ART LIBRARIES SOCIETY OF NORTH AMERICA

35th Annual Conference
Sheraton Colony Square, Atlanta, Georgia, April 26-May 1

Session 18: New Voices in the Profession 2: Papers on Arts and Visual Resources Librarianship
Sunday, April 29, 2007 11:00am-12:30pm

Presentation notes for Developing a Shape-and-Composition CBIR Thesaurus for the Traditional Chinese Landscape

Speaker: Tang Li, University of Maryland

This paper is developed from a term paper for a seminar on access to images in the Spring 2006. At that class I was very interested in the content-based image retrieval and wanted to do my final project on this topic. Before the library school I had been studying in the history of Chinese art, especially Chinese painting for several years, so that's why I Chinese painting as a case study for CBIR. I was able to refresh my knowledge of Chinese art and also learn new things while working on the project.

Content-Based Image Retrieval (CBIR)

Content-based image retrieval (CBIR) was developed in the early 1990s. It uses automatic extraction of lower level image features, such as texture, color, shape, and structure, to catalog and retrieve images. CBIR classifies and searches images according to similarities of automatically extracted visual features. A number of representative generic CBIR systems have been developed in the past ten years, including, for example, IBM's Query by Image Content (QBIC) search engine. The State Hermitage Museum in St.

Petersburg, Russia (http://hermitagemuseum.org/html_En/index.html, accessed 11/3/2006), was an early adopter of QBIC for its digital image collections. The QBIC system allows users to query image databases based on color percentages and layout, textures and shape occurring in the images. Recent study of end users' image query behaviors by Hsin-Liang Chen has suggested practicability of CBIR-based tools in the field of art history. Particularly, the CBIR approach is potentially useful to investigate relationships among paintings based on objective visual facts.

CBIR and the Chinese Landscape

Previous research by Danqing Zhang, Binh Pham and Yuefeng Li in applying CBIR to Chinese paintings has suggested that the two elemental issues in CBIR, feature extraction and similarity measures, tend to be domain-specific because image content is identified not only by data but also by 'context' (i.e. domain-specific knowledge). They have demonstrated that a case study of a controlled and well-defined domain of images is useful to validate and further enhance existing CBIR techniques. Therefore, this paper will focus on application of CBIR to traditional Chinese paintings, specifically Chinese landscapes.

In terms of subject matter, Chinese painting can be categorized into three sets: landscape, bird-and-flower, and figure paintings. Of them Chinese landscape is regarded as the most important set because it was a central topic for literati painters who produced most of extant Chinese paintings.

As stated by Zhang, Pham and Li, CBIR is potentially an excellent and feasible retrieval mechanism for the Chinese landscape in terms of its visual features. Chinese landscapes promote capturing the essence rather than the real shape of nature to express the painters' ideas and feelings. As a result, painters use relatively simple forms and textures, only a few colors, and a small number of brush strokes. Moreover, a limited number of object types are depicted in the Chinese landscape, commonly mountains, rocks, water, clouds, woods and trees, and sometimes dwellings, pavilions, bridges, figures, and animals. For each object type, there are usually a small number of varieties. Furthermore, compositions of Chinese landscapes follow certain perspectives and models.

So this paper develops a shape-and-composition CBIR thesaurus for the Chinese landscape dated from the Song to Qing periods (960-1911). It is designed to facilitate extracting and indexing image content data for effective retrieval performance. Approximately 1000 paintings were studied to extract the features to be included in the thesaurus. One thing I'd like to mention here is that the thesaurus is still in its theoretical stage. It has not been applied into practice.

The thesaurus consists of two parts: Part 1 which focuses on shapes and lines and Part 2 which provides composition templates. Both of the two parts are arranged in an alphabetic and hierarchical order. The thesaurus emphasizes discrimination among object types to improve retrieval of relevant images, so it adopts not only shapes but also lines and shape combinations. Shapes can be divided into basic shapes such as circle,

rectangle, and triangle; and special shapes such as irregular polygon, cloud icon, and U-shape. Special shapes are developed for those object types that either are unique to Chinese arts and culture or are peculiarly shaped so that they cannot be abstracted easily into basic shapes. Lines include straight lines, arc lines and wavy lines. Composition templates characterize visual layout structures that were commonly adopted in the Chinese landscape.

Shape-and-Composition CBIR Thesaurus for the Chinese Landscape (960-1911)

Part I is the core of the thesaurus. It includes object types and their varieties that frequently appear in the Chinese landscape. Object types are divided into two main sets: A. *primary elements* and B. *secondary elements*.

Part I Shapes and Lines

A. *Primary Elements* set refers to object types that are always found in the Chinese landscape. It consists of five categories, namely *clouds* (A1)*, mountains* (A2)*, plants* (A3)*, rocks* (A4)*,* and *water* (A5).

A1. Clouds: Clouds in the Chinese landscape are very difficult to abstract into basic shapes because they were either delineated randomly out of artists' imagination or represented by leaving irregular empty spaces of various sizes. Therefore, I use a universal cloud icon to stand for this category, but the one for *with contours* (A1.1) in solid lines while the one for *without contours* (A1.2) in dashed lines to demonstrate their difference.

A2. Mountains refers to two main subcategories: *distant mountains* (A2.1) which are usually painted in ink washes and a few brushstrokes and *mountains in a close view* (A2.2) which are often meticulously delineated

with skylines and texture. Similar to the way I differentiate clouds *with contours* (A1.1) from clouds *without contours* (A1.2), I use shapes in solid outlines to stand for *mountains in a close view* (A2.2) and shapes in dashed lines to represent *distant mountains* (A2.1).

Whether it is depicted in a close view or in the far distance, a mountain in the Chinese landscape is usually composed of peaks and crags. Accordingly, each subcategory of the mountains category is divided into two main types: *crags* interpreted as right trapezoids here (A2.1.1/A2.2.1) and *peaks* interpreted as triangles here (A2.1.2/A2.2.2). These two main types are further classified hierarchically into their varieties on the basis of their spatial relationships. So here is an example of horizontally isolated distant peaks. And here is an example of isolated crags in a close view.

A3. The Plants category is pretty self-explanatory here. One thing I want to draw your attention is the way I treat reeds and grass. Since *reeds and grass* in the Chinese landscape are usually depicted with piles of simple brushstrokes, a group of five straight vertical lines are adopted to symbolize them, but lines for *reeds* are longer than those for *grass* to exemplify their difference in nature.

A4. Rocks category: I want to draw your attention to the special *linglong*-shaped rocks here because they are unique to the Chinese culture. They were once very popular in ancient Chinese gardens. Their grotesque forms were mostly carved by Chinese artists or artisans. The icon here serves as a generalized one because *linglong*-shaped rocks in the Chinese landscape

were usually rendered out of artists' imagination and thus they have no standard forms.

A5. For the Water category, I want to point out that the subcategory *non-waterfalls* (A5.1) include all kinds of bodies of water such as rivers, springs, lakes, and seas because they are essentially portrayed in a same way in the Chinese landscape.

B. *Secondary Elements* refer to object types that sometimes appear in a Chinese landscape. It contains four categories, namely *animals* (B1)*, architecture* (B2)*, persons* (B3), and *transportation facilities* (B4).

B1. The Animals category is composed of birds, classified by their various motions, and mammals, classified according to popularity of their appearance.

B2. The Architecture category here is pretty straightforward**.** Here I show you a picture of the Yellow Crane tower in my hometown Wuhan, China. Evidently, the tower falls into the multi-storied subcategory. It is one of the four prestigious towers in China and thus occurred in many Chinese landscapes.

B3. In the person category, a single Person is interpreted as composition of two basic shapes, namely a circle standing for the head, and a rectangle or triangle for the body. The rectangle is used when the body is stretching (vertically or horizontally). Otherwise different types of triangle are used to represent the body.

B4. For the Transportation Facilities category, I want to draw your attention to the dragon boat subcategory. It is unique to Chinese culture. The

dragon boat is a very long and narrow human powered boat, rigged with decorative Chinese dragon heads and tails. Dragon boat races are traditionally held at the Duanwu festival also known as the Dragon Boat festival on the fifth day of May in the Chinese calendar. The races are intended to commemorate the death of an ancient patriot poet Qu Yuan (338-278 BC).

Part II of the thesaurus consists of 14 composition templates that are commonly applied to the Chinese landscape. Based on overall layout of objects, the composition templates are classified into two main categories, namely *fully filled* (C1) and *one part* (C2). Abstract rectangles are used to symbolize objects in the painting and illustrate their spatial relationships.

For the 'fully filled' category, I'd like to say a little bit about the two types *fragmented* (C1.1.2) and *vertically superimposed* (C1.1.3). 'Fragmented' means objects are spread all over the painting without evident spatial relationships between one another. 'Vertically superimposed' means objects are usually lofty and thus superimposed in the vertical dimension. Here I show you an example of landscape with fragmented composition on the left. On the right is an example in vertically superimposed composition.

The 'one part' (C2) category is pretty self-explanatory here. The example I show you here is in the right composition.

Testing the Thesaurus

The thesaurus was tested by applying it to index four randomly selected Chinese landscape which are dated to different time periods and had not been used previously in developing the thesaurus. One example is shown

here. No attempt was made to test retrieval, which would require access to an entire collection indexed with the thesaurus.

Example 1:

Yuan. Wu, Zhen. *Fisherman*. 1341. Source: Freer Gallery of Art, Washington DC. http://www.asia.si.edu/collections/singleObject.cfm?ObjectId=10144 [accessed 4/26/07]

This is a Yuan landscape painted by Wu Zhen from the Freer Gallery of Art. We can see the painting consists of both *primary elements* and *secondary elements*, specifically distant peaks (isolated and adjacent), rocks, river, fishing boat and fisherman. Here is a slide with imposed diagrams of major object types and relationships from the thesaurus. The composition of this Song landscape corresponds to the fragmented type.

Summary

1. This paper develops a shape-and-composition CBIR thesaurus for the traditional Chinese landscape dated from the Song to Qing periods (960-1911). Development of the thesaurus is based on visual features of Chinese landscapes, including less complexity of forms and textures, a few colors, and a certain number of object types, varieties and composition structures.
2. This thesaurus emphasizes discrimination among object types to improve recall of relevant images. It adopts not only basic shapes (such as circle, triangle, and rectangle) but also lines and shape combinations. Furthermore, special shapes are developed to represent object types with unique forms.

3. Results from the thesaurus testing demonstrate that shape features and composition templates are sufficient to represent the content of the paintings. Therefore, this shape-and-composition CBIR thesaurus has the potential to be a feasible and effective means to index and retrieve the Chinese landscape. It may be applied to those museums and libraries that have large collections of Chinese landscapes, such as the Freer Gallery of Art and the Metropolitan Museum of Art.

Further research includes processing additional paintings to test and ensure the completeness of the thesaurus with regard to shapes and composition abstractions in Chinese landscape paintings. In addition, it may be necessary to develop some computing algorithm and techniques to automatically process the feature extraction. Furthermore, to make the thesaurus more solid and effective for a CBIR system, the thesaurus needs to incorporate color and texture as well. Finally, the thesaurus should be tested to determine its usefulness for adequately representing users' information needs and its retrieval effectiveness in differentiating among various elements of Chinese landscape p