Abstract

Content based retrieval of digital video is a key component of any Interactive Multimedia System. This requires the ability to retrieve, classify and analyze video sequences based on the properties of objects and their actions. Hypervideo takes interactivity in digital video to its peak. It adds excitement and life to otherwise passive video viewing. In Hypervideo the user marks an object of interest ‘hot’ by clicking on any video object and then, one can track the ‘hot’ object as it moves. By hyperlinking such ‘hot’ objects to secondary videos one can add interactivity to the video.

These two applications bring out the importance of processing digital video for object recovery and feature extraction.

Digital Video is becoming increasingly available in MPEG compressed format. While the compression reduces the storage and network costs, it increases the cost of processing since the video data has to be decompressed before it can be processed. The overhead of decompression is enormous. One way to circumvent this problem is to process the video in compressed form. Motivated by the above considerations, in this Thesis, we consider the following operations in compressed domain in MPEG-1 video streams:

1. Object Segmentation
2. Object Tracking
3. Object Skeletonization

While the MPEG video compression standard is optimized to achieve good compression, it is a real challenge to perform robust video processing operations in compressed domain.

Object Segmentation is done making use of the encoded motion information. MPEG video streams have motion vectors coded for every block of 16x16 pixels for the predictively coded video frames. We first extract the block motion vectors from the coded video stream and do a smoothing in order to eliminate isolated, noisy motion vectors. Then a dense motion field (one in which every pixel has a motion vector) is
obtained by means of interpolation. We have used both a linear interpolation technique and a fast surface interpolation technique using wavelets as preconditioning transform (later proposed by Pentland).

After the interpolation, segmentation is accomplished by clustering of the pixels making use simple object motion heuristics. We have presented three clustering algorithms for object segmentation and discussed their working under different conditions of real world motion.

The previous technique segments the object of interest in a frame. As the object moves across frames, it is necessary to track it. We have presented two object tracking algorithms for tracking the object through the frame sequence. One of these tracks the nearest motion centroid, while the other uses the MPEG motion vectors for tracking.

Object Skeletonization involves compact representation of an object by its skeleton. We have used two approaches for object skeletonization, namely Medial Axis Transformation and Object Thinning. Both these methods are applied after the segmentation of the object. The Medial Axis Transformation uses the distance transform and the Thinning Algorithm iteratively deletes the unwanted edge points of the object to obtain the object skeleton.

We have demonstrated the above three operations in the form of a Hyper Video in which, when a user clicks on an object, a separate video showing object skeletons is played. This has applications in sports and choreography videos.