Volume 4, Issue 4, April 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Gesture Extraction using Depth Information

Ratika Pradhan, Ronak Pradhan, Jhuma Sunuwar
CSE Department, SMIT
India

Abstract— Gesture recognition has been a subject of much study lately as a promising technology for man-machine communication. Various methods have been proposed to locate and track body parts (e.g., hands and arms) including markers, colors, and gloves. In this paper, we have proposed a low cost, efficient hand gesture extraction technique that extracts gesture based on skin-colour segmentation followed by depth segmentation. In an ideal situation when human communicates with the computer using gesture, the depth of the gesture information and the other body parts differs; this information can be used to separate gesture from other complex background.

Keywords— Gesture, Color segmentation, Depth Map, Local Shading.

I. INTRODUCTION

Gesture recognition pertains to recognizing meaningful expression of motion by a human involving hand, arm, eyes, face and/or body [1]. A hand gesture is form of nonverbal communication, which allows a person to communicate a range of thoughts and feelings with or without speech. Hand gestures can be classified as the static gesture and dynamic gesture. Static gesture also known as posture is defined solely by the static hand configuration and location without any movements involved. Dynamic gesture refers to a sequence of posture connected by continuous motions over a short time span.

Research on Hand Gesture recognition has gained a lot of attentions due to its wide range of applications ranging from designing aid for hearing impaired person, recognizing sign language to designing gesture enable electronic gadgets. Hand Gesture Recognition Systems consists of three basic steps: detection, tracking (for dynamic gesture only) and recognition. The detection layer is responsible for extracting visual features that can be attributed to the presence of hands. The tracking layer is responsible for performing temporal data association between successive image frames and the recognition layer is responsible for grouping the spatiotemporal data extracted to particular classes of gestures. Out of the three steps: Detection or extraction of the gesture from the real world scene has been a tedious task without using the data glove. Recent works in gesture recognition are broadly classified as: Data Glove approach and Vision based approach. Data Glove approach uses set of sensors that are used to capture feature points for a specific static gestures and can also be used for motion tracker which can later be used for recognition. Vision based approach is cost effective but efficiency relies on the computation involved in it.

Many researchers are working on vision based approach for gesture recognition. Lee et al. [2] have proposed a method to recognize band gestures extracted from images with a complex background for a more natural interface in HCI (human computer interaction). The proposed method obtains the image by subtracting one image from another sequential image, measures the entropy, separates hand region from images, tracks the hand region and recognizes hand gestures. Hasan et al. [3] have presented a novel technique for hand gesture recognition through human–computer interaction based on shape analysis. The main objective of their work is to explore the utility of two algorithms that are used for extracting the features to recognize the hand gestures-the hand contour and the complex moment algorithm. Liu [4] have used active statistical model for hand gesture extraction and recognition where a set of feature vectors is normalized and aligned and then trained by Principle Component Analysis (PCA). Victor et al. [5] have used Skin Modeling based on a cylindrical model in the CIELab color space and Hand modeling based on elliptical skin blobs. Initially all the skin pixels form a single blob and then it is processed to determine the one corresponding to hand. Most of the approaches uses skin color segmentation to extract the gesture followed by some refinement to it but fails completely if the background contains the same color and many application uses well set up background to capture the gesture image or frame. In this paper, an approach based on the depth is discussed to extract gesture form the complex background. At first, for background removal, skin color based segmentation is used to extract the multiple skin regions in an image frame, for example, the face and the hands, which may also overlap with each other. The second step is to use depth segmentation separates the hand region and other skin colored regions, for eg. face.

II. PROPOSED METHODOLOGY

The schematic diagram illustrating the workflow for the detection of Hand gesture in image frame is given in Figure 1.

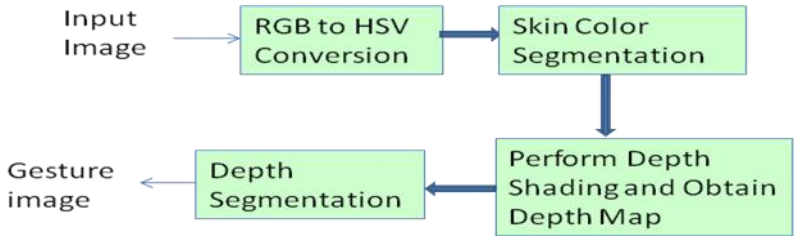


Figure 1 : Schematic diagram of the proposed methodology.

A. RGB to HSV Conversion:

The RGB color space expresses colors as an additive combination of three additive primary colors of light: red, green, and blue. A commonly used color space that corresponds more naturally to human perception is the HSV color space whose three components are hue, saturation, and value. The formulas used to convert RGB to HSV depend on which of the RGB components is largest and which is smallest.

- Let r, g, b ∈ [0,1] be the red, green, and blue coordinates, respectively, of a colour in RGB space.
- Let max be the greatest of r, g, and b, and min the least.
- To find the hue angle h ∈ [0, 360], the saturation, s, and the value, v for HSV space, we compute:

$$h = \begin{cases} 0 & \text{if } \max = \min \\ (60° \times \frac{g-b}{\max - \min} + 0°) \bmod 360°, & \text{if } \max = r \\ 60° \times \frac{b-r}{\max - \min} + 120°, & \text{if } \max = g \\ 60° \times \frac{r-g}{\max - \min} + 240°, & \text{if } \max = b \end{cases}$$

$$s = \begin{cases} 0, & \text{if } \max = 0 \\ \frac{\max - \min}{\max} = 1 - \frac{\min}{\max}, & \text{otherwise} \end{cases}$$

$$v = \max$$

B. *Skin Color Segmentation*

Once the HSV image is obtained, next step is to extract skin using colour segmentation. In this process, all pixels whose *saturation* and hue *values* are within a pre defined range of skin are not altered, whereas values of the remaining pixels are set to 0 (color black). The Hue range for skin color was determined as 0 - 0.11. The Saturation values were within the range 0.2 - 0.7. This process extracts the entire region in image frame that are in skin color for eg. Face, hand and other region covered with skin color.

C. Gesture extraction using Shape from Shading

The next process is to identify or extract the gesture for recognition. Here intensity plays the vital role; it is obvious that the region closer to the sensor at the time of gesture acquisition appears brighter than the other region. The same difference in intensity is used for gesture extraction. In this process, we have used Pentlands local shading [6] approach to generate the depth map shown in figure 2.

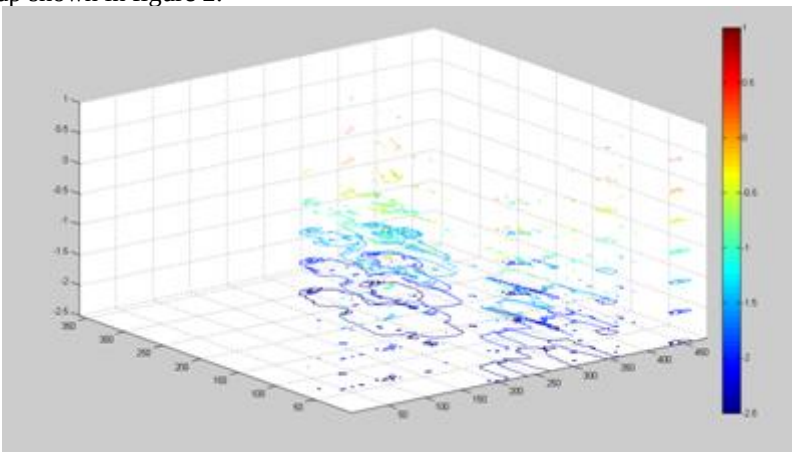


Figure 2: Depth Map

D. Depth Segmentation

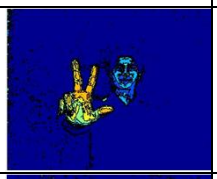
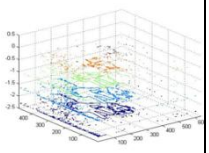
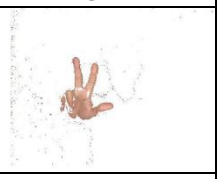
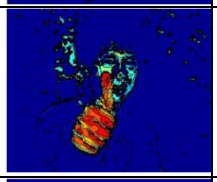
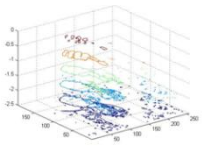
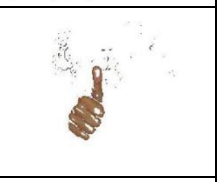
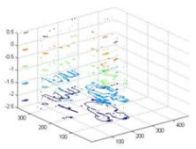
A relatively simple process, wherein the pixels within predefined depth threshold are left unaltered, whereas values of

those pixels not within the threshold value, are set to 255 (color white). The process is similar to color segmentation but differs in the fact that it takes into consideration the depth of the gesture from the depth map obtained from the depth shading procedure.

III. RESULT AND DISCUSSION

The table below shows the result obtained using the proposed methodology:

TABLE I : RESULTS OBTAINED FOR GESTURE EXTRACTION

INPUT IMAGE	HSV IMAGE	SKIN EXTRACTION	2D DEPTH MAP	3D DEPTH MAP	GESTURE EXTRACTED
					
					
					

The required gesture regions were extracted successfully when the gesture region is located as a closer depth than the other skin colored regions of the body. The methods fails when the depth of the required region and other skin colored regions are the same as shown in third image. At the same time, multiple gesture regions may also be extracted if they are at the same depth.

IV. CONCLUSION

Gesture extraction is one of the primary tasks in hand gesture recognition. This paper suggests low-cost and efficient gesture extraction method that can be used to extract static gesture form the image or frame for recognition. The intensity of the skin color region is taken into account for the gesture extraction. The result shows that gesture can be extracted efficiently using the depth of the various region in an image or frame.

REFERENCES

[1] Sushmita Mitra, Tinku Acharya, Gesture Recognition: A Survey, IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews, Vol. 37, No. 3, May 2007.

[2] Lee J, Lee Y, Lee E, Hong S, Hand Region Extraction and Gesture Recognition From Video Stream With Complex Background Through Entropy Analysis, Conf. Proc. IEEE Eng. Med. Biol. Soc. 2:1513-6, 2004.

[3] Haitham Sabah Hasan, Sameem Binti Abdul Kareem, Gesture Feature Extraction For Static Gesture Recognition, Arabian Journal For Science And Engineering, Vol. 38, Issue 12, Pp 3349-3366, 2013.

[4] Nianjun Liu And Brian C. Lovell, Hand Gesture Extraction By Active Shape Models, In Proc. Digital Image Computing: Techniques And Applications, 2005.

[5] Victor Ayala-Ramirez, Sergio A. Mota-Gutierrez, Uriel H. Hernandez-Belmonte and Raul E. Sanchez-Yanez, A Hand Gesture Recognition System Based On Geometric Features And Color Information For Human Computer Interaction Tasks, Xalapa, Ver., Mexico, 2011.

[6] Pentland, Alex P. ; Artificial Intelligence Center, SRI International, Menlo Park, CA 94025; Departments of Computer Science and Psychology, Stanford University, Stanford, CA 94305.

