

# Extracting the trajectory of writing brush in Chinese character calligraphy

Fenghui Yao<sup>a,b,\*</sup>, Guifeng Shao<sup>a</sup>, Jianqiang Yi<sup>b</sup>

<sup>a</sup>*Department of Computer Science, College of Engineering, Technology and Computer Science, Tennessee State University, 3500 John A. Merritt Blvd., Nashville, Tennessee 37209*

<sup>b</sup>*Laboratory of Complex System and Intelligence Science, Institute of Automation, Chinese Academy of Sciences, P.O. Box 2728, Beijing 100080, PR China*

Received 28 July 2003; accepted 6 August 2004

## Abstract

This paper describes the extraction of the trajectory of the writing brush in Chinese character calligraphy (CCC), based on image and curve processing techniques and the calligraphy knowledge. This trajectory is used in a CCC robot which is developed to inherit CCC techniques. In CCC, the writing styles can be roughly classified into five different styles—ancient, angular, block, semi-cursive, and cursive style. This paper is limited to discuss the trajectory extraction from the character image written in block style. Firstly, for a given Chinese character, its image patterns in block style are retrieved from CCC database which contains 29,456 characters written by different famous calligraphers in Chinese history. Then the image of the designated writing is thinned. The coordinates of the line passing the centers of each stroke are detected from the thinned image with aid of writing order information. These coordinates represent the thinned-center-line of the stroke (TCLS, for short). And then, TCLS is separated into several curve segments according the calligraphy knowledge. The trajectory of the writing brush is considered as B-spline curves determined by the points on curve segments. The trajectory and the pressure control information are sent to the CCC robot to imitate calligrapher's behavior. The experiment results show that the proposed method obtains very good trajectories of the writing brush for CCC robot. © 2004 Elsevier Ltd. All rights reserved.

*Keywords:* Chinese character calligraphy; Character image pattern; Thinning; Trajectory of writing brush; B-spline; Calligraphy robot

## 1. Introduction

The Chinese character calligraphy (CCC) culture has more than 4000 years' history. Chinese characters are divided into six categories: pictographic characters, pictophonetic characters, associative compounds characters, self-explanatory characters, phonetic loan characters and synonymous characters. These characters are constructed by strokes. There are 28 strokes used to construct all characters. Table 1 shows their image patterns and names. To make the description easy, these strokes are coded from  $S_0$  to  $S_{27}$ , i.e., the strokes are

denoted by a set  $S = \{S_0, \dots, S_{27}\}$  where strokes  $S_{20}$  and  $S_{27}$  are used only to constructed the simplified Chinese characters. CCC expresses these characters by using the writing brush, Chinese ink, paper and stone slab, which are called four treasures for study. Fig. 1(a) shows three writing brushes; (b) a piece of Chinese ink; (c) a stone slab; (d) a piece of artwork of a character for “dragon” written by using the writing brush and Chinese ink. Though the writing brush in CCC is similar to the brush used for watercolor painting in the West, it has a finer tip suitable for dealing with a wide range of subjects and for producing the variations in line required by different styles. Since the materials used for CCC and Chinese paintings are essentially the same, developments in calligraphic styles and techniques can also be used

\*Corresponding author. Tel./fax: +1-61596 35875; +1-61596 35847.  
E-mail address: [fyao@tnstate.edu](mailto:fyao@tnstate.edu) (F. Yao).

Table 1  
Stroke patterns to construct Chinese characters

Stroke code	Pattern	Stroke name	Stroke code	Pattern	Name
$S_0$	、	Dot stroke	$S_{14}$	㇇	Horizontal and fold stroke
$S_1$	一	Horizontal stroke	$S_{15}$	㇇	Horizontal, fold and hook stroke
$S_2$	丨	Vertical stroke	$S_{16}$	㇇	Horizontal and left-falling stroke
$S_3$	丿	Leftfalling stroke	$S_{17}$	㇇	Left-falling and fold stroke
$S_4$	㇇	Rightfalling stroke	$S_{18}$	㇇	Left-falling and dot stroke
$S_5$	㇇	Rising stroke	$S_{19}$	㇇	Vertical and fold stroke
$S_6$	㇇	Vertical and hook stroke	$S_{20}$	㇇	Horizontal, fold and rising stroke
$S_7$	㇇	Curved hook stroke	$S_{21}$	㇇	Horizontal, fold and curved-hook stroke
$S_8$	㇇	Inclined hook stroke	$S_{22}$	㇇	Vertical, fold and curved-hook stroke
$S_9$	㇇	Lying hook stroke	$S_{23}$	㇇	Horizontal, fold, fold and left-falling stroke
$S_{10}$	㇇	Vertical and turn-right stroke	$S_{24}$	㇇	Horizontal, left-falling and curved-hook stroke
$S_{11}$	㇇	Vertical, turn-right and hook stroke	$S_{25}$	㇇	Horizontal, fold, fold and curved-hook stroke
$S_{12}$	㇇	Vertical and rising stroke	$S_{26}$	㇇	Horizontal, fold and turn-right stroke
$S_{13}$	㇇	Horizontal and hook stroke	$S_{27}$	㇇	Horizontal, fold and left-falling stroke

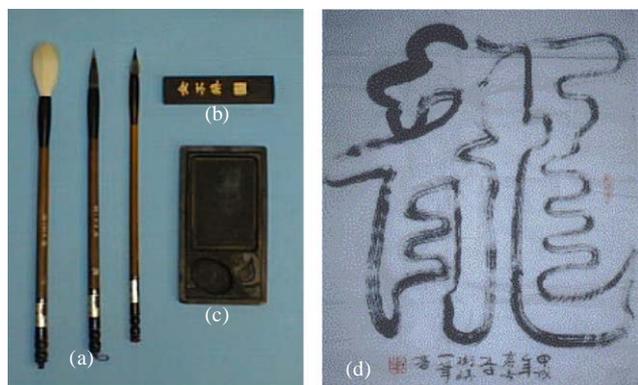


Fig. 1. (a) Writing brush, (b) Chinese ink, (c) stone slab, (d) an artwork of a character for “dragon” written by using the writing brush and Chinese ink.

in Chinese painting. The Chinese ink has been used in CCC and Chinese painting for over 2000 years. When the ink cake is ground on the painter’s stone slab with fresh water, ink of various consistencies can be prepared depending on the amount of water used. Thick ink is very deep and glossy when applied to paper or silk. Thin

ink appears lively and translucent. As a result, in ink-and-wash paintings it is possible to use ink alone to create a rhythmic balance between brightness and darkness, and density and lightness, and to create an impression of the subject’s texture, weight and coloring. The CCC and Chinese painting may be done either on Chinese paper or silk. The paper is very absorbent and the amount of size in it will dictate the quantity of ink used for strokes on the paper. Different types of paper produce different results; some are rough and absorb ink quickly like a sponge, others have a smooth surface which resists ink. Chinese paper is usually known as rice paper in English. The silk is less absorbent than paper.

Brushstroke is best shown on paper. Because of this reason and the paper’s variety of texture and finish, the paper is favored by artists and calligraphers. The calligraphers in Chinese history created and developed a variety of writing styles to write characters by using writing brushes. These styles can be roughly categorized into (i) ancient, (ii) angular, (iii) block, (iv) semi-cursive, and (v) cursive. Fig. 2 shows five different writings of the Chinese character for “wind” (Matsuda, 1996). The ancient style as shown in Fig. 2(a) employs thin curved



Fig. 2. Five different writings of the Chinese character for “wind”: (a) ancient, (b) angular, (c) block, (d) semi-cursive, (e) cursive.

lines to express strokes. The ends of these thin curved lines are sharp, and the lines are vigorous. This style was prospering very well in the Qin Dynasty (221–207 BC). The angular style was invented by Chengmao—a prison officer in the Qin Dynasty, to deal with the complicated official documents rapidly. In this style, as shown in Fig. 2(b), the horizontal strokes are written horizontally, and the vertical strokes vertically, and the combination among strokes are emphasized. The angular style is much easier to write than the ancient style. The block style was developed in the Han Dynasty (206 BC–AD 220). It is an improved style based on the angular style. The block style becomes further easier to write than the angular style. This style was vigorous in the Tang Dynasty (618–907). The semi-cursive style lies between the block style and the cursive style. In the semi-cursive style, as shown in Fig. 2(d), the corner is not as squarish as in angular style, and is not as round as in the ancient style. This style is a variation of the block style. In the cursive style, as shown in Fig. 2(e), the structure is simple, and there is no break in the movement track of the writing brush. It is possible to write quickly. There are many variations about the cursive style, and there are many artworks written in the cursive style. The CCC culture was also imported by Japan, Korea, and Southeast Asian countries, and was developed in their own way in these countries and became a part of the cultures in these countries.

However, since the hard writing tools such as pencil, pen and ball-pen were introduced into these countries, more and more people put away the writing brushes and took the pencils and pens. And in most schools, the students are being taught to write only by pencils. Even for those schools where there is a subject to teach students to write by writing brushes, the amount of time for such a subject is very small in comparison with other subjects. This situation has lasted more than half of a century. Therefore, generally speaking today, the young people in China, Japan, Korea and so on cannot write characters by writing brushes except those people who received a special training to write characters by writing brushes. Furthermore, with the spread of the computer word processor, such as Word, Ichitaro, and so on, more and more people get used to them and do not like to write characters even by pen or pencil, not to mention the writing brush. Therefore, the

number of competent calligraphers is becoming smaller and smaller day by day. If this situation continues for several decades, the CCC culture may face the crisis of extinction.

The purpose of this work tries to preserve and inherit the CCC culture by a robot system (Bradbeer and Billingsley, 2002). The whole work consists of three steps. The first step is to preserve the characters written by famous calligraphers in Chinese history. Because many famous CCC artworks were written on planks or engraved on stone tablets, many of them are not in complete status after several thousands of years. Although many of them were made rubbed copies, the copies are not in good status. It is difficult to restore these characters on paper copies. Here, these characters are extracted and put into a CCC database. This makes it easier to search and restore the CCC artworks on computer. The second step is to inherit the CCC techniques by a CCC robot. The CCC means not only the static artwork but also the dynamic process to produce the artwork, which contains the control and the movement of the writing brush. This dynamic processing contains a lot of techniques such as pressure control, speed control, turn control and so on. These are very difficult to master and imitate although the people may know the principles or rules to write by writing brushes because there exists a big difference between the CCC theory and the practice and because the top of the writing brush is very soft. The hand and arm of the people may not move as precisely as he imagines, but the robot arm can. So we propose to employ the robot to inherit the CCC techniques. The third step is to let the robot to teach CCC techniques to the human calligraphy learners. In this way the robot can preserve and inherit the CCC culture. At present we constructed the prototype system for the first two steps. This paper focuses on the second step, that is, to inherit the calligraphy skills of human calligraphers by a robot. This contains the trajectory extraction and pressure control of the writing brush, control of speed to move the writing brush, interconnection among strokes, and so on. Among these factors, the trajectory control of the writing brush is the most important. Even for the same stroke, the trajectory of the writing brush is different in five different styles as shown in Fig. 2. The trajectory depends on what style the character is written. And it does exist in the character image pattern itself. This paper mainly discusses the extraction of the trajectory of the writing brush from the character image patterns written in block style.

The remainder of this paper is as follows. Section 2 describes the basic knowledge in CCC. Section 3 relates the trajectory extraction of the writing brush from the image pattern. Section 4 shows the experiment results. And the paper finishes with the conclusions and remarks.

## 2. Basic knowledge in CCC

Among the styles of the Chinese characters in daily use, the block style is used most frequently and is most accurate. To write the block style characters by the writing brush, each stroke has its own composition method specific for writing brush. The composition method of each stroke consists of three parts: the start of the stroke, the movement of the writing brush and the stop of the stroke (Suzuki, 2002), which are summarized in the following basic knowledge rules,  $BR_0$ ,  $BR_1$  and  $BR_2$ .

### 2.1. $BR_0$ : Start of the stroke

The natural direction of the hand holding the writing brush is that the tip of the hand is toward the left front of the body, and the elbow toward the right rear. When falling down the writing brush in this orientation, the tip of the writing brush is toward the upper left, and the stomach of it the lower right. In this case, the direction of the tip of the writing brush referring to the line vertical to the center line of the stroke, i.e., the angle  $\theta$  as shown in Fig. 3(a) (1), is about  $60^\circ$ . This angle is called the *start stroke angle* hereafter. Although in fact to write the character vigorously, the start stroke angle is controlled in the range from  $0^\circ$  to almost vertical to the direction in which the writing brush will move. Generally, it is alright to control this angle at about  $60^\circ$  regardless of the vertical stroke, the horizontal stroke or others.

### 2.2. $BR_1$ : Movement of the writing brush

The next is the movement of the writing brush. From the start of the stroke, the writing brush is moved to the right to form a horizontal stroke, downward to form a vertical stroke, curved lower left to form a left falling stroke and others. Fig. 3(a) (2) shows the drawing of the horizontal stroke. In this stroke, the tip of the writing stroke lies on the upper side of the stroke, and the

stomach of it the lower. Similarly, in vertical stroke, the tip of the writing stroke lies on the left side of the stroke, and the stomach of it the right. The route of the tip of the writing brush is called the *front of the stroke*, and the route of the stomach the *back of the stroke*. When the writing brush is used naturally, the front and back of the stroke appear apparently.

### 2.3. $BR_2$ : End of the stroke

The direction of the stop of the stroke is naturally the same with that of the start as shown in Fig. 3(a) (3). The stop of the left falling stroke and right falling stroke changes to sweep.

Fig. 3(b) shows a horizontal stroke written with the writing brush, and (c) shows the trajectory of the writing brush in 3D space. With these composition methods, a stroke is not simply a symbol, or a dot or a line, but a strong, vigorous and emotional life. As demonstrated in Fig. 3(d), although the strokes in a character are independent and are not connected, when starting to write the next stroke from the present stroke, the writing brush is controlled to make a smooth trajectory in 3D space (refer to the dot line). The end of the present stroke makes contact with the start of the next stroke on this trajectory. This is called the *rhythmic coherence of the strokes*. Because all characters are constructed by strokes with their own rhythmic coherence, naturally the movement of the writing brush may be slow in some cases and fast in others according to this rhythmic coherence. Generally speaking, the writing brush is controlled to move slowly at the start, end, and turn point of the stroke, and fast at other places such as the straight and curved part of the stroke.

These three basic CCC knowledge rules are suitable to all strokes. The knowledge rules for specific strokes are briefly summarized in Table 2, where  $P_i$  shows the points to determine the shape of the stroke,  $B_i$  to determine the start of it,  $E_i$  the end.

## 3. Trajectory extraction of writing brush

### 3.1. Taking CCC as a problem in robot system

The CCC contains both physical factors and emotional factors. The physical factors are latent in the shape of the character or the shape of the stroke. The emotional factors change upon the emotion of the calligrapher, and there are a lot of variations. Here we limit the discussion to the realization of the physical factors in general cases, i.e., the calligrapher is with calm emotion. The physical factors are as follows: the start and the end of the stroke, the route through which the writing brush will pass, the changes of the width and so on. However, the basic operation is to control the

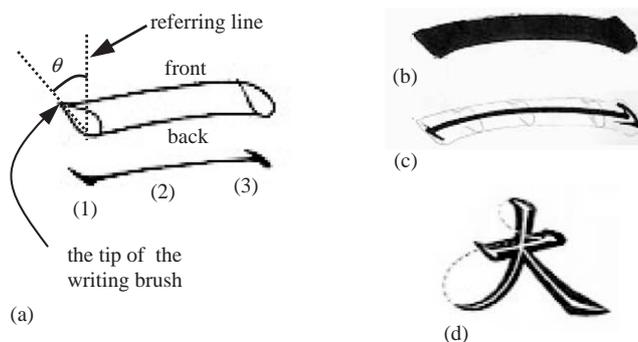
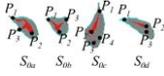
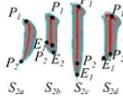
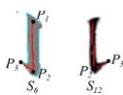
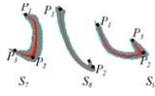
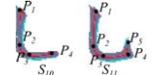
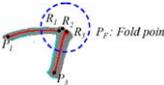
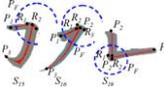
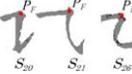
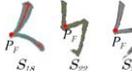


Fig. 3. (a)–(d) Basic techniques for writing.

Table 2  
Knowledge to write the strokes in CCC

Stroke code	Image pattern	Rule code	Description
$S_0$		$R_0$	There are four variations for $S_0$ , which are coded as $S_{0a}, \dots, S_{0d}$ . (1) For $S_{0a}, S_{0b}$ , and $S_{0d}$ , starting from $P_1$ where the tip of the writing brush just touches the paper, the writing brush is pushed down gradually and moved to $P_2$ , then taken up slightly and moved to $P_3$ , and held up at $P_3$ . (2) For $S_{0c}$ , the movement from $P_1$ to $P_2$ is the same as above. The writing brush is taken up slightly and moved to $P_3$ and then to $P_4$ , and held up at $P_4$ .
$S_1$		$R_1$	(1) According to $BR_0$ , starting from $B_1$ where the tip of the writing brush just touches the paper, the writing brush is pushed down gradually and moved to $B_2$ , then taken up slightly and moved back to $P_1$ . (2) The writing brush is moved to $P_4$ smoothly. (3) It is taken up slightly at $P_4$ and moved to $E_1$ , then moved to $E_2$ and $E_3$ successively, and held at $E_3$ .
$S_2$		$R_2$	There are four variations for $S_2$ , which are coded as $S_{2a}, \dots, S_{2d}$ . (1) The movement of the writing brush at $P_1$ is the same with that related in $R_1$ (1). (2) For $S_{2a}$ , the writing brush is moved from $P_1$ to $P_2$ on a smooth curved trajectory, and held up at $P_2$ . (3) For $S_{2b}$ , the writing brush is moved from $P_1$ to $P_2$ straightly, then taken up slightly and moved to $E_1$ and $E_2$ , and is held up at $E_2$ . (4) For $S_{2c}$ and $S_{2d}$ , the writing brush is moved from $P_1$ to $P_2$ straightly, then taken up gradually from $P_2$ and moved to $E_1$ , and held up at $E_1$ .
$S_3$		$R_3$	(1) The movement of the writing brush at $P_1$ is the same with that related in $R_1$ (1). (2) The writing brush is moved from $P_1$ to $P_2$ on a smooth curved trajectory, and held up at $P_2$ .
$S_4$		$R_4$	(1) At $P_1$ , the writing brush just touches the paper. (2) The writing brush is moved from $P_1$ to $P_2$ on a smooth and curved trajectory swelling out. (3) It is moved from $P_1$ to $P_2$ along a nearly straight line. (4) From $P_3$ to $P_5$ , it is moved on a smooth and curved trajectory concaving in, and held up at $P_4$ . (5) When moving from $P_1$ to $P_4$ , the writing brush is pushed down gradually and deepest at $P_4$ .
$S_5$		$R_5$	(1) The movement of the writing brush at $P_1$ is the same with that related in $R_1$ (1). (2) From $P_1$ to $P_2$ , the writing brush is taken up gradually and moved along a nearly straight line, and held up at $P_2$ .
$S_6, S_{12}$		$R_6$	(1) The movement of the writing brush at $P_1$ is the same with that related in $R_1$ (1). (2) The writing brush is moved straightly to $P_2$ and folded to $P_3$ , and held up at $P_3$ .
$S_7, S_8, S_9$		$R_7$	Although these three strokes take different kinds of shapes, they have similar characteristics. (3) For $S_7$ and $S_9$ , the writing brush just touches the paper at $P_1$ . It is pushed down gradually and moved to $P_2$ along a smooth curved trajectory, and then taken up gradually and folded to $P_3$ , and held up at $P_3$ . (4) For $S_8$ , the movement of the writing brush at $P_1$ is the same with that related in $R_1$ (1). Then it is moved to $P_2$ along a smooth curved trajectory and is folded to $P_3$ , and held up at $P_3$ .
$S_{10}, S_{11}$		$R_8$	(4) The movement of the writing brush at $P_1$ is the same with that related in $R_1$ (1). (5) The writing brush is moved to $P_2$ straightly, and then moved to $P_4$ along a smooth curved trajectory. For $S_{11}$ , it is taken up slightly and moved to $P_4$ , and held up at $P_5$ . For $S_{10}$ , it is held up at $P_4$ .
$S_{13}$		$R_9$	(5) The movement of the writing brush at $P_1$ is the same with that related in $R_1$ (1). (6) The writing brush is moved from $P_1$ to $P_2$ on a smooth curved trajectory, and taken up slightly at $P_2$ , then moved back to $P_3$ and then folded to $P_4$ . It is held up at $P_4$ .
$S_{14}$		$R_{10}$	(3) It can be thought of as that $S_1$ and $S_2$ are combined at $P_F$ . $P_F$ is called <i>fold point</i> or <i>joint point</i> , here and after. (4) At $P_F$ , the writing brush is moved from $R_1$ to $R_2$ then to $R_3$ , as shown in the circled area in the second column in this row.
$S_{15}, S_{16}, S_{19}$		$R_{11}$	(6) $P_F$ is considered as the joint point. (7) $S_{15}$ can be considered as the combination of $S_1$ and $S_6, S_{16}$ of $S_1$ and $S_3$ , and $S_{19}$ of $S_2$ and $S_1$ . (8) The movement of the writing brush at the joint point $P_F$ is the same with that described in $R_{10}$ (2).
$S_{20}, S_{21}, S_{26}$		$R_{12}$	(3) $P_F$ is taken as the joint point. (4) Each of these strokes can be considered as the combination of two strokes as given in the following: $S_{20} = S_1 + S_{12}$ ; $S_{21} = S_1 + S_{11}$ ; $S_{26} = S_1 + S_{10}$ ;
$S_{23}, S_{24}, S_{25}$			$S_{23} = S_{16} + S_{16}$ ; $S_{22} = S_{16} + S_7$ ; $S_{25} = S_{16} + S_{15}$ ;
$S_{18}, S_{22}, S_{27}$			$S_{18} = S_3 + S_4$ ; $S_{22} = S_2 + S_{15}$ ; $S_{27} = S_2 + S_{16}$ .

writing brush move on the paper and control the pressure to the writing brush. This can be robotized as the problem to generate the 3D trajectory of the writing brush in the robot coordinate system  $\Sigma_R$ . Let  $P(x, y, z)$  express a point on the 3D trajectory of the writing brush. The change of  $z$ -coordinate means the change of the pressure to the writing brush,  $x$ - and  $y$ -coordinates form a 2D trajectory of the writing brush in  $XY$ -plane. If this plane is set on the writing table as shown in Fig. 4, the robot hand holding a writing brush will write the character on the paper that is placed on the writing table, when it is controlled to move in the 2D trajectory. Fig. 4 shows the prototype of CCC robot system, which consists of a calligraphy dictionary, (a) robot arm, (b) robot hand, (c) writing brush, (d) writing brush holder, (e) Chinese ink holder, (f) writing table, (g) system controller and CCC database. As shown in Fig. 4,  $\Sigma_R$  is a right-hand Cartesian coordinate system. The area for CCC robot writing lies on  $XY$ -plane in  $\Sigma_R$ , and is determined by  $P_1(x_R, y_R, 0)$  and  $P_2(x_R + T_R, y_R - T_R, 0)$ , where  $T_R$  is size of the writing area. The center of the writing area is given by  $(x_R + T_R/2, y_R - T_R/2, 0)$ .

In the following, we first relate the extraction of 2D trajectory of the writing brush in  $XY$ -plane, then discuss the generation of the pressure of the writing brush along the 2D trajectory.

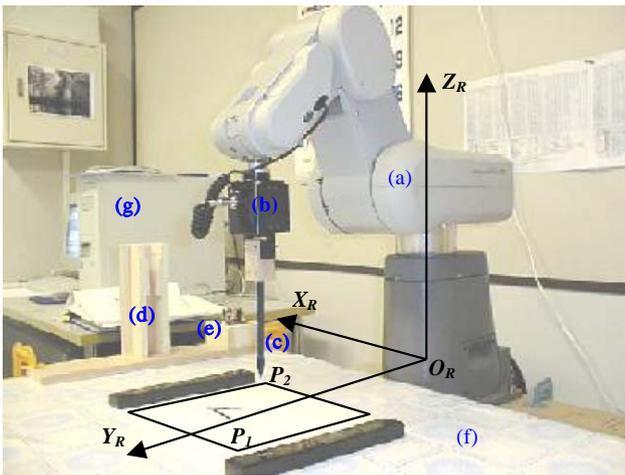


Fig. 4. (a)–(g) Prototype of CCC robot.

### 3.2. Extraction of 2D trajectory of the writing brush

The Chinese character written by a writing brush is an image. This image contains the trajectory of the writing brush that the calligrapher generates. The trajectory extraction means to search the moving route of the writing brush. The proposed algorithm contains the extraction of the skeleton of the character and the curve fitting of the skeleton, which are given below in detail.

#### 3.2.1. Extraction of the skeleton of the character

The image pattern of the input character is thinned by employing the thinning algorithm (Hasegawa et al., 1986). Then the thinned image is obtained. Fig. 5(b) shows the thinned character image of the image pattern in (a). The thinned image is then represented by the coordinate points. To do this, the writing order information of the Chinese character is needed. The writing order of the character in Fig. 5(a) is given in (c). Generally, the writing order is saved in the writing order dictionary (WOD). The WOD is constructed based on the points as shown in Fig. 5(c), which are assigned in character coordinate system  $\Sigma_C$ . The fields of WOD are as follows (refer to Fig. 5(c)).

- (i) *Code*: Chinese character code, 2 bytes.
- (ii) *Style*: 1 byte, 0, 1, ..., 4 for ancient, angular, block, semi-cursive style and cursive, respectively, and 5 for all styles.
- (iii) *Ratio*: 1 byte. To keep the original shape of the characters written by different calligraphers, the ratio of the height and width of the character, i.e.,  $H/W$ , is necessary. To be able to record this ratio with 1-byte integer, the integer  $H/W \times 100$  is used.
- (iv) *Pointer to character image pattern dictionary (CPD)*: 4 bytes, position in the CPD for the character in the code field.
- (v) *Length*: 2 bytes, length of the folded-line, that is, the number of points in the coordinate system  $\Sigma_C$  (refer to Fig. 5(c)).
- (vi) *Points*:  $x$ - and  $y$ -coordinates of points in  $\Sigma_C$ . Each point takes 4 bytes. Therefore, the size of a record is  $8 + \text{length} \times 4$  bytes.

It is necessary to note that  $y$ -coordinate of the end of the stroke in WOD is marked by “-1”, and  $x$ -coordinate by the stroke code as given in Table 1.

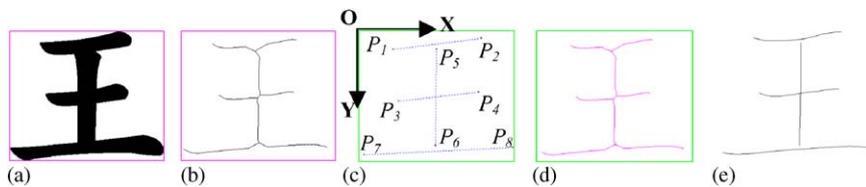


Fig. 5. (a) Image pattern of the Chinese character for “king”, (b) thinned image of (a), writing order of the character in (a), (d) skeleton of the character in (a), (e) the trajectory of the writing brush of the character in (a).

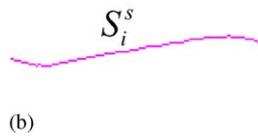
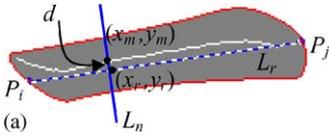


Fig. 6. (a) A thinned stroke and its writing order, (b) the skeleton of the stroke in (a) represented by the coordinates in  $\Sigma_C$ .

To express the thinned image by the coordinate point in  $\Sigma_C$ , a searching procedure is necessary. Let us employ Fig. 6(a) to explain this searching procedure. In Fig. 6(a), two points,  $P_i$  and  $P_j$ , represent the writing order of the stroke. The searching is performed along the reference line,  $L_r$ , which is determined by these two points and is shown by white and blue dotted line.  $L_n$  is a line passing  $(x_r, y_r)$  on  $L_r$  and is perpendicular to  $L_r$ . The coordinates of the point deviated  $d$  dots from  $L_r$  is given by

$$\begin{aligned} x_m &= \pm \frac{d}{\|P_i - P_j\|} (y_j - y_i) + x_r, \\ y_m &= \mu \frac{d}{\|P_i - P_j\|} (x_j - x_i) + y_r, \end{aligned} \quad (1)$$

where when  $x_m$  takes “+” and  $y_m$  “-”,  $(x_m, y_m)$  lies on the right side of  $\overline{P_i P_j}$ , and when  $x_m$  takes “-” and  $y_m$  “+”,  $(x_m, y_m)$  on the left side of  $\overline{P_i P_j}$ , tracing from  $P_i$  to  $P_j$ . The searching is performed in the range of  $0 \leq d \leq d_T$ . If the pixel at  $(x_m, y_m)$  is white, i.e., it falls on the thinned image pattern, it is thought of as a dot of the skeletonized image. If such point cannot be determined when  $d$  changes in the range of  $[0, d_T]$ , the searching procedure fails, and  $x_m$  and  $y_m$  are given a value “-1”. This processing is done for all points on  $L_r$  from  $P_i$  to  $P_j$ . Then the skeletonized stroke is obtained as shown in Fig. 6(b). Let  $S_i^s$  represent this skeletonized stroke,  $S_i^s = \{P_{i1}^s, P_{i2}^s, \dots, P_{iL_i}^s\}$ , where  $P_{ik}^s$  is the point in  $\Sigma_C$  ( $k = 1, 2, \dots, L_i$ ),  $i$  is the stroke code shown in Table 1, and  $L_i$  is the length of  $S_i^s$ . In the following, the start point  $P_{i1}^s$  and the end point  $P_{iL_i}^s$  of  $S_i^s$  are simply written as  $P_s$  and  $P_e$  if no confusion.

For the character image pattern in Fig. 5(a), the skeletonized character is shown in Fig. 5(d) and is represented by  $C_{\text{king}} = \{S_1^s, S_1^s, S_2^s, S_1^s\}$ . It is worth noticing that although  $C_{\text{king}}$  contains three  $S_1^s$ 's, generally speaking these three skeletonized strokes have different lengths.

### 3.2.2. Curve fitting of the skeletonized character

Although the skeletonized character is obtained as a set of coordinate points in  $\Sigma_C$ , the trajectory of the writing brush is not simply connecting these points along the writing order. In other words, these points cannot be used as the trajectory of the writing brush directly because simply connecting these points forms a notched trajectory, as shown in Fig. 5(d). Generally, in ordinary CCC, the trajectory of the writing brush is a

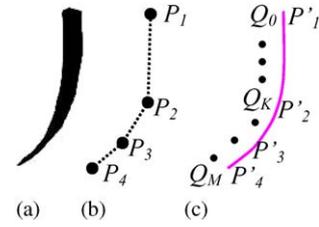


Fig. 7. (a) A left falling stroke, (b) points of the writing order for the stroke in (a), (c) control points and points of B-spline curve.

smooth curve, except the fold points (joint points). Therefore, a curve fitting technique is necessary to generate a smooth trajectory from the points of the skeletonized character. There are many techniques for curve fitting, such as curve fitting by least-squares approximation, curve fitting by composite polynomials, curve fitting by splines (Unser, 1999), and so on. Here, the B-spline fitting is employed. The following, firstly, introduce B-spline fitting in general, and then explain the B-spline fitting techniques for specific strokes.

Fig. 7(a) shows the image of  $S_3$  in Chinese character. Let us employ it to describe 2D trajectory generation of the writing brush. The points representing the writing order of the stroke in (a) are given in (b). If these points are simply connected by a straight line, the stroke is a folded line, and is not as beautiful as the one in (a). To solve this problem, firstly, the image pattern in (a) is skeletonized with the method related in Section 3.2.1. The skeletonized stroke,  $S_3^s$ , is shown in (c). By taking points  $P_j^s$  (here  $j = 1, \dots, 4$ ) on  $S_3^s$  as the control points, B-spline curve determined by these control points is calculated (Kakazu and Furukawa, 1995; Sakurai, 1993), as shown in Fig. 7(c). This B-spline is considered as the 2D trajectory of the writing brush. In detail, B-spline curve is obtained according to

$$x_{i,m}(t) = \sum_{j=0}^m x'_j B_{i+j,m+1}(t), \quad (2)$$

$$y_{i,m}(t) = \sum_{j=0}^m y'_j B_{i+j,m+1}(t)$$

and

$$\begin{aligned} B_{i,k}(t) &= \frac{t - q_i}{q_{i+k-1} - q_i} B_{i,k-1}(t) + \frac{q_{i+k} - t}{q_{i+k} - q_{i-1}} B_{i+1,k-1}(t), \\ B_{i,1}(t) &= \begin{cases} 1 & (q_i \leq t < q_{i+1}), \\ 0 & (t < q_i, t \geq q_{i+1}) \end{cases} \end{aligned} \quad (3)$$

and

$$\begin{aligned} q_0 &= q_1 = \dots = q_{m+1} = 0, \\ q_N &= q_{N+1} = \dots = q_{N-1+m+1} = 0, \end{aligned} \quad (4)$$

$$q_{i+m+1} = i + m/2 \quad (i = 0, 1, \dots, N - m - 1), \quad (5)$$

where  $x'_i$  and  $y'_i$  are coordinates of the control points,  $x_{i,m}(t)$  and  $y_{i,m}(t)$  are coordinates of the points on

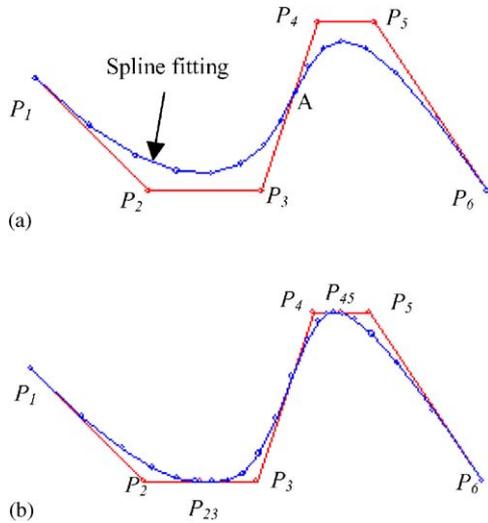


Fig. 8. (a) Fold line determined by six control points and its B-spline fitting, (b) the fitting result is improved by inserting two control points  $P_{23}$  and  $P_{45}$  into  $P_2$  and  $P_3$ ,  $P_4$  and  $P_5$ , respectively, where the curvature takes bigger value.

B-spline curve,  $m$  is the degree of B-spline, and  $N$  is the number of the control points, i.e., knots for B-spline. Note that the number of the points on B-spline curve is  $M = N \times K + 1$ , where  $K$  is the number of divisions between two control points. Hereafter, the point on B-spline curve (node points) is denoted by  $Q_j$  ( $j = 0, \dots, M$ ).

However, the fitting curve obtained in this way is always in the interior of the partial conve hull determined by the control points, as shown in Fig. 8(a). And if the nubner of control points is not enough, the fitting curve will deviate from the original fold curve (refer to (a)). This can be improved by increasing the number of the control points. Fig. 8(b) shows the fitting result by inserting another two control points  $P_{23}$  and  $P_{45}$  into  $P_2$  and  $P_3$ ,  $P_4$  and  $P_5$ , respectively. It is clear that the fitting result is much better than that in (a). Keeping this in mind, let us explain the curve fitting for specific strokes.

3.2.3. Curve fitting for specific strokes

3.2.3.1.  $S_1$ . Fig. 9(a) shows the image pattern of  $S_1$ . Usually, according to CCC knowledge, this stroke is bent upward at point A, and downward at B. The skeletonized  $S_1^s$  is shown in Fig. 9(b). To obtain a good fitting,  $S_1^s$  is separated into three partial curves,  $\{P_{i1}^s, \dots, P_{i,L_i/4}^s\}$ ,  $\{P_{i,L_i/4}^s, \dots, P_{i,3L_i/4}^s\}$ , and  $\{P_{i,3L_i/4}^s, \dots, P_{i,L_i}^s\}$  ( $i = 1$ ). For the first and third partial curves, they are fitted with dense control points, and the second with sparse control points. The number of dense control points and that of sparse control points are denoted by  $N_{dense}$  and  $N_{sparse}$ , separately. The control points are taken from the first and third partial curves at the interval of  $L_1/(4N_{dense})$ , and from the

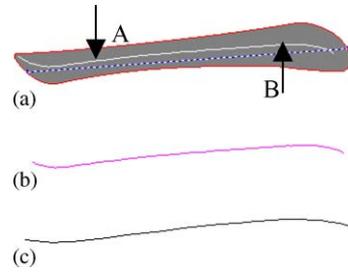


Fig. 9. (a) Image pattern of  $S_1$ , (b) skeletonized pattern of  $S_1$ , (c) trajectory of the writing brush for  $S_1$ .

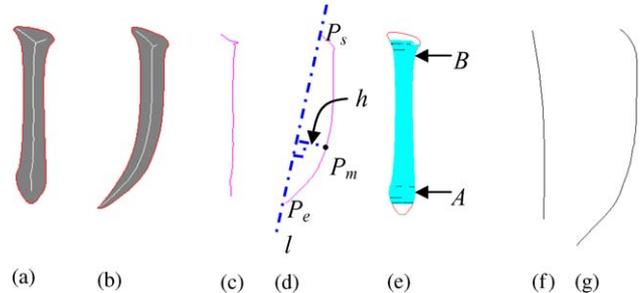


Fig. 10. (a,b) Two kinds of image patterns for  $S_2$ , (c) and (d) are the skeletonized patterns, (e) shows the width of the image pattern in (a), (f) and (g) are trajectories of the writing brush for  $S_2$ .

second at  $L_1/(2N_{sparse})$ . The value of  $N_{dense}$  is about twice that of  $N_{sparse}$ . The curve fitting result for  $S_1$  is given in Fig. 9(c), which is the trajectory of the writing brush for  $S_1$ .

3.2.3.2.  $S_2$ . Fig. 10(a) and (b) shows two kinds of image patterns for  $S_2$ . Here, the processing includes the classification of  $S_{2a}$  and others, and the curve fitting. As shown in Fig. 10(d),  $l$  is a line determined by the start point  $P_s$  and end point  $P_e$  of the skeletonized pattern, and  $P_M$  is the most deviated point from  $l$ . If the deviated distance  $h$  at  $P_m$  is larger than the threshold value  $D_{S_2}$ , the corresponding pattern is considered as  $S_{2a}$  and its trajectory is the same with  $S_3$  (details are refered to Section 3.2.3.3). Otherwise it is thought of as  $S_{2b}$ ,  $S_{2c}$  or  $S_{2d}$ . In this case it is necessary to check the width of the image pattern. As displayed in Fig. 10(e), let  $w_A$  and  $w_B$  denote the width at points A and B, respectively, if  $w_A - w_B$  is smaller than the threshold value  $D_w$ , it is thought of as  $S_{2b}$ , otherwise it is considered as  $S_{2c}$  or  $S_{2d}$ . The curve fitting for  $S_2$  is done by using sparse control points. The trajectories of the writing brush are given in Fig. 10(f) and (g), respectively.

3.2.3.3.  $S_3$ . As mentioned above, Fig. 10(b) also shows the image pattern for  $S_3$ . The most deviated point  $P_m$  is serached in the following way. The distance  $h$  from all points on  $S_3^s$  to the line  $l$  determined by  $P_s$  and  $P_e$  are calculated.  $P_M$  is the point in which  $h$  takes the

maximum, i.e.,

$$h = \max \left\{ \frac{(y_e - y_s)x_i - (x_e - x_s)y_i + ((x_e - x_s)y_s - (y_e - y_s)x_s)}{\|P_e - P_s\|} \right\},$$

$$i = 1, 2, \dots, L_3, \quad (6)$$

where  $(x_s, y_s)$  and  $(x_e, y_e)$  are the coordinates of  $P_s$  and  $P_e$ , respectively, and  $(x_i, y_i)$  of the  $i$ th point on  $S_3^s$ .

$S_3^s$  is then separated into two partial curves  $\{P_{i1}^s, \dots, P_{i,m-\delta}^s\}$  and  $\{P_{i,m-\delta}^s, \dots, P_{i,L_i}^s\}$  ( $i = 3$ ). The first curve is fitted with sparse control points and the second with dense control points. The trajectory of the writing brush for  $S_3$  is given in Fig. 10(g).

It is necessary to note that the  $P_m$  is not taken as the separation point directly. The point deviated to  $P_s$  with a distance  $\delta$ , i.e.,  $P_{m-\delta}$ , is taken as the separation point. This is because the curvature at  $P_m$  is bigger than other points, it is necessary to do curve fitting with dense control points at the local region including  $P_m$ . The value of  $\delta$  is set at  $(L_3 - m)/4$ .

**3.2.3.4.  $S_4$ .** Fig. 11(a) and (b) shows two kinds of the image patterns for  $S_4$ . The processing here involves the recognition of these two patterns and the curve fitting. When tracing from the start point  $P_s$  to the end point  $P_e$  along  $S_4^s$ , if the most deviated point  $P_m$  from the line  $l$  determined by  $P_s$  and  $P_e$  is on the left side of  $l$ , the corresponding pattern is thought of as the one in (b). This case is the same with processing of  $P_{0d}$  (details are given in Section 3.2.3.6). The following mainly discusses the processing for the image pattern in (a). Based on the CCC knowledge, usually this stroke is bent toward upper right at point A, and lower left at B. To obtain a good fitting for  $S_4$ ,  $S_4^s$  is separated into three partial curves,  $\{P_{i1}^s, \dots, P_{i,m_1}^s\}$ ,  $\{P_{i,m_1}^s, \dots, P_{i,m_2-\delta}^s\}$  and  $\{P_{i,m_2-\delta}^s, \dots, P_{i,L_i}^s\}$  ( $i = 4$ ). The first and third partial curves are fitted with dense control points, and the second with sparse control points.  $m_1$  is set at  $L_4/4$  and  $m_2$  is obtained according to Eq. (6). The trajectory of the writing brush for  $S_4$  is shown in Fig. 11(d).

**3.2.3.5.  $S_5$ .** Fig. 12(a) shows the image patterns for  $S_5$ . The processing technique is the same as that for  $S_3$  (refer to Section 3.2.3.3), except that the first partial curve is fitted by using dense control points and the second by sparse control points. The trajectory of the writing brush obtained is shown in Fig. 12(c).

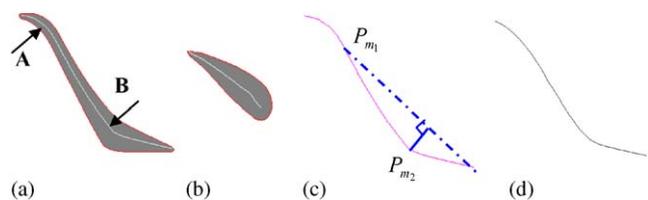


Fig. 11. (a,b) Two kinds of image patterns for  $S_4$ , (c) skeletonized pattern of (a), (d) trajectory of writing brush for  $S_4$ .



Fig. 12. (a) Image pattern of  $S_5$ , (b) skeletonized pattern of  $S_5$ , (c) trajectory of the writing brush for  $S_5$ .

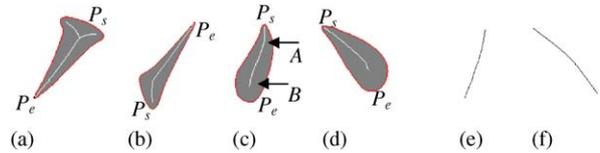


Fig. 13. (a)–(d) Four kinds of image patterns of  $S_0$ , (e)–(f) trajectories of the writing brush for the image pattern in (c) and (d).

It is necessary to note that the start point  $P_s$  and end point  $P_e$  of  $S_5$  are in opposite positions with respect to those of  $S_3$ .

**3.2.3.6.  $S_0$ .** Fig. 13(a)–(d) shows four kinds of the image patterns for  $S_0$ . Here, the processing includes the recognition of these patterns and the curve fitting. The recognition is based on the inclination of  $S_0^s$  and the stroke width. Let  $\theta$  denote the inclination of  $S_0^s$ ,  $\theta = \tan^{-1}((y_s - y_e)/(x_s - x_e))$ . The criteria of discretion is based on CCC knowledge. If  $90^\circ < \theta < 180^\circ$ , the stroke is recognized as  $S_{0d}$ ; if  $-180^\circ < \theta < -90^\circ$ , it is thought of as  $S_{0b}$ . And if  $0^\circ < \theta < 90^\circ$  and  $w_A < w_B$ , it is considered as  $S_{0c}$ , otherwise it is  $S_{0a}$ . The curve fitting technique for  $S_{0a}$  is the same with that of  $S_3$ , and for  $S_{0b}$  the same with  $S_5$ . The curve fitting for  $S_{0c}$  and  $S_{0d}$  are done by using dense control points. The trajectories of the writing brush for these two kinds of dot strokes are shown in Fig. 13(e) and (f), respectively.

It is worth noting that  $w_A$  is the average width of the stroke in the local region near point A, this is the same to  $w_B$ . The point A is set at  $L_0/5$  and B at  $4L_0/5$  (refer to Fig. 13(c)).

**3.2.3.7.  $S_6$  and  $S_{12}$ .** Fig. 14(a) and (b) show image patterns for  $S_6$  and  $S_{12}$ , respectively. The processing for these two strokes includes the fold point detection and the curve fitting. The fold point exists in the region in which the interior angle  $\phi$  of the point of the writing order takes the minimal value, i.e.,

$$\phi = \min \left\{ \cos^{-1} \frac{\|P_i - P_{i-1}\|^2 + \|P_{i+1} - P_i\|^2 - \|P_{i+1} - P_{i-1}\|^2}{2 \cdot \|P_i - P_{i-1}\| \cdot \|P_{i+1} - P_i\|} \right\},$$

$$i = 1, 2, \dots, W, \quad (7)$$

where  $W$  is the number of points of the writing order.

Suppose  $P_i$  is the point of the writing order with minimal interior angle, its inclination angle  $\theta$  is

determined as  $\theta = \tan^{-1}((y_{i+1} - y_i)/(x_{i+1} - x_i))$ . The fold point is considered as the cross point of the line  $l_2$  and the skeletonized stroke, where  $l_2$  is the line passing  $P_i$  and with inclination angle  $\alpha$ . The relation between  $\alpha$  and  $\phi$ ,  $\theta$  is divided into four cases as shown in Fig. 15, where  $\beta = \varepsilon\phi$  and  $\varepsilon$  is the weight coefficient. These four cases can be written as

(i)  $\alpha = \theta - \varepsilon\phi$ , (8)

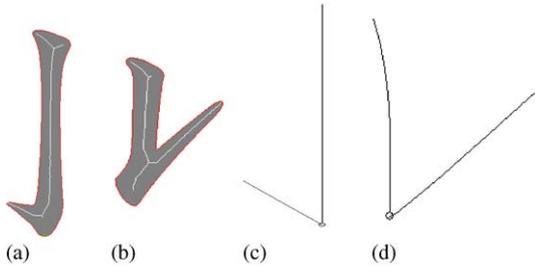


Fig. 14. (a,b) Image patterns of  $S_6$  and  $S_{12}$ , (c)–(d) trajectories of the writing brush for the image patterns in (a) and (b).

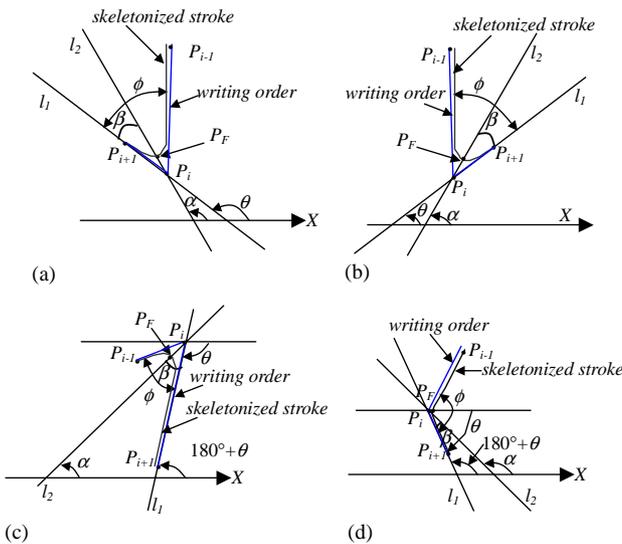


Fig. 15. The position of fold (joint) point and the shape of the writing order when tracing from the start point  $P_s$  to the end point  $P_e$ . (a,c) Show that the partial skeletonized stroke around the fold point is concave to right side but  $\theta$  is positive in (a) and negative in (c); (b) and (d) show that the partial skeletonized stroke around the fold point is concave to left side but  $\theta$  is positive in (b) and negative in (d).

(ii)  $\alpha = \theta + \varepsilon\phi$ , (9)

(iii)  $\alpha = 180^\circ + \theta - \varepsilon\phi$ , (10)

(iv)  $\alpha = 180^\circ + \theta + \varepsilon\phi$ . (11)

$S_6$  belongs to case (i) and  $S_{12}$  to (ii). The searching for the fold point is performed in the following way. Tracing along  $S_i^s$  ( $i = 6, 12$ ), if  $P_{i,m-1}^s$  is above or on  $l_2$  and  $P_{i,m+1}^s$  is below or on  $l_2$ ,  $P_{im}^s$  is thought of as the fold point  $P_F$ . Then  $S_i^s$  ( $i = 6, 12$ ) is separated into two partial curves,  $\{P_{i1}^s, \dots, P_{im}^s\}$  and  $\{P_{im}^s, \dots, P_{iLi}^s\}$ , at the fold point. The fitting for the first partial curve is the same with that of  $S_2$ , and for the second is simply connecting  $P_F$  and its end point. The fitting result for  $S_i^s$  ( $i = 6, 12$ ) is shown in Fig. 14(c) and (d), where the small “o” indicates the fold point. This is the same hereafter.

3.2.3.8.  $S_7, S_8$  and  $S_9$ . Fig. 16(a)–(c) shows image patterns for and  $S_7, S_8$  and  $S_9$ , respectively. The processing for these three strokes is similar to  $S_6$  and  $S_{12}$  related in last section. To find the fold point, it is necessary to determine the inclination angle of the line  $l_2$  (refer to Fig. 15). The relation between the writing order and the skeletonized stroke for  $S_7$  is shown in Fig. 15(a) and Eq. (8), and those for  $S_8$  and  $S_9$  are in Fig. 15(b) and Eq. (9). The searching procedure for the fold point is the same with method related in last section. After the fold point is obtained,  $S_i^s$  ( $i = 7, 8, 9$ ) is separated into two partial curves,  $\{P_{i1}^s, \dots, P_{im}^s\}$  and  $\{P_{im}^s, \dots, P_{iLi}^s\}$ , at the fold point ( $P_{im}^s$  is thought of as the fold point  $P_F$ ). The fitting for the first partial curve is the same with that of  $S_3$ , and for the second is simply connecting  $P_F$  and its end point. The fitting result for  $S_i^s$  ( $i = 7, 8, 9$ ) is shown in Fig. 16(d)–(f), where the small “o” indicates the fold point.

3.2.3.9.  $S_{10}$  and  $S_{11}$ . Fig. 17(a) and (b) shows image patterns for and  $S_{10}$  and  $S_{11}$ , respectively. These two strokes are similar except that  $S_{11}$  has a fold point as shown in (a) but  $S_{10}$  does not. Here,  $S_{11}$  is employed to explain the processing for these two strokes.

As shown in Fig. 17(c),  $S_i^s$  ( $i = 10, 11$ ) is separated into two partial curves at  $P_{i,m1}^s$ .  $P_{i,m1}^s$  is the most deviated point from the reference line  $l_1$  determined by the start

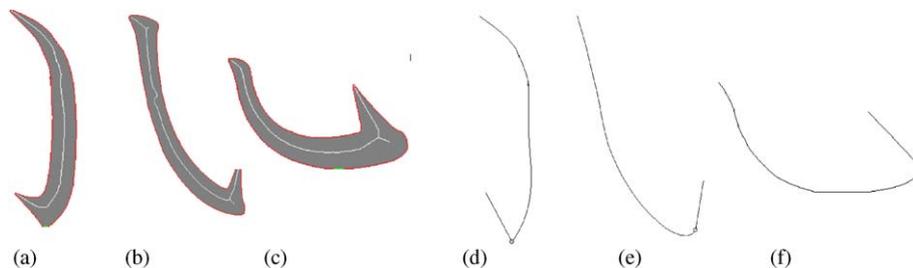


Fig. 16. (a)–(c) Image patterns of  $S_7, S_8$  and  $S_9$ , respectively, (d)–(f) trajectories of the writing brush for the image patterns in (a)–(c), accordingly.

point  $P_s$  and end point  $P_e$  of  $S_i^s$ .  $P_{i,m_1}^s$  is determined according to Eq. (6).  $P_{i,m_2}^s$  can be determined according to the method related in Section 3.2.3.7 (this processing is not necessary for  $S_{10}$ ), which is the fold point  $P_F$ . Then  $S_i^s$  is re-separated into two curves,  $\{P_{i1}^s, \dots, P_{i,m_1-\delta}^s\}$  and  $\{P_{i,m_1-\delta}^s, \dots, P_{i,m_2}^s\}$ . The first curve is fitted with sparse control points, the second with dense control points. From  $P_F$  to the end point, it is simply to connect these two points. The fitting results for these two strokes are shown in Fig. 17 (d) and (e).

3.2.3.10.  $S_{13}$  and  $S_{14}$ . Fig. 18(a) and (b) shows image patterns for and  $S_{13}$  and  $S_{14}$ , respectively. The fold (joint) points are the types as shown in Fig. 15(c), and it can be determined according to Eqs. (7) and (10). The curve fitting for  $S_{13}$  from the start point to the fold point is the same with that of  $S_1$ , and then connecting the fold point and the end point.  $S_{14}$  is considered as the combination of  $S_1$  and  $S_2$ . The trajectories for these two strokes are shown in Fig. 18(c) and (d).

3.2.3.11.  $S_{15}$ ,  $S_{16}$  and  $S_{19}$ . Fig. 19(a)–(c) shows image patterns for and  $S_{15}$ ,  $S_{16}$  and  $S_{19}$ , respectively. The fold (joint) points for  $S_{15}$  and  $S_{16}$  are the types as shown in

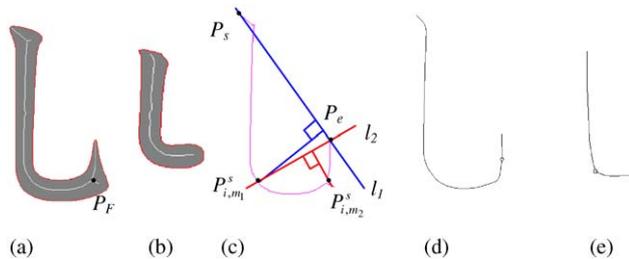


Fig. 17. (a)–(c) Image patterns of  $S_7$ ,  $S_8$  and  $S_9$ , respectively, (d)–(f) trajectories of the writing brush for the image patterns in (a)–(c), accordingly.

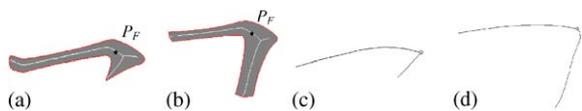


Fig. 18. (a,b) Image patterns of  $S_{13}$  and  $S_{14}$ , respectively, (c) and (d) trajectories of the writing brush for the image patterns in (a) and (b), accordingly.

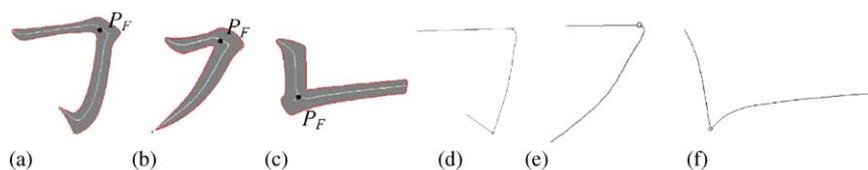


Fig. 19. (a)–(c) Image patterns of  $S_{15}$ ,  $S_{16}$  and  $S_{19}$ , respectively, (d)–(f) trajectories of the writing brush for the image patterns in (a)–(c), accordingly.

Fig. 15(c), and can be obtained based on Eqs. (7) and (10). The fold point for  $S_{19}$  is the type shown in Fig. 15(b) and can be decided according to Eqs. (7) and (9). The curve fitting for  $S_{15}$  can be considered as the combination of  $S_1$  and  $S_6$ , the fitting for  $S_{16}$  of  $S_1$  and  $S_3$ , and that for  $S_{19}$  of  $S_2$  and  $S_1$ . The trajectories for these three strokes are shown in Fig. 19(d)–(f).

3.2.3.12.  $S_{20}$ ,  $S_{21}$  and  $S_{26}$ . The processing for these three strokes are similar.  $S_{20}$  is only employed in simplified Chinese character. Fig. 20(a) and (b) shows image patterns for and  $S_{21}$  and  $S_{26}$ , respectively. The fold (joint) points for them are on the left of the line  $l_1$  when tracing from the start to the end point of  $l_1$  and is the most deviated point from  $l_1$ , as shown in Fig. 20(c). The distances from  $l_1$  is defined in Eq. (6).  $l_1$  is the line determined by the start point and end point of  $S_i^s$  ( $i = 20, 21, 26$ ). The curve fitting for  $S_{20}$  can be considered as the combination of  $S_1$  and  $S_{12}$ , the fitting for  $S_{21}$  of  $S_1$  and  $S_8$ , and that for  $S_{26}$  of  $S_1$  and  $S_{10}$ . The trajectories for  $S_{21}$  and  $S_{26}$  are shown in Fig. 20(d) and (e), respectively.

3.2.3.13.  $S_{18}$ ,  $S_{22}$  and  $S_{27}$ . The processing for these three strokes are similar to those for  $S_{20}$ ,  $S_{21}$  and  $S_{26}$  related in last section.  $S_{27}$  is only employed in simplified Chinese character. Fig. 21(a) and (b) show image patterns for and  $S_{18}$  and  $S_{22}$ , respectively. The difference is that the fold (joint) points for them are on the right of the line  $l_1$  and is the most deviated point from  $l_1$ , as shown in Fig. 21(c). The curve fitting for  $S_{18}$  can be thought of as the combination of  $S_3$  and  $S_0$ , the fitting for  $S_{22}$  of  $S_2$  and  $S_{15}$ , and that for  $S_{27}$  of  $S_2$  and  $S_{16}$ . The trajectories

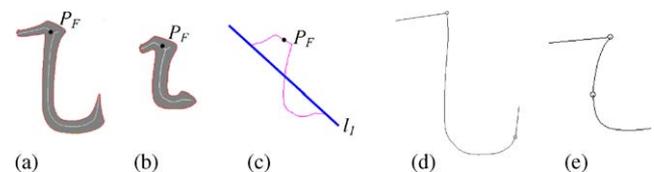


Fig. 20. (a,b) Image patterns of  $S_{21}$  and  $S_{26}$ , respectively, (c) relationship between the fold point and reference line  $l_1$ , (d) and (e) trajectories of the writing brush for the image patterns in (a) and (b), accordingly.

for  $S_{18}$  and  $S_{22}$  are shown in Fig. 21(d) and (e), respectively.

3.2.3.14.  $S_{23}$ ,  $S_{24}$  and  $S_{25}$ . The processing for these three strokes are similar. Fig. 22(a)–(c) shows image patterns for  $S_{23}$ ,  $S_{24}$  and  $S_{25}$ , respectively. The fold (joint) points for them lie in the valley on the right of the line  $l_1$  when tracing from the start to the end point of  $l_1$ , as shown in Fig. 22(d). The fold point  $P_F$  is determined by searching the two peak points  $P_{m1}$  and  $P_{m2}$ , and then searching the point between  $P_{m1}$  and  $P_{m2}$  whose deviation distance from  $l_1$  is minimal. The curve fitting for  $S_{23}$  can be thought of as the combination of  $S_{16}$  and  $S_{16}$ , the fitting for  $S_{24}$  of  $S_{16}$  and  $S_7$ , and that for  $S_{25}$  of  $S_{16}$  and  $S_{15}$ . The trajectories for  $S_{23}$ ,  $S_{24}$  and  $S_{25}$  are given in Fig. 22(e)–(g), respectively.

Above, we discussed the trajectory extraction for the strokes to construct the characters. However, only with trajectory the robot cannot write the characters as human calligraphers do in CCC. It is necessary to control the pressure of the writing brush in order to control the width of the strokes. This will be related below.

#### 4. Pressure control of the writing brush

The control of the pressure to the writing brush is, in fact, to control the width of the stroke. This is realized by controlling  $z$ -coordinates of the writing brush. After taking the Chinese ink, the writing brush is moved to its default position  $P_d(x_d, y_d, z_d)$ . The changes of  $x$ - and  $y$ -coordinates are determined by coordinates of points on B-spline curve, as related in Section 3.2.3. Here we

discuss how to control the changes of  $z$ -coordinates. At the default position, the distance from the tip of the writing brush to the paper is  $z_d$ , as shown in Fig. 23(a). The  $z$ -coordinate of the writing brush is controlled in the range of  $z_d - d_{\min}$  to  $z_d - d_{\max}$ , where  $d_{\min}$  is the distance that the writing brush is moved downward along the  $z$ -axis so that the tip of the writing brush just touches the paper, as shown in Fig. 23(b), and  $d_{\max}$  is the length of the head of the writing brush. Let  $d$  denote the distance that the writing brush is moved downward from its default position along  $z$ -axis, which is called width of the stroke, and  $d_{\min} \leq d \leq d_{\max}$ . Further, let us introduce  $d_x = d/F_x$ , and  $d_y = d/F_y$  ( $F_x$  and  $F_y$  are constant ratios to  $d$ ), which are used to control the writing of the start and end of the stroke. When the writing brush is moved from  $(x_d - 2d_x, y_d + d_y, z_d - d_{\min})$  to  $(x_d + 2d_x, y_d - d_y, z_d - 1.1d)$ , the tip of the writing brush is in the shape as shown in Fig. 23(c). Then it is moved to  $(x_d, y_d, z_d - d)$ , and then it is controlled to write a stroke. This operation is necessary for writing the start of the stroke as that in human calligrapher. The width of the stroke is controlled according to

$$z_k = z_d - \left( d - k\eta \frac{d - d_{\min}}{M} \right), \tag{12}$$

where  $k = 0, 1, \dots, M$ ,  $\eta$  controls the degree of the inclination of the line in (12), i.e., controls the speed so that the stroke gets thin or thick when the writing brush is moved on B-spline curve.

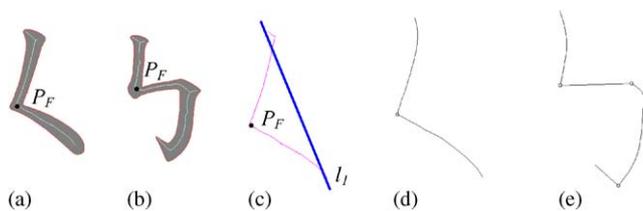


Fig. 21. (a,b) Image patterns of  $S_{18}$  and  $S_{22}$ , respectively, (c) relationship between the fold point and reference line  $l_1$ , (d) and (e) trajectories of the writing brush for the image patterns in (a) and (b), accordingly.

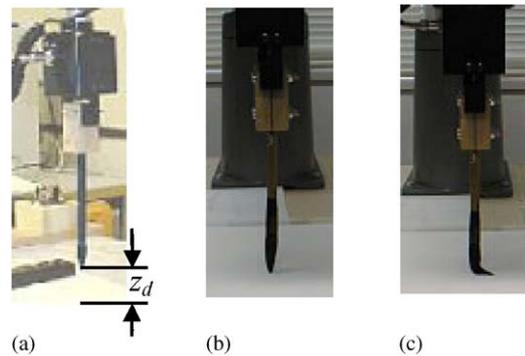


Fig. 23. (a) Default position of the writing brush; (b) the position that the writing brush just touches the paper; (c) the shape of the tip of the writing brush at width  $d$ .

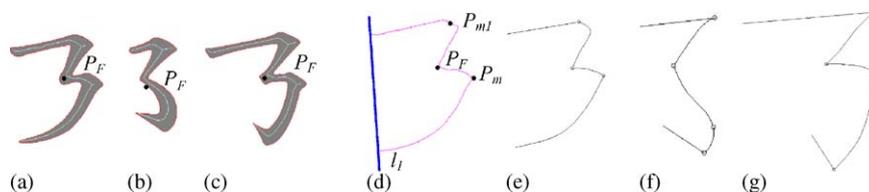


Fig. 22. (a)–(c) Image patterns of  $S_{23}$ ,  $S_{24}$  and  $S_{25}$ , respectively, (d) relationship between the fold point and reference line  $l_1$ , (e)–(g) trajectories of the writing brush for the image patterns in (a)–(c), accordingly.

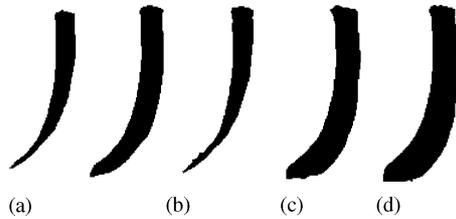


Fig. 24. (a) A left falling stroke ( $\eta = 0.9$ ), (b)–(e)  $\eta = 0.3$ , (c)  $\eta = 1.1$ , (d)  $\eta = -0.6$ , (e)  $\eta = -0.9$ , respectively.

By combining the techniques related above, a stroke can be written smoothly and beautifully by a robot hand holding a writing brush. Fig. 24(b)–(e) shows the same stroke in (a) by changing the value of  $\eta$ , (b) and (c) show that a stroke is getting thin by setting  $\eta$  a positive value. The larger  $\eta$  is, the quicker the stroke gets thin, (d) and (e) show the case that  $\eta$  is negative. The smaller  $\eta$  is, the faster the stroke gets thick.

## 5. Experiment results

The whole system is implemented on Windows platform, and the programming language is C++. The value of the number of the sparse control points,  $N_{\text{sparse}}$ , is set at 4–6, and  $N_{\text{dense}}$  is set at 8–12. The size of the writing area  $T_R$  is set at 200 mm. The searching range,  $d_T$ , is set at 15 dots. The values of  $D_{S_2}$  and  $D_w$  for the classification of  $S_{2a}$  and others is 20 and 10 dots, respectively. The number of divisions,  $K$ , between two control points is set at 3. The values of  $d_{\text{min}}$  and  $d_{\text{max}}$  are dependent on the length of the writing brush tip. For the present writing brush in use, they are 12 and 30 mm, respectively. The width of the stroke,  $d$ , is set at 19 mm. Note that  $d$  is the displacement of the robot hand along  $z$ -axis from its default position. The ratios to the width of the stroke, i.e.,  $F_x$  and  $F_y$ , are set at 5. The value of the degree of inclination,  $\eta$ , is summarized in Table 3. Some experimental results are shown in Fig. 25. The first column from left shows the input image patterns of characters for “mountain”, “water”, “flat”, “heart”, “minute”, in order from top to bottom. The second column gives the writing order denoted by the points in  $\Sigma_C$ . The third column lists the trajectories of the writing brush for the characters in the first column. The fourth column shows characters written by CCC robot based on the trajectories in the third column and the CCC knowledge related in Section 2.

## 6. Conclusions

This paper related the trajectory extraction of the writing brush in CCC based on image and curve processing techniques and the CCC knowledge. This trajectory is used in a CCC robot which is developed to

Table 3  
Value of the coefficient of the degree of inclination

Stroke code	$\eta$
$S_0$	
Right inclined dot	2.2
Rising dot	2.0
Right falling dot	2.4
$S_3$	0.7
$S_4$	2.0
$S_7$	2.0
$S_9$	2.0
$S_{16}$	0.3
$S_{18}$	
First half	0.3
Second half	-0.5

inherit CCC techniques. Because the CCC is not static, but a dynamic process of an activity concerning a lot of complicated factors such as the pressure control to the writing brush, speed control of the writing brush, how to write the start and end of the stroke, how to write the turn and fold on the way of the stroke, and so on, we proposed to inherit this dynamic process by a robot system. We developed the writing techniques for a robot arm to write block style character with a writing brush. The total number of Chinese characters is more than 800,000, and that in daily use is about 3500 (URL: [wenxue.tom.com](http://wenxue.tom.com)). No matter how complicated the Chinese character is, it can be constructed by the strokes in Table 1. Therefore, we mainly related the writing techniques for a CCC robot to write the strokes in Table 1. At present, the robot can write any character in block style.

A character can be written in five different styles. At present, the robot can write characters in block style with the techniques related in Section 3. And also, it is necessary to make the robot be able to write the characters in other styles (ancient, angular, semi-cursive, and cursive). These are our future works. The direct application of this system is the design and printing of signboards. However, the main purpose of this research is to preserve and develop the CCC culture. Nowadays, with the spread of the computer word processor, such as Word, Ichitaro, and so on, more and more people get used to them and do not like to write characters even by pen or pencil, not to mention the writing brush. Therefore, the number of competent calligraphers is becoming smaller and smaller day by day. If this situation continues for several decades, the CCC culture may face the crisis of extinction. If the robot can master all the skills of a professional calligrapher, it can do creative jobs such as making new CCC artworks. Further, the robot can instruct people in the study of calligraphy. In this way, the robot can preserve, inherit and develop the CCC culture. This is the final goal we are working toward.

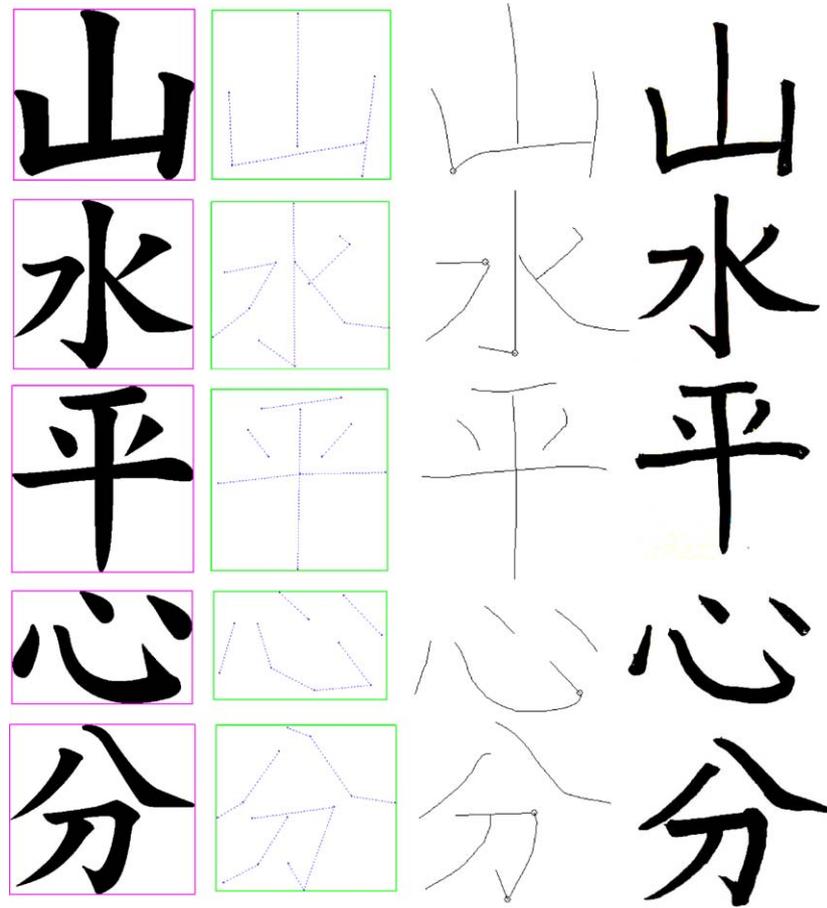


Fig. 25. First column from left: input characters, second column: writing order for the characters in the first row, third column: extracted trajectories of the writing brush for the characters in the first row, fourth column: characters generated by a CCC robot.

## Reference

- Bradbeer, R.S., Billingsley, J. (Eds.), 2002, *Mechatronics and Machine Vision 2002: Current Practice*. Research Studies Press Ltd., England.
- Hasegawa, J., et al. (Eds.), 1986, *Image Processing on Personal Computer*. Gijyutsu-Hyoron Co. Ltd., Tokyo.
- Kakazu, Y., Furukawa, M. (Eds.), 1995, *Fundamentals of Shape Processing Engineering*. Morikita Shupan Co., Ltd., Tokyo.
- Matsuda, N. (Ed.), 1996, *Gotaijikan*. Kashiwashobo Publishing Co., Ltd., Tokyo.
- Sakurai, A. (Ed.), 1993, *Spline Function Based on C*. Tokyo Denki University Press, Tokyo.
- Suzuki, S. (Ed.), 2002, *Fundamentals of Chinese Character Calligraphy (Block Style)*. KIN-ENSHA, Tokyo.
- Unser, M., 1999. Splines—a perfect fit for signal and image processing. *IEEE Signal Processing* 16 (6), 22–38.

**Fenghui Yao** He received his B.E. degree from Dalian Maritime University, China, in 1984, M.E. degree and Doctor of Engineering degree from Kyushu Institute of Technology, Kitakyushu, in 1988 and 1992, respectively. He worked as a research associate in Faculty of Engineering, Kyushu Institute of Technology, from 1994 to 1996. He joined the Faculty of Engineering, University of East Asia, Shimono-seki, in 1997. He served as an associated professor at General Information Processing Center, Shimane University, from 2002 to 2003. Currently he is an associate professor at Department of

Computer Science, Tennessee State University. His current research interests are image processing, mobile robot, sensor fusion, multimedia information processing, computer architecture, and parallel processing. He is a member of IEEE, IPSJ, IEICE and RSJ.

**Guifeng Shao** She received her B.E. degree from Wuhan University of Technology on Surveying and Mapping, China, in 1986, and M.E. degree and Doctor of Engineering degree from Kyushu Institute of Technology, Kitakyushu, in 1991 and 1995, respectively. She served as an associate professor at Department of Commercial Science, Seinan Gakuin University, Fukuoka, from 1994 to 2003. Currently she is an assistant professor at Department of Computer Science, Tennessee State University. Her research interests are natural language understanding and knowledge representation. She is a member of IPSJ.

**Jianqiang Yi** He received his B.E. degree from Beijing Institute of Technology, China, in 1985, and M.E. and Doctor of Engineering degrees from Kyushu Institute of Technology, Kitakyushu, Japan, in 1989 and 1992. From 1992 to 1994, he worked at Computer Software Development Company, Tokyo. From 1994 to 2001, he worked as a chief researcher at Mycom, Inc., Kyoto. Currently, he is a Professor in the Laboratory of Complex Systems and Intelligence Science, Institute of Automation, Chinese Academy of Sciences. His research interests include theories and applications of fuzzy control, neural networks, intelligent control, intelligent robotics, and underactuated systems.