A UNIFIED PROBABILISTIC FRAMEWORK FOR SPONTANEOUS FACIAL ACTIVITY MODELING AND UNDERSTANDING

By

Yan Tong

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: Department of Electrical, Computer, and Systems Engineering

Examining Committee:

Dr. Qiang Ji, Thesis Adviser
Dr. Katherine Isbister, Member
Dr. George Nagy, Member
Dr. Badrinath Roysam, Member
Dr. Birsen Yazici, Member

Rensselaer Polytechnic Institute
Troy, New York

September 2007
(For Graduation December 2007)

ABSTRACT

Facial activities are the most natural and powerful means of human communication. Facial activity is characterized by rigid head movement, facial muscular movements, and their interactions. Automatic inference of 3D facial activities from 2D images captured in real time has applications in a wide range of areas. However, developing such a system encounters several challenges. As a result of these challenges, understanding facial activity has been limited to recognizing rigid and nonrigid facial motion separately and statically, often ignoring their semantic and dynamic interactions.

This research aims to explicitly exploit the prior knowledge about facial activities and to systematically combine the prior knowledge with image measurements in order to achieve accurate, robust, and consistent understanding of facial activity. Robust computer vision methods are introduced to obtain measurements of both rigid head movement and nonrigid facial muscular movements. A hierarchical multistate pose-dependent approach is developed for facial feature detection and tracking in order to extract rigid and nonrigid facial motions accurately. A dynamic Bayesian network (DBN) model is built to capture the spatio-temporal relationships among different facial muscular movements. The DBN model is then expanded to a unified probabilistic framework to simultaneously and coherently represent rigid head movement, nonrigid facial muscular movements, their interactions, and the uncertainties in their image observations. This framework not only captures the temporal evolution of rigid and nonrigid facial motions, but also characterizes the dynamic dependencies among nonrigid facial muscular movements, which are especially important for interpreting spontaneous facial activities. Advanced learning methods are proposed to construct the framework systematically from both subjective knowledge and training data. Finally, facial activity recognition is accomplished through probabilistic inference by systemically integrating the visual measurements with the facial activity model.

The proposed framework is validated on posed facial expression databases and,

more importantly, on spontaneous facial expression databases, which are created under realistic environment including illumination variation, face pose variation, and occlusion. The experiments show that the proposed system yields significant improvements in recognizing both rigid and nonrigid facial activities from posed facial expressions and spontaneous facial expressions. The performance improvement is especially significant for spontaneous facial activity recognition.