Proceedings of the 6th IWFHR, pp.587-596 (1998.8)

HANDWRITING-BASED USER INTERFACES EMPLOYING ON-LINE HANDWRITING RECOGNITION

M. NAKAGAWA, K. AKIYAMA, T. OGUNI and N. KATO

Dept. of Computer Science, Tokyo Univ. of Agri. & Tech.
2-24-16 Naka-cho, Koganei, Tokyo, 184, Japan.
phone: +81-423-88-7144
fax: +81-423-87-4604
E-mail: nakagawa@cc.tuat.ac.jp

This paper describes our research to enhance handwriting-based user interfaces. It consists of building infrastructures, advancing handwriting recognition technology, studying human interface and developing applications. Our goal is to provide consistent and creative human interfaces among a wide range of pen input devices such as PDA, desktop tablet and electronic whiteboard. The on-line handwritten character recognition is the key technology. Our method has marked 90 to 95 % correct recognition rates without learning to a large database of on-line handwritten Japanese text. Its recognition speed is about 0.02 sec./character on a Pentium 200 Mhz processor and roughly 0.2 sec./char. on a small PDA machine. The method is not only robust to stroke connections and pattern distortions but also highly customizable for personal use. Upon the request of learning an input pattern, it identifies a deformed subpattern (radical), registers the (sub)pattern and extends the effect to all the character categories whose shapes include it. This paper also describes educational applications which benefit from pen interfaces and the handwriting recognition engine.

Keywords: pen interfaces, on-line recognition, educational applications.

1. Introduction

Although character recognition is language dependent, handwriting itself is universal. English, Chinese, Arabic, Japanese, etc. or even pictorial languages can be expressed using a single pen. By writing with a pen, one can express one's thinking most easily using pen-trace patterns. Moreover, thinking is not interrupted by the actions for writing. Thinking and writing form a positive feedback loop system to grow and clarify one's idea. This nature of writing is suited for creative work rather than labor-intensive tasks.

To realize creative human computer interfaces using handwriting while preserving the above characteristics, handwriting recognition technology, handwriting-based human interfaces and applications must be studied and enhanced. Even without machine recognition, however, there are many applications where the pen is mightier than the mouse. With recognition, the power of handwriting is greatly extended.

2. Research Strategy and Structure

Our basic stance in this research is to study handwriting recognition, handwriting-based human interfaces and applications all based on common stable infrastructures and to pursue improvement and enhancement in each of them. Of course, it is difficult to perfect each of these subjects, but even if each of them is neither perfect nor generic enough, their well-balanced combination may provide a platform to work on and enhance each of them. Moreover, even if each of them has some weak points, a combination that would compensate each other's weak points could be found if they are studied totally. Therefore, our research is composed of four layers as shown in Figure 1.

After the boom of pen computing ended in a short period, it has been suffering from the disappointment that the premature products for inappropriate applications have brought to the users. Nevertheless, we think that the pen has enormous potentials for future applications that is yet to be exploited.

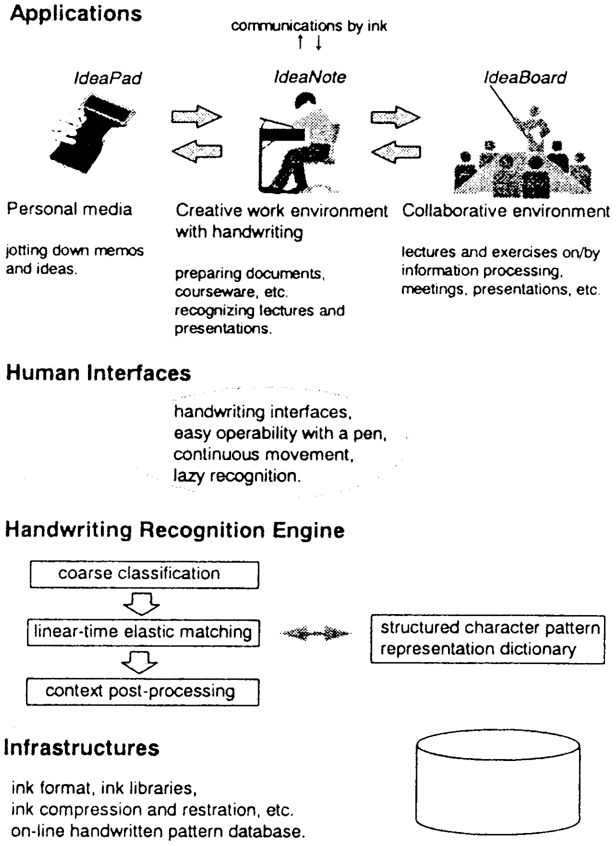


Figure 1. Overall structure of the handwriting interfaces.

3. Infrastructures

The infrastructures to promote handwriting-based human interfaces include efforts to define an electronic ink format [1], development of library routines to process electronic ink [2] and provision for a large database of on-line handwritten character patterns storing about 3 million patterns. Its details are described in [3].

4. Character Recognition Engine

4.1 Basic Approach

In the trend of personalization of information processing systems, we have been seeking for on-line recognition oriented to personal use. We have the following knowledge on character pattern distortions and variations in our previous study:

(1) There are wide variations in the stroke order to form a character pattern, but it is considerably stable in each individual's writing. Moreover, it is stable in writing of each radical among kanji patterns which include the radical.

(2) Stroke connection is unstable even in writing of one person. It is hard to guess where strokes are likely connected.

(3) It is often that a fairly long invisible pen movement between strokes is turned visible and forms excessive part to a normal character pattern.

For personal use, stroke order is stable so that the variations can be prepared in the character pattern dictionary. Even for anonymous use, stroke order is never at random and registration of various stroke orders to a limited number of radicals improve the overall performance significantly. They may be effectively shared as subpatterns. On the other hand, stroke linkage and appearance of pen movement into visible strokes are unstable even in writing of one person. If the stroke order is not assumed, provision must be made for possible stroke connections and pen movement appearances in the character pattern dictionary. Since they are unstable, however, this provision cannot be adequate and it is often violated. If the stroke order is assumed, on the other hand, their occurrences only have to be considered between successive strokes.

These considerations have lead to our current approach:

(1) Stroke order is represented in the character pattern dictionary and a matching algorithm solves stroke connections and pen movement appearances with the real-time constraint on a low performance CPU satisfied.

(2) In character pattern representation, common subpatterns are shared so that registering variations to a subpattern extends its effect to all the character categories whose shapes include the subpattern.

4.2 On-line Character Recognition System

The on-line handwritten character recognition system is composed of a coarse classifier, linear-time elastic matcher, structured character pattern representation (SCPR) dictionary and context post-processor.

The coarse classifier first employs on-line features such as X-, Y- projection lengths to screen candidates and then matches principal components of directional features between input and candidate patterns so that it reduces the candidates from more than 4,000 to 200 with nearly 99 % inclusion of the correct answer in the candidates.

The linear-time elastic matcher accepts stroke connections and considerable distortions robustly. It is 6 or 7 times faster than DP-matching using beam search with comparable recognition performance.

The SCPR dictionary represents more that 4,000 categories by roughly 700 subpatterns so that its size is reduced to about 150 Kbyte which is far smaller than those of conventional systems. Not only the small size, it provides the structural learning described later.

The context postprocessor employs a bigram table derived from a large volume of Japanese sentences and applies the Viterbi search to find the best sequence of characters. Roughly speaking, it can correct the half of recognition errors. Its processing time is 1/1,000 of that for character recognition so that its processing time is completely negligible. Its problem has been the size of the bigram table for more than 4,000 categories. However, our recent study has shown that it can be reduced to 100 Kbyte without decreasing its effect significantly.

The basic algorithms of the coarse classification, linear-time elastic matching and context post-processing are described in [4, 5, 6]. Each subsystem and total combination has been tuned using the learning sets in the database. The total engine has marked 90 to 95 % correct recognition rates without learning to the test sets of the database of on-line handwritten Japanese text. The overall recognition time is about 0.02 sec./input character on a Pentium 200 Mhz processor and roughly 0.2 sec./char. on a small PDA machine.

4.3 Structural Learning

In the case of on-line recognition, demands for customization are especially high since it is most likely employed for personal use. Even if writer dependent, distorted, deformed or simplified patterns might not be recognized at first but made recognized by registration or machine learning, usability is improved and user's satisfaction will be derived. For this reason, many systems are equipped with the capability of learning such patterns. However, each unrecognized pattern has to be registered or learned one by one.

We made a much more effective method. When the system is requested to learn an input pattern, it investigates which subpattern (radical) or the pattern as a whole is non-standard and registers the (sub)pattern to the SCPR dictionary. Since common subpatterns are shared by characters, the effect of registering a (sub)pattern extends to all the character categories whose shapes include it.

Figure 2 shows deformed handwritten patterns of the "言" family and Figure 3 shows their recognition results. At first, the system has no way to recognize them correctly. Then, an arbitrary pattern (say, first pattern) is selected and designated to be learned. The system identifies the deformed radical (shown in gray) of "言", although it is connected to the right radical, as shown in Figure 4. After this learning, the system can recognize deformed patterns of the "言" family as shown in Figure 5.

The structural learning not only registers radicals to the SCPR dictionary but also updates the dictionary for the coarse classification. When a deformed radical is added to the SCPR dictionary, template character patterns which may include it are generated and each is tested whether the coarse classification accepts or rejects it. If it is rejected, the coarse classification dictionary is updated so that it is included in the candidates. The basic mechanism was described in [7].

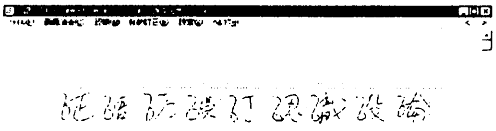
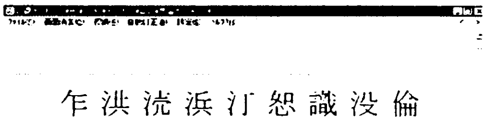


Figure 2. Deformed handwritten patterns of the "言"family.



乍 洪 浣 浜 汀 怨 識 没 倫

Figure 3. Recognition results before the structural learning.

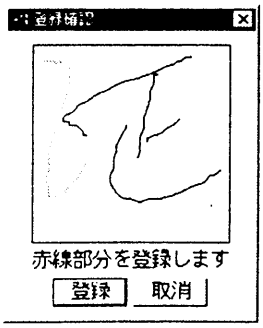
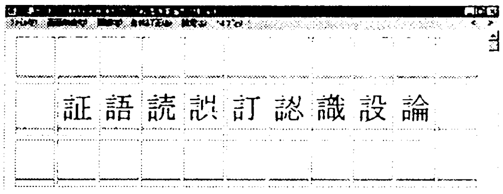


Figure 4. Identification of a non-standard radical.



証 語 読 訳 訂 認 識 設 論

Figure 5. Recognition results after the structural learning.

Since there are about 100 categories in the JIS (Japanese Industrial Standard) 1st set whose patterns include "言", it seems important that the learning of a single character pattern extends to all the 100 categories. On previous systems, people had to register his patterns to the 100 categories. On the other hand, people only have to teach once on our system.

From our database study, each individual has his own deformations. They appear in several hundreds character categories but deformed are often common radicals. This implies that people can teach these deformations to the system by just teaching several patterns which include his style of radicals.

5. Human Interfaces

Human interfaces are as important as handwriting recognition technology so that handwriting-based systems can be accepted by users. We have been proposing lazy recognition scheme in which the display of recognition is delayed until needed [8, 9]. Until this proposal, experimental systems and products had been applying handwriting recognition immediately after each pattern was written, causing the user's thinking to be interrupted due to confirmation of correct recognition and correction of misrecognition. One's thought is better developed by working with one's handwriting. Pattern recognition must be applied only when handwriting is viewed as a group of codes. The lazy recognition scheme also facilitates context post-processing.

So far, several specific interfaces have been developed for a small PDA [10], a desktop tablet [11] and a large interactive electronic whiteboard [12].

6. Applications

Without meaningful applications, nobody would use handwriting-based user interfaces. We think that a group of such applications is those that involves creative thinking. Moreover, a pen is a scalable interface, i.e., it can be used for PDAs, desktop tablets and large interactive boards. The complete system may be used in the following fashion: You write down memos and ideas on a PDA while you are away from your office, download them to the desktop machine and apply pattern recognition when necessary. You communicate with others by handwritten e-mail and search handwritten mail using a search engine. You may invoke recognition engines to process handwriting and utilize text processing. You may also prepare material for presentations on a desktop and present it on an electronic board. The live presentations can be recorded and reviewed on a desktop or a PDA.

A combination of small, medium, and large pen input devices may seem similar to the ubiquitous computing at PARC [13], but our aim is to provide the power of computing to our activities of writing on various sizes of papers, boards and so on. In contrast to this, the research at PARC is oriented towards collaboration where a small device (called ParcTab) is mainly used to identify the user's location rather than for jotting down memos or ideas.

The following sections describe applications which benefit from pen interfaces and the robust handwriting recognition engine.

6.1 Programming Education on the IdeaBoard

A teacher writes a program, explains it, then makes the system recognize and execute it. The teacher can capture audience's attention to his writing as well as show the execution of the program in front of the audience.

This application runs on a large interactive electronic whiteboard named *IdeaBoard*. Figure 6 shows the latest hardware. It has 680 cd/m^2 brightness, 180:1 contrast rate and 1600 x 1280 resolution based on the rear-projection 70-inch High Vision TV technology. Brightness of the display is most important. If a class room has to be dark like a theater, the demerits overwhelm than the merits. The display must be usable in a well-lighted environment so that students can see the teacher's face and actions as well as the contents in the screen.

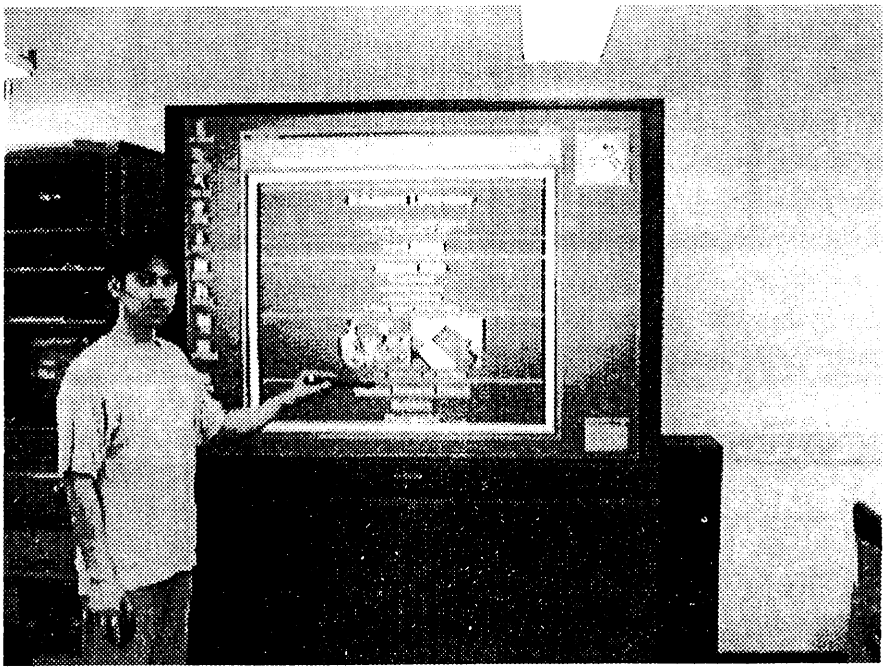


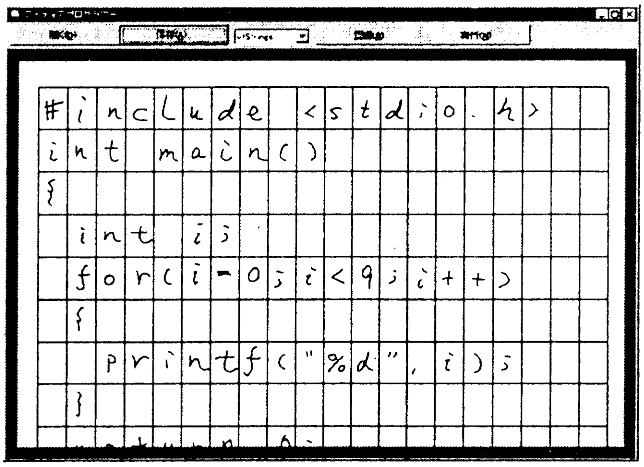
Figure 6. *IdeaBoard* in use for WWW browsing

Figures 7, 8, and 9 show a handwritten program, the result of its recognition using the context of C programs and the result of its execution, respectively.

Here, we follow the lazy recognition scheme which delays the display of recognition until needed.

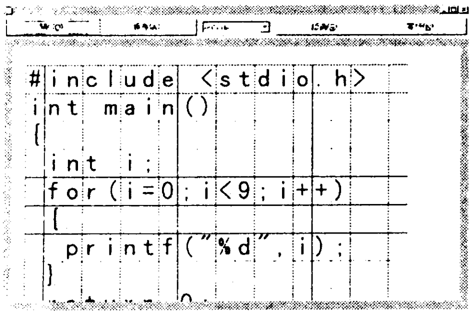
When a teacher is writing a program and explaining it, machine recognition is not only unnecessary for the teacher and students, but it is even worse than when no recognition is employed. Recognition immediately after a pattern is written causes interruption because of the correction of incorrect recognition and the verification of recognition even if a pattern is correctly recognized.

When the teacher wants to show the execution of a handwritten program, however, program recognition is highly convenient and effective. Although the teacher might have to correct some misrecognitions, it is worth the trouble since the machine can show the execution of the program that the teacher has just written in front of the students.



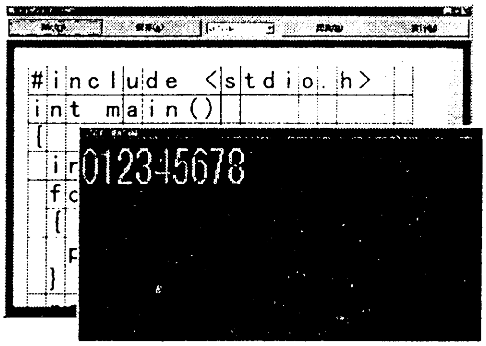
```
#include <stdio.h>
int main()
{
 int i;
 for(i=0;i<9;i++)
 {
  printf("%d", i);
 }
```

Figure 7. A handwritten program.



```
#include <stdio.h>
int main()
{
 int i;
 for(i=0;i<9;i++)
 {
  printf("%d", i);
 }
```

Figure 8. Result of program recognition.



```
#include <stdio.h>
int main()
{
```

```
012345678
```

Figure 9. Result of program execution.

6.2 *Arithmetic Education*

Children write arithmetic equations on the board and the system answers whether they are correct or not. They can share calculations with others, teachers and the system. Figure 10 shows the *IdeaBoard* with this application running.

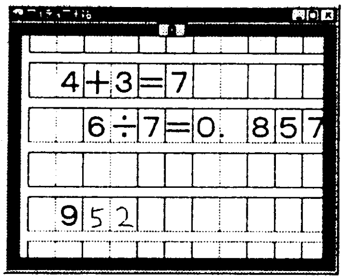


Figure 10. Arithmetic education on *IdeaBoard*.

7. Conclusion

This paper has described our research to provide coherent and creative human computer interfaces based on handwriting. Main emphases have been laid on an on-line handwritten character recognition engine and applications which may add essential values to character recognition.

The on-line handwritten character recognition engine has marked 90 to 95 % correct recognition rates without learning to a large database of on-line handwritten Japanese text. It is small and quick enough to be employed for a small PDA machine. The system is not only robust to stroke connections and pattern distortions but also highly customizable for personal use due to the structural learning capability.

The educational applications described runs on a large interactive electronic whiteboard. They seems to be an important group of applications which benefit from pen interfaces and the robust handwriting recognition engine.

This research has been supported by the Advanced Software Enrichment Project of IPA under MITI of Japan.

References

[1] H. Tanaka, N. Kato and M. Nakagawa, Prototyping of a Digital Ink E-mail System based on a Common Ink Format, Proc. HCI International '97, Vol. 2, pp.435-438 (1997.8).

[2] A. Naito and M. Nakagawa, A Method of Compression and Restoration of Digital-Ink, (in Japanese) SIG notes, IPS Japan, **96**, 94, pp.9-16 (1996).

[3] M. Nakagawa, T. Higashiyama, Y. Yamanaka, S. Sawada, L. Higashigawa and K. Akiyama, On-line Handwritten Character Pattern Database Sampled in a Sequence of Sentences without Any Writing Instructions, Proc. ICDAR '97, Vol.1, pp.376-381 (1997.8).

[4] M. Nakagawa and K. Akiyama: "A Linear-time Elastic Matching for Stroke Number Free Recognition of On-line Handwritten Characters," Proc. 4th IWFHR, pp.48-56 (1994.12).

[5] M. Nakagawa, K. Akiyama, L. V. Tu, A. Homma and T. Higashiyama, Robust and Highly Customizable Recognition of On-line Handwritten Japanese Characters, Proc. 13th Int'l Conf. on Pattern Recognition , Ⅲ, pp.269-273 (1996).

[6] M. Nakagawa, T. Oguni and A. Homma: "A coarse classification of on-line handwritten Japanese characters," *Progress in Handwriting Recognition,* World Scientific, pp.357-362 (1997).

[7] M. Nakagawa and L. V. Tu, Structural Learning of Character Patterns for On-line Recognition of Handwritten Japanese Characters, P. Perner, P. Wang and A. Rosenfeld eds. Advances in Structural and Syntactic Pattern Recognition, Springer-Verlag, pp.180-188 (1996).

[8] M. Nakagawa, K. Machii, N. Kato and T. Souya, Lazy Recognition as a Principle of Pen Interfaces, INTERCHI'93 Adjunct Proc. pp.89-90 (1993.4).

[9] M. Nakagawa, N. Kato, K. Machii and T. Souya, Principles of Pen Interface Design for Creative Work, Proc. 2nd Int'l Conf. on Document Analysis and Recognition, pp.718-721 (1993).

[10] K. Oomika, A. Naito and M. Nakagawa: "Idea memo PDA in scalable handwriting interfaces," Proc. HCI International '97, Vol.2, pp.455-458 (1997.8).

[11] N. Kato, N. Fukuda and M. Nakagawa: "Adjustment Mechanism for a Drawing System with Stationary Metaphors," Proc. HCI International '95, pp.155-160 (1995.7).

[12] M. Nakagawa, T. Oguni and T. Yoshino, Human Interface and Applications on IdeaBoard, Proc. IFIP TC13 Int'l Conf. on Human-Computer Interaction, pp.501-508 (1997.7).

[13] M. Weiser, The Computer for the 21st Century, Scientific American, pp.94-104 (1991).