Handprinted Character Recognition Using Relative Relation of Plural Classification Methods

Daisuke NISHIWAKI, Keiji YAMADA, Jun TSUKUMO and Tsutomu TEMMA
C&C Information Technology Research Laboratories, NEC Corporation
1-1, Miyazaki 4-chome, Miyamae-ku, Kawasaki, Kanagawa 216 JAPAN

Abstract

This paper describes a new handprinted character recognition method. Plural recognition methods, each of which has mutual complement properties, are integrated in order to achieve higher recognition performance, compared with using individual recognition methods. A reliability measure is proposed to select the most certain answer out of answers by several recognition methods. The measure expresses how much the answer, which is given as a result for each recognition method, is reliable for an input pattern.

The efficiency of the reliability measure is discussed through experiments on Japanese "HIRAGANA" character recognition, using two kinds of primitive recognition methods. It was confirmed that the error ratio was reduced by half, compared with using primitive recognition methods individually.

1 Introduction

Recently, OCRs are becoming popular as office automation equipment. However, many restrictions still exist in using OCRs. To maintain the practical recognition performance for OCRs, their users have to print characters carefully. This is because no perfect character recognition method exists, as yet. Any recognition method misclassifies some of the sampled patterns. It is very important to solve this problem, in order to develop OCRs which can be used more human-friendly.

There are two approaches to achieving an efficient character recognition method, using a feature extraction method and a discrimination method already developed. One approach is to extract rules from patterns, which were misclassified by a recognition method, based on experience and intuition on the part of the developer. However, this approach produces too many exceptional processings, so that the recognition dictionary might be difficult to maintain.

The other approach is to integrate several character recognition methods. In this approach, it is assumed that several patterns, which would be misrecognized by one method, would be correctly recognized by another method. These recognition properties, here, are called as mutual complement properties. C. Y. Suen et al have reported recent efforts corresponding to this approach for handwritten numeral recognition[7][8]. Several integrating methods on a feature extraction level have been proposed for handprinted Chinese character[1][2]. The methods combine feature vectors, individually obtained by each feature extraction method, into one feature vector. This kind of integration method is effective for object character patterns, which consist of many categories.

This paper discusses integrating plural recognition methods on the decision level. In order to explain the concept easily, two character recognition methods were selected here. First, the mutual complement properties for each recognition method were investigated. Then, a reliability measure was proposed for the recognition results, which were obtained by each recognition method.

This paper describes how to calculate the reliability in Section 2, and the general processing flow and the final decision algorithm are presented in Section 3. Then, results of Japanese "HIRAGANA" character recognition experiments are presented in Section 4. The effectiveness of the proposed integrating method is indicated in Section 5. The proposed integrating method is summarized in Section 6.

Table 1. The outputs obtained from two recognition methods; Broken lines express combinations of the same category outputs.

	1st candidate	2nd	3rd	. . .
Method#1	Ci (d1(i))	Ch (d1(h))	Cj (d1(j))	. . .
Method#2	Cj (d2(j))	Ci (d2(i))	Ch (d2(h))	. . .

Category name(Distance)

2 Classification Method Integration

2.1 Discussion on Results by Recognition Methods

To integrate two recognition methods, the authors first investigated outputs obtained from two kinds of primitive recognition methods.

When two recognition methods are used, each individual method arranges category names for candidates in sequence according to distance, as shown in Table 1. In this case, the result assortment is as shown in Table 2.

Table 2. Assortment of the answers for the outputs.

	Case-1	Case-2	Case-3	Case-4
Method#1	Correct	Correct	Wrong	Wrong
Method#2	Correct	Wrong	Correct	Wrong

In Case1, both answers are correct. Here, the answers are the category names of primary candidates obtained by the individual methods. The answers are the same in Case1. In Case2, a correct answer is obtained in Method#1, while a wrong answer is obtained in Method#2. On the contrary, in Case3, a correct answer is obtained in Method#2, and a wrong answer is obtained in Method#1. In Case4, both answers are wrong. Furthermore, Case4 consists of two sub-cases, Case4.1 and Case4.2. In Case4.1, the same answers are obtained by Method#1 and Method#2, and the answers are wrong. In Case4.2, different answers are obtained by the methods, and the answers are wrong.

To reduce the error ratio, only the answers, which belong to Case1, should be selected as a final decision, and the other patterns which have different answers as primary candidate categories, should be rejected. However, the recognition ratio is obviously less than the recognition ratios individually obtained by primitive recognition methods.

Therefore, to retain the recognition ratio and to reduce the error ratio at the same time, it is necessary to select a correct answer obtained by one of the recognition methods on Case2 and Case3. The authors propose a method of selecting the correct answers in Section 2.3. It is essential to use recognition methods which have mutual complement properties for the integration, as shown in Fig. 1.

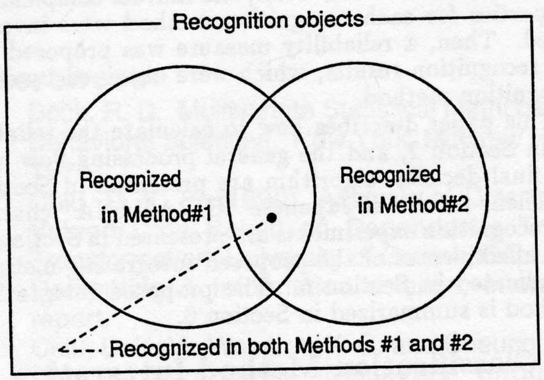

Fig. 1 Mutual complement recognition properties.

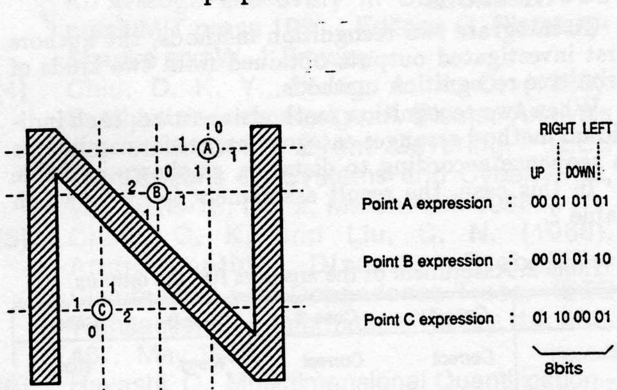

Fig. 2 Characteristic Loci example.

2.2 Individual Classification Method Properties

Here, to make the discussion easy, the authors selected two kinds of well-known primitive character recognition methods, in which the same discrimination method was used, and two kinds of feature extraction methods were used.

The constitutions of feature extraction methods and the discrimination method are as follows;

- Method#1:Each pixel's value and Projection Distance Method[3]
- Method#2:Characteristic Loci[4] and Projection Distance Method

The feature extraction method in Method#1 uses positional information. The existence of the character stroke is investigated on each pixel. Therefore, the feature dimension is equal to the picture size of the input pattern.

The feature extraction method in Method#2 uses structural information. This method scans from each background pixel upward to count the crossings over strokes. Downward crossing number, left crossing number and right crossing number are examined in the same way. This feature is called Characteristic Loci[4]. Many feature extraction methods, which concentrate extracted information on background pixels, is based on Characteristic Loci. This feature extraction method is not sensitive, but it is flexible for multifont printed characters and handwritten characters.

Figure 2 shows an example of Characteristic Loci. For all pixels belonging to the background, the number of crossings for the character strokes is counted for each direction.

Here, the number of crossings is expressed by 2 bits. Therefore, the maximum number of crossings is 3. If the crossing number becomes more than 3, it is treated as 3. Therefore, Method#2's feature space dimension is 256.

Both recognition methods use the Projection Distance Method for discrimination, which is a sort of the subspace method[6]. Figure 3 shows the concept involved in the method. The subspaces are determined by Eigen vectors, in ascending order of Eigen values,

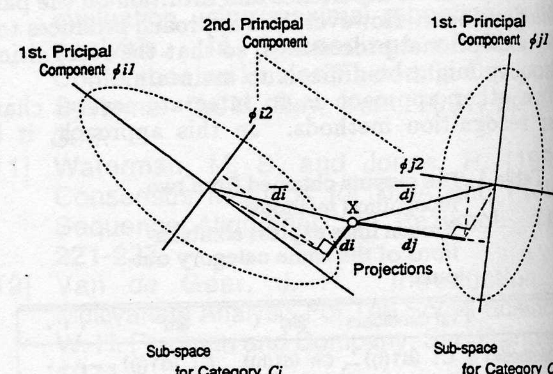

Fig. 3 Projection Distance Method.

which are obtained by the Principle Component Analysis Method(PCA), for each object category. In Figure 2, a subspace is determined by the Eigen vectors, ϕ_{i1} and ϕ_{i2}, for category C_i. For category C_j, a subspace is determined by ϕ_{j1} and ϕ_{j2}. The projection distance is determined as the length of a projection line from an input pattern to the subspace. The bold lines show the distances in Fig. 3. An unknown input pattern is expressed with X, and the projection distances for two categories are d_i for C_i, and d_j for C_j. Here, d_i is shorter than d_j. Therefore, input pattern X belongs to Category C_i. In this case, it is impossible to decide which category the input pattern X belongs to, because both Euclidean distances, $\bar{d}_i$ and $\bar{d}_j$, are the same. As described above, in the projection distance method, pattern distributions on the subspaces are taken into consideration to define the distance as a similarity between an input pattern X and each category.

To make recognition properties clear for the above-mentioned recognition methods, two experiments on Japanese "HIRAGANA" character recognition were carried out, individually. Table 3 shows the object set for the recognition experiments. They consist of 46 characters. They were picked from a handprinted Chinese character data base ETL-9B[5], which was released by the Electrotechnical Laboratory, Japan.

Table 3. Japanese "HIRAGANA" character set. They are cursivly written combined constant and vowel symbols used in written Japanese language texts.

あ(a)	い(i)	う(u)	え(e)	お(o)	は(ha)	ひ(hi)	ふ(hu)	へ(he)	ほ(ho)
か(ka)	き(ki)	く(ku)	け(ke)	こ(ko)	ま(ma)	み(mi)	む(mu)	め(me)	も(mo)
さ(sa)	し(shi)	す(su)	せ(se)	そ(so)	や(ya)		ゆ(yu)		よ(yo)
た(ta)	ち(chi)	つ(tsu)	て(te)	と(to)	ら(ra)	り(ri)	る(ru)	れ(re)	ろ(ro)
な(na)	に(ni)	ぬ(nu)	ね(ne)	の(no)	わ(wa)				を(wo)
					ん(n)				

One hundred samples per category were used for training, and 100 other samples were used as unknown patterns. Four thousand and six hundred patterns were used as training, and 4600 other patterns were used as unknown input patterns. Each experiment varied the picture sizes during the feature extractions, and varied the number of principal components.

Table 4 shows the results without rejections. An error ratio can be reduced for unknown patterns by using rejection, generally. However, recognition ratios decrease at the same time. Therefore, the maximum recognition ratio for each recognition method is individually obtained in these cases. The picture sizes for the feature extraction are 32 by 32 pixels in Method#1, and 48 by 48 pixels in Method#2. Ten principal components are used to calculate projection distance in both discriminations. The recognition ratios are 92.22% for Method#1, and 91% for Method#2.

Table 4. Recognition results for individual classification methods.

		Correct	Error	Reject
Method#1	Training	99.15%	0.85%	—
	Unknown	92.22%	7.78%	—
Method#2	Training	96.35%	3.65%	—
	Unknown	91%	9%	—

Table 5 shows the results corresponded to the assortment in Table 2. There are 3947 unknown patterns belonging to Case1. There are 295 unknown patterns in Case2, and 239 unknown patterns in Case3. One hundred and nineteen other unknown patterns belong to Case4. These results show that 12% of unknown patterns were recognized by one of the methods. From these evaluations, it was ascertained that these two recognition methods have mutual complement properties, as shown in Fig. 1.

If only answers belonging to Case1 as final recognition results are selected, the recognition ratio is obtained at 85.80% for unknown patterns. On the other hand, the recognition ratio is 97.41%, if a correct answer is selected from one of the methods, for Case2 and Case3. It is possible to increase the recognition ratio, by selecting correct answers in Case2 and Case3.

Table 5. Recognition result details corresponded to the assortment in Table 2.

	Case-1	Case-2	Case-3	Case-4
Method#1	Correct	Correct	Wrong	Wrong
Method#2	Correct	Wrong	Correct	Wrong
Training	4400	161	32	7
Unknown	3947	295	239	119

2.3 Reliability Evaluation Method

As described in Section 2.2, appropriate ways to cope with Case2 and Case3 are important to achieve higher recognition performance. The following proposes how to select a correct answer from either recognition method on Case2 or Case3.

A scattering diagram of training patterns for a category is shown in Fig. 4. The horizontal axis means the distance scale for Method#1, and the vertical axis means the distance scale for Method#2. Individual plots belong to Case1. In the case of training patterns, occurrence probabilities for Case2 and Case3 are small, as shown in Table 5. It is difficult to

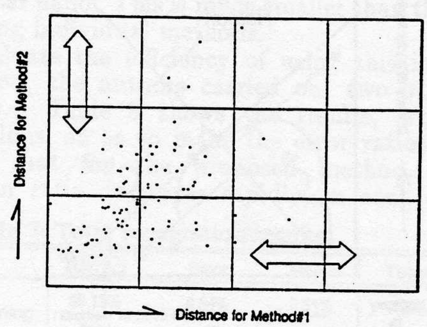

Fig. 4 A scatter diagram, showing relative relations.

precisely enclose existence regions, corresponding to Case2 and Case3, for scarce occurrences. Therefore, the authors assume that these regions exist around the arrow marks in Fig. 4, based on the following assumptions. In Case2, the distance value for Method#1 is apt to be small, and is apt to large for Method#2. The opposite tendency is applicable in Case3.

Figure 5 is used as a brief model for Figure 4 to explain the reliability evaluation method easily. According to the above assumptions, the authors model an acceptance area as an approximate triangular area, shown in Figure 5, to enclose the region corresponding to Case2 and Case3 as well as Case1, simultaneously. When unknown patterns can be accepted by the category as a final decision result, the unknown patterns exist in the triangular acceptance area.

Here, two kinds of recognition methods are used. Therefore, two primary candidates appear in Case2 or Case3. One of these candidates is assumed as a correct answer. A triangular area is made for each candidate. At least, one triangular area is selected, in which an unknown pattern exists, in order to recognize an unknown pattern.

To define the triangular areas automatically for individual categories, the first principal component axis for the Case1 distribution is used, as shown in Fig. 5. For one candidate category C_k, the principal component axis V_k is determined by an Eigen vector $(v_1(k), v_2(k))$. By using a perpendicular line L, the triangular area is defined. Perpendicular line L is drawn so as to be in contact with Case1 distribution S for the training patterns.

A decision can easily be made regarding whether or not an unknown input pattern exists in the triangular area, by calculating the first principal component score for the input pattern. Score E_i for candidate category C_i, determined according to Method#1, is described as

$$E_i = d_1(i)v_1(i) + d_2(i)v_2(i) \tag{1}$$

Here, $d_1(i)$ is the C_i distance value, obtained as the Method#1 answer. Also, $d_2(i)$ is the distance value for category C_i, obtained by Method#2. Score E_j for candidate category C_j, answered by Method#2, is also calculated in the same way.

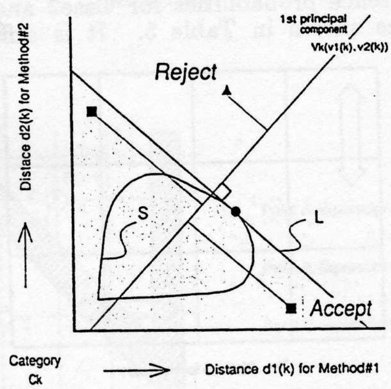

Fig. 5 Acceptance area definition, 1st principal axis for Category Ck.

Two threshold values, E_{imax} and E_{jmax}, were previously defined to decide rejections. Here, E_{imax} is a maximum value E_i calculated from training patterns according to C_i, and E_{jmax} is a maximum value of E_j. By comparing the score E_i with E_{imax}, the existence of an unknown pattern for the candidate category C_i is judged immediately, instead of defining the function of L.

3 Processing Flow and Final Decision Algorithm

3.1 Processing Flow

A recognition system was constructed, using the reliability evaluation method described in Section 2.3. Figure 6 shows the processing flow. The system consists of two recognition methods, a comparing and integrating part and a final decision part. The recognition methods can be operated independently.

First, the answers, C_i and C_j, which are individually obtained from both methods as the primary candidates, are compared (as shown in Fig. 6 (a)). If both answers are the same, the candidate category name is decideded as a final decision. If they are different, two distance vectors $(d_1(i), d_2(i))$ and $(d_1(j), d_2(j))$ are generated for the Method#1 answer, C_i, and the Method#2 answer, C_j, as shown in Table 1 (Fig. 6 (b)). Then, scores, E_i and E_j, are calculated, for C_i and C_j, according to Equation (1), (as shown in Fig. 6 (c)). They are compared with the maximum values, E_{imax} and E_{jmax}, individually to decide on the rejections (as shown in Fig. 6(c)). If candidate were remaining, the final decision result is obtained in the final decision part, by using E_i and E_j (as shown in Fig. 6 (d)).

3.2 Final Decision Algorithm

The principal component score, described in Section 2.3, has the following properties;

- The shorter the distance vector becomes, the smaller the score gets.
- Assuming two distance vectors, having the same Euclidean distance value from the center, the closer the distance vector gets to one of the axes, the smaller the score vector gets.

The second property is suitable to pick up a point, which belongs to Case2 or Case3 and locates near the axes, on each candidate triangular area. Therefore, the scores can be used as a reliability evaluation value, to decide which answer should be selected, if both distance vectors for the answers are not rejected.

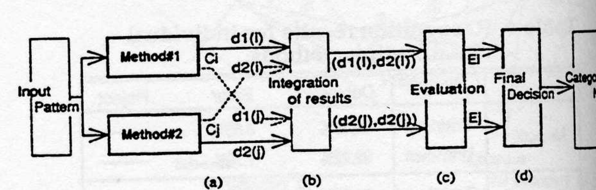

Fig. 6 Processing flow.

The decision algorithm is shown Fig. 7. The bold line indicates the final decision part. First, E_i and E_j are compared with E_{imax} and E_{jmax}, individually (as shown in Fig. 7 (c)). If both of the evaluation values are greater than the threshold value, the system rejects the input pattern (as shown in Fig.7 (c)). If not, a candidate answer, which has a smaller evaluation value, is selected (as shown in Fig. 7 (d)). If the evaluation value is greater than the corresponding threshold value, as shown in Fig. 7(e), the candidate answer is rejected (as shown in Fig. 7 (g)). If they are not greater, the candidate answer is selected as a final decision answer (as shown in Fig. 7 (f)).

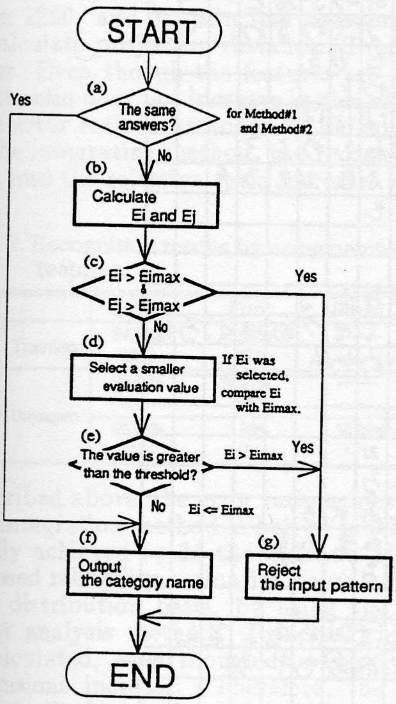

Fig. 7 Final decision algorithm.

4 An Experiment in Japanese "HIRAGANA" Character Recognition

An experiment was carried out to investigate the effectiveness of the reliability evaluation method. The final decision objects are for Case2, Case3 and Case4.1. They were previously picked up. The total of 196 objects was used for training patterns, and a total of 628 for unknown patterns. Four training samples correspond to Case4.2, and 25 unknown patterns correspond to Case4.2.

The system regards the primary candidates as a final recognition result, for input patterns having the same answers. It is impossible to find patterns belonging to Case4.2 in the processing flow. This is because the two kind of answers, for Casel or Case4.2, can not be discriminated between. The number of patterns belonging to Case4.2 is small. Therefore, the countermeasures for Case4.2 are shelved this time, and are added to the total error sum.

Table 6 shows details regarding the decision results. 48.89% of the objects are discriminated correctly, and 32.48% of the objects are rejected, for the unknown pattern examination. Object pattern images are shown in Fig. 8. The bold line frames indicate the mistakenly recognized patterns on the final decision.

Table 6. Details of the final decision result.

Correct	Error	Reject
307	117	204
48.89%	18.63%	18.63%

5 Discussion

On the final decision, the mistakenly recognized 117 patterns, shown in Table 6, are segregated into 2 groups as follows;

- Type A:Neither of the two recognition methods has a correct answer.
- Type B:A wrong answer is selected as the final decision, even though one of the methods has a correct answer

The patterns belonging to Type-A should be rejected. Here, 35 patterns fall into Type A. In this experiment, about 70% of all pattern in Type-A were rejected. One of the solutions for Type-A is to change the threshold values for individual categories, so that the rejection ratio might increase. Eighty two patterns fall into Type B. These patterns should be selected as the correct answer output by one recognition method. In this paper, the authors used primitive recognition methods, having mutual complement properties, in order to easily discuss the effectiveness of the integrating methods. The two recognition methods, used here, are very simple and cannot discriminate between similar characters. For similar character recognition, it is necessary to use efficient primitive recognition methods and then integrate them.

The total recognition results, using this integrating method, are shown in Table 7. The recognition ratio is kept at 92%, which is as good as the recognition ratio for each recognition method individually experimented upon. The error ratio was reduced to 3.09%, on the other hand. This is much smaller than the error ratios using individual methods.

To evaluate the efficiency of using this integrating method, the authors carried out two other experiments. Table 8 shows the results, when using rejections, so as to make the error ratios equivalent to that for the proposed method. The recognition ratio decreases rapidly, in each case of

Table 7. Total recognition results.

	Correct	Error	Reject	Total
Training	99.15% 4561	0.54% 25	0.31% 14	100% 4600
Unknown	92.48% 4254	3.09% 142	4.43% 204	100% 4600

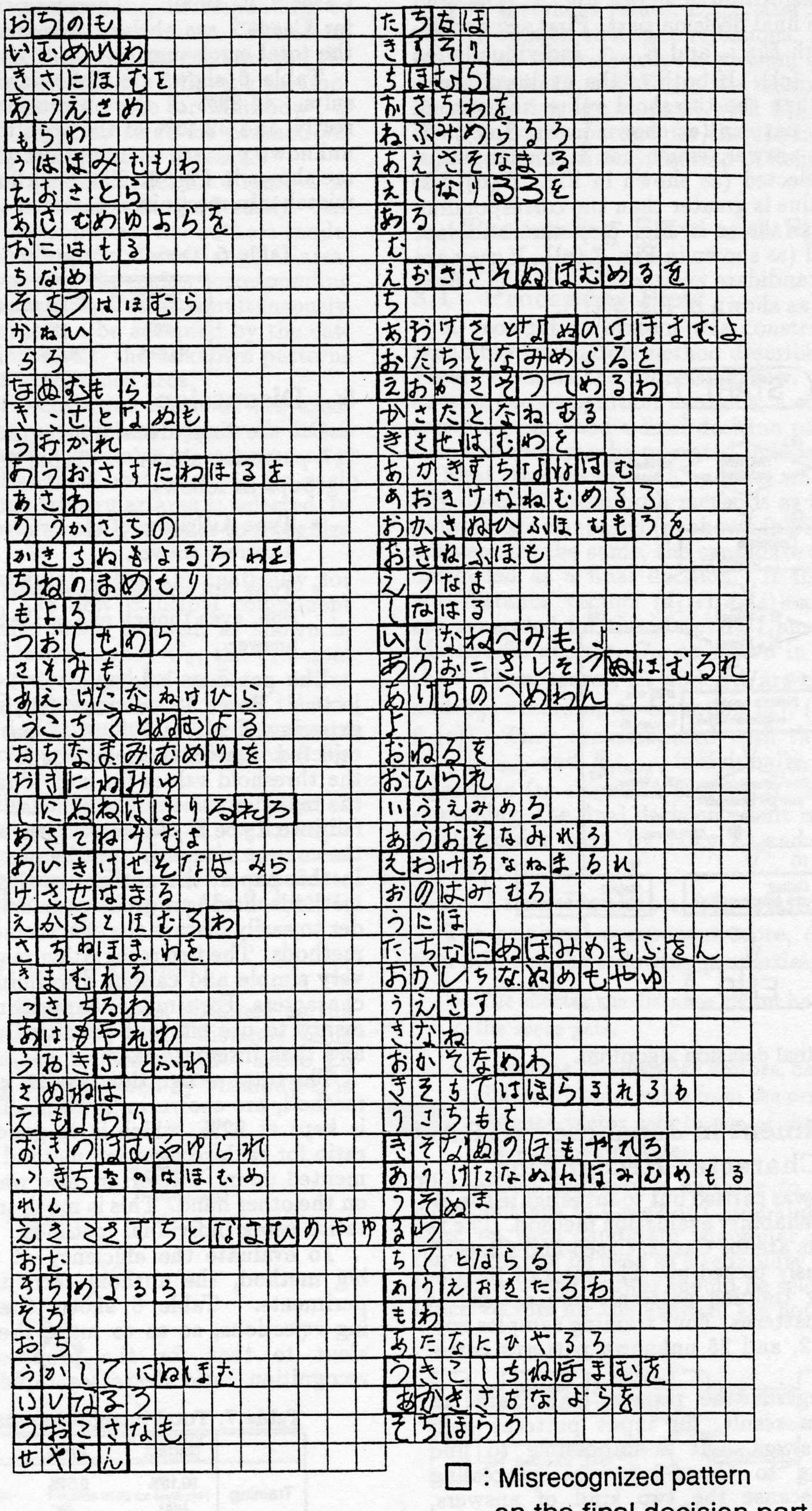

Fig. 8 Final decision objects.

Method#1 and Method#2. Then, rejection ratios also increase.

Table 8. Recognition results for each classification methods using thresholding for rejection.

	Correct	Error	Reject	Total
Method#1	69.13% 3180	2.85% 131	28.07% 1289	100% 4600
Method#2	67.61% 3110	2.83% 130	29.56% 1360	100% 4600

In the other experiment, two kinds of features, which are used in Method#1 and Method#2, are combined into a feature vector. This is an integration method on the feature extraction level. The dimensions were 1280, and 10 principal components were used to calculate projection distance. Table 9 shows the results. Even though the features are combined, the correct ratio does not increase in this experiment. When the error ratio is arranged to be equivalent to that for the integrating method, the recognition ratio decreases, and the rejection ratio increase.

Table 9. Recognition results by using combined features.

	Correct	Error	Reject
Training	94.11% 4329	5.89% 271	——
Unknown	89.59% 4254	10.41% 142	——
Unknown	66.33% 3051	2.85% 131	30.82% 1418

As described above, the error ratio achieved by the proposed integration method is half as much as those individually achieved by Method#1 and Method#2. The proposed reliability evaluation value is calculated on Case1 distribution basis, by using the principal component analysis method. Principal components can be calculated, even though the Case1 distribution dimensions increase. Therefore, the proposed method is applicable to when more than three recognition methods are integrated. The authors integrated mutual complement recognition methods, which have the same recognition object, in order to achieve more efficient performance. In order to expand recognition objects, it is necessary to integrate recognition methods, which have different recognition objects, like numerals and Chinese characters, on the contrary.

6 Conclusion

A new handwritten character recognition method was described. Two different recognition methods were integrated, in order to achieve a higher performance. The authors found the properties needed on the integration methods, and made it clear that mutual complement properties were necessary for the primitive methods. Then, on the basis of the investigation results, an integration method for recognition methods was presented. An evaluation measure, which expresses the reliability for each category, was proposed in the integration. The evaluation value was used to select correct answers for inputs, if different answers were obtained by the recognition methods. The value was also used to reject input patterns, if an assured correct answer wasn't obtained by the recognition methods. To estimate the evaluation value efficiencies, a recognition system was constructed.

The integration method efficiencies are summarized as follows;

- Both error ratio and rejection ratio were reduced simultaneously. Especially, about 50% reduction was obtained for the error ratio.
- The input patterns, where correct answers weren't obtained from both recognition methods, were rejected at 70% rate.

The rejection ratio reduction was also confirmed in the experiment using the integration method. No efficiencies were obtained, even though the two feature vectors, which were used in Method#1 and Method#2, were combined.

As described above, basic effectiveness of the proposed recognition method was confirmed through an experiment in Japanese "HIRAGANA" Character Recognition. It is necessary to study a method to presume the distributions corresponding to Case2 and Case3 more precisely. The performance of the integrated recognition method can also be raised by studying on mutual complement properties of recognition methods theoretically. The recognition ratio can be increased to select suitable methods for recognition objects.

Acknowledgement

The authors are indebted to their colleagues in Pattern Recognition Research Laboratory for their helpful discussion and cooperation.

References

[1] Hagita et al; Classification of Handprinted Chinese Characters by Global and Local Stroke Density, Trans. IECE Japan (in Japanese), vol. J65-D, No. 6 1982.

[2] T. Hai et al; A Method for Handwritten Kanji Character Recognition -Recognition Method by Multiple Standpoints and Particular Shape Extraction-, Trans. IEICE Japan (in Japanese), vol. J68-D, No. 4, 1985.

[3] Ikeda et al; Projection Distance Method for Recognition of Hand-written Characters, Trans. IPS Japan (in Japanese), vol. 24, No.1, 1983.

[4] H. A. Glucksman; Classification of Mixed-font Alphabetics by Characteristic Loci, 1967 Digest of 1st Ann. IEEE Computer Conf. 1967.

[5] T. Saito et al; On the Data Base ETL9 of Handprinted Characters in JIS Chinese Characters and Its Analysis, Trans. IEICE Japan (in Japanese), vol. J68-D, No. 4, 1985.

[6] E. Oja; Subspace Methods of Pattern Recognition, Research Studies Press Ltd. 1983.

[7] C. Y. Suen et al; Streching the Limits of Handwriting Recognition Systems, Proc. of 2nd IPTP Conf. on Postal Processing Systems in the 21st Century and Character Recognition 1992.

[8] C. Nadal, R. Legault and C. Y. Suen; Complementary algorithms for the recognition of totally unconstrained handwritten numerals, Proc. 10th Int. Conf. on Pattern Recognition 1990.