Human Robotic Communication System for Health Care

1 Introduction

Social interaction allows knowledge sharing and cooperation. Many efforts are devoted to allow close interaction of robots with people from different perspectives. Research efforts seeking to provide more natural, human-centered communication and human interaction with computers have gained rapidly growing interests. An important perceptive of this research is the user interfaces, where the computer is endowed with perceptive capabilities that allow it to acquire the information about the users and the environment [1]. Human actions recognition and event detection have gained more interest of late[2], among video processing community because they find various applications in interactive systems ,automatic surveillance [3] and health monitoring systems [4], video indexing and retrieval, robot motion, human computer interaction and segmentation [5].

One of the important applications of human action recognition is gesture recognition which is helpful in human robotics interaction. This feature is useful to develop sensor-based communication-enabled autonomous cyber physical systems which can be used in the health care for elderly people and to interact with people having disability. The economic and societal potential of these systems is immensely greater than what has been realized. In order to develop the technology major investments are being made worldwide. An interface which provides the human robot communication system to communicate with deaf, deaf-blind and elderly people is proposed. The system will also be studied and different methodologies will be adopted and tested to create the additional sensory features (haptic) and speech recognition component which can take audio instructions to make the communication system natural (like humans interact with each other in daily life).

2 The Problem

People need support from others in order to thrive, but in our society many cannot get the support they need. As our population becomes more mobile and fractionated, and especially as it ages, individuals may find it more and more difficult to get the emotional, social and instrumental support they need to lead healthy, happy and productive lives. Computer agents or robots may be able to fill this growing void, by providing many of the psychological benefits to their users that are known to accrue from having trusting, caring relationships with other people. Deaf-blind people are not able to see and to hear any kind of sounds. Deaf blindness is a severe disability and deaf-blind people suffer from depression and isolation. The disability not only implies limitations in the communication capabilities, personal and learning autonomy but also causes difficulties in perception of the environment and interpersonal relations. Deaf-blind communicate with others or among each other using Tactile Sign Language. The idea is to do research in this area to provide a communication system which is able to make deaf-blind people to communicate among them. It could improve their access to other deaf-blind persons and decrease their level of isolation.

3 System

The approach proposed in artificial intelligence (e.g. supervised learning, evolutionary adaptation, engineering approach, etc.) by taking inspiration from developmental psychology and cognitive and neural sciences. The goal of the proposed methodology is the design and development of a remote communication system involving the robot. The system will provide the platform for the communication between two deaf-blind, deaf-blind and deaf and deaf-blind and hearing person having the knowledge of the sign language because the robot will convey the message in sign language. This system can also be used for health care especially for elderly people who can give oral or gestural instructions.

Figure 1 shows a diagram of the system architecture. As shown in the figure 1 on the left hand side would be the user which will perform the sign while on the right side would be the robot to receive instructions. The signer will perform the signs using the sign language which will be acquired and recognized using the acquisition hardware (i.e. Kinect). While on the receiver side a humanoid robot will be used, which will perform the signs on the basis of instructions provided to the system from the user which can send instructions from remote position. Instructions provided from remote interfaces are transmitted using worldwide web. The users can understand the signs on the basis of prior sign languages knowledge or using tactile sign language (in case of deaf blind). The signs features such as location, coarse shape, color and velocity are segmented and recognized through pattern recognition techniques. This information about the signs is stored in the database which allows the storage and retrieval of the needed information for each single sign. The data is sent through internet, received by a target user (receiver) that can touch an interface that actually reproduce the signing in tactual sign language form. This interface is based on anthropomorphic robotic hands and arms. Through this interface the user can understand the meaning of what the signer (sender) is signing.

In the next step the acquisition and the recognition of signs performed by the signer that wants to convey the message is targeted. The sign will be acquired by means of dedicated hardware. The use of both cyber gloves and Microsoft XBOX Kinect will be studied first in a separate and then in an integrated mode. The integrated approach with both devices will be studied. The cyber gloves for finger and palm acquisition, the Kinect for the relative position of the hands in the 3D space.

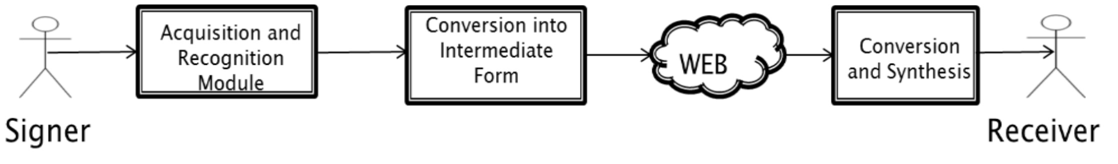

Figure 1 System Architecture

3.1 Acquisition and recognition

This step targets the acquisition and the recognition of a sign performed by the signer that wants to send the message. The sign will be acquired by means of dedicated hardware. The use of both Cyber gloves and Microsoft Kinect will be studied first in a separate and then in an integrated mode. The system would also be able to take the audio instructions as well. If the person who lost the hearing capability somehow on the later stage in the life but can communicate using audio instructions.

Automated sign language recognition from video has been studied for at least about twenty years. However, there still exist problems in the field of sign language, mainly because of the following factors; (a) extraction of sub-words automatically from a sign data stream; (b) How to tackle the phenomenon of movement epenthesis in a continuous gesturing flow; (c) Extracting signer position and invariant features; (d) identification the and classification of gestures from video or set of movements. Firstly, after detection of the hand as shown in figure 2, we attempt to model the image frames of each video independently and perform image recognition by segmenting movements, trying to decrease the effects of intersession variability that occur among signer's videos and finally to model the sequence of those input frames for each signer in order to classify the gesture. It suggests a way in order to model a signer-dependent model for accurate gesture recognition using state of the art techniques used in the field of speaker recognition. While gesture recognition has evolved considerably in the past years, modern approaches still suffer from increased errors in the presence of intersession variability.

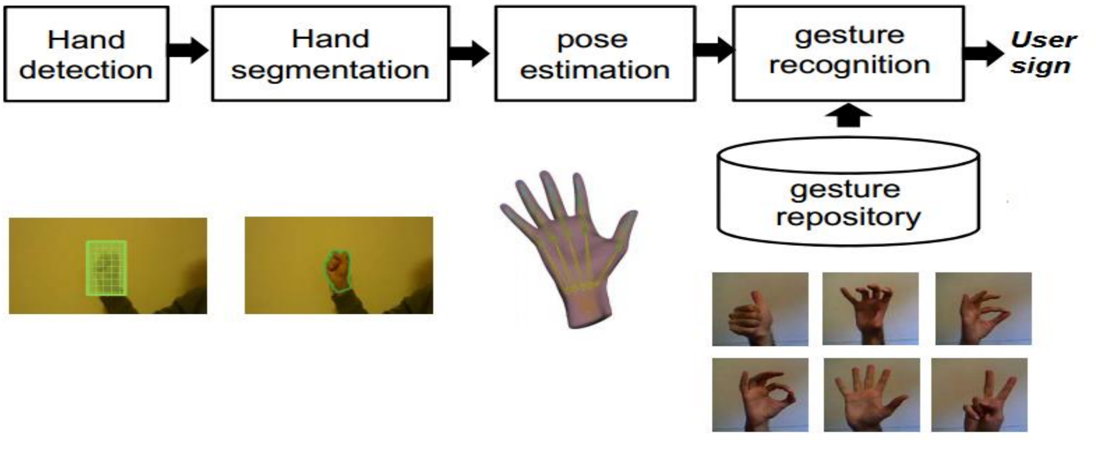

Figure 2 Gesture recognition

3.2 Conversion into an intermediate form

This step aims at converting the acquired data. The output of the conversion is the intermediate form named AEWLIS in the figure1. This is the formal representation of information collected by the robot. The AEWLIS files are sent through communication channel to receiving end where there is Synthesis Module which will perform synthesis of the information received. The detailed flow is shown in the figure 2. In the figure the acquisition mechanism is shown. Different steps involved in the acquisition are shown. First of all the system detect the hand and recognizes the segment. Then its estimates the pose and compares it with the existing corpus and provides the details of the sign as output.

3.3 Synthesis Module

This module converts intermediate into movement of the robotic hands. The data to be sent to the hands is generated from a subset of the intermediary information. This module is made of two sub modules: the intermediate form to robot parameters converter and the actuator that manages the transmission to the physical robotic device and the timing between the signs. The final device is composed by robotic hands equipped with two robotic arms. The availability of the arms is to be verified.

References

1. Zabulis, X., Baltzakis, H., Argyros, A.: Vision-based Hand Gesture Recognition for Human-Computer Interaction. In: Stefanides,C. (ed.) The Universal Access Handbook, Human Factors and Ergonomics., pp. 34.1–34.30. Lawrence Erlbaum Associates, Inc. (LEA), New Jersey (2009)
2. VijayaLakshmi, D., M. Spandana, and L. Ravi Kumar. "Human action Recognition from Homogenous Motions based on Artificial Intelligence."International Journal 3.12 (2013).
3. D. Ayers, M. Shah, Monitoring human behavior from video taken in an office environment, Image and Vision Computing 19 (12-1) (2001) 833 – 846.
4. Hornfeck, K., Zhang, Y., & Lee, K. (2012, March). Philos: a sociable robot for human robot interactions and wireless health monitoring. In Proceedings of the 27th Annual ACM Symposium on Applied Computing (pp. 293-294). ACM.
5. T. Starner, A. Pentland, Real-time American sign language recognition from video using hidden Markov models, MIT Media Lab, TR-375, 1995.