ABSTRACT

Recently, human-robot symbiotic systems have been studied extensively due to the increasing demand of welfare service for the aged and handicapped under the situation of decreasing of the younger generation. In the future, it will be difficult to provide help to the aged and disable persons such as taking care, nursing, informing important information, recreation, etc., by trained human. To build human-robot symbiotic society, where robots are able to support elderly or disable people, the robot should be capable of recognizing user, gesture, gaze, speech and text commands.

To use gestures is a natural way for human to interact with robot. However, gestures are varying among individuals or varying from instance to instance for a given individual. The hand shape and human skin-color are different for different persons. The gesture meanings are also different in different cultures. A significant issue in building visual gesture-based human-robot interactive system is to utilize user-specific knowledge for gesture interpretation and adapt to new gestures and users over time.

In this work, we propose a vision and knowledge-based gesture recognition system for human–robot interaction. In the proposed method, a frame-based knowledge model is defined for the person-centric gesture interpretation and human-robot interaction. In this knowledge model, necessary frames are defined for the known users, robots, poses, gestures and robot behaviors. The system first detects a human face using a combination of template-based and feature-invariant pattern matching approaches and identifies the user using the eigenface method. Then, using the skin-color information of the identified user three larger skin-like regions are segmented from the YIQ color spaces, and face and hand poses are classified by the subspace method. The system is capable of recognizing static gestures comprised of face and hand poses, and dynamic gestures of face or hand in motion. It is implemented using the frame-based Software Platform for Agent and Knowledge Management (SPAK). Known gestures are defined as frames in SPAK knowledge base using the combination of face and hand pose frames. If the required combination of the pose components is found then corresponding gesture frame will be activated. Dynamic gestures are recognized using the state transition diagram of face poses in a sequence of time steps. The person-centric interpretation of gesture is achieved using the frame-based approach and the interaction between humans and robots is determined using the hierarchical frame-based knowledge model.

In this work adaptation method for new users, hand poses, gestures and robot behaviors by combining computer vision and knowledge-based approaches is proposed. In this method one or more clusters are built to learn new user or new pose. New robot behavior can be learned according to the generalization of multiple occurrence of same gesture with minimum user interaction. An experimental human-robot interaction system comprised of an entertainment robot 'Aibo' and a

humanoid robot 'Robovie' has demonstrated the effectiveness of the proposed method. In this research, we also compare three pattern-matching approaches for face and hand poses classification: general PCA, person-specific subspace and pose-specific subspace methods. In the pose-specific subspace method, training images are grouped based on pose and eigenvectors for each pose are generated separately. In the person-specific subspace method, hand poses are grouped based on each person and for each person one PCA is used.

This dissertation makes three major contributions in the filed of human-robot symbiotic systems. First, the proposed system utilizes user specific knowledge for person-centric gesture interpretation that supports cultural variation of gesture meanings. The knowledge-based approach allows user to define or edit gestures and robot behaviors frames easily using SPAK knowledge Editor. The segmented skin regions are more noise free for known person because of using person-centric threshold values for the YIQ components. The performance of person-specific subspace method is better than the general PCA method in the same environment for face and hand pose classification.

Second, this thesis proposes a new user and hand pose adaptation method using a multi-cluster based incremental learning method, which supports different orientations of the images of a person's face and different orientations of same hand poses.

Third, this work combines several system components as an integrated system by means of a knowledge-based software platform SPAK. The system components include vision based face and gesture recognizer, knowledge manager, and various robot software and hardware components. We have demonstrated that it is simple and intuitive to develop a natural human-robot interactive system by integrating computer vision with a knowledge-based platform in the internet-based distributed environment.

The ability to understand hand gestures and person-centric meanings will improve the naturalness of human interaction with robot, and allow the user to communicate in complex tasks without using tedious sets of detailed instructions. We believe that integration of vision and knowledge-based approaches is a major step towards achieving intelligent and natural human-robot interaction.