# LETTER Efficient Hand Segmentation and Fingertip Detection Using Color Features of Skin and Fingernail

Yaming WANG<sup>†\*</sup>, Student Member, Jiansheng CHEN<sup>†</sup>, and Guangda SU<sup>†a)</sup>, Nonmembers

key words: hand segmentation, color space, color temperature, fingertip detection, fingernail

## 1. Introduction

In recent years, vision-based methods have become the mainstream in the area of gesture recognition, which can be divided into three steps: gesture modeling, gesture analysis and gesture recognition [1]. For a certain model, features are extracted in the analysis stage for parameter estimation, which are further classified and interpreted in the recognition step.

The hand gesture recognition system, developed for human-computer interaction use in our previous work [2], consisted of four modules, namely hand segmentation, contour extraction, fingertip detection and gesture recognition. Among first three modules belonging to gesture analysis stage, the contour extraction algorithm based on [3] worked well. In this paper, the goal is to improve the hand segmentation and fingertip detection, whose high accuracy can benefit the recognition of some basic but fingertip-sensitive gestures, such as the digits expressed by the number of fingertips.

In terms of hand segmentation, we studied the skin color of Asian people as the hue of the sample image changes due to different light sources or camera settings. Based on the regularity, we designed the  $YU_{skin}V_{skin}$  color space from traditional YUV space, which is optimal for hand segmentation if linear thresholds are adopted. Further research aimed to perform robust hand segmentation without the need for white balance.

Our previous system used geometric characteristic for fingertip detection, which might easily be interfered by skin

DOI: 10.1587/transinf.E96.D.1894

area with similar geometric feature. To improve this, we originally studied the concomitance of finger and fingernail as important feature of fingertip, together with geometric feature to achieve multi-feature judgement.

After introducing the new techniques mentioned above, our system could extract smooth hand contour and detect fingertip with accuracy rate larger than 98.7% and false alarm rate less than 3.9%, which laid solid foundations for high-level recognition of fingertip-sensitive gestures.

## 2. Optimal Linear Color Space for Hand Segmentation

## 2.1 The Feature of Skin Color in YUV Color Space

In the literature of hand segmentation, the most widely used color spaces are YUV and YIQ which separate chromatic components from lightness component. The two color spaces are essentially the same since YIQ is obtained just by rotating the U - V coordinate axes 33°.

The transformation from RGB to YUV is as follows:

$$Y = W_{\rm R}R + W_{\rm G}G + W_{\rm B}B = 0.299R + 0.587G + 0.114B,$$
  

$$U = U_{\rm max}\frac{B-Y}{1-W_{\rm B}} = -0.147R + 0.289G + 0.436B,$$
  

$$V = V_{\rm max}\frac{R-Y}{1-W_{\rm R}} = 0.615R - 0.515G - 0.100B,$$
 (1)

where  $W_{\rm R} = 0.299$ ,  $W_{\rm G} = 0.587$ ,  $W_{\rm B} = 0.114$  are weight of R, G, B when calculating the lightness *Y*;  $U_{\rm max} =$ 0.436 and  $V_{\rm max} = 0.615$  are determined by requirements in practice. When  $R, G, B \in [0, 255]$ ,  $Y \in [0, 255]$ ,  $U \in$  $[-255U_{\rm max}, 255U_{\rm max}]$  and  $V \in [-255V_{\rm max}, 255V_{\rm max}]$ 

[4] showed that regardless of change of lightness, the area of skin color is compact on U - V plane, which brings about convenience to hand segmentation. [5], [6] obtained similar result in YIQ space.

Before proceeding to the color space for hand segmentation, for simplification and symmetry, we modify U and V by setting  $U_{\text{max}} = V_{\text{max}} = 0.5$ . In the rest of this paper, U and V refer to the modified version.

## 2.2 The Design of YU<sub>skin</sub>V<sub>skin</sub> Color Space

When the color temperature of light source changes from high to low, the hue of the recorded image, which can be represented by the color appeared in that image corresponding to a white object, will change from bluish to white, and

Manuscript received January 30, 2013.

Manuscript revised April 1, 2013.

<sup>&</sup>lt;sup>†</sup>The authors are with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China.

<sup>\*</sup>Presently, with the Department of Electrical and Computer Engineering, University of Maryland, MD 20742, USA.

a) E-mail: susu@tsinghua.edu.cn

then to reddish [7]. To find the effective color space, we studied the distribution of skin color under different hues on U - V plane, and found that, as the hue of the whole image changes from cold to warm (the white area from blue to red), the hand color area keeps the compact characteristic mentioned in Sect. 2.1, and shifts approximately linearly on U - V plane. Based on this principle, it is convenient to design color space suitable for hand segmentation.

The color space  $YU_{skin}V_{skin}$  for hand segmentation was designed as follows. Firstly, sample images of hands with different hues were taken. For each image, we calculated  $\bar{U}$  and  $\bar{V}$ , which are the arithmetic mean of U and V components of all the hand area pixels, respectively. On U – V plane,  $(\bar{U}, \bar{V})$  is the centroid of hand color area. Then we drew the trajectory of  $(\bar{U}, \bar{V})$ . As depicted in Fig. 1, it has excellent linear property. Utilizing straight line V = kU + bto approximate the trajectory, we took this line as V<sub>skin</sub> axis, the projection of O(0,0) to V<sub>skin</sub> axis as the new origin O'. The U<sub>skin</sub> axis is the line which passes O' and orthogonal to V<sub>skin</sub> axis. As shown in Fig. 2, the transformation from YUV to YU<sub>skin</sub>V<sub>skin</sub> is

$$U_{\rm skin} = U\cos\theta + V\sin\theta - b\sin\theta$$
  
$$V_{\rm skin} = -U\sin\theta + V\cos\theta,$$
 (2)



**Fig. 1** The red dots denote the trajectory of  $(\overline{U}, \overline{V})$ , the centroid of hand skin area on U–V plane. It moves closer to V axis when the image becomes bluish, while closer to U axis when the image becomes reddish. The blue line denotes V = kU + b, the linear approximation of the trajectory.



Fig. 2 YU<sub>skin</sub>V<sub>skin</sub> and its comparison with YS<sub>a</sub>S<sub>b</sub>.

where  $\theta$  is the acute angle between V axis and the line V = kU + b. In practice, we got k = -0.9469, b = 9.422,  $\theta = 46.56^{\circ}$ . Further combine the YUV – RGB transformation, then

$$U_{\rm skin} = 0.2468R - 0.5318G + 0.2850B - 7.3738$$
$$V_{\rm skin} = 0.4665R - 0.0478G - 0.4178B,$$
(3)

where  $R, G, B \in [0, 255]$ .

#### 2.3 Hand Segmentation in YU<sub>skin</sub>V<sub>skin</sub>

When the image has proper lightness and achieves white balance, choose threshold  $\{U_1, U_2, V_1, V_2\}$  such that

$$P\{U_{1} \le U_{skin} \le U_{2}\} \ge 99\%$$

$$P\{V_{1} \le V_{skin} \le V_{2}\} \ge 99\%.$$
(4)

Practically we obtained  $(V_1, V_2) = (8, 33), (U_1, U_2) = (-10, 10).$ 

Furthermore, when the hue of the whole image changes, we found that  $U_1$ ,  $U_2$  remains unchanged and  $V_1$ ,  $V_2$  always satisfies  $V_2 - V_1 \approx 25$ , which means that  $V_{skin}$ is the approximate symmetric axis of hand area. Therefore, when hand area takes a relatively large percentage of space in the whole image, it is able to conduct robust hand segmentation. Firstly calculate  $\bar{V}_{skin}$ , the mean of  $V_{skin}$  of the whole image. Secondly design dynamic thresholds  $V_1$ ,  $V_2$ such that  $V_2 - V_1 = 25$  and  $\bar{V}_{skin} \in [V_1, V_2]$ . This approach worked effectively according to our experiment.

## 2.4 Brief Comparison with Previous Work

[4] utilized YUV for hand segmentation and the threshold is  $\arctan(V/U) \in [105^\circ, 135^\circ]$ . The area of  $\arctan(V/U) \in [105^\circ, 135^\circ]$  on U – V plane is much larger than the rectangular area determined by (4). A larger area means that more extra pixels other than hand area pixels might be included during segmentation.

[5] designed  $YS_aS_b$  on U – V plane which was represented by I – Q coordinate system in the paper. Since the linear shift property mentioned in Sect. 2.2 was not considered, it directly connected the origin and the centroid *M* to acquire the  $S_a$  axis, which failed to be the approximate symmetric axis of hand area. As depicted in Fig. 2, it is obvious that the linear thresholds determined in  $YU_{skin}V_{skin}$  (the inner rectangular) have a more compact constrains on U – V plane than that of  $YS_aS_b$  (the outer rectangular).

As can be seen from the analysis above,  $YU_{skin}V_{skin}$  outperforms the previous YUV based hand segmentation approaches in the sense of linear thresholds, and can be adjusted easily when the image fail to achieve white balance.

## 3. Fingertip Detection Utilizing Its Concomitance with Fingernail

## 3.1 Color Based Distinction between Fingernail and Skin

So far, few research has focused on this issue. [5], [6] designed  $YS_aS_b$  color space.  $S_a$  was used to recognize and



Fig. 3 Finding the best coordinate axis to distinct skin and fingernail.

its orthogonal component,  $S_b$  was used to detect fingernail. However,  $S_b$  was not designed for fingernail extraction specifically.

Similar to Sect. 2, we need to find out an axis on U - V plane, such that the projection of skin and fingernail area to this axis has the maximum difference. As displayed in Fig. 3, denote the center of mass point of skin and fingernail area as M, N respectively. To simplify the problem, we can use M and N to represent the skin and fingernail area. Among all the lines which passes the origin O, the coordinate axis on which the projection of segment MN reaches the maximum is the one parallel to MN. The axis was obtained through statistic approach as follows. We manually extracted fingernail and skin areas from 30 hand images and calculated the coordinates of M and N respectively. Then

$$(\bar{U}_M, \bar{V}_M) = (-11.78, 28.41),$$
  
 $(\bar{U}_N, \bar{V}_N) = (-6.91, 28.11),$  (5)

where  $\bar{U}_M$  and  $\bar{V}_M$  denote the means of U and V components of skin respectively, while  $\bar{U}_N$  and  $\bar{V}_N$  are the corresponding means of fingernail.

Let  $\theta_1$  denote the acute angle between V axis and *MN*, as depicted in Fig. 3, then

$$\theta_1 = \arctan\left(\frac{\bar{U}_N - \bar{U}_M}{\bar{V}_N - \bar{V}_M}\right) = 86.41^{\circ} \tag{6}$$

Notice that  $\theta_1$  is approximate to 90°. Therefore it is reasonable to directly use U component to distinct fingernail from skin in practice. The frequency distribution of skin's and fingernail's U component is shown in Fig. 4. As can be seen, the U components of skin and fingernail have distinctive distribution such that they can be classified easily. Further experiment indicated that the following experimental formula holds.

$$\bar{U}_N \approx \bar{U}_M + 5. \tag{7}$$

During the experiment, (7) was robust when the image became reddish or bluish. Thus we can conclude that, although the color of skin and fingernail both changes, the relationship between them remains stable.

## 3.2 Realization of Fingertip Detection

Based on the result of Sect. 3.1, we combined the geometric



Fig. 4 Frequency distribution of U component of skin and fingernail.



**Fig.5** Experiment result of distinguishing fingertip and pseudo-fingertip with the information of fingernail.

feature of fingertip and the information provided by fingernail to conduct multiple feature detection.

Firstly, after hand segmentation, calculate  $\bar{U}_1$ , the mean of U component of hand area. Secondly, for each point along the hand contour, calculate the radius and the center of osculating circle and choose those points whose geometric characteristic is similar to fingertip. Then for each chosen point in the last step, the mean of the U component inside the corresponding osculating circle is calculated, which is denoted by  $\bar{U}_2$ . If  $\bar{U}_1$  and  $\bar{U}_2$  satisfy

$$\bar{U}_2 \ge \bar{U}_1 + U_{\rm th},\tag{8}$$

the area inside the osculating circle is regarded as fingernail and the corresponding point along the hand contour is chosen as the candidate of fingertip. Lastly, the adjacent candidate points are clustered and thus we obtain the final fingertip point. In practice, the threshold  $U_{\text{th}}$  can be chosen according to (7) and experiment environment.

Figure 5 confirms the effectiveness of utilizing the information of fingernail. The red point denotes the real fingertip detected by the system. The black points inside the hand area denote all the possible centers of osculating circles, which means the geometric characteristic of contour near them is similar to that of fingertip. These areas include acute angles caused by the watch, the joint of finger and the real fingertip. As can be seen from the result, only the real fingertip is recognized by the system because the information provided by the fingernail is considered.

## 4. Experiment Result

In this section, experiments were conducted to test the combination of all the techniques mentioned in the previous sections. Test environment, methods and results are described sequentially.

The hardware environment of the experiments is listed in Table 1.

Two experiments were conducted. 16 Asian people took part in both experiments. The source of light was fluorescent. Before each experiment began, the camera was adjusted such that the lightness and the white balance were proper. During each experiment, the settings of the camera remained the same. The resolution of images taken by the camera was  $640 \times 480$ .

Test 1 was conducted under white background, aiming to test the accuracy of hand segmentation and fingernail extraction. 3 images were taken for each participant, including the images with skin areas whose geometric characteristic was similar to fingertip, as described in Sect. 3.2. In total 48 images from 16 participants were tested.

Test 2 intended to test the robustness of the system in the sense of complex background. The experiment was conducted in an office. In the sample images, the hand area was smaller than previous experiment, and human face was included in the background, which was eliminated by our system according to the concomitance of face and eyes. 2 images were taken from each participants such that 32 images were tested in total.

We used detection rate  $\eta$  and false alarm rate  $\epsilon$  to quantitatively describe the performance of fingertip detection. Suppose the total number of fingertips in all the images is S, the number of fingertips missed by the system is  $e_1$  and the number of false detection is  $e_2$ . Then

$$\eta = 1 - \frac{e_1}{S}, \quad \epsilon = \frac{e_2}{S}.$$
(9)

The test results are displayed in Table 2. The experiment results suggest that our system can extract hand contour and fingertip location with high accuracy in real time, when the camera is adjusted properly.

 Table 1
 Hardware environment of the experiments.

| Camera | Logitech HD Pro Webcam C920 |
|--------|-----------------------------|
| CPU    | Intel Xeon 2.8 GHz          |
| Memory | 2 GB DDR2 SDRAM             |

Table 2 Experiment result.

| Test | Detection Rate $\eta$ | False Alarm Rate $\epsilon$ | Speed (s/image) |
|------|-----------------------|-----------------------------|-----------------|
| 1    | 98.7%                 | 3.9%                        | 0.194           |
| 2    | 99.3%                 | 1.9%                        | 0.298           |

#### 5. Conclusion

In this paper, novel hand segmentation and fingertip detection methods have been proposed. Based on the principle of skin color when the color temperature changes, the color space  $YU_{skin}V_{skin}$  has been designed. The novel color space performed well when the camera was properly adjusted, and it can resist the change of hue when the hand area is relatively large in the image. In terms of fingertip detection, the color feature of fingernail has been studied and the concomitance of fingertip and fingernail has been adopted to enhance the performance of fingertip detection. In practice, the fingertip detection rate achieved more than 98.7% with false alarm rate less than 3.9%.

# 6. Future Work

To completely achieve hand segmentation which is robust to the change of hues, one possible approach is to combine our result with automatic white balance. According to [7], the common process of automatic white balance is to obtain the reference white point, calculate image's deviation from white balance and adjust the color of the whole image. If we can find the relationship between Fig. 1 and the deviation of reference white point, we can estimate the centroid point of skin color area on U – V plane from the reference white point and thus set dynamic thresholds. Therefore, robust hand segmentation regardless of the hue of the image might probably be achieved. After that, the accuracies of the evolved method with dynamic thresholds under various light conditions remain to be tested systematically.

## References

- V.I. Pavlovic, R. Sharma, and T.S. Huang, "Visual interpretation of hand gestures for human-computer interaction: A review," IEEE Trans. Pattern Anal. Mach. Intell., vol.19, no.7, pp.677–695, July 1997.
- [2] F. Fang, "Interactive computing based on high speed image processing technology," Thesis of Bachelor's Degree, Tsinghua University, 2011. Available from https://sites.google.com/site/fangfeihomepage/ Fang11a.pdf?attredirects=0&d=1
- [3] G. Su, "Contour filling with chain construction," J. Computer Research and Development, vol.24, no.9, pp.61–64, Sept. 1987.
- [4] H. Yao, M. Liu, W. Gao, X. Fan, H. Zhang, and Y. Lu, "Method of face locating and tracking based on chromatic coordinates transformation of color images," Chinese J. Computers, vol.23, no.2, pp.158–165, Feb. 2000.
- [5] E. Tamaki and K. Hoshino, "Personalized color system for robust extraction of skin-color," The Virtual Reality Society of Japan, vol.12, no.4, pp.471–478, 2007.
- [6] E. Tamaki, T. Miyaki, and J. Rekimoto, "A robust and accurate 3D hand posture estimation method for interactive systems," IPSJ J., vol.51, no.2, pp.229–239, Feb. 2010.
- [7] C.C. Weng, H. Chen, and C.S. Fuh, "A novel automatic white balance method for digital still cameras," Proc. IEEE Int. Symposium on Circuits and Syst., vol.4, pp.3801–3804, May 2005.