ABSTRACT

Nowadays, Content-based search and retrieval of video data have become a challenging and important issue. Several types of audio and visual informations are included in the video, which are difficult to extract, combine or trade-off in common video information retrieval. Most of the video retrieval techniques are inefficient because they enable only a specific kind of feature set. In this thesis enhanced feature extraction techniques are proposed for content-based video retrieval system, based on the integration of several visual features. The proposed system analyzes all frames within a shot to extract more visual features for shot representation, which integrates multiple visual features to capture the retrieval information more accurately.

The proposed content-based video retrieval system has been worked with three different kinds of integrated visual feature sets. The first kind of visual feature set includes motion, edge and color features, which are extracted using Fast Fourier Transform (FFT), edge histogram and color histogram, respectively. The spatial distribution of edges is captured by the edge histogram with the help of sobel operators. Color histogram is the most extensively used method, which is recognized by using the HSV (Hue, Saturation, and Value) color space. In order to minimize the dimensionality of
the data, the features to be extracted are preferably compact and discriminative. In order to symbolize the temporal information of videos, motion is considered to be the key feature. Motion feature improves the performance of content-based video retrieval when used with the other low-level features of edge histogram and color histogram.

The second kind of visual feature set includes four different kinds of video features namely, motion, texture, color and contour for every video shot. For the content-based video retrieval, efficient motion feature extraction is a significant step. In the proposed system motion feature extraction has been done using morphological and gradient operation. Texture feature is extracted from Gaussian Mixture Model (GMM). Color histogram, more robust to changes due to scaling, orientation, perspective, and occlusion of images, is recognized after the segmentation process using anisotropic diffusion. As a final feature, contour of all the objects in the shots are extracted by using the sparse field method. Finally, all these features of a video, motion, texture, color and contour are combined as a feature set and stored in the feature library.

The third kind of visual feature set includes object features, location of the detected objects and object movement direction detection features which are extracted using fuzzy c-means (FCM) clustering, spatial mask and binary representation, respectively. The main advantages of these enhanced
visual feature set representations can increase the retrieval efficiency. Particularly, the proposed content-based video retrieval system illustrates that a variety of feature sets with a combination of exact similarity measure can enhance the retrieval efficiency. The proposed system retrieves similar video shots for a query video shot from a collective set of videos.

The performance of the proposed video retrieval algorithms has been improved in comparison with the existing techniques. Kullback-Leibler Distance (KLD) and Latent Semantic Indexing (LSI) are used to measure the similarity between the query and database video shots of video retrieval system.