AUTOMATED VIDEO ANALYSIS USING LOW-LEVEL MOTION FEATURE TRACKS

By

Anil Cheriyadat

An Abstract of a Thesis Submitted to the Graduate

Faculty of Rensselaer Polytechnic Institute

in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: ELECTRICAL ENGINEERING

The original of the complete thesis is on file
in the Rensselaer Polytechnic Institute Library

Examining Committee:

Prof. Richard J. Radke, Thesis Adviser

Prof. Kim L. Boyer, Member

Prof. Qiang Ji, Member

Prof. Brett R. Fajen, Member

Rensselaer Polytechnic Institute
Troy, New York

April 2009
(For Graduation May 2009)

ABSTRACT

Automated video understanding of dynamic scenes is a challenging problem, but is the key for the future smart vision systems. This thesis develops robust techniques for automated video understanding based on low-level motion primitives- i.e, automatically tracked feature points. The techniques proposed in this thesis can successfully address crowded environments and other operating challenges such as variations in ambient light or movements of sensor platforms. Our proposed approach take motivation from the innate biological capabilities for perceptual organisation demonstrated by the human vision system.

This thesis has four main techniques for automated video understanding. First, we propose an algorithm for detecting dominant motions in dense crowds based on low-level motion tracks. Accurately tracking individual objects in such scenes is difficult due to inter- and intra-object occlusions that cannot be easily resolved. We overcome the limitations posed by dense crowded scenes by automatically detecting and clustering low-level feature point tracks to identify dominant motion patterns in crowd movements. Second, we address the problem of motion segmentation to automatically identify motions generated by independent objects, articulated parts of the same object, or the camera itself using a new technique that can segment different motions from partial, incomplete and noisy motion data. We show that the proposed algorithm can be directly applied to video sequences generated by a variety of sources, from hand-held to airborne cameras. Third, we introduce a method for independent motion detection in dynamic scenes containing many similar moving objects. Our object detection system takes locations of tracked low-level feature points as input, and produces a set of independent coherent motion regions representing individual objects. The proposed approach is a trade-off between algorithms that require object specific models and algorithms that do not require any object specific information. We show that by applying spatial constraints to the statistical clustering, we can efficiently detect independently moving objects belonging to a single class. Finally, we extend this technique by combining low-level motion infor-

mation with high-level spatial information to detect and track objects belonging to multiple classes in challenging video sequences. All of our algorithms are tested on several real video sequences, and results are compared with other recently published related work.