A Method of Gesture Segmentation Based on Skin Color and Background Difference Method

Duan Hong*
Software School of Xiamen University
Xiamen city,China
*corresponding author: hduan@xmu.edu.cn

Luo Yang
Software School of Xiamen University
Xiamen city,China
e-mail:779446731@qq.com

***Abstract*—This paper proposes a method of segmentation of hand gesture, based on the static skin color model and background subtraction under complex background. This paper discussed the reasonable threshold selection of Cb, Cr in the skin color model and the segmentation selection of skin area combining the background segmentation. Finally, an inequality of hand's outline feature is proposed to complete the division processing of the palm part of the area of skin. Experiments show the accurate segmentation of gesture under a complex static background.**

Keywords-Gesture Segmentation, Skin Color, Background Difference

I. INTRODUCTION

The development of human-computer interaction technology makes some gesture recognition technology gradually become a research hotspot. Among them, vision-based monocular bare hand gesture recognition relative to other HCI methods with natural, direct, simple features, has a wide application prospect in the computer game, robot control and home appliances control. While in this gesture recognition method, because of the multiple meaning of gestures, separating the staff quickly and accurately from video image sequence become the key premise [5].

Color is important information of the hand. It has relative stability and can be distinguished from other objects, so the color characteristic is a common method in the detection of manpower. When researching the clustering feature of color in the YCbCr space, Chai [2] judges the CbCr color values of the input pixel, which fall into the range of Cb=[77,127] and Cr=[133,173], for the skin color pixels. The study of Jianguo Wang et al. present a YcgCr color space which is similar to the YcbCr one. They make the points whose pixel values of three components Y, Cg, Cr are respectively in the range of [60, 235], [90, 127], [129, 173], as the color pixel. According to the statistical results of skin color, which says that skin color with different luminance component Y has a different aggregation range in CbCr space, Lei Ming and others project all color components of Y, Cb and Cr onto YCb and YCr spaces, and get the Cb and Cr value corresponding to different Y value. Although these methods are simple and rapid, they have some problems: (1) Different color space has different clustering degree on skin color. We need to select a better color space by experiment. (2) The threshold which is too small will make some color region to be mistaken for an non-skin color region, and the too large one will make non-skin color region as a skin color region. It needs to select a large number of samples for testing in statistics. At the same time, methods above cannot directly divide up the hand from the hand skin area.

This paper presents a method of gesture segmentation based on static skin color model and background difference, in a complex background. Select the reasonable Cb, Cr threshold from the traditional skin color through the experiment, and do the segment selection on skin region combined with the background difference method. Then separate out the palm portion from the skin region by a hand contour feature inequality.

II. BASED ON THE STATISTICAL STATIC SKIN COLOR SPACE MODEL

Rein - Lien Hsu [1] found that people with different skin color can present a clustering feature in YCbCr space through the experimental analysis, as shown in Figure 1.

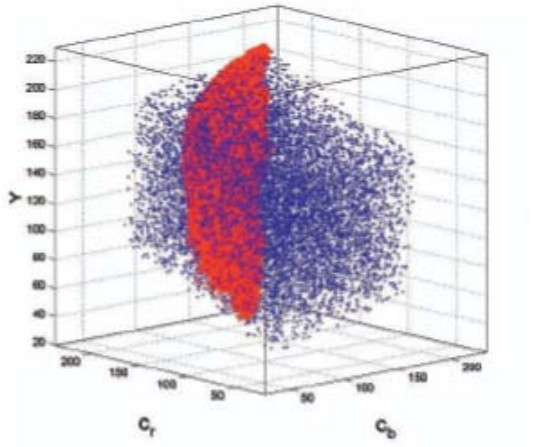

Figure 1. The distribution in YCbCr space of skin color [1]

Among them, the blue represent the point which can be simulated by the display, the red represent the point in the skin color area. We can screen them according to these gathered rules. When the outside environment is steady, formulating the skin color range directly with the mathematical expression is a simple and efficient method.

But when using the formulating skin color method, we find that the value between Cb and Cr given by Chai can not detect the Asian skin well according to the experimental results. We chose 200 Asian skin color's pictures of different illumination conditions and we found that the value of Cb is

Published by Atlantis Press, Paris, France.

from 96 to 143 and the value of Cr is from132 to 164. So we take them as the skin color model and using them to divide the skin. The preliminary experimental results are shown in figure 2.

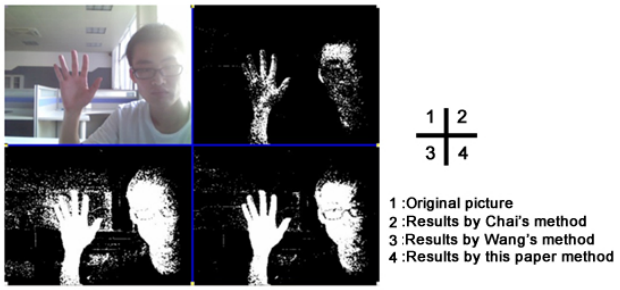

Figure 2. Comparision chart of the experimental results under complex background

III. DIFFERENCING MOVEMENT ANALYSES

Dynamic gesture is composed with a series of sequence of image. We can segment the moving part by using the method of moving detection. Then in gesture configuration we can further segment the hand area. Now there exist several differential detection methods: Temporal differencing, optical flow and background subtraction.

Temporal differencing is to select continues image sequence, doing the difference operation of color between adjacent frames, and at last binarize it to extract the moving part.

Moving detection based on the optical flow use the optical feature of moving object which changes by the time. It can also detect the moving object when the camera is moving as well.

Background subtraction is a widely-used method, it can be used for segmenting the moving object in the static situation. It requires getting the background of environment, and then we can get the foreground by doing the difference of current image and the background, which is the moving part. It is similar to the temporal differencing, which also need binarization to get rid of the part that the difference value is relatively small.

For our application background capture the video for the fixed camera, the background of image is relatively stable. And the current background subtraction can deal with most situations including noise. And also, is easier to calculate, so we decide to take the method of background subtraction.

First of all, we need to do the modeling of background, we need to initialize the background model. It should initialize n frames for the modeling of background model right after the camera is started (n can be bigger than 30 in the lab). Background updating equation (1) is:

$$F_{background} = F_{background} * a + F_{new} * (1 - a) \quad (1)$$

a is the coefficient of background updating, after testing, we find that we can get a better result when the value of a is between 0.7~0.8. is a newly-captured frame.

In the stage of configuring, we take the method of differencing to get the foreground. As shown as equation (2):

$$F_{foreground} = F_{new} - F_{background} \quad (2)$$

Because now is RGB tri-channel image, it need to be transform to a gray level image which is single-channel image. And then do the binarization. The equation (3) is:

$$f(x, y) = \begin{cases} 0, & if\ (diff < c) \\ 1, & other \end{cases} \quad (3)$$

c is the threshold .when the differencing result is bigger than c means the difference is relatively big, which is foreground ,the moving object.

IV. FACE EXCLUSION

After processing the image by the foregoing instructions, theoretically, the only part of the image that would be interrupting the division of hands is the Face. For easier detection, we make a rule that a single hand should be wide opened with its five fingers spread as its initializing gesture when it comes into the framing zone of the camera for the first time, and all we need to do is detecting this gesture now. It has been found clear after dozens of tests, if the feature of a contour matches the equation below [9]:

$$4<=ftips<=5\&\&3<=gros<=4\&\&\ 4<=defects<=5 \quad (4)$$

Then we consider it to be a hand, ftips here means the number of fingers, gros means the number of the concave parts between fingers, defects means the characteristic value of the gesture contour. In this way we can exclude the face and obtain the initializing window for out gesture. And we only need to track this specific zone in the coming algorithm for the detection.

V. EXPERIMENT AND CONCLUSION

We combine the skin color model and the background subtraction method to achieve the hand segmentation in this experiment. Raging skin color method required by the color skin model can meet the real-time needs. On the other hand, in the actual tests, the color of some objects is like the color of skin (such as the wood desk), but those objects are generally fixed in the viewfinder, so we can use the background subtraction method to exclude this part of the region . Finally, screen the region of hand by using the Equation 4. Figure 3, Figure 4, Figure 5 shows the results of experiments under different environments.

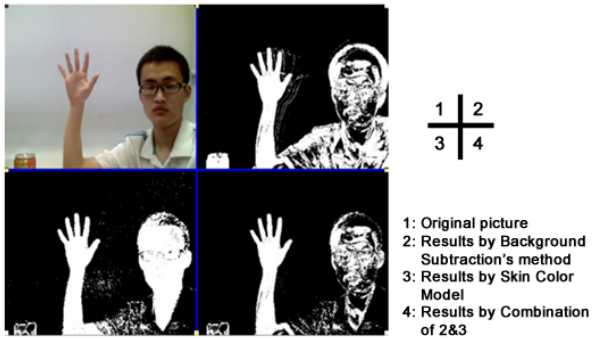

Figure 3. Contrast effects under simple background and normal light

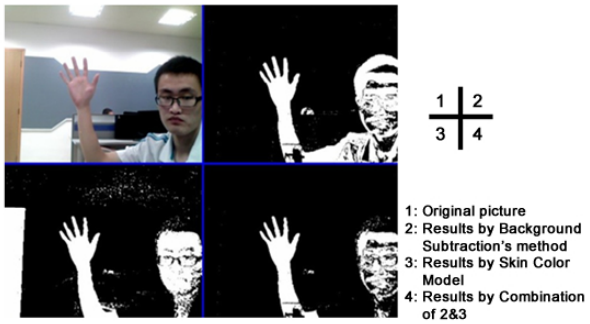

Figure 4. Contrast effects under complex background and normal light

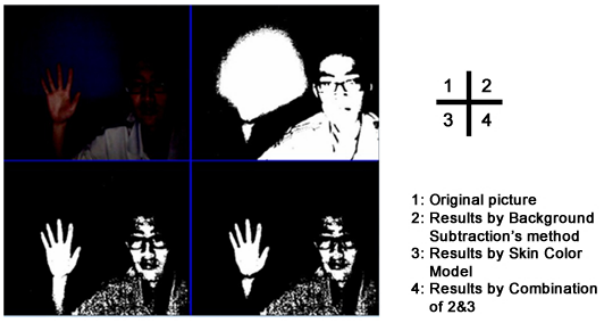

Figure 5. Contrast effects under dark light

According to the contrast effects of the experiments, the method mentioned in this paper can separate the hand gesture better, even under the dark light.

REFERENCES

[1] Rein-Lien Hus, Mohamed A-M, Anil K. Jain, "Face Detection in color Images," IEEE Transactions On Pattern Analysis and Machine Intelligence, vol.24, No.5, pp.696-706, 2002.

[2] Chai D, Ngan K.N, "Locating Facial Region of a Head-and-shoulders color image," In :Proceedings of the 3rd International Conference on Automatic Face and Gesture Recognition. Nara, Japan, pp.124~129, 1998.

[3] Jianguo Wang, Jizhao Hua, Jingyu Yang, "Detection for complexion under complex background, based on the new color space YcgCr," Computer Science, vol.34, No.8, pp.229~231, 2007.

[4] Ming Lei, Junying Zhang, Jiyang Dong, "An algorithm for complexion detection under variable light condition," Computer engineering and applications, vol.38, No.24, pp.123~125, 2002.

[5] Huaping Yu, Jiaoyan Ai, "Demarcation based on dynamic complexion of gesture zone," Computer technology and automation, vol.29, No.2, pp.86~88, 2010.

[6] Kalman, "A New Approach to Linear Filtering and Prediction Problems," Transaction of the ASME Journal of Basic Engineering, pp.35-45, 1960.

[7] A.Elgammal, D.Harwood and L.Davis, "Non-parametric model for background subtraction," In: Proceedings of IEEE ICCV'99 Frame-rate workshop, 1999.

[8] Anderson C, Bert P and Vander Wal G, "Change detection and tracking using pyramids transformation techniques," In:Proc SPIE Conference on Intelligent Robots and Computer Hsion, Cambridge, pp.72-78, 1985.

[9] Yang Luo, "Gesture Recognition based on Skin Color and Background Difference," Xiamen University, 2012.

Published by Atlantis Press, Paris, France.