Chapter 6
Clustering Techniques for Content-Based Feature Extraction From Image

Madan U. Kharat
MET Institute of Engineering, India

Ranjana P. Dahake
MET Institute of Engineering, India

Kalpana V. Metre
MET Institute of Engineering, India

ABSTRACT
Image retrieval is gaining significant attention in areas such as surveillance, access control, etc. The content-based feature extraction plays a very crucial role in image retrieval. For the characterization of a specific image, mainly three features (i.e., color, texture, and shape) are used. Multimedia can store text, image, audio, and video which can be processed and retrieved. The various techniques are used for image retrieval such as textual annotations, content-based image retrieval in many application areas like medical imaging, satellite imaging, etc. However, most of these techniques were designed for specific domains and universally accepted method is yet to be designed; hence, CBIR is a field of active research. Similar output images indicate efficiency of search and retrieval process. In this chapter, the authors have discussed various image feature extraction techniques and clustering approaches for content-based feature extraction from image and specifically focused on color based CBIR techniques.

INTRODUCTION
In past few decades limitations in the metadata-based system resulted into more research in content-based image retrieval as the usage of digital images increased in huge amount. In CBIR the two major processes are feature extraction and similarity measurement. CBIR clearly defines use of visual content of image like color, texture, shape etc. The contents of any image are explored by the visual feature.

DOI: 10.4018/978-1-5225-5775-3.ch006
These visual features/contents are retrieved from the images and are described by multidimensional vectors. The distance/similarity between the feature vectors of data images is calculated and relevant images are retrieved. Image feature representation, extraction and indexing are the fundamental bases of content-based image retrieval. Mainly three features are used to characterize a specific image: color, texture and shape. Wide ranges of algorithms were developed to handle image retrieval based on these factors. Feature extraction is one of the essential requirements in image retrieval process. The color feature is one of the most popular visual features.

The simplest approach for content-based image retrieval is pixel to pixel comparison and the color difference between images is computed. The images need to resize to fixed dimension before comparison. Generally, color feature is represented using the color histogram. The approach based on global color histogram depicting the number of pixels belonging to each of specific set of colors. This approach does not consider the similarity of colors in neighborhood pixels. In the local histogram-based method, an image is partitioned into different rectangular cells and histograms for each cell are computed separately and the color difference between images is computed. A color signature-based method partitions an image and the color signature is calculated for each partition. The difference between the color signatures of the images decides the similarity of the images (Parekh, 2006).

To increase the accuracy of CBIR system, focus of research is shifting from designing sophisticated low-level feature extraction algorithms to reduce the ‘semantic gap’ between the visual features and the richness of human semantics.

The clustering groups the set of similar objects. Various clustering approaches are used to group, identify and reduce the content-based image features. Accuracy of clustering is based upon the algorithm, which is selected, and the use of parameters like distance function, density threshold and required number of clusters. General cluster models which are used includes: Connectivity models, Centroid models, Distribution models, Density models, and Graph-based models. In color clustering large patches of color in comparison with small scattered spots leads to clustering-based feature extraction. Two images appear to be similar if they contain large similar color patches roughly at the similar locations. This cluster-based approach isolates single color clusters in an image instead of image partitioning. A color space is mathematical representation of a set of color. There are several color spaces for specific area of applications. Different Color spaces like RGB, CMYK, CIE Lab, HSV emphasis three color variants that characterize color: hue, saturation, intensity. Lighting and shading are related to intensity and variation of this factor can affect the performance. Clustering of other attributes with color for feature extraction is essential for content-based image retrieval.

Clustering of color pixels in the image plays an important role in segmentation process. Various clusters can be created considering the color difference between them. The concept of connected component is utilized to form a specific color cluster. Decision can be taken based on dominant cluster and overlapping of them (Solomon, & Breckon, 2011). Feature extraction based on clustering provides more efficiency without any need to have the prior knowledge regarding the colors present in application. This approach can be combined with knowledge-based interpretation technique resulting in accuracy of content-based image retrieval system. The other features of image like texture as well as shape are also important for retrieval of similar query images from database. Various hybrid techniques of feature extraction and clustering are also required to get the accurate retrieval of images.

Further in various sections we have discussed about multimedia database, Image low and high-level features, different techniques of feature extraction, selection and clustering methods for image in CBIR system, last section provides applications of CBIR in various field.
20 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the product's webpage:
www.igi-global.com/chapter/clustering-techniques-for-content-based-feature-extraction-from-image/207230?camid=4v1

www.igi-global.com/e-resources/library-recommendation/?id=105

Related Content

Predicting Key Recognition Difficulty in Music Using Statistical Learning Techniques
Ching-Hua Chuan and Aleksey Charapko (2014). International Journal of Multimedia Data Engineering and Management (pp. 54-69).
www.igi-global.com/article/predicting-key-recognition-difficulty-in-music-using-statistical-learning-techniques/113307?camid=4v1a

Optimizing Quality-of-Experience for HTTP-based Adaptive Video Streaming: An SDN-based Approach
Sangeeta Ramakrishnan, Xiaoqing Zhu, Frank Chan, Kashyap Kodanda Ram Kambhatla, Zheng Lu, Cindy Chan and Bhanu Krishnamurthy (2016). International Journal of Multimedia Data Engineering and Management (pp. 22-44).
www.igi-global.com/article/optimizing-quality-of-experience-for-http-based-adaptive-video-streaming/170570?camid=4v1a

Interactive Television Research Opportunities
www.igi-global.com/chapter/interactive-television-research-opportunities/17477?camid=4v1a

The Decision Making Process of Integrating Wireless Technology into Organizations
www.igi-global.com/chapter/decision-making-process-integrating-wireless/17243?camid=4v1a