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.