PRINCIPLES OF PEN INTERFACE DESIGN FOR CREATIVE WORK

Masaki Nakagawa
Naoki Kato
Kimiyoshi Machii
Toshio Souya

Reprinted from PROCEEDINGS OF THE SECOND INTERNATIONAL
CONFERENCE ON DOCUMENT ANALYSIS AND RECOGNITION,
Tsukuba Science City, Japan, October 20-22, 1993

IEEE Computer Society
10662 Los Vaqueros Circle
P.O. Box 3014
Los Alamitos, CA 90720-1264

Washington, DC • Los Alamitos • Brussels • Tokyo

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.
IEEE

IEEE COMPUTER SOCIETY

Principles of Pen Interface Design for Creative Work

Masaki Nakagawa Naoki Kato Kimiyoshi Machii *Toshio Souya

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

Abstract

This paper presents basic design concepts for a creative writing environment with pen and display integrated tablet to promote human thinking rather than to bother a user with the chores of input. The pen is suitable for creative work since one can express almost everything and is not bothered by the method to use. Experimental pen-based systems and products have not exploited the 'automated' nature of handwriting. They try to recognize handwriting immediately after each pattern is written with the result of frequent misrecognition and thus interrupt user's thinking. This paper presents a structure of representation to handle the ambiguity of handwritten patterns, as well as lazy recognition scheme which delays the display of recognition until needed. One's thought is better developed by working with one's handwriting. This paper also presents a prototype for text preparation, automatic segmentation of text and diagrams and mathematical formula recognition. They are all consolidated with lazy recognition.

Keywords: pen interface, on-line recognition, lazy recognition, creativity support.

1. Introduction

Pen computers have been gathering attention since the beginning of the '90's. Often cited merits are direct manipulation, writing and drawing capability rather than pointing, scalability, mobility and so on.

Nevertheless, these are not the only merits that the pen brings. As the writing (drawing) process can be done subconsciously without being bothered by troublesome details, this helps promote creativity.

Moreover, text, diagrams, tables, equations, etc. are written with a single pen without changing the device so that no interruption is made to the thought process. This fact, however, implies that a handwritten pattern is ambiguous.

Since prefect pattern recognition is nearly impossible, there is a need to rethink the interface cycle of recognition, display, correction, and re-input. In the traditional on-line interfaces immediate recognition of the input is sought[2,6,10]. However, there is an interruption to the thinking process owing to the checking of the recognition result and the correction of any incorrect recognition. Here we present lazy recognition, where the display of the recognition result is delayed until it is needed, founded on the idea that it is better for the creativity and development of the thinking process to continue working with the handwritten pattern. We proposed lazy recognition first in[7] showed a prototype for text writing/recognition in[8], and validated the design in [9]. A similar concept is presented in [3]. This paper presents why lazy recognition matches pen interfaces.

To apply computer processing to freely written patterns, the segmentation of characters and diagrams is prerequisite. Numerous attempts have been made on this problem in off-line researches, but none in on-line probably because the need has not been recognized. According to the common sense of on-line interfaces, mode selection is easily employed to input characters and diagrams without confusion. But, this is not taking creative interaction into account.

2.Why creative with pen and paper

2.1 The importance of creative input support

In creative document preparation the roles of creator and typist cannot be separated. This is because by putting pen to paper our thought becomes clear and our creativity is stimulated.

Observation on the actual styles of document preparation were made. Only a small number of people depend entirely on computers from the initial stage of thinking to the final stage of print out. The majority of people work on paper with a pen at creative stages to prepare contents, to structure, to brush up, to restructure and so on. On the other hand, they also go into the trouble of inputting text to attain the conveniences of the computer processing. This fact shows

---

*presently at Tokyo Reserch Lab. IBM Japan Ltd.

This reserch is supported in part by the Grant-in-aid for Scientific Reserch from the Japanese Ministry of Education under the project number 0558027.

that both the advantages of pen and paper and the advantages of computers are indispensable to creating documents. The former is especially large for creative stages.

2.2 From the viewpoint of cognitive psychology

According to what cognitive psychology teaches, even when humans just perceive (whether hearing or seeing) stimulus (voices, images, etc.) from the outside world, 'attention' is necessary [1]. Creative writing is a much higher-level of intellectual work than perception. Therefore, there would be a need to pay even more attention to that content. Generally, if 'attention' is simply divided amongst two actions, not only do these actions quality and efficiency fall, but the mental burden also increases.

On the other hand, there is work that requires no attention what so ever. Highly dexterous work is like this and is called 'automatic' work. A good example of this can be seen in some English typists. Since their typing is automatic, there is no need for 'attention' and it is possible to simultaneously type and have a completely unrelated conversation.

Therefore, to pay 'attention' to the manuscript contents in creative input, the work of inputting into the computer must be made 'automatic'.

Nevertheless, for work to be made 'automatic' there is a need to spend an excessively large amount of time on training. If it is only the input of characters this may be achieved through the automation of keyboard input, but it is extremely difficult to automate the use of software for creating diagrams or layout.

Here we shall consider handwriting. The work of handwriting is excessively trained in everyday life, and can be thought of as being 'automated'. Because handwriting work has been 'automated', 'attention' can always be given to the contents when doing work using pen and paper. For this reason there is no burden to the train of thought. This can be thought of as the reason why one can concentrate on thinking.

2.3 State of the art of pen interfaces

In Japan up until now there has been much study of on-line handwritten input systems. There has also recently been a boom in pen computers, and several pen computers are being marketed.

In most of the systems, however, one is able to preserve just one's handwriting, or pen-trace patterns, but unable to then pass them onto the recognition engine. Even it one already has the data right in front of oneself, it has to be re-entered. Also, in the recognition mode, once the recognition result is displayed one can no longer see the original pattern. One has to re-enter without any reference to why it was incorrectly recognized in the first place.

In these systems, the pen-trace pattern has been used only to determine the character code. However, even if it is unrecognizable to the recognition engine, the pen-trace pattern is easily understood by the person that wrote it. Moreover, it contains various information, such as would be found if one were to view a note or handwritten original draft, beyond the information of just the character code. The handwritten pattern must be utilized regardless of them being unrecognizable to the recognition engine.

3 Representation for handwritten objects

3.1 Ambiguity of the handwritten pattern

A handwritten pattern may represent a code(function), figure or just a pattern as it is. One of the merits of handwriting is that various kinds of objects can be written with a single pen. Given the context they are in, humans can distinguish between these. Taken from a different point of view, a handwritten pattern is ambiguous. For instance, character patterns are usually intended to denote character codes. Similarly, flowchart or circuit diagram symbols as well as their interconnections express functions to be performed rather than just figures. On the other hand, diagrams that represent a systems organization or concept, rough diagrams of processes, or abstracts of real things have no necessity to be interpreted as functions, they may simply be formatted as figures. Also, a diagram jotted down inside a diagram representing a function cannot be said to be a function. The reason being that the domain of interpretation may differ. Moreover, a signature written in that diagram should be kept as its pen-trace pattern, translating it to character codes takes away its meaning. A representational structure for them must be provided.

3.2 Layered representation

As shown in fig.1, when we handle handwritten objects as patterns, countering the fact that we have the same flexibility as with pen and paper, there is little merit of applying editing operations. On the other hand, when we handle them as figure descriptions, editing and formatting is effective, but there is a decrease in the flexibility of expression. Taking it a step further, if we handle them as functions, the logical merit becomes the greatest, for example, simulation or conversion to programs becomes possible, but, the flexibility in regards to expressing handwritten objects becomes even smaller.

Computers cannot judge whether a piece of writing denotes a function, a diagram to be formatted, or just a pattern which needs no processing, so we take as the internal representation a hierarchical representation of: functional representation; figure representation; and handwritten pattern. Then, when a process has been specified, the decision is made as to which level of representation is most appropriate (for example, if an eraser is chosen then the original pattern is appropriate, if character string search is chosen then the character code). Recognition is the process of moving from the lower representation towards higher ones in this hierarchy, while

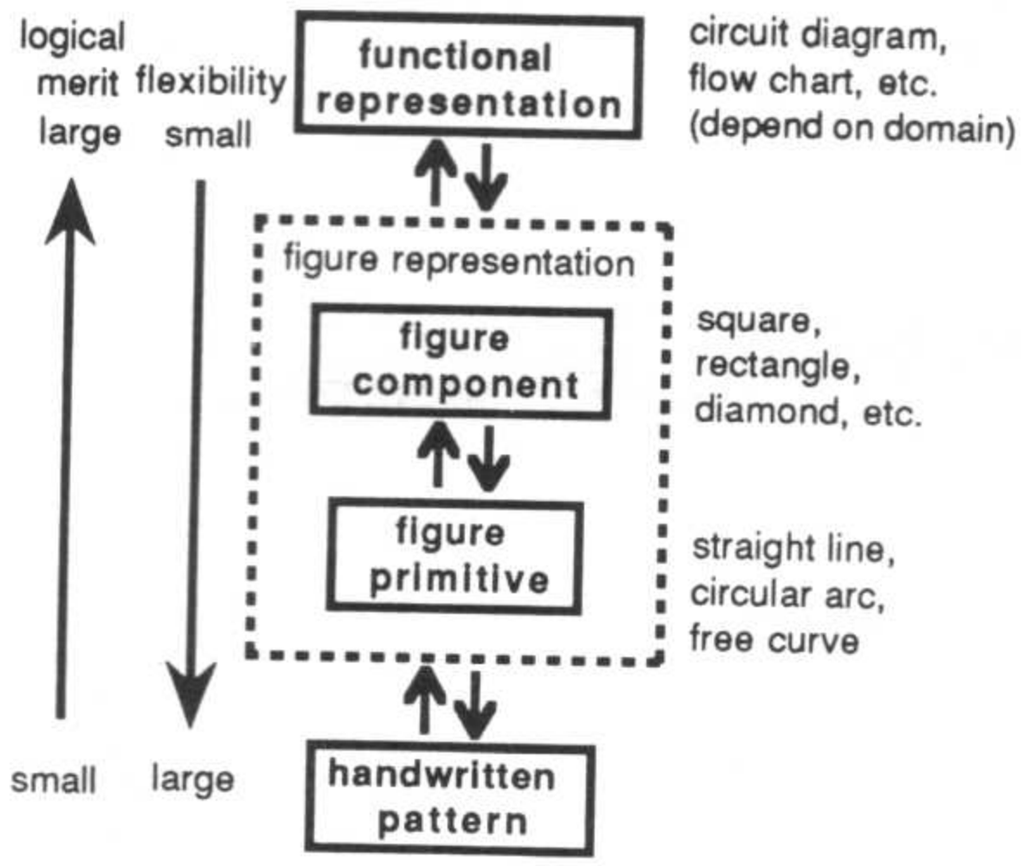


Fig. 1 Hierarchical representation for the handwritten pattern.

drawing provides lower representations up to the specified representation.

4. Lazy recognition for creative input

When an emphasis is laid on the human factor for creative work, immediate feedback is not required or even bothering. Coarse or lazy interaction may be more effective than fine or busy interaction.

In lazy recognition, the recognition process, or display of the recognition result is not done straight after each individual pattern has been written. The recognition is delayed or is run in the background or in a separate window. In any case, the important point is not to make the user conscious of the recognition during the creative phase. At the point in time when the input of one document has been completed, or when the user decides 'this is one section', in other words when the thinking process has naturally run dry, the confirmation and correction of the recognition result can be done. At this phase, only the task of confirmation and correction need be concentrated on.

Lazy recognition removes the obstruction to the thought process interleaved for the convenience of computer processing. In handwritten input even the pre-recognition pattern can be easily read. Rather, when considering the disruption to the thought process caused by continual replacement of handwritten patterns by fonts, working with the handwritten pattern has better continuity, even if it is poorly written.

In the case where the pattern recognition engine is poor, there is the concern that the correction work after input may come to an enormous amount. However, when using the same recognition system, even if it is a real-time method, there will be no difference in the amount of correction work. The amount of correction work is a problem that stems from the recognition system. Rather, by using lazy recognition, the thinking environment can be preserved by the correction work not being mixed in at the time of input.

5. Research problems motivated by the design

5.1 A prototype for text preparation

We have prototyped a creative writing environment for text where only the recognition of characters has been included using our own recognition engine [5, 8], i.e., diagrams and equations can be inputted as patterns but left unrecognized. In this prototype, we have taken manuscript paper, usually used in the creation of manuscripts in Japanese, as a model for the writing/display format. Fig.2 shows an example of text insertion.

Gestures are basically lazily executed, except for those whose actions are related to the manuscript paper itself, which are performed immediately. Amongst these gestures, page turning and addition/insertion are included.

On the other hand, gestures to change the content should not be executed immediately. For instance, text designated to be deleted by a gesture can be further modified and turned back into text if deletion is delayed. This is what we are doing on paper. Another need of lazy execution of gestures is found in the CSCW environment. A gesture to do some action must be approved by others before it is really executed. Undo cannot replace the lazy execution of gestures.

The merits of this prototype as a text editor are limited as the recognition of diagrams and equations has not yet been realized, but from the members of our laboratories use of it we have gained reactions such as, "it's not as good as paper, but as I used it the feeling of awkwardness went," or "even just the ability to draw diagrams while working on the text is very helpful."

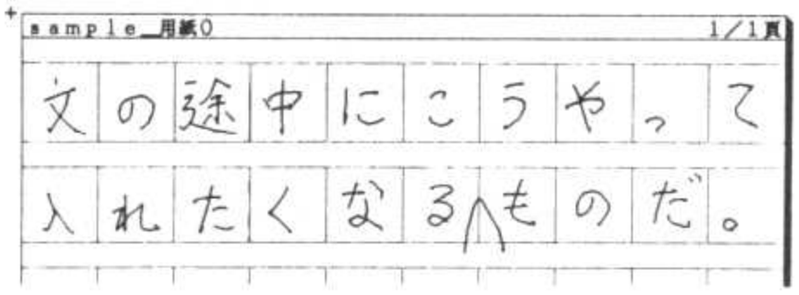
5.2 On-line character and diagram segmentation

Research into the automation of on-line character pattern and diagram pattern segmentation has been done. A graph of the log of character stroke lengths shows that they are normally distributed. On the other hand, the lengths of strokes that make up diagrams and tables plot evenly in a wide range. This nature has been mainly used to perform the segmentation.

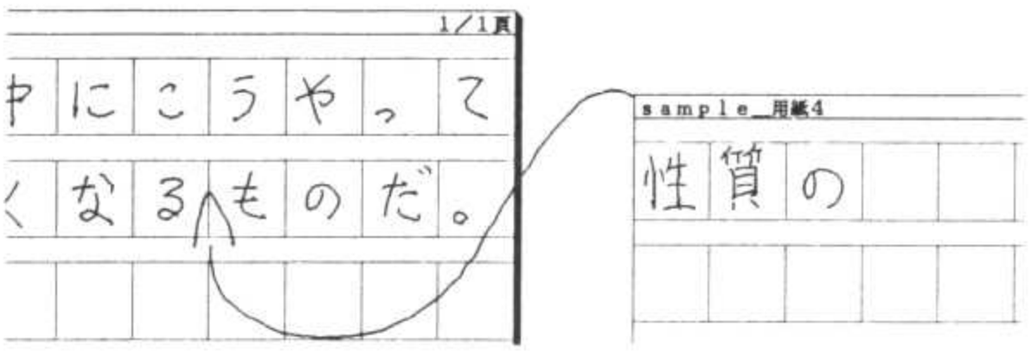
Here, the promptness is important. When the user wants to start confirmation and correction, it must be able to be started immediately.

As the details are described in[4], only the result is shown. The result, applying this to cases where the characters are many or few, shows that segmentation is made 88% correctly on average within 0.1 sec/page.

Here, the segmentation is made using the overall pattern written on the tablet, and because of this, it is favorable to have the interface consolidated with lazy recognition.



Writing of insertion gesture.



After a linkage line is drawn, a new manuscript paper is displayed on which additional writing is made.

Fig. 2 An example of text insertion.

5.3 Mathematical formula recognition

Mathematical formula recognition is studied. Here again, handwritten formulas are left unprocessed until the user requires. The recognition is divided into three stages: symbol segmentation, symbol recognition and structure recognition. After writing, the user concentrates on helping the system process patterns correctly. The result of each stage is displayed as shown in fig.3 and the user's confirmation and correction can be reflected by pen gestures.

References

[1] Anderson, J.R. Cognitive Psychology and its Implications. 2nd ed. W.H.Freeman, New York (1985).

[2] Goldberg, D. and Goodisman, A. Stylus use interfaces for manipulating text. Proc. UIST'91 (Nov.1991) 127-135.

[3] Guerfali, W. and Plamondon, R. An editing tool for handwritten text. 5th Handwriting Conf. of IGS, Arizona (Oct. 1991) 151-153.

[4] Machii, K., Fukushima, H. and Nakagawa, M. On-line text/drawing segmentation of handwritten patterns. Proc. ICDAR '93 to appear.

[5] Nakagawa, M., Aizawa, T., Komoda, C., Ikeda, Y. and Takahashi, N. Syntactic pattern recognition with stochastic dissimilarity in Japanese on-line input system (JOLIS) 1/1.5. Proc. 8th ICPR, Paris (Oct. 1986) 1059-1062.

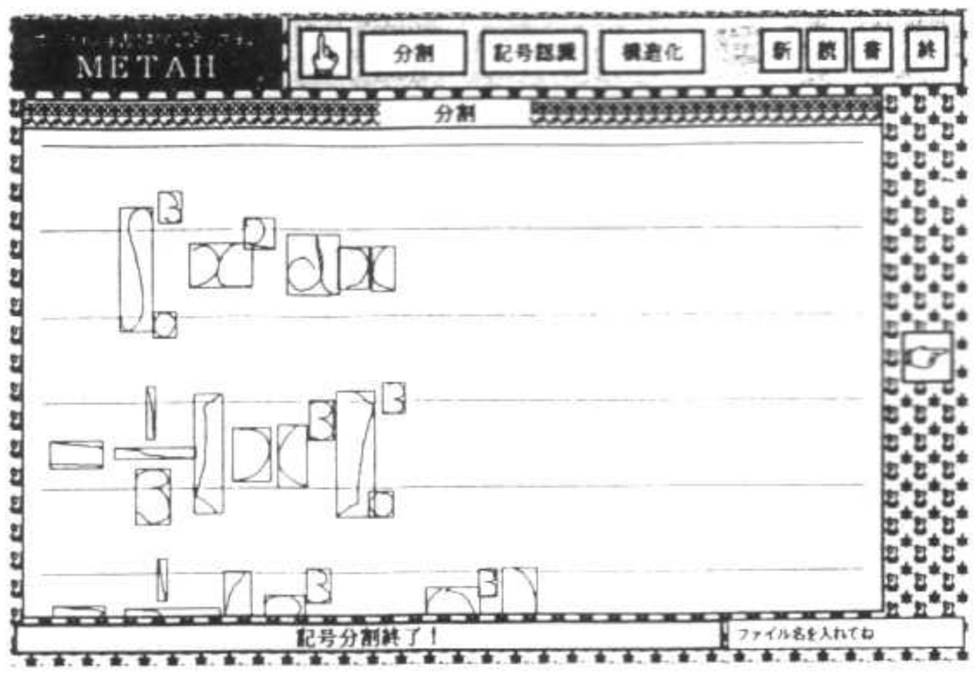
[6] Odaka, K., Wakahara, T. and Hashimoto, S. On-line handwritten character recognizer (in Japanese). Trans. IECE Japan, J65-D, 8 (Aug. 1982) 951-958.

[7] Souya, T. et al. User interface through handwriting(in Japanese). Proc. 30th Programming Symposium, IPS Japan (Jan. 1990) 1-10.

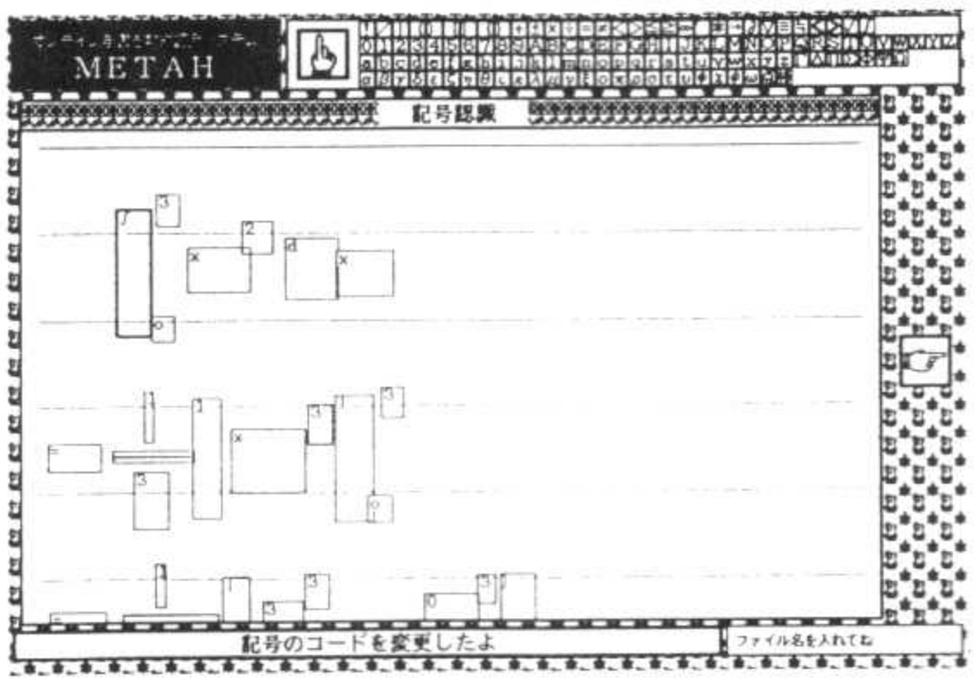
[8] Souya, T., Fukushima, H., Takahashi, N. and Nakagawa, M. Handwriting interface for a large character set. Proc.5th Handwriting Conf. of IGS, Arizona (Oct. 1991) 166-168.

[9] Souya, T., Fukushima, H., Takahashi, N. and Nakagawa, M. A basic design of handwriting user interface with razy recognition, Trans. IPS Japan, 34, 1 (Jan . 1993) 158-166.

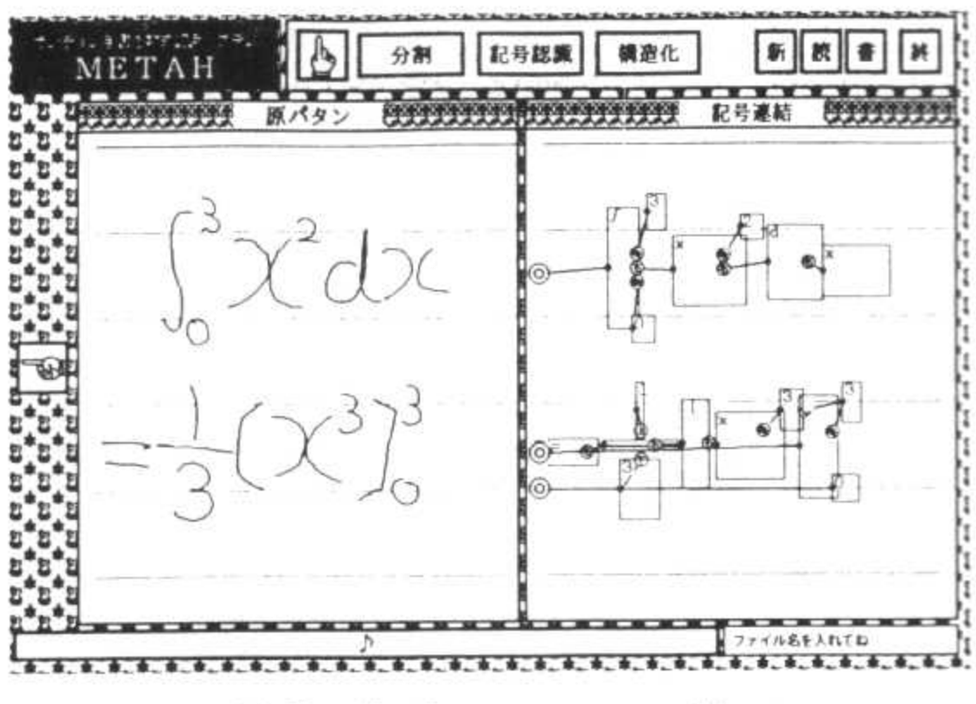
[10] Tamori, H., Chatani, M., Sakano, A. and Yoshida, K. A method of on-line handwritten character recognition by fuzzy inference(in Japanese). Trans. IEICE Japan , J74-D-II, 2 (Feb. 1991) 166-174.



(a) Result of segmentation.



(b) Result of symbol recognition.



(c) Result of structure recognition.

Fig. 3 Mathematical formula recognition.