk}$$

$$dx_{eik} = x_{ei} - x_{ek} \quad , \quad dy_{eik} = y_{ei} - y_{ek} \tag{6}$$

$$c_2 = \begin{cases} 20 & (\text{If } L_i \le 10) \\ 10 & (\text{If } L_i > 10) \end{cases}$$

$$c_3 = 10$$

In the case of a long segment, the position of one side point is stable, but the position of the other side point is unstable. Therefore, if the length of the mask segment i is long ($L_i \ge 20$), it also evaluates the association function on the success of the following condition of the position:

$$\sqrt{(x_{p1i} - x_{sk})^2 + (y_{p1i} - y_{sk})^2} \le c_1 \quad (\text{If } de_{ik} \le c_1) \tag{7}$$

$$\sqrt{(x_{p3i} - x_{ek})^2 + (y_{p3i} - y_{ek})^2} \le c_1 \quad (\text{If } ds_{ik} \le c_1)$$

If the length of the input segment k is long, it evaluates on the success of the analogously condition. Similarity $P'_{i(k)}$ of the pair (i,k) is assigned as follows:

$$\begin{aligned} P'_{i(k)} = \max[\quad & 1 - W_1 \max(|\theta_i - \theta_k| - c_{21}, 0) \\ & - W_2 \max(|L_i - L_k| - c_{22}, 0) \\ & - W_3 \max(ds_{ik} - c_{23}, 0) \\ & - W_3 \max(de_{ik} - c_{23}, 0) \quad , 0] \end{aligned} \tag{8}$$

Where c_{21}=6, c_{22}=10, and c_{23}=2 are capacities for minute differences; W_1=0.1, W_2=0.01, and W_3=0.01 are weight of each feature's axes.

The initial mask probability $P^{(0)}_{i(k)}$ and the initial input probability $P^{I(0)}_{i(k)}$ are evaluated with similarity $P'_{i(k)}$:

$$P^{(0)}_{i(k)} = \frac{P'_{i(k)}}{\max_{k'} P'_{i(k')}} \tag{9}$$

$$P^{I(0)}_{i(k)} = \frac{P'_{i(k)}}{\max_{i'} P'_{i'(k)}} \tag{10}$$

Where the k' are all the input segments which are associated with $\mathbf{T}_i$ and the i' are all the mask segments which are associated with $\mathbf{O}_k$.

2.4 Relaxation operation

The relaxation operation is used to dynamically change the probability according to the compatibility between the configurations of mask segments and the input segments. We assume that there are spring connections[5] between the starting point of $\mathbf{T}_i$ and the ending point of $\mathbf{T}_j$ which is the neighboring segment. Where segment $\mathbf{T}_j$ is the neighboring segment of the start point of the segment $\mathbf{T}_i$. Let $r^s_{ij(k,l)}$ be the start point tension in spring connection between $\mathbf{T}_i$ and $\mathbf{T}_j$. If $\mathbf{T}_i$ is associated with the input segment $\mathbf{O}_k$ and $\mathbf{T}_j$ is associated with the input segment $\mathbf{O}_l$, then $r^s_{ij}(k,l)$ is evaluated as follows:

$$\begin{aligned} & r^s_{ij(k,l)} \\ & = \max\{\min(1 - W_5|dx_{sik} - dx_{ejl}| + c_{25}, 1), 0\} \\ & + \max\{\min(1 - W_5|dy_{sik} - dy_{ejl}| + c_{25}, 1), 0\} \end{aligned} \tag{11}$$

Where W_5=1/16 is a weight; c_{25}=1/8 is a capacity for minute differences. The end point tension $r^e_{ij(k,l)}$ is defines analogously.

The probability $P^{(t)}_{i(k)}$ is evaluated by the tension in the spring connection between the segments as follows:

$$P^{(t+1)}_{i(k)} = \frac{q_{i(k)} * P^{(t)}_{i(k)} * P'_{i(k)}}{\max_k (q_{i(k)} * P^{(t)}_{i(k)} * P'_{i(k)})} \tag{12}$$

$$P^{I(t+1)}_{i(k)} = \frac{q_{i(k)} * P^{I(t)}_{i(k)} * P'_{i(k)}}{\max_k (q_{i(k)} * P^{I(t)}_{i(k)} * P'_{i(k)})} \tag{13}$$

$$q_{i(k)} = \frac{q^{s}_{i(k)} + q^{e}_{i(k)}}{\sum_{j=1}^{n_s} L_j + \sum_{j'=1}^{n_e} L_{j'}} \tag{14}$$

$$q^{s}_{i(k)} = \sum_{j=1}^{n_s} \left\{ L_j * \max_{l} \left(r^{s}_{ij(k,l)} * P^{(t)}_{j(l)} * P^{1(t)}_{j(l)} \right) \right\} \tag{15}$$

$$q^{e}_{i(k)} = \sum_{j=1}^{n_e} \left\{ L_{j'} * \max_{l'} \left(r^{e}_{ij'(k,l')} * P^{(t)}_{j'(l')} * P^{1(t)}_{j'(l')} \right) \right\} \tag{16}$$

Where t is a number of iteration.

2.5 Decision

Most suitable associations between mask segment i and input segment $\tilde{k}$ from all candidates are fixed by the obtained probability $P^{(t)}{}_{i(k)}$ and $P^{1(t)}{}_{i(k)}$.

The distance D_J between the mask J is as follows:

$$D_J = \left[\sum_{i=1}^{\varphi} \left\{ \left(1 - R_{i(\tilde{k})}\right) * L_i \right\} + \sum L'_i * P'_d + \sum L'_k \right] * c_6 / \sum_{i=1}^{\varphi} L_i \tag{17}$$

$$R_{i(\tilde{k})} = \left[\left\{ \sum_{j=1}^{n_s} \left(L_j * r^{s}_{ij(\tilde{k},\tilde{l})} * P'_{j(\tilde{l})} \right) + \sum_{j'=1}^{n_e} \left(L_{j'} * r^{e}_{ij'(\tilde{k},\tilde{l}')} * P'_{j'(\tilde{l}')} \right) \right\} * P'_{i(\tilde{k})} \right] / \left\{ \left(\sum_{j=1}^{n_s} L_j + \sum_{j'=1}^{n_e} L_{j'} \right) * 2 \right\} \tag{18}$$

$$P'_d = \max\left\{ \left(P_d - 0.5\right) * 2, 0 \right\} \tag{19}$$

$$c_6 = 1000 \tag{20}$$

Where L'_i is the length of unmatched mask line segments; L'_k is the length of unmatched input line segments.

The result is the category J which has the shortest distance D_J of all candidate categories.

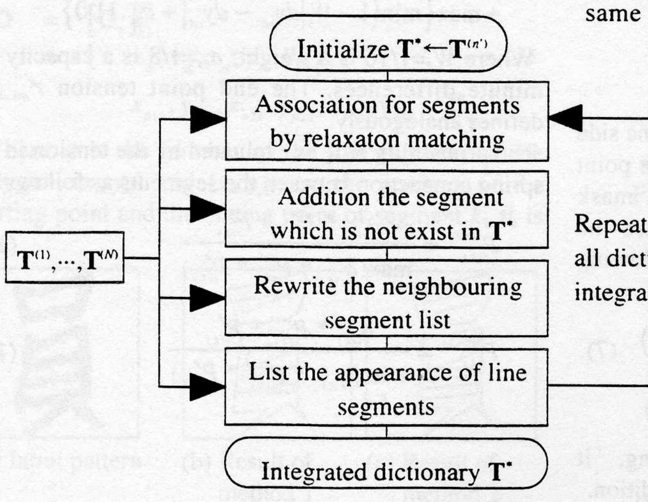


Fig.5 Dictionary Integration.

3 Integrated dictionary

3.1 Making the plural Dictionaries

The polygonal approximation which was described in **2.1** extracts different shapes which are in the same category, because the handprinted character changes the shape, and tracing the edge is sensitive to lines that touch. A solution for this problem is a multi-template which prepares a plural template for one category.

In the learning method which was described in **2.2**, the feature of each line segment takes in statistical property. The number of segments is used in the initial data, therefore the influence of polygonal approximation of initial data is strong. If the initial data is different, the created dictionary has the feature of a different polygonal approximation. To absorb the diversity of polygonal approximation, it is suitable that it prepares plural mask features which are made by different initial data.

Different N dictionaries $\mathbf{T}^{(1)},\cdots,\mathbf{T}^{(N)}$ are made by this way.

3.2 Dictionary integration

A multi-template is effective for character recognition. If the number of templates increases, the performance of recognition rises. The size of the dictionary and the process time becomes huge.

We propose the integrated dictionary which is made by integrating plural dictionaries. There is one template for one category, but its performance is near that of the multi template for one category.

Fig.5 shows a method of integration of the dictionaries which is as follows.

(1) Initialize the integrated dictionary

The average of the square of the distance $\mathbf{D}^{*(n)}$ from the dictionaries $\mathbf{T}^{(1)},\cdots,\mathbf{T}^{(N)}$ to all learning data of the same category are calculated. A dictionary $\mathbf{T}^{(n')}$ is

Table 1. Segment's appearance list.

name of segment	Dictionary number 1	2	3	4
32	1	1	1	1
33	1	1	1	1
34	1	1	1	1
35	1	0	1	1
36	1	1	1	1
37	1	1	1	1
38	0	1	1	0
39	0	0	0	1
40	0	0	0	1
41	0	0	0	1

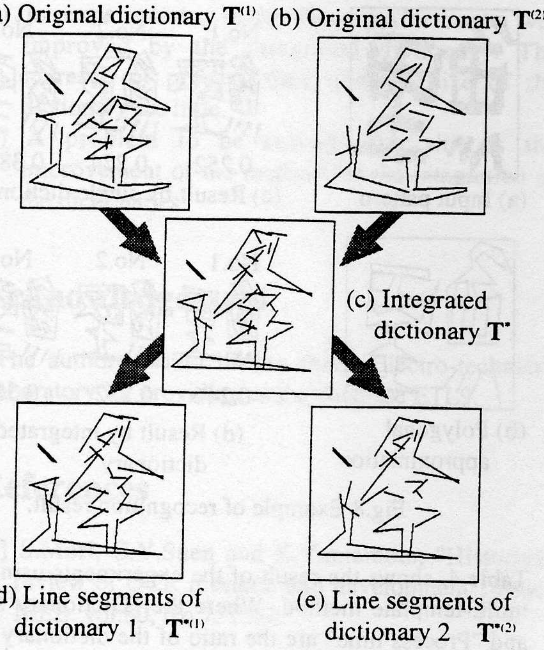


Fig.6 Example of integrated dictionary.

selected from $T^{(1)},\cdots,T^{(N)}$ as the most suitable data which has the minimum $D^{*(n)}$. It substitute a dictionary $T^{(n)}$ for the integrated dictionary T^{*}.

(2) Associations and addition the segment

The associations of each segment between T^{*} and $T^{(n)}$ (n=1,2,$\cdots$,N :$n \neq n'$) are obtained by relaxation matching which is the same as the recognition system. If a segment of an additional dictionary has a corresponding segment, it is a common segment. The integrated dictionary uses a segment of the T^{*}. If a segment doesn't have a correspondence, this segment doesn't exist in the T^{*}. It adds this segment to the integrated dictionary. It repeats until all dictionaries integrate.

(3) Rewrite the neighboring segment list

List of neighboring segments of additional segments are rewriting to name of integrated dictionary's segment which is correspondence of these segment.

(4) List the appearance of the segments

Segment's appearance of each dictionary is listed to express which of the dictionary do each segment appear. Table 1 shows an example of this list which integrate four dictionaries. Where "1" shows that segment i is appeared in the dictionary $T^{(n)}$; "0" shows that segment i is not appeared in the dictionary $T^{(n)}$. This list is used at the recognition. Using this list can express various shapes as deformation of handprinted character, vanishing the feature by lines that touch, and the mistake of polygonal approximation.

Fig.6 shows an example of an integrated dictionary which integrates Fig.3(a) and Fig.3(b). There is described that Fig.3(a) isn't suitable, but it expresses a

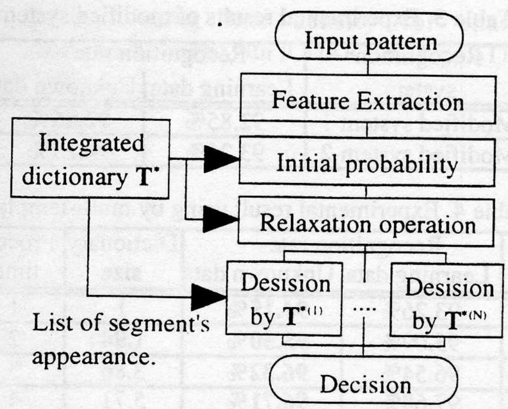


Fig.7 Recognition by integrated dictionary.

subcategory, because some input patterns touch the lines as this data.

3.3 Recognition by integrated dictionary

Fig.7 shows the recognition method by an integrated dictionary. It gets the initial probability and operates the relaxation by recognition method as the same as a single dictionary's one. Deciding the correspondence and calculating the distance which are using segments which appear only one dictionary $T^{*(n)}$ within the integrated dictionary are acted. They are acted as each $T^{*(1)},\cdots,T^{*(N)}$ by segment's appearance list. The distance D_J is the minimum value of the obtained distance. The result is the category J which has the shortest distance D_J of all candidate categories.

4 Experiments

4.1 Experiment data

A recognition result depends on the data. At this time, the handprinted character database ETL9B[8] was used which is binary and normalized to 64X63 pixels in size. The ETL9B consists of 200 subsets and each subset consists of 3036 categories (with 2965 Kanji characters and 71 Hiragana characters). The dictionary was created with 100 learning subsets which are odd

Table 2. Initial data of the number of dictionaries N.

Initial data of Dictionary	N 1	2	4	6	8
Printed data	○	○	○	○	○
Subset 0		○	○	○	○
Subset 2			○	○	○
Subset 4			○	○	○
Subset 6				○	○
Subset 8				○	○
Subset 10					○
Subset 12					○

Table 3. Experimental results of modified system

Recognition system	Recognition rate	
	Learning data	Unknown data
Modified system 1	92.85%	**92.86%**
Modified system 2	93.26%	**94.15%**

Table 4. Experimental result using by multi-template

N	Recognition rate		Dictionary size	Process time
	Learning data	Unknown data		
1	93.26%	**94.15%**	1	1
2	95.09%	**95.50%**	1.94	2
4	96.54%	**96.32%**	3.86	3
6	97.68%	**96.71%**	5.71	4

Table 5. Experimental results using by integrated dictionary

N	Recognition rate		Dictionary size	Process time
	Learning data	Unknown data		
1	93.26%	**94.15%**	1	1
2	94.29%	**94.83%**	1.09	1.11
4	95.37%	**95.15%**	1.14	1.29
6	96.23%	**95.48%**	1.17	1.46
8	96.60%	**95.81%**	1.19	1.62

subsets.

The printed character data which was used in the initial value of the dictionary was one subset of JEIDA Fuji format printed character database. The name of the font of this subset is "Iwata chu-futo-kyokasho-type", and its size is 12 points.

Eight dictionaries were using in the experiments of the multi-template method and integrated dictionary method. These dictionaries were made by the different initial data. Using dictionaries of each number of dictionaries N was shown in the table 2.

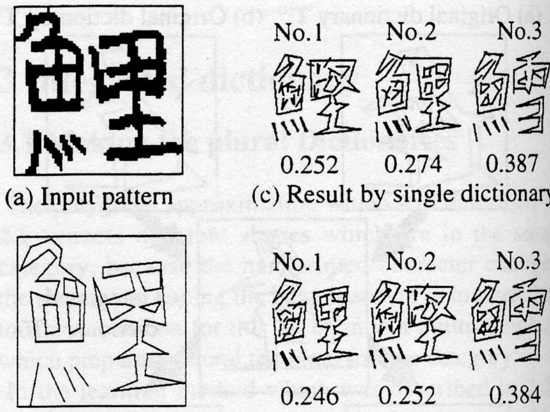
The first 10 subsets were used for recognition tests. We used the first 100 candidates of preclassification with the the identical category for main matching. The directional pattern matching method[9] was used for preclassification. The probability which exists for the right answer in these candidates is 99.82% for the unknown subset, and 99.59% for the learning subset.

4.2 Experiments for evaluating a modified recognition system

Table 3. shows the result of the experiments of modified system. Modified system 1 is improved in the feature extraction, initial probability and association with plural segments. Modified system 2 is use the high quality initial data of the dictionary.

These results show that these modifies are effective.

4.3 Experiments on multi-template



(a) Input pattern (c) Result by single dictionary

(b) Polygonal approximation (d) Result by integrated dictionary

Fig.8 Example of recognition result.

Table 4. shows the result of the experiments using by multi-template method. Where the "Dictionary size" and "Process time" are the ratio of the dictionary size and the process time of the recognition which defines each of single dictionary (N=1) as one.

Recognition rate is improved with the increase of the number of dictionary N. but the process time and the dictionary size are in proportion to increase of N. Unprepared increase is of no practical use.

4.4 Experiments on integrated dictionary

Table 5. shows the result of the experiments using by integrated dictionary. Recognition performance of integrated dictionary is near the multi dictionary.

The increase of process time by the integrated dictionary which is proposed in this paper is less than using the multi-dictionary. Increased size of the dictionary is so little. because many segments are held in common.

Fig.8 shows a example of the recognition result. This input example(Fig.8(a)) is mistaken by the single dictionary (Fig.8(c)). but the integrated dictionary gives a right answer.(Fig.8(d))

5 Conclusion

Some concluding remarks are listed as follows:

1) We proposed a modified recognition system.
   - 1.1) Segmentation for extracting the shape of the inner holes.
   - 1.2) Using the high quality data for the initial data of the dictionary.
   - 1.3) Neglecting the no suitable case in association with plural segments.
   - 1.4) Expanding the allowance for the position of long segments by path points.

2) We proposed the integrated dictionary.

3) In experiments, recognition performance was improved by the integrated dictionary. The increase of process time and the size of the dictionary are little.
4) A problem to be solved after this is the improvement of the method for the integration of the dictionaries.

Acknowledgement

The authors would like to thank Electro-technical Laboratory for providing us the database ETL9.

References

[1] S.Mori, C.Y.Suen and K.Yamamoto, "Historical review of OCR research and development", Proc. IEEE, vol.80, No.7, 1992.

[2] S.Omachi, F.Sun and H.Aso, "An algorithm for construction of multi-template dictionary for character recognition considering between-class variation", IEICE Trans., vol.J79-DII, No.9, pp.1525-1533, 1996. (in Japanese)

[3] T.Nagasaki, T,Yamamoto and M.Nakagawa, "The behavior of dynamic relaxation in an elastic stroke model for character recognition", Proc. 4th ICDAR, vol.1, pp.16-22, 1997.

[4] A.Rosenfeld, R.A.Hummel and S.W.Zucker, "Scene labeling by relaxation operations", IEEE Trans. SMC, 6, pp.420-433, 1976.

[5] K.Yamamoto and A.Rosenfeld, "Recognition of handprinted Kanji characters by a relaxation", Proc. 6th ICPR, pp.395-398, 1982.

[6] T.Horiuchi, K.Toraichi, K.Yamamoto and H.Yamada, "Relaxation optimizing processes in extended probabilistic space", Proc. 3rd ICDAR, pp.266-269, 1995.

[7] I.Sakaga, T.Saito, H.Yamada and K.Yamamoto. "Recognition of handprinted characters in the first level of JIS Chinese characters", Proc. 8th ICPR, pp.570-572, 1986.

[8] K.Yamamoto, H.Yamada and T.Saito. "Current state of recognition method for Japanese characters and database for research of handprinted character recognition", From Pixels to Features III, North-Holland, pp.105-116, 1992.

[9] T.Horiuchi, R.Haruki, H.Yamada, K.Yamamoto, "Two-dimensional extension of normalization method using line density for character recognition", Proc. 4th ICDAR, vol.2, pp.511-514, 1997.