Tracking of Fingertips and Centers of Palm using KINECT

Jagdish L. Raheja  
Machine Vision Lab  
Digital Systems Group  
CEERI/CSIR, Pilani  
Rajasthan, INDIA  
jagdish@ceeri.ernet.in

Ankit Chaudhary  
ankitc.bitspilani@gmail.com

Kunal Singal  
Punjab Engineering College  
Chandigarh, India  
kunal5596@gmail.com

Abstract— Hand Gesture is a popular way to interact or control machines and it has been implemented in many applications. The geometry of hand is such that it is hard to construct in virtual environment and control the joints but the functionality and DOF encourage Researchers to make a hand like instrument. This paper presents a novel method for fingertips detection and centers of palms detection distinctly for both hands using MS KINECT in 3D from the input image. KINECT facilitate us by providing the depth information of foreground objects. The hands were segmented using the depth vector and centers of palms were detected using distance transformation on inverse image. This result would be used to feed the input to the robotic hands to emulate human hands operation.

Keywords- Human Computer Interface, Image Processing, Image Segmentation, 3D Segmentation, Depth Vector, Natural Computing, KINECT

I. INTRODUCTION

Many researchers have proposed different methods for dynamic hand gesture recognition using fingertip detection dedicated to different applications. It begins a new era to see the existing problems in the area of Hand Gesture recognition, as the Sensors which can give depth information, are available. Microsoft KINECT® is one of the examples and it is able to detect individual finger motions. There are several limitations in existing approaches, see [3]. Garg [4] used 3D images in his method to recognize the hand gesture, but this process was complicated and inefficient. Processing time is very critical factor in real time applications. Yang [16] analyses the hand contour to select fingertip candidates, then finds peaks in their spatial distribution and checks local variance to locate fingertips. This method was not invariant to the orientation of the hand. There are other methods, which are using directionally Variant templates to detect fingertips [6][12]. Few other methods are dependent on specialized instruments and setup like the use of infrared camera [6], stereo camera [17], a fixed background [2][8] or use of markers on hand. Raheja [11] showed an efficient real time technique for natural hand with simple background, where orientation constraints was covered.

This paper describes a novel method of motion patterns recognition generated by both hands without any kind of sensor or marker. We call it Natural Computing as no sensor, marker or color is used to detect skin specifically. User would be able to do operations with natural hand. The moving fingertips detection in real time video needs a fast and robust implementation of method. Many fingertip detection methods are based on hand Segmentation technique because it decreases pixel area which is going to process, by selecting only areas of interest. However most hand segmentation methods can’t do a clearly hand segmentation under some conditions like fast hand motion, cluttered background, poor light condition. Poor hand segmentation performance usually invalidates fingertip detection methods. Researchers [6][7][13] used infrared camera to get a reliable segmentation.

Few researchers [2][5][8][9][14][15] in their work limit the degree of the background clutter, finger motion speed or light conditions to get a reliable segmentation. Some of fingertip detection methods can’t localize accurately multidirectional fingertips. Researchers [1][2][8][14] assumed that the hand is always pointing upward to get precise localization. In our previous approaches [18][20], we detected fingertips and calculated bended fingers angles in 2D for one hand. This method is more robust and reliable as it is based on depth information given by the sensor, while in 2D it was based on segmentation methods.

![Figure 1. MS KINECT Architecture](image)
II. FINGERTIP DETECTION IN 3D

A. Getting the Depth Image from KINECT

The internal architecture of MS KINECT is shown in Figure 1. It have infrared camera and PrimeSense sensor to compute the depth of the object while the RGB camera is used to capture the images. As Frati [19] stated “It has a webcam–like structure and allows users to control and interact with a virtual world through a natural user interface, using gestures, spoken commands or presented objects and images”, it is clear that KINECT is a robust device and could be used in different complex applications. The depth images and RGB image of the object could be getting at the same time. This 3D scanner system called Light Coding which employs a variant of image-based 3D reconstruction. The depth output of KINECT is of 11 bit with 2048 levels of sensitivity [21]. The depth value $d_{raw}$ of a point in 3D can be defined as calibration procedure [19]

\[ d = K \tan(Hd_{raw} + L) - O \]

where $d$ is the depth of that point in cm, $H$ is $3.5 \times 10^{-4}$ rad, $K=12.36$ cm, $L = 1.18$ rad and $O=3.7$ cm.

![Figure 2: Depth image acquired using MS KINECT](image)

B. Hands Tracking and detecting Hand Point

Hands tracking and detection of Hands Points were done using NITE modules, which uses Bayesian Object Localization for hand detection. We use OpenNI modules which provide C++ based API for hand detection and tracking. The image with depth information is shown in Figure 2.

![Figure 3: (a) Threshold image, (b) Image of one hand](image)

C. Segmentation by Depth

The depth image was segmented after putting a threshold on depth of hand points and the hands were detected by choosing the blob which contains hand point as shown in fig 3.

![Figure 4: (a) Palm in one hand image, (b) Fingers mask for one hand](image)
D. Fingertips detections

Fingertips were detected by first finding the palm of the hands which were obtained by applying a big circular filter on the image, so that all the fingers in the images would be removed. Now palm of the hand would be subtracted from the original hand image to get the segmented finger masks as shown in figure 4 and figure 5. The finger masks were multiplied with original depth image to get the depth map of fingers. After examining the depth map, it was clear that at fingertips the value of depth was minimum that imply they are closest to camera then remaining fingers. So, the fingertips would be calculated by finding the minimum depth in each finger. This process was used for both hands simultaneously also and results were very encouraging. The results of fingertips detection in real-time for single hand and both hands are shown in figure 6 and figure 9. Here orientation of hand is not a constraint, as 3D sensor is able to detect it in any direction.

III. CENTRE OF PALM DETECTION

Centers of the palms were detected by applying the distance transform in the inverted binary images of hand as shown in figure 7. It was clearly visible that the maximum of the distances transform was giving the center of the palm. As this experiment can use both hands also, we made distinctness to recognition of both the hands. As shown in figure 9, if it is single hand the fingertips and center of palm would be detected in white color while if both hands are detected in the image, the right hand would be detected as white and the left hand would be detected as pink. This color difference make sure that center of palm of one hand should not match with the other hand.

IV. RESULTS

From our lab setup, we are able to identify fingertips and center of palm very accurately, even when the fingers are bent by large angle. The whole system was implemented in real time and the results were very encouraging as shown in figure 8 and 9.
V. CONCLUSIONS

The detection of fingertips and centres of palm has been discussed which will be used in our project “Controlling the robotic hand using hand gesture”. In this approach the user can show one hand at a time or both hands. This project is still in progress and it will make a significant change in applications which can be harmful to human life. The Movement of user’s finger will control the robotic hand by moving hand in front of camera without wearing any gloves or markers.

ACKNOWLEDGMENT

This research is being carried out at Machine Vision Lab, Central Electronics Engineering Research Institute (CEERI/CSIR), A Govt. of India Laboratory, Pilani, INDIA. It research was a part of current project “Controlling the electro-mechanical hand using hand gesture”. Authors would like to thank The Director, CEERI Pilani for providing research facilities and for his active encouragement and support.

REFERENCES


[10] "The teardown,” Engineering & Technology , vol.6, no.3, pp.94-95, April 2011


Figure 9: Results of fingertip detection in real time