

Robust Hand Gesture Recognition using Kinect

Motivation :

Hand gesture is one of the most natural ways to give commands to the computer or communicate with a robot. Speech recognition being the most obvious method for interaction requires a lot of learning to be done on the part of the computer/robot as it is user specific.

Traditional hand gesture recognition methods used data from optical sensors to identify the posture of the hand. Demarcating the portion in the image which has the hand is quite a non trivial task. However, in most of HCI systems, it is a reasonable assumption to make that the hand is the foremost object in front of the camera. So we can exploit the depth information of the objects in the scene to effectively segment out the hand. The computational complexity of this process would be quite low if we have the depth information of the objects. Ever since the launching of low cost Microsoft's Kinect depth camera in November, 2010, lot of research is being done to design a robust hand gesture recognition algorithm using depth mapping.

Previous Work :

Detecting the gestures which can be differentiated just by comparing only the orientation and number of fingers in a hand gesture has already been done with high accuracy using a modification of Earth Mover's Distance [1]. Work has also been done on skeletal tracking[2]. Kinect can track 20 body joints in real time.

Innovation :

We plan to use both the algorithms [1] and [2] to come up a gesture recognition algorithm that uses both : the knowledge of shoulder joint, elbow joint and wrist joint [2] and the information conveyed by the fingers and their orientation[1] .

This will greatly increase the cardinality of the command set that can be understood by the computer/robot in real time and that too, just using hands.

Approach :

We will try to merge both the approaches in [1] and [2], so that our algorithm can not only understand the spatial information of the whole hand but also the sign pattern indicated by fingers.

References :

- [1]Z. Ren, J. Yuan, and Z. Zhang. Robust hand gesture recognition based on finger-earth mover's distance with a commodity depth camera. In Proc. of ACM MM, 2011.
- [2]Charles, J., Everingham, M.: Learning shape models for monocular human pose estimation from the Microsoft Xbox Kinect. In: Proc. IEEE Workshop on Consumer Depth Cameras for Computer Vision (2011)