Sharif University of Technology
Computer Engineering Department

PhD Thesis

Object Extraction and Tracking in Video Processing

By
Mohammad Khansari

Supervisor
Dr. Hamid R. Rabiee

DISSERTATION
Submitted to the department of Computer Engineering of
Sharif University of Technology
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

March 2008

Abstract

During the last two decades, object-based video processing has been a universal challenge to researchers both in academia and industry. One the main problem in this area is the proper use of pre-processing algorithms for object segmentation/extraction. In many applications, the exact boundary pixels of objects must be extracted and tracked throughout a video clip. The extracted objects can be used in many emerging multimedia systems such as MPEG-4 for object based coding or MPEG-7 for creation of the required descriptors.

In this thesis, a novel feature based algorithm in the wavelet domain for extraction of exact boundary points in image sequences is presented. After surveying existing object extraction and temporal tracking algorithms, a novel transform domain algorithm is presented. In the proposed algorithm, a subset of the exact object's boundary points is first specified by the user. The amplitude of coefficients in the best basis tree expansion of the undecimated wavelet packet transform (UWPT) is used to create a Feature Vector (FV) corresponding to each selected boundary pixels, at the reference frame. This FV can be tracked temporally through the image sequence. Full search for the best match is then performed by using the FVs within an adaptive search window. The algorithm can be used to extract the contours of both rigid and non-rigid objects in image sequences. Moreover, we have defined an objective performance evaluation scheme for object's shape tracking. Experimental results on the new algorithm and objective and subjective comparisons with block matching and particle filters show good performance in presence of object translation, small rotation, scaling and various type noise (Gaussian white noise and quantization distortion).

Moreover, the introduced UWPT pixel-wise FVs is used for object tracking in crowded video scenes. In this case, the algorithm is again initialized by the user through specifying a region around the object of interest at the reference frame. Then, coefficients of the UWPT of the region are used to construct a FV for every pixel in that region. Optimal search for the best match is then performed by using the generated FVs inside an adaptive search window. Adaptation of the search window is achieved by inter-frame texture analysis to find the direction and speed of the object motion. This temporal texture analysis also assists in tracking of the object under partial or short-term full occlusion. Moreover, the tracking algorithm is robust to Gaussian and quantization noise processes. Experimental results show that the proposed algorithm has good performance for object tracking in crowded scenes on stairs in airports or train stations in the presence of object translation, rotation, small scaling and occlusion compared to color-histogram based algorithms, Discrete Wavelet Transform and Particle filters.

Keywords: *Video object extraction, MPEG-4, Undecimated Wavelet Packet Transform, Video object segmentation, Object tracking, inter-frame texture analysis, occlusion handling*