2. PRESERVATION OF CCC ARTWORKS

To make it easier to restore and search the characters written by famous calligraphers in Chinese history, it is necessary to construct the CCC database. The CCC database consists of a Chinese character pattern dictionary (CPD) and an index dictionary of the CPD (ICPD).

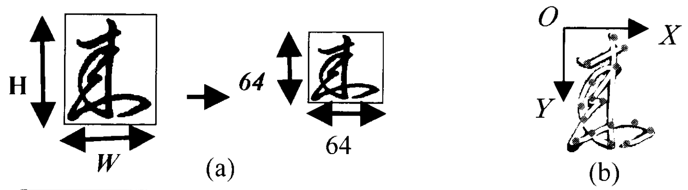
2.1. CPD

The CPD is constructed on the basis of characters written by famous calligraphers in Chinese history since the Tang Dynasty [2–4]. The image patterns of all characters with different styles written by different calligraphers are registered to the CPD with a fixed record length. The image patterns of characters are extracted in the following way. (i) A scanner is used as the input device. The input image is binarized by the fixed thresholding method. (ii) All contours are detected from the binarized image [5] and their circumscribed rectangles are calculated. (iii) For a rectangle, if its size (larger one between width and height) is smaller than T_n, it is thought of as a noise clump and is discarded. If it is included in another rectangle, it is also discarded because the corresponding area of the given rectangle can be represented by another rectangle. If it intersects with another rectangle, they are merged into one rectangle. After this processing, the left rectangles represent the area of a character or a part of a character. Figure 3b shows the extracted characters from the input image in Fig. 3a. (iv) The extracted characters are registered in the CPD, after being normalized into $T_{char} \times T_{char}$ dots (at present $T_{char} = 64$). The record structure of the CPD consists of four fields: code, style, ratio and character pattern image, as given in Fig. 4a, which are described in the following.

- (i) *Code*. Chinese character code, 2 bytes.
- (ii) *Style*. One byte, 0, 1, ..., 4 for ancient, angular, block, semi-cursive style and cursive, respectively, and 5 for all styles.
- (iii) *Ratio*. One 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 one-byte integer, the integer $H/W \times 100$ is used.
- (iv) *Character pattern image*. Because the normalized character pattern is a binary image, 1 byte can represent 8 dots. The total size of one character pattern image is 8×64 bytes. Therefore, the size of one record in the CPD is fixed; its length is 516 bytes.

2.2. ICPD

The index dictionary is necessary to be able to manage the CPD. The ICPD provides the information for CPD searches. Its record consists of three fields: code, number and pointers to the CPD, as shown in Fig. 4b. Its fields are as follows.



Code	*Style*	*Ratio*	*Character pattern image*	(c)

Code	*Number*	*Pointers to CPD*	(d)

Code	*Style*	*Ratio*	*Pointer to CPD*	*Length*	*Points*	(e)

Figure 4. (a) Normalization of the extracted character for 'come'. (b) Writing order represented by points in the coordinate system *XOY*. (c) Record structure of the CPD. (d) Record structure of the ICPD. (e) Record structure of the WOD.

(i) *Code.* Chinese character code, 2 bytes.

(ii) *Number.* Two bytes, number of character patterns corresponding to the character in the code field, registered in the CPD, i.e. different versions of the same character are all registered in the CPD.

(iii) *Pointers to the CPD.* Pointers to the character patterns in the CPD, corresponding to the character in the code field. Each pointer takes 4 bytes, so the length of a record is $Number \times 4 + 4$ bytes.

In this way, any new writing of a character that is already registered in the CPD or the writings for a new character can be added to the end of the CPD. Note that it is necessary to update the ICPD if the CPD is updated by an operation such as adding new writings for a new character.

Furthermore, to let a robot do CCC activity, the writing order must be given. The writing order is saved to the Writing Order Dictionary (WOD). The WOD is constructed based on the points of the folded line as shown in Fig. 4b, which are assigned in coordinate system XOY. The fields of the WOD are as follows (refer to Fig. 4e).

(i) *Code.* Chinese character code, 2 bytes.

(ii) *Style.* Same as the style field in the CPD.

(iii) *Ratio.* Same as the ratio field in the CPD.

(iv) *Pointer to the CPD.* Four bytes, position in the CPD for the character in the code field.

(v) *Length.* Two bytes, length of the folded line, i.e. the number of points in the coordinate system *XOY* (refer to Fig. 4b).

(vi) *Points.* x- and y-coordinates of points in *XOY*. Each point takes 4 bytes. Therefore, the size of a record is $8 + Length \times 4$ bytes.

Figure 5. Image patterns of the character for 'water' registered in the CPD, written by different calligraphers in different styles.

2.3. CCC database searching and image pattern restoration

Two main operations of CCC database management are searching and image pattern restoration. For a given Chinese character, the CCC database is searched at high speed by using the character code or the character code and its writing style. From the output results, the user can designate the character to restore. The designated character is enlarged and displayed on the screen. The restoration is preformed manually because it needs the knowledge about the Chinese character. The restored image pattern is then written back to the CPD. Figure 3c shows the restored image patterns for the characters extracted from the input image in Fig. 3a. Figure 5 shows the image patterns of a character for 'water' registered in the CPD, written by different calligraphers in different styles.

3. INHERITANCE OF CCC BY A ROBOT

CCC is a dynamic activity which concerns the pressure control, movement speed control and turn angle control of the writing brush, etc. The more difficult and important thing is to restore and inherit this activity. It is very difficult for a person to master CCC techniques because the top of the writing brush is soft and hard to control. Generally speaking, it needs several to tens of years of training for a person to reach a good level of CCC. Here, we employed a robot system to inherit the CCC techniques.

3.1. CCC robot system

Figure 6 shows the prototype of the 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 the CCC database. Let $\sum_R$ represent the robot coordinate system, $\sum_R$ is a right-hand Cartesian coordinate system. $\sum_R$ is given in Fig. 6, the area for CCC robot writing lies on the XY-plane in $\sum_R$, 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 the size of the writing area. The center of the writing area is given by $(x_R + T_R/2, y_R - T_R/2, 0)$.

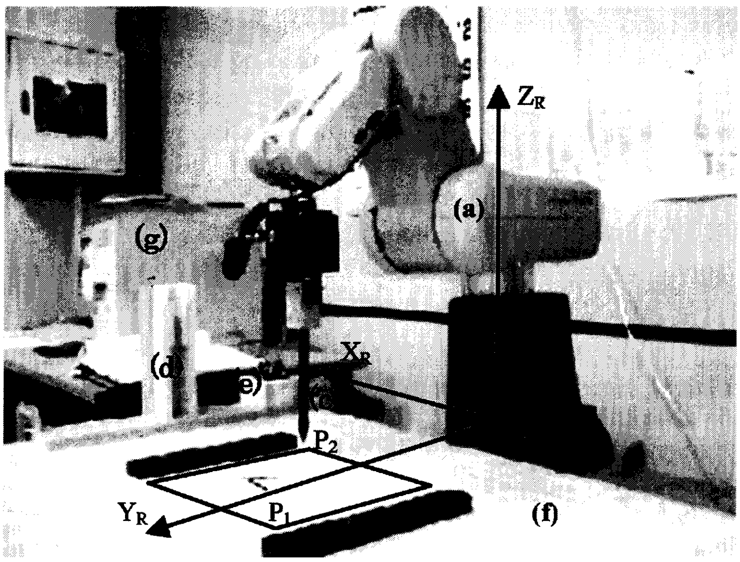


Figure 6. Prototype of the CCC robot. (a) Robot arm. (b) Robot hand. (c) Writing bruch. (d) Writing brush holder. (e) Chinese ink holder. (f) Writing table. (g) System controller and CCC database.

3.2. Basic techniques for writing block style characters in CCC

Among the styles of Chinese characters in daily use, the block style is used most frequently and most accurately. To write the block style characters by the writing brush, each stroke has its own composition method specific for the 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 end of the stroke [2, 3].

(i) *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 lowering the writing brush in this orientation, the tip of the writing brush is toward the upper left and the body of it toward 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 θ as shown in Fig. 7(a, 1), is about 60°. This angle is called the start stroke angle hereafter. However, in fact, to write the character vigorously, the start stroke angle is controlled in the range from 0° to almost vertical to the direction in which the writing brush will move. Generally, it is alright to control this angle at about 60° regardless of the vertical stroke, the horizontal stroke or others.

(ii) *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, lower left to form a left falling stroke and others. Figure 7(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 body of it lies on the lower. Similarly, in the vertical stroke, the

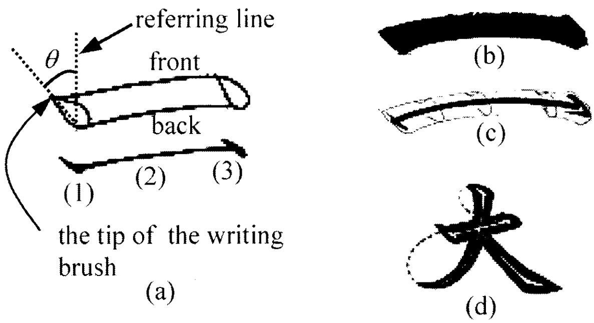


Figure 7. Basic techiques for writing strokes. (a) (1) Start of the stroke. (2) Movement of the writing brush. (3) End of the stroke. (b) Sample of the horizontal stroke. (c) Trajectory of the writing brush in 3D space for the horizontal stroke. (d) Rhythmic coherence of the strokes in the character for 'big'.

tip of the writing stroke lies on the left side of the stroke and the body of it lies on the right. The route of the tip of the writing brush is called the front of the stroke and the route of the body is called the back of the stroke. When the writing brush is used naturally, the front and back of the stroke appear together.

(iii) *End of the stroke.* The direction of the end of the stroke is naturally the same with the direction of the start as shown in Fig. 7(a, 3). The end of the left falling stroke and right falling stroke changes to a sweep.

Figure 7b shows a horizontal stroke written with the writing brush and Fig. 7c 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. 7d, 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 dotted line). The end of the present stroke contacts 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 these rhythmic coherences. Generally speaking, the writing brush is controlled to move slowly at the start, end and turn points of the stroke, and fast at other places such as the straight and curved parts of the stroke.

3.3. Taking CCC as a problem in a robot system

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 a general case, i.e. the calligrapher has '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 writing brush movement 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 $\sum_R$. Let $P(x, y, z)$ express a point on the 3D trajectory of the writing brush. The change of the z-coordinate means the change of the pressure to the writing brush, and x- and y-coordinates form a 2D trajectory of the writing brush in the XY-plane. If this plane is set on the writing table as shown in Fig. 6, 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. The following gives the description in detail.

3.3.1. 2D trajectory generation of the writing brush. Figure 8a shows one of the strokes in a Chinese character. Let us employ it to describe the 2D trajectory generation of the writing brush. Usually, Chinese characters are written by arranging strokes in the correct order. To let the robot do CCC, the writing order of the strokes must be given. This information is put into the WOD. In the WOD the stroke is expressed by the control points. The control points of the stroke in Fig. 8a are given in (b), i.e. the writing brush should pass $P_1(x_1, y_1)$ to $P_4(x_4, y_4)$. If these four points are simply connected by straight lines, the stroke is a folded line and is not as beautiful as the one in (a). Another problem is that the control points may not lie on the center of the stroke. To solve these problems, first, the image pattern in (a) is thinned as shown in (c) and the control points are adjusted to P'_j $(j = 1, \ldots, 4)$ on the thinned image. The processing is as follows. Let l denote the line vertical to the line $P_j P_{j+1}$. P'_j is searched in the range from the point P^l_j to P^r_j along l. When the searching encounters a white dot on the thinned image, that point is thought of as P'_j. If this searching fails, P_j is set as P'_j. P^l_j and P^r_j lie on the left side and right side of $P_j P_{j+1}$, respectively, when moving from P_j to P_{j+1}. Their coordinates are given by:

$$
\begin{aligned}
x &= \mp(y_{j+1} - y_j) \times T_W / \|P_j P_{j+1}\| + x_j, \\
y &= \pm(x_{j+1} - x_j) \times T_W / \|P_j P_{j+1}\| + y_j,
\end{aligned}
\tag{1}
$$

where T_W is the range to search, i.e. the distance from P_j to P^l_j or P^r_j. When x takes '$-$' and y '$+$' in the first item in (1), (x, y) lies in the left side of line $P_j P_{j+1}$ when

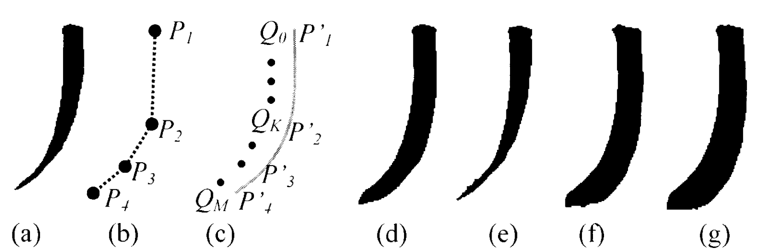


Figure 8. (a) A left falling stroke ($\delta = 0.9$). (b) Control points for the stroke in (a). (c) Adjusted control points of the B-spline curve. (d) $\delta = 0.3$; (e) $\delta = 1.1$; (f) $\delta = -0.6$; (g) $\delta = -0.9$.

moving from P_j to P_{j+1}. In the opposite case, (x, y) lies in the right side of line $P_j P_{j+1}$.

Second, the B-spline curve determined by P'_j $(j = 1, \ldots, 4)$ is calculated [6–8]. This B-spline is considered as the 2D trajectory of the writing brush. In detail, the B-spline curve is obtained according to:

$$x_{i,m}(t) = \sum_{j=0}^{m} x'_i B_{i+j,m+1}(t), \qquad y_{i,m}(t) = \sum_{j=0}^{m} y'_i B_{i+j,m+1}(t), \tag{2}$$

and:

$$\begin{aligned} B_{i,k}(t) &= \frac{t - q_i}{q_{i+k-1}} B_{i,K-1}(t) + \frac{q_{i+k} - t}{q_{i+k} - q_{i-1}} B_{i+1,K-1}(t), \\ B_{i,1} &= \begin{cases} 1 & (q_i \leqslant t < q_{i+1}) \\ 0 & (t < q_i, or\ t \geqslant q_{i+1}), \end{cases} \end{aligned} \tag{3}$$

and:

$$q_0 = q_1 = \cdots = q_{m+1} = 0, \qquad q_N = q_{N+1} = \cdots = q_{N-1+m+1} = 0, \tag{4}$$

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

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

3.3.2. Moving speed of the writing brush. When the writing brush is controlled to move along the trajectory determined in the way as shown above, it will move faster in the interval from P'_1 to P'_2 than from P'_2 to P'_3, although the number of points of the B-spline curve in these two intervals is the same. This is because the B-spline curve in the interval from P'_2 to P'_3 is shorter than that in the interval from P'_1 to P'_2. This is similar to the hand movement of a human calligrapher.

3.3.3. 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 the 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 the x- and y-coordinates are determined by the coordinates of points on the B-spline curve. Here, we discuss how to control the changes of the 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. 9a. 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

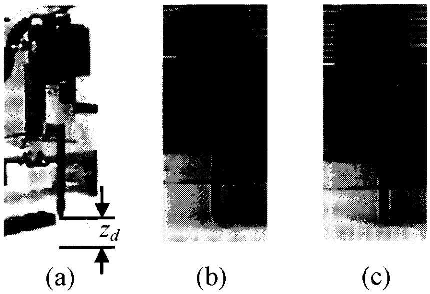


Figure 9. (a) Default position of the writing brush. (b) The position where the writing brush just touches the paper. (c) The shape of the tip of the writing brush at width d.

touches the paper, as shown in Fig. 9b, 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 the $-Z$-axis, it is called the width of the stroke, and $d_{\min} \leqslant d \leqslant d_{\max}$. 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 $(x_{\rm d} - 2d_x, y_{\rm d} + d_y, z_{\rm d} - d_{\min})$ to $(x_{\rm d} + 2d_x, y_{\rm d} - d_y, z_{\rm d} - 1.1d)$, the tip of the writing brush is in the shape as shown in Fig. 9c. Then it is moved to $(x_{\rm d}, y_{\rm d}, z_{\rm d} - d)$ and controlled to write a stroke. This operation is necessary for writing the start of the stroke as done by a human calligrapher. Details will be given in Section 3.4. The width of the stroke is controlled according to:

$$z_k = z_{\rm d} - \left(d - k\delta\frac{d - d_{\min}}{M}\right), \tag{6}$$

where $k = 0, 1, \ldots, M$, δ controls the degree of the inclination of the line in (6), i.e. controls the speed that gives thin or thick strokes when the writing brush is moved on the B-spline curve.

By combining the techniques related above, a stroke can be written smoothly and beautifully by a robot hand holding a writing brush. Figure 8d–g shows the same stroke in (a) by changing the value of δ. Figure 8d and e shows that a stroke gets thinner by setting δ a positive value. The larger δ is, the quicker the stroke gets thinner. Figure 8f and g shows the case that δ is negative. The smaller δ is, the faster the stroke gets thicker. Other control techniques for different strokes will be related in the next section.

3.4. Control techniques to write strokes

All Chinese characters are constructed by one to several of the strokes in Table 1. The following will discuss the control techniques for a robot hand holding a writing brush to write these strokes. In the following, the coordinates of the control points are changed to the coordinates in the robot coordinate system $\sum_{\rm R}$ by:

$$x_{\rm r} = (x_{\rm R} + T_{\rm R}/2) + x_{\rm w} - X_{\rm char}, \quad y_{\rm r} = (y_{\rm R} - T_{\rm R}/2) + (H_{\rm char} - y_{\rm w}) - Y_{\rm char}, \tag{7}$$