

Structural Learning of Character Patterns for On-Line Recognition of Handwritten Japanese Characters

Masaki Nakagawa and Le Van Tu

Dept. of Computer Science, Tokyo Univ. of Agriculture and Technology
2-24-16 Naka-cho, Koganei, Tokyo, 184, Japan

Abstract. This paper describes the structural learning of Kanji patterns in on-line handwritten character recognition. Upon the request to learn an input pattern, the system investigates which subpattern or the pattern as a whole is non-standard, registers the (sub)pattern and extends the effect of the registration to all the character categories whose shapes include it. A character pattern representation dictionary stores character patterns as combinations of subpatterns so that a common subpattern is shared by all the character categories which include it in their shapes. The recognizer constructs all template patterns from its constituent subpatterns for matching them with an input pattern. Registration of a character pattern invokes identification and registration of a non-standard subpattern in it so that the effect extends to all the characters whose shapes include it. A preliminary evaluation shows it is highly effective without any bad side effect.

1 Introduction

The on-line recognition of handwritten Japanese characters has been studied for nearly 30 years toward the relaxation of constraints on character patterns to be recognized. Stroke-number-free and stroke-order-free methods have been sought[1, 2]. Nevertheless, heavily deformed patterns, patterns of writer's own style, etc. are still hard to recognize without registering templates for them in the pattern dictionary.

In the case of on-line recognition, demands for customization are especially high since it is most likely employed for personal use. If writer-dependent distorted, deformed or simplified patterns might not be recognized at first but made recognizable by registration or machine learning, usability is improved and user satisfaction will be derived. Therefore, functions to register non-standard patterns on the spot have been provided. This registration, however, has still been unsatisfactory.

When people write Kanji or Chinese character patterns, their writing is generally consistent in the sense that they write a radical (subpattern) in the same way in almost all Kanji (Chinese) characters whose shapes include the radical. For instance, if one has the habit to write the radical “言” in one's own non-standard style, one usually employs that style in writing “語”, “論”, “証”,

“說”，“誤”，“認”，“識” and almost all the characters whose shapes include the “言” radical.

In conventional on-line recognition systems, a user who writes a non-standard style of “言” has to register not only a pattern for it, but also non-standard character patterns to all the character categories of the “言” family.

This paper presents a solution to this problem which automatically identifies which subpattern or the pattern as a whole is non-standard, registers the (sub)pattern and extends the effect of the registration to all the character categories whose shapes include it. Consequently, work for pattern registration is greatly reduced and user satisfaction will be increased.

Even if automatic identification of a non-standard subpattern is not made, structured character pattern representation where common radicals are shared enables the registration of a radical pattern to have effect on all the character categories whose shapes include the radical [3]. Without the identification of a non-standard subpattern by the system, however, one must investigate which subpattern is erroneous and register it to the subpattern category. Here the following problems exist. First, the human idea of radicals does not necessarily correspond to machine defined radicals. Second, human registration might not be the best registration for the system.

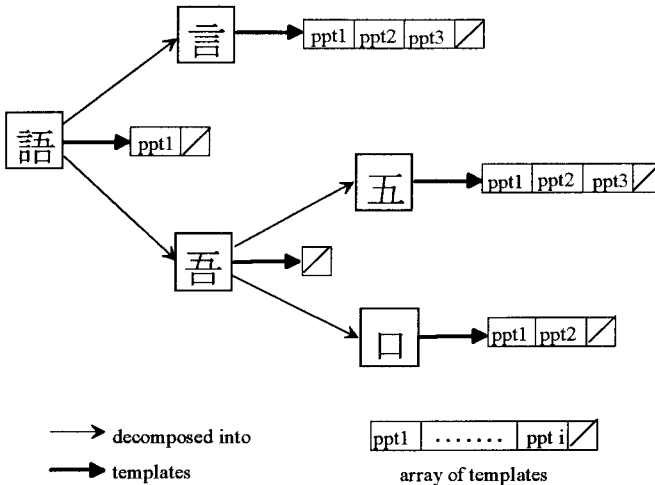


Fig. 1. Structured character pattern representation.

2 Structured character pattern representation

A character pattern category is hierarchically decomposed into composite subpattern categories and basic subpattern categories. A composite subpattern category is further decomposed into composite subpattern categories or basic subpattern categories, while a basic subpattern category is a terminal structure. Actual templates are usually registered to basic subpattern categories but may be registered to composite subpattern categories or character categories. Fig.1 shows an example. In this case, $3 \times 2 \times 3 + 1 = 19$ kinds of character pattern templates for “語” can be produced from this representation. A basic or composite subpattern category is shared by character or composite subpattern categories whose shapes include it.

3 Total learning strategy

The total learning strategy is shown in fig. 2. A character pattern which does not have a subpattern structure is treated by the non-structural learning where a character pattern is registered as a whole, while a character pattern which is supposed to have subpatterns is processed by the automatic structural learning. It will be detailed later. A character pattern which has subpatterns might be deformed so badly that the system could not find any constituent subpatterns. In such a case, the system judges that the automatic structural learning may not work correctly and asks the user to separate the pattern into constituent subpatterns.

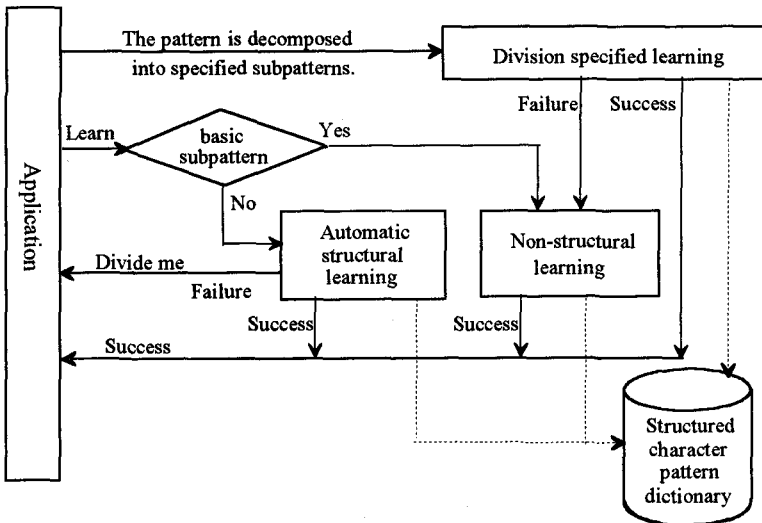


Fig. 2. Total learning strategy.

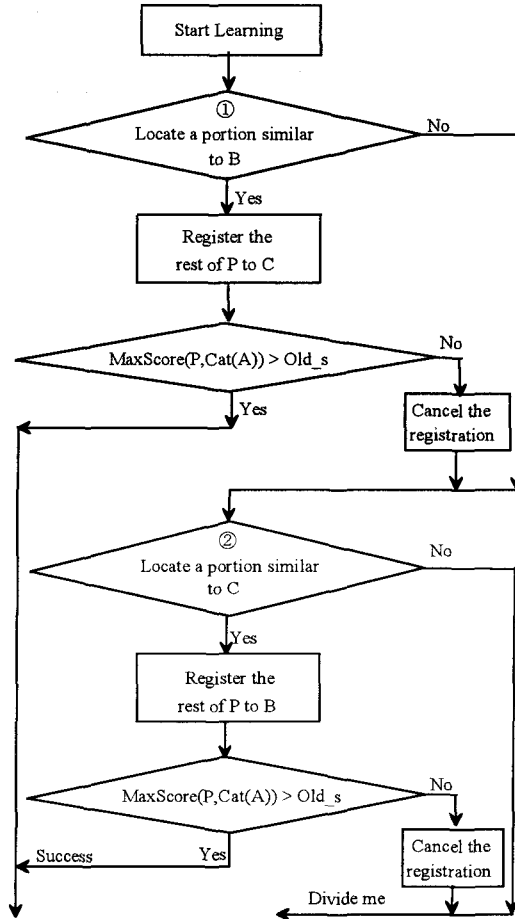


Fig. 3. Automatic structural learning.

4 Automatic structural learning

The structural learning is automatically invoked by the user's indication to learn an input pattern for a character pattern category which is composed of subpattern categories. Assume an input pattern P is misrecognized as X by the similarity value Old_s , although a user wants P to be recognized as a character category A . Then, the user will tell the system to learn P as A . Further assume that A is composed of B and C . $Cat(A)$ denotes a set of template patterns that can be generated for the category A and $MaxScore(P, Cat(A))$ denotes the highest similarity between P and members of $Cat(A)$.

The learning process is shown in fig. 3. Assume that P is already normalized, i.e., expanded linearly to a square with the horizontal and vertical ratio

preserved. First a portion in P similar to B is searched for. If it is not found, a portion similar to C is searched for. If a portion similar to B is found, the rest of P is hypothetically registered to the C category. If $MaxScore(P, Cat(A))$ becomes greater than $Old.s$ after the registration, the learning is completed. Otherwise, the process goes to find a portion similar to C . Again, if it is found, the rest is hypothetically registered to B . If $MaxScore(P, Cat(A))$ becomes greater than $Old.s$ after the registration, the learning is completed. Otherwise, the input pattern is subjected to the division specified learning in fig. 2 where the input pattern is divided into subpatterns by a user.

5 Effect

The recognition system to which this learning capability has been incorporated is working on the Windows 3.1 operating system. The combination of a linear-time elastic matching [4] with coarse classification and context post-processing has realized 90 to 95 % recognition rates without learning for unconstrained character patterns of 3345 kinds. The recognition time is about 0.5 sec. for an



Fig. 4. Structural learning for the “言” family and the “木” family.

input pattern on i486 DX2/66Mhz processor. The system is far more robust than conventional systems, but the customization capability is still essential to recognize writer dependent patterns.

The upper window in fig. 4 shows misrecognitions of character patterns which have deformed “言” radicals and those of character patterns which have pictorial “木” radicals. At first, the system has no way to recognize them. Then, an arbitrary pattern is selected in each group (the first pattern in each case) and designated to be learned. Effect is shown in the lower window in fig. 4.

Arbitrary selection of a pattern to be learned reflects the real use of on-line recognition. When a system does not recognize some character pattern, a user would designate the learning of that pattern and expect its effect to spread. Although the registration of an averaged pattern is generally most effective for the system, to realize this in on-line recognition for personal use is very difficult.

Next, our database of 12,000 character patterns/person which have been collected in a sequence of sentences without any writing instructions, was used in the evaluation of this structural learning[5]. Precisely speaking, the text from which character patterns are sampled is composed of 10,154 characters of sentences which include 1227 frequently used character categories and 310 accompanying character categories, as well as 1808 less frequently used characters with one pattern for each category, thus covering the JIS 1st set of 3345 categories.

Table 1. Number of character categories and patterns which include “木”, “田”, “隹”.

Radical	No. of char. categories	No. of char. patterns
木	205	304
田	70	112
隹	50	70

Table 2. Recognition rates without learning for 6 people's patterns of the “木”, “田” and “隹” families.

Set	木 family	田 family	隹 family
	1st hit rate- 10th cumulative rate(%)	1st hit rate- 10th cumulative rate(%)	1st hit rate- 10th cumulative rate(%)
1	86.18-93.09	36.61-65.18	45.71-60.00
2	85.83-88.82	49.11-66.96	65.71-77.14
3	75.66-78.62	78.95-84.21	78.57-81.25
4	89.14-94.08	94.64-94.64	90.00-95.71
5	83.88-86.84	72.32-84.82	24.29-32.86
6	88.16-91.78	88.39-92.86	81.43-82.86

Before learning	75.66-78.62	36.61-65.18	24.29-32.86
After learning	81.25-83.88	91.96-96.43	80.00-87.14

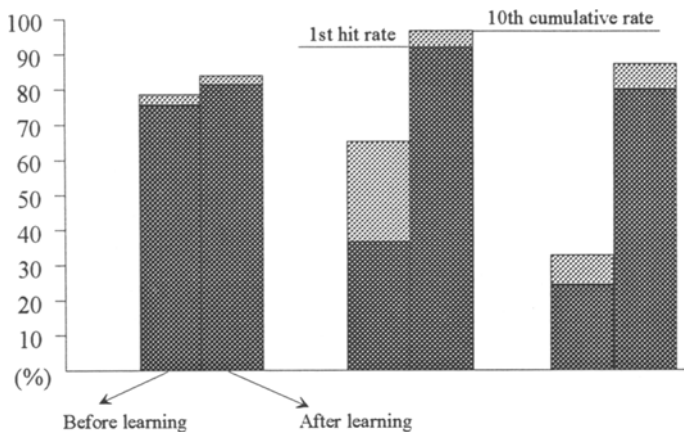


Fig. 5. Effect of learning for the worst sets in table 2.

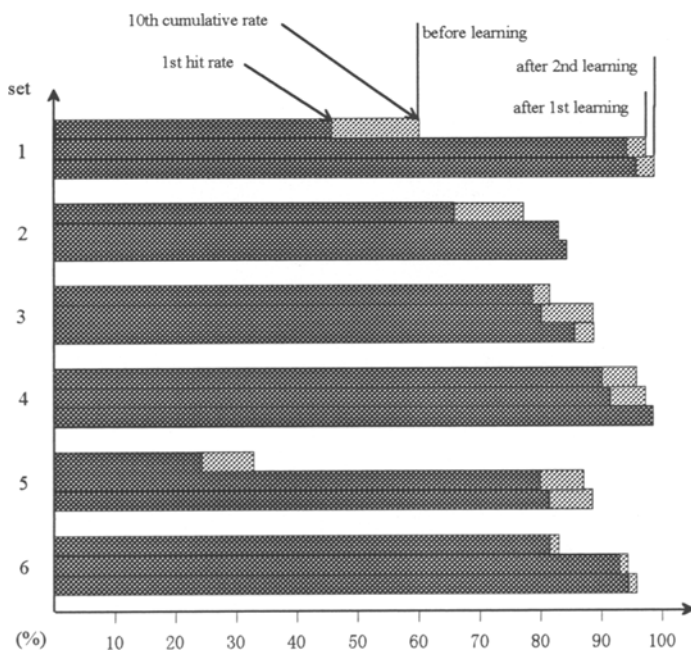


Fig. 6. Effect of learning for the 6 people's “佳” family characters.

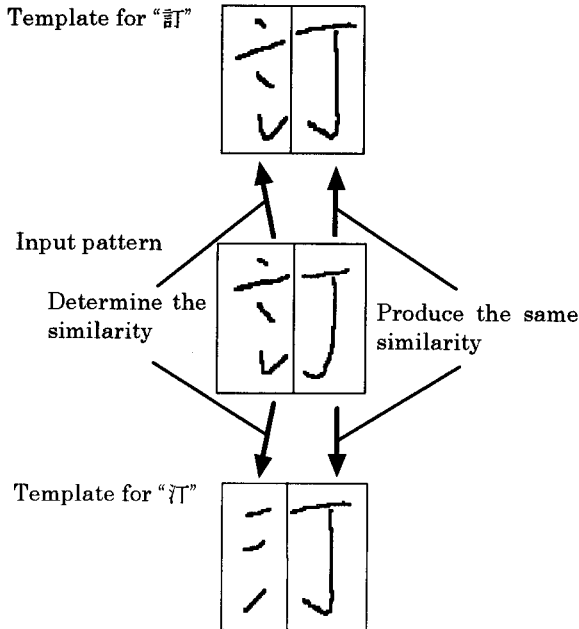


Fig. 7. Distinction by the distinctive part.

Although patterns of 80 people are available at the moment, this paper reports the preliminary evaluation examining 6 people's patterns simply because we have just started this evaluation.

Table 1 shows the number of character categories and patterns whose shapes include "木", "田", "隹". These character patterns were extracted from the 6 people's patterns and 3 x 6 sets were examined.

In the structured character pattern representation dictionary, one neat pattern has been prepared for each of "木", "田" and "隹". Except that it has representations for all the 3345 character categories including common variations. This dictionary is referred to as *Dict*. Recognition rates for 3 x 6 sets without learning are shown in table 2, where the 1st hit rate denotes the rate that the correct category is nominated as the top candidate while the 10th cumulative rate denotes the rate that the correct category is nominated within the 10th candidate.

In each column of table 2, the worst set was chosen for learning. The first misrecognized pattern was learned. The recognition rates have been improved as shown in fig. 5.

Next, learning was attempted for each of the third column of table 2, i.e., 6 sets of the "隹" family character patterns. The radical "隹" is called *Hurutori* and known as one of the radicals that people write in many ways. For each set, the first misrecognized pattern was learned so that the initial dictionary *Dict*

was individually updated and the recognition rate was evaluated. Then, with each updated dictionary, again the first misrecognized pattern was learned so that each dictionary was further individually updated and the recognition rate was evaluated. The effect is shown in fig. 6. The increase from poor recognition rates is high since structural learning extends very effectively.

The side effect of this learning on other characters has not been observed in the preliminary evaluation. This is due to the fact that our recognition method is highly distinctive. Even if a side effect exists, it may only change candidate order in a few cases but the disorder will be canceled by the context post-processing.

One of the reasons that the recognition method is distinctive is due to the structural character pattern representation. When the deformed radical for “言” is registered as shown in fig. 4, the template patterns for “訂”, “計”, “証” and so on may become closer to those for “汀”, “汁”, “江” and so on, respectively. Because the right radical patterns are shared by each pair such as “訂” and “汀”, however, distinction is focused to the left radical. That is, distinction is made by the distinctive part as shown in fig. 7.

6 Conclusion

Structural learning of kanji patterns for on-line recognition has been described. A preliminary evaluation shows it is highly effective without any bad side effect.

References

1. Tappert, C.C., Suen, C.Y. and Wakahara, T.: The State of the Art in On-line Handwriting Recognition, IEEE Trans. PAMI, 12, 8 (1990) 787-808
2. Nakagawa, M.: Non-keyboard Input of Japanese Text — On-line Recognition of Handwritten Characters as the Most Hopeful Approach—, Japanese Information Processing, IPS Japan, 13, 1 (1990) 15-34
3. Komoda, C., Nakagawa, M. and Takahashi, N.: Grammatical Representation of Hierarchical Structure of Kanji Patterns and Its Advantage for On-line Recognition of Simplified Distorted and Wrong Stroke Order Patterns (in Japanese), Trans. of IEICE, J70-D, 4 (1987) 777-784
4. Nakagawa, M. and Akiyama, K.: A Linear-time Elastic Matching for Stroke Number Free Recognition of On-line Handwritten Characters, Proc. 4th IWFHR (1994) 48-56
5. Nakagawa, M., Higashiyama, T., Yamanaka, Y., Sawada, S., Le, T.V. and Akiyama, K.: Collection and Utilization of On-line Handwritten Character Patterns Sampled in a Sequence of Sentences without Any Writing Instructions, Technical Report of IEICE, PRU95-110 (1995) 43-48