

Ph. D. Thesis

Content-Based Image Retrieval of Color Image

Jung-Dong Jang

Department of Electronic Engineering

Graduate School

Yeungnam University

(Advised by Processor Taihong Yi)

Abstract

Consider that everyday we view various forms of visual media, such as photographs, graphics, animations and videos. These media are provided increasingly in digital form. For example, the WorldWideWeb (WWW) is one such source for viewing hundreds of gigabytes of digital visual information. While there is great accessibility to large stores of digital imagery, new systems need to be developed to manage and search better for the visual information. To cope with this problem, there have been two major approaches to study image retrieval from somewhat different angles.

The one is to annotate each image manually with text. This approach has however insuperable problems that creating annotations by hand can become hopeless because databases can be very large. Manual indexing is that the content of an image is a highly subjective concept, that is, different people may perceive it differently.

Another approach is to index images directly based on the visual image content.

Content-based image retrieval was proposed in the early 90's and it has been under intensive research from then on. In general, visual feature extraction is fundamental bases in the content-based image retrieval paradigm. Visual features are usually divided into four classes depending on what kinds of properties they describe. The classes are color, texture, shape, and spatial relationship.

Automatic feature extraction would be desired in large-scale image retrieval system. Besides, the manual annotation of text is likely to be error prone with incompatible or inconsistent text description. In addition, some of the image databases are not confined to specific domains. In these situations, content-based retrieval approaches are preferred.

However, the use of color information has been used widely in content-based retrieval systems. It is clear that color plays a very important role in defining an image. But color alone is insufficient to characterize an image. In order to overcome these problems, we discuss a technique of considering color-spatial information with spatial knowledge to obtain an overall impression of the image.

In this paper, content-based image retrieval considering color spatial information is proposed. Our goal is to improve the above mentioned shortcomings through following process. In the first, Sequential-Clustering schemes, which are quite straightforward and fast method, are introduced to cluster the homogeneous property of any images. Then, clustered color regions are used as the basis for computing distance histogram with relevant color-spatial information. Finally, Color mean of each regions and distance histogram computed from the center of each clustered color regions are used as feature vectors.

Our proposed method needs only 18 features and required storage space is smaller than with conventional methods. However, our experimental results showed not only that retrieval efficiency improved 8.5 percent over conventional methods but also much computation and data storage was saved. In addition, this method showed excellent Precision vs. Recall evaluations and was judged better by human perception.