Controlling Slides using Hand tracking and Gesture Recognition

Rabia Fayyaz*, Eun Joo Rhee*
*Department of Computer Engineering, Hanbat National University
*rabiafyz@hotmail.com

손의 추적과 제스쳐 인식에 의한 슬라이드 제어

Rabia Fayyaz*, 이은주*
*한밭대학교 컴퓨터공학과
*rabiafyz@hotmail.com

요 약

The work is to the control the desktop Computers based on hand gesture recognition. This paper is worked on real time tracking and recognizes the hand gesture for controlling the slides based on hand direction such as right and left using a real time camera

Key Words: Segmentation, Tracking, Gesture Recognition

1. Introduction

In human communications, the movement of hands is used to provide gestures that form a major part of the "body language" which delivers the mood of a person. Thus, the detection and tracking of hands can provide the clues to a person's mode of thought.

Many researchers in the computer vision and robotics have tried to control mouse using hand and video devices. One approach, by Cristina, Javier, Ramon and Francisco, used HSV color space to segment hand and convex hull defects to define the hand. He used average of depths to define hand gesture transition from Stop to start, from Start to Move and from No Hand to start [1].

Another approach was developed by, Hojoon Park [2]. He used Ycrcb space for segmenting hand and convex hull with defects used to define the fingers tips. The presented algorithm in this paper is a small part of Hojoon Park's method. This project work is related to Francisco and Park's work. But this project work is done with simple method and applicable in limited area that's to move slides by hand in fixed environment.

In this paper, the use of hand tracking is presented in controlling slides presentation that is based on three main steps: hand segmentation, hand tracking, and gestures recognition from hand orientation.

First Step, the hand segmentation is done in HSV and lab color space. HSV color cue is used to characterize the skin color but its much depend on lightning condition. And also its break the hand contours that effect the hand segmented result. To prevent this, morphological approach is used so that contours of hand preserved in a good way but not smooth.

Second step is to do the tracking on biggest object that's hand using OpenCV. Third step is to classify the hand gestures sequence that work in two directions Right or Left side.

Finally, work is shown with the good and bad segmentation that effect the hand tracking The algorithm flow of this paper work is shown in given below figure 1.

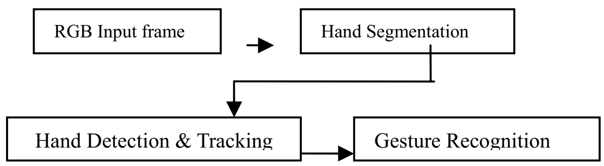


Figure 1. System Flow

2. Hand Segmentation

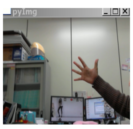
The purpose of this paper research is to perform the mouse tasks (right and left move) by hand as user control desktop computer system. Skin values are tested to segments hand in HSV, Ycrcb and Lab color spaces.

2.1. Skin in HSV, Ycrcb and Lab color spaces

Non-geometric features that is hand segmented based on pixels color. For color detection, it is very difficult to work with RGB space. Because RGB is more likely depended on lightning condition. Skin color is very much affected in lightning condition like dim or dark light or bright light. Mathematically, yet skin is not defined in any color space. So need to pick skin color range carefully. Many methods have developed by researchers to segment the skin in different color spaces [3,4].

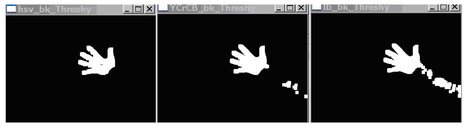
HSV, Ycrcb and Lab color spaces are tested to segment good skin values. HSV is good for skin detection but HSV work at a specific distance for color detection that distance is defined by HSV values. Lab and Ycrcb is also good to detect skin than HSV but hard to pick the skin value in both color space. Because skin is related to pink (light and dark) and red color as a result may be with hand, user' dress color could be detect. Also lab and Ycrcb are camera dependent instead of lightning condition. Infarct, lab worked out for various skin colors.

As a consequence, in this paper, HSV and Lab aggregation is preferred which could detect only face or hand that can be seen in given below figure 2. Also Hsv skin color is related to light pink and red color which could also detect other user cloths but not much as lab and Ycrcb do. That is why morphological components are used to tackle the little noise and hand is detected in HSV and Lab color space to make it good for segmentation.

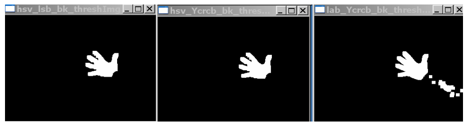


(a) Original Image

(b) In HSV, Ycrcb and LAB without background subtract



(c) In HSV, Ycrcb and LAB with background subtraction



(d) In HSV_Lab_background, hsv_Ycrcb_background and lab_Ycrcb_background

Figure 2. Skin segmentation in hsv, Ycrcb and lab

In the (b) of Figure 2, hand is segmented in hsv, lab and Ycrcb color space without background subtraction. In (c), hand is segmented in hsv, lab and Ycrcb with background subtraction and in (d) segmentation is done with the mixture of hsv-lab, hsv-Ycrcb and lab-Ycrcb color spaces with background subtraction.

The flow of hand segmentation is given Figure 3. Mathematically, background subtraction can write equation (2.1).

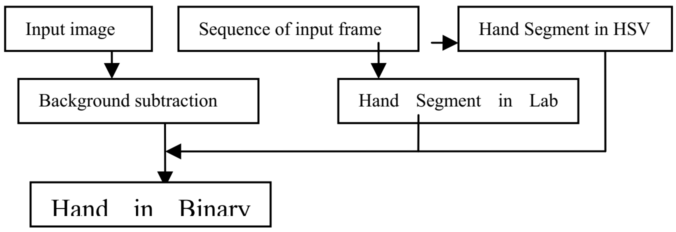


Figure 3. Hand Segmentation

$$\left|(f(x,y)\ ?\ b(x\ y)\right| \quad \text{---- (2.1)}$$

Where $b(x,y)$ = gray level background frame,
$f(x,y)$ = gray level continuous frame

3. Hand Tracking

After hand segmentation[2,5], hand is tracked with central moments of $f(x,y)$ defined in equation (3.1). Since the f(x, y) is the actual image and is assumed to be continuous. For this purpose, a discrete way to describe moments is used as shown below equation (3.2).

$$\mu_{pq} = \int_{a1}^{a2}\int_{b1}^{b2}(x-\bar{x})^p(y-\bar{y})^q f(x,y)\,dxdy, \quad - (3.1)$$

$$\mu_{m,n} = \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} (x - c_x)^m (y - c_y)^n f(x, y) \quad \text{--- (3.2)}$$

The integrals have been replaced by summations. The order of the moment is m + n. usually, moments are calculated about (0, 0). So simply the constants c_x and c_y can be ignore. Hand area is found by the zeroth order moment μ_{00} of $f(x,y)$ as binary image in equation (3.3). To calculate the center of segmented hand, first and second moment order are used as shown below equation (3.4) and (3.5).

$$\mu_{0,0} = \sum_{x=0}^{w} \sum_{y=0}^{h} f(x, y) \quad \text{--- (3.3)}$$

First moment: $sum_x = \sum_{x=0}^{w} \sum_{y=0}^{h} x f(x,y)$ --- (3.4)

Second moment: $sum_y = \sum_{x=0}^{w} \sum_{y=0}^{h} y f(x,y)$ --- (3.5)

To get the average, divide each by the number of pixels. The number of pixels is the area of the image that's defined by the zeroth moment as shown in equation (3.3). So centeriod of segmented hand is calculated as shown in the equation (3.6).

$$\mu_{1,0} = \frac{sum_x}{\mu_{0,0}} \quad , \quad \mu_{0,1} = \frac{sum_y}{\mu_{0,0}} \quad \text{--- (3.6)}$$

One interesting thing about this technique is that it is not very sensitive to noise. The centroid might move a little bit but not much. After locating hand, hand is detected as biggest contour which could be a hand or face as shown in figure 4.

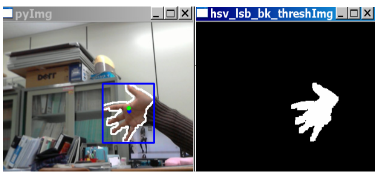


Figure 4. Hand tracked

In figure 4, the blue rectangle shows the detected hand and the blue circle points are the current location, and the green point is the previous location. The idea here is that if hand does not move in the image then current and previous location should be same. That's not happening in this case because previous and current positions change by a small amount each time, even the object is stationary. The small change is due to a wide variety of factors such as lighting, color camera resolution. So, according to statistical data, the hand is restricted in a static state.

<u>The static hand and no hand condition</u> are specified with the change in current and previous hand location. The condition for static hand state is given below algorithm [3.1]. The condition may change depending on the background and lightning condition.

[Algorithm 3.1]

{ $if((-1 < dx < 1) \&\& (-1 < dy < 1))$

Then Hand not move

Else $if(dx == 0 \&\& dy == 0)$

Then No hand in the video}

,Where $dx = x_i - x_{i-1}$, $dy = y_i - y_{i-1}$.

To find <u>angle of moving</u> hand, it works with the change in current and previous location in x-axis and y - axis that could be positive or negative will help to find angle with respect to x-axis i-e arc tan (dx , dy), where dx = $x_i - x_{i-1}$, $dy = y_{i-1} - y_i$, Using angle, right and left direction can be identified by applying test condition.

<u>Direction of motion</u> sets by the following given figure 7. Color shows direction with respect to its quadrant. The direction of right and left side is defined in algorithm [3.2].

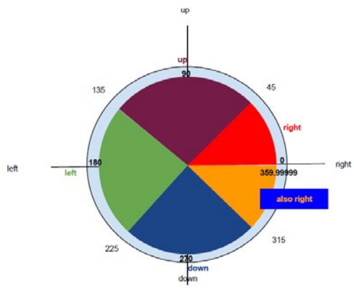


Figure 7. Direction for sides in circle

[Algorithm 3.2]

{If ((90 > angle > 315) or (90>angle >0) or (angle == 0) or (angle == 90) or (angle == 360))

Then move right

Else if ((angle >= 180) && (angle <=269))

Then hand move left}

4. Hand Gesture Recognition

Hand gesture is recognized based on its defined hand direction. The recognized gesture results are shown in figure 9 where direction is represented as a red clock. The direction is specified by angle as shown in the algorithm [3.2].

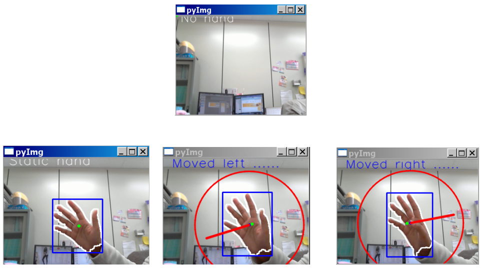
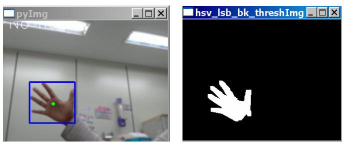


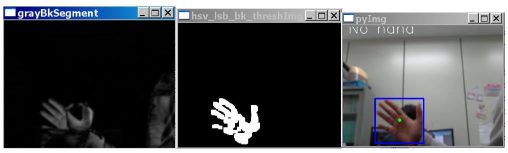
Figure 9: hand state direction

5. Conclusion and Discussion

In the experiment, we tested skin segmentation with the fixed and bright background in HSV and Lab color space.



(a) Good Segmentation on bright background



(b) Bad Segmentation on dark background

Figure10.Hand Segmentation and Tracking Results

In Figure 10, hand is segmented on bright background which means hand and background in gray scale is at high intensity, separate the hand from the background. But hand segmentation is not good on dark background intensity level. As a result hand segmentation effected badly which affects the tracking.

But this skin detection problem can be solve using two camera Then hand could track in an efficient way to recognize right and left move gestures in the context of slides movement like next and previous slides motion. More modification is listed to make it more applicable.

1) Good Background subtraction is required because the problem with simple background subtraction is that it creates black patches in binary hand (object). This happened if object' s intensity nears to black is same with dark background intensity, which creates black patch in moving hand, and can fix using two cameras.

2) Gesture Recognition needs to improve that can move the slides only in one direction either left or right in an effective way. Using a 3D camera, with a depth map, these tasks would be substantially easier. However these results, from a 2D image provide sufficient recognition and tracking to facilitate Slides gesture operation.

References

[1] Cristina Manresa, Javier Varona, Ramon Mas and Francisco J.Perales, "Real-Time Hand Tracking and Gesture Recognition for Human-Computer Interaction" , ELCVIA ISSN:1577-5097, Computer Vision Center/Universitat Autonoma de BarceLona,Barcelona,Spain, January 2000.

[2] Hojoon Park, "A Method for Controlling Mouse Movement using a Real-Time Camera" , Brown University, Providence, RI, USA, Conduit CS Dept Newsletter, 2010.

[3] Vladimir* Vassili Sazonov Alla Andreeva, "A Survey on Pixel-Based Skin Color Detection Techniques" , Faculty of Computational Mathematics and Cybernetics, Moscow State University, Moscow Russia, V Vezhnevets, V Sazonov , Proc. Graphicon, 2003.

[4] C. W. Kit J. See N. A. Abdul Rahim," RGB-H-CbCr Skin Colour Model for Human Face Detection " , *MMU* International Symposium on Information & Communications Technologies (M2USIC), 2006.

[5] Simon Xinmeng Liao, Image Analysis by Moments, Department of Electrical and Computer Engineering, University of Manitoba Winnipeg, Manitoba, Canada, 1983.