ABSTRACT

Chinese brush calligraphy is a valuable civilization legacy and a high art of scholarship. It is still popular in Chinese banners, newspaper mastheads, university names, and celebration gifts. There are Web sites that try to help people enjoy and learn Chinese calligraphy. However, there lacks advanced services such as content-based retrieval or vivid writing process simulation for calligraphy learning. This article proposes a novel Chinese calligraphy learning approach: First, the scanned calligraphy pages were segmented into individual calligraphy characters using minimum-bounding box. Second, the individual character’s feature information was extracted and kept. Then, a corresponding database was built to serve as a map between the feature data and the original raw data. Finally, a retrieval engine was constructed and a dynamic writing process was simulated to help learners get the calligraphy character they were interested in and watch how it was generated step by step.
INTRODUCTION

When computers and the Internet become more and more popular to the general public, less and less people have chances to write with a pen and to enjoy the beauty of writing. Calligraphy is a kind of writing, and a popular communication tool in ancient China. It is not only delightful to the eye and an inspiration to the spirit, but also a creative art. Yet, you do not have to be an “artist” to learn calligraphy, you can learn the skills and write them every time you want. According to thousands of years of learning experience, Chinese calligraphy learning process can be divided into three main consecutive steps: reading, understanding, and simulating.

In terms of Web-based learning, key issues in such process are: how to manage all the data to display the beauty of the different styles of the same calligraphy to learners; how to help learners find the context of an interesting character; and how to help learners follow good writing examples since it is impossible to trace the entire history and show how a particular calligraphy character was written. Correspondingly, our system consists of a large database managing all the scanned original data and the corresponding feature data, a retrieval engine helping learners find the same calligraphy character written in different styles by different people in different dynasties, and a simulator helping learners get a vivid idea about how a calligraphy character was written.

The remainder of this paper is organized as follows: the second section discusses the related works. The third section presents the system architecture of our system. The fourth section gives the data structure. In the fifth section, the main functions of our Web-based calligraphy learning system were described in detail. In the sixth section, the implementation and evaluation were done. And in the final part, conclusions and future works are given.

RELATED WORKS

Numerous researches have been done on exploring techniques for Web-based learning (such as Zhuang, 2002; Zhuang, 2004). But, these techniques do not fit Web-based Chinese calligraphy learning well. Some Web sites have been developed to try to fit learners’ needs to enjoy and learn Chinese calligraphy (such as http://www.wenyi.com/art/shufa; http://www.shw.cn/93jxsd/jxsd.htm). They do provide some basic information and many useful learning materials; however, they provide no advanced dynamic services such as content-based search, and they do not tell the vivid writing process of an individual calligraphy character that may be of interest by a learner.

If it is a text query, Google is the biggest and fastest search engine. It also provides image-searching function based on the name of the image. Yet, you cannot submit a text query and retrieve character images similar to it. Many previous content-based image retrieval works used low-level features such as colors, textures, and regions. However, such features cannot represent shape properties of character, hence irrelevant images are frequently retrieved. Recently, there has been work done to handle shape features effectively (Belongie, Malik, & Puzicha, 2002; Park, 2004). Still, they do not work well for calligraphy character image retrieval. Our previous work (Yueting Zhuang, 2004) has proposed a new approach to retrieve calligraphy character.

SYSTEM ARCHITECTURE

Figure 1 gives out an overview of our system architecture of Web-based Chinese calligraphy learning. Its infrastructure mainly includes data collection, segmentation, and feature extraction, which serve for advanced learning purposes.
Data Collection

The original books, mostly ancient, were scanned at 600 dpi (dots per inch) and kept in DjVu format by researchers of our China-U.S. Million Book Digital Library Project (Zhao & Huang, 2004). These digitized resources, together with their corresponding metadata are saved and packaged. The metadata standard (Edocument Metadata, Version 2.0) we used is released by Zhejiang University Libraries. It combines two kinds of metadata: DC and MARC.

Segmentation

When digitized page images are obtained, segmentation is needed in order to get feature information of an individual calligraphy character. Much research has been done on segmentation of printed page (such as Breuel, 2002; Manmatha, Han, Riseman, & Croft, 1996). Yet no published paper has been done successfully on Chinese calligraphy pages’ segmentation. It is mainly because calligraphy characters have more connection, and the background has more noise such as man-made seals. Our proposed segmentation approach first adjusts color luminance to get rid of red seals and smooth them to take out some noises. Then binarization was done followed by projecting. After that, pages were cut into columns according to the projecting histogram, and columns continued to be cut into individual characters using minimum-bounding box as used in Manmatha et al. (1996). Figure 2 gives an example of our experiment, showing how a calligraphy page was cut into individual calligraphy characters.

Compared with Manmatha et al. (1996), our segmentation approach made special constrained parameters to fit the characteristics of Chinese calligraphy. Let $x_{i,s}$ and $x_{i,e}$ denote the start and the end position of the $i^{th}$ cutting block, then according to our long term segmentation experiences it subjects to the following constrains:

$$x_{i,e} - x_{i,s} \geq 5, \ i = 0,1,2,3,\ldots, n$$  \hspace{1cm} (1)$$

$$2.5 \times \frac{1}{n} \sum_{i=1}^{n} (x_{i,e} - x_{i,s}) \geq x_{i+1,s} - x_{i,e} \geq 0.35 \times \frac{1}{n} \sum_{i=1}^{n} (x_{i,e} - x_{i,s})$$  \hspace{1cm} (2)$$

This is because according to thousands of years of calligraphy writing experience, the width of individual character images on the same page tend to be similar, that is to say they have a minimum and maximum threshold for width, as described in formula 2. Let $\text{width}_i$ and $\text{height}_i$ be the width and height of a cutting block, then

$$0.6 \leq \frac{\text{height}_i}{\text{width}_i} \leq 1.2$$  \hspace{1cm} (3)$$
Formula 3 tells the story that Chinese characters are always in square as introduced in You-Shou and Xiao-Qing (1992). With the aforementioned idea, most of the characters are segmented correctly. But still, there are few man-made connections that cannot be correctly segmented automatically such as in Figure 2, the fourth column from the right. Then, we draw the minimum-bounding box, which can be dragged and dropped manually.

Feature Extraction

After the segmentation was done, the next step is to extract features of individual calligraphy characters. In our approach, a calligraphy character is represented by its contour points instead of its skeleton. This is because skeleton representation is very sensitive to noise. As a result, it produces distorted strokes and the proper shape of the character cannot be detected.

According to the minimum-bounding box, we first normalize the individual character to 32×32 in pixels. Then, canny edge detector as introduced in http://homepages.inf.ed.ac.uk/rbf/HIPR2/canny.htm was employed to get its contour point’s positions in Cartesian coordinates. Finally the values that denote the position of contour points were serialized to a string and kept in the database.

For learning purposes, a learner may want to know where an individual calligraphy character comes from and who wrote it. So, the original location of individual calligraphy characters, that is to say the location of the minimum-bounding box, should be kept too.

DATA STRUCTURE

The scanned original data are large and in disorder, and they need further management. We built a special data structure to map the extracted feature data to the original raw data. The map consists of four tables: book, works, character, and author, as shown in Figure 3. Many individual calligraphy characters compose a calligraphy works created by a calligraphist, namely an author. And many calligraphy works build up a calligraphy book.

Figure 2. An example of segmenting a page into individual characters using minimum-bounding box
The arrows show how these four tables are related by special elements. In the table of character, co-points are a string produced by the feature data of an individual calligraphy character. Then, it is natural that when we get an individual calligraphy character, we can tell in which works and in which page it is located by checking the table of character to find the worksID, the page and the minimum-bounding box (represented by “top_X,” “top_Y,” “bottom_X,” and “bottom_Y”). We can also tell in which book to find this character by searching the table of works using the key of “bookID,” and also who wrote this calligraphy character by checking the table of author using the key of “authorID.”

**KEY CALLIGRAPHY LEARNING SERVICES**

**Learning Object Retrieval**

For personalized learning purposes, different learners may be interested in different styles of the same calligraphy character. In our system, we use our new content-based calligraphy character image retrieval approach (see detailed description in Yueting Zhuang, 2004). In that, we use inexact shape context matching. However, in terms of Web-based learning its response time is beyond endurance: the average retrieval time is about 3.6 minutes for 336 isolated characters. In this paper, we develop a new architecture fit
for Web-based calligraphy character learning, as shown in Figure 4.

Three ways are proposed to speed up the retrieval time, one is preprocessing, another is classification, and the third is dimensionality reduction to reduce computing time. In preprocessing, shape features of every individual calligraphy character were extracted in advance, then serialized to a string and stored in the database, together with their corresponding metadata. Thus, only the features of the query image need to be extracted dynamically.

For classification, complexity index was used besides the metadata. Let \( f(x, y) \) be the gray value of a pixel. If a pixel belongs to the background, then let \( f(x, y) = 0 \), else \( f(x, y) = 1 \). The moment \( m_{ij} \) can be defined as:

\[
m_{ij} = \int \int x^i y^j f(x, y) \, dx \, dy
\]

And the root of second-order central moment in X and Y direction are defined as follows:

\[
\sigma_x = \sqrt{(m_{20} - m_{00}^2)/m_{00}}
\]

\[
\sigma_y = \sqrt{(m_{02} - m_{00}^2)/m_{00}}
\]

Then complexity index can be computed as:

\[
C = (L_x + L_y)/\sqrt{\sigma_x^2 + \sigma_y^2}
\]

Where \( L_x \) and \( L_y \) are the length of the longest stroke in X direction and Y direction respectively. If \(|C_i - C_j| \leq 7\), then the character \( i \) and \( j \) are considered in the same complexity degree range, in which the retrieval function search.

The number of sampled points dominates the computing time of each shape-matching process. So dimensionality reduction is needed. The Number of Connected Points (NCP) is defined as the number of contour points existed in its 8-neighbourhood as introduced in (Lau, Yuen, & Tang, 2002). If \( NCP \geq 2 \) and three consecutive points are in the same direction, then they are considered as parts of the same stroke. The middle point was taken out, and the reminder two points keep the structure information.

In order to measure the efficiency of these three speeding up approaches, we compare our proposed shape corresponding approach after implemented speeding approaches with projecting approach (Manmatha et al., 1996) and Earth movers’ distance approach (Cohen & Guibas, 1999), as shown in Table 1. All of the tests are performed on a regular Intel(R)/256RAM personal computer. Compared with our early method, the database in this paper is enlarged from 336 individual calligraphy characters to 1,750 calligraphy characters, which are segmented from works of about 60 calligraphists living in different dynasties. Table 1 indicates that an average single calligraphy character matching takes about 640 ms (336 character, 3.6minutes) in our early method, while here only 52 ms (1,750 character, 1.52 minutes).

### Dynamic Writing Process Simulating

After an individual calligraphy character is displayed before the learners, together with corresponding information of where it comes from and who wrote it, the next service is to offer a
visualization of how such a calligraphy character was written step by step, which may help immersion learning.

In order to simulate the dynamic writing process, stroke extraction and stroke sequence estimation should be made. We use contour segments to extract strokes of calligraphy character as introduced in Lee, Wu, and Huang (1997). Strokes in a handwritten calligraphy character are often connected together (see Figure 5), and they are not necessarily corresponding to those of well printed characters. Yet it does not matter for extracting connected strokes, as long as the correct writing sequence is estimated. One assumption for estimating the order of the stroke sequence is that people always write a calligraphy character from left to right and from top to bottom (You-Shou & Xiao-Qing, 1992, p.14). The other assumption is that people always write a calligraphy character as fast and convenient as possible. So, if strokes are connected, total distance traveled in the writing process should be minimized as introduced in Lau et al. (2002), and when a cross corner is encountered, we choose to follow the most straightforward contour segment, which has the biggest angle—“people write it as convenient as possible.” Based on the aforementioned assumptions, we estimated the stroke sequence and develop a video to simulate how a calligraphist wrote a calligraphy character step by step, as show in Figure 5.

Figure 5. A calligraphy character example and the corresponding video simulating its writing process

Figure 6. Screen shot of a retrieval example
IMPLEMENTATION AND EVALUATION

In the experiment, approaches described previously are used and tested with the database and it consists of 1,750 individual calligraphy characters. These characters are segmented from a book named Chinese Calligraphy Collections, which consist of 12 volumes. Figure 6 shows a retrieval example.

If a learner interested in one style of the character, for example, the last one in the second row in Figure 6, then this particular calligraphy character can be clicked and a new page (see Figure 7) will pop up showing its original page with a minimum-bounding box marked out where it is, and also who wrote it. In Figure 7, when the name of the author is clicked, a portrait of the author accompanied by a brief resume will be shown. And if the individual character is clicked, a plug-in video will show up playing the estimated and visualized writing process.

Recall and precision are the basic measures used to quantitatively speculate the effectiveness of retrieval approach. Recall is the ratio of the number of relevant records retrieved to the total number of relevant records in the database, and precision is the ratio of the number of relevant records retrieved to the total number of irrelevant and relevant records retrieved. Here, we use average recall and average precision. They are defined as:

$$\text{recall}_{\text{average}} = \frac{1}{C} \sum_{i=1}^{C} \frac{\text{recall}}{n_i}$$

$$\text{precision}_{\text{average}} = \frac{1}{C} \sum_{i=1}^{C} \frac{\text{precision}}{n_i}$$

Where $C$ is the number of character, $n_i$ is the number of total styles of the same character $i$. We randomly chose 20 characters (each has more than six different styles) from the database, that is to say when $C=20$ and $n_i \geq 6$, then Figure 8 can be
Chinese Brush Calligraphy Character Retrieval and Learning

Figure 8. Comparison of three approaches of their average recall and precision ratio on 20 characters with each having more than six different styles

drawn. It is obvious that the recall ratio is higher than traditional content-based image retrieval.

CONCLUSION AND FUTURE WORK

We proposed a new system to help people who are interested in calligraphy to enjoy the beauty of different styles of the same Chinese character, to learn its detailed corresponding information of a particular style of a character (for example, who wrote it, and in what environment), and also to learn how it can be generated step by step. While the experiment is somewhat preliminary, it works efficiently and clearly demonstrates the applicability of our system to enhance Web-based Chinese calligraphic learning.

Our further development of this system will include speeding up the retrieval, developing a bigger and robust database, and offering more convenient ways for query submitting, such as scratching a query or typing in a text query by keyboard.

ACKNOWLEDGMENT

This research is supported by the China-U.S. Million Book Digital Library Project (see http://www.cadal.net), and the National Natural Science Foundation of China (No. 60272031).

REFERENCES


Lau, K. K., Yuen, P. C., & Tang, Y. Y. (2002). Stroke extraction and stroke sequence estimation


