

Hand Signal Recognition Aiming at Robot Control

Yukihiro MIYOSHI

Abstract: A method to recognize hand shape, i.e., numbers of fingers stuck out; using a color image is presented. A mean and dispersion of hue values and a mean of saturation values, applying the HSL color model, in a small square area of an image are revealed to be useful features to distinguish skin-area and contour of the hand. The contour of the hand is traced by the border following method using the small square area. Changes of tangential directions of the contour of the hand are used to identify the number of fingers stuck out. The wavelet shrinkage is used to disclose the feature patterns of the tangential directions and make an automatic hand-recognition possible. Although only a few images evaluated, numbers of the fingers were recognized automatically.

Keywords: color image recognition, mean of hue, dispersion of hue, skin-area, hand recognition, wavelet shrinkage

1. Introduction

Recognition of hand signals using digital image processing is expected as a method for human-computer interface¹⁾ and a method to instruct a robot on a work²⁾. In this paper a method to recognize hand posture, especially a number of fingers stuck out, is developed. Color information, e.g., RGB values and HSL values of pixels include variation and the direct use of those values tends to result in an unstable recognition. A mean and some kind of dispersion are adequate as the cue to detect hand area in an image. In order to realize a natural signaling, neither usage of the glove nor simple color backgrounds are supposed for the method in this paper. And aiming at the human-robot communication, time-consuming and sophisticated calculation like an artificial neural network is not supposed.

2. Skin area cognition in an image

Fig. 1 is a 640×480 24-bit RGB color image showing a hand and a background. This image is captured using a color video camera under fluorescent ceiling lights in a room. Fig. 2 shows R , G , B values of pixels in the rectangular area A, i.e., the skin area, and in the rectangular area B, i.e., the non-skin area, of the image shown in Fig. 1. Integer numbers between 0 and 255 represents red, green and blue intensities of the pixels. The pixels in the skin area A and those in the non-skin area B occupy different RGB space, as far as for the image in Fig. 1, so classification into two classes, i.e., a skin area and a non-skin area, is possible using some sort of classifier, e.g., a neural network³⁾. It is desirable, however, that features to recognize the skin area are intuitively understandable. Hue is one of those features.

R , G , B values of pixels can be mapped to hue, saturation and lightness components referring to the double hex-cone HSL color model⁴⁾. Hue, saturation and lightness



Fig. 1 Image showing a hand and a background

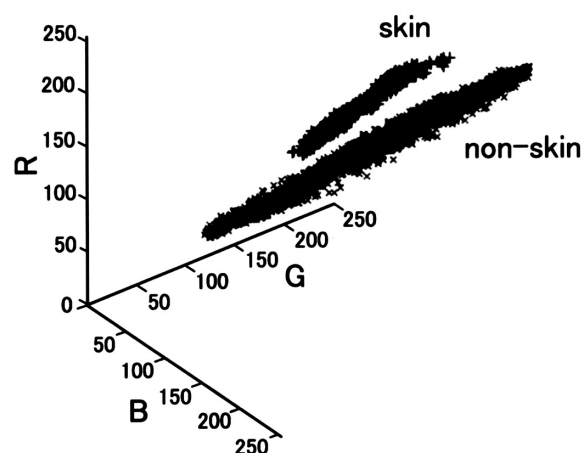


Fig. 2 RGB intensity of pixels in the skin area A and the non-skin area B

components are calculated using the following relations. In the following relations, Max and Min denote maximum and minimum values of R , G and B components respectively. H , L and S denote the same hue, saturation and lightness functions as the HSL color model. The L and S are, however, integer numbers between 0 and 512 while H is an integer number between 0 and 360. H is defined as the value shifted from the usual HSL color model in order to prevent H values related to skin area taking discontinuous values around 0 and 360. In what follows, H value of 120 represents red color.

$$L = Max + Min \quad 0 \leq L \leq 512 \quad (1)$$

$$H = \begin{cases} 120 + \frac{60(G-B)}{Max-Min} & Max=R \\ 240 + \frac{60(B-R)}{Max-Min} & Max=G \\ \frac{60(R-G)}{Max-Min} & Max=B \end{cases} \quad 0 \leq H \leq 360 \quad (2)$$

$$S = \begin{cases} \frac{512(Max-Min)}{L} & 0 < L \leq 255 \\ \frac{512(Max-Min)}{512-L} & 256 \leq L \leq 512 \end{cases} \quad 0 \leq S \leq 512 \quad (3)$$

Fig. 3 shows the hue, H , and saturation, S , values of the pixels in the skin area **A** and the non-skin area **B** depicted in Fig. 1. Hue values calculated from the pixels in the skin area **A** fall in narrow range, i.e., between 125 and 155, while those in the non-skin area **B** spread all the range, i.e., between 0 and 360, but with low saturation (spreading all the range is not always the case for other background area). Saturation values in the skin area **A** do not take lower values in this image: S values for the skin area almost greater than 50.

Fig. 3 gives a strategy to recognize a skin area in a captured image which includes a hand and a background: (1) hue values for a skin area take values within narrow range, i.e., between 125 and 155; (2) their dispersion is very

small; (3) saturation values for a skin area are not so small. Hue alone, however, is insufficient to recognize the skin because of the overlap in the range between 125 and 155.

Three features (a mean of hue values, H_m , dispersion of hue values, H_d , and a mean of saturation values, S_m , in a square area) can be used to realize above-mentioned strategy. The H_m and the S_m are estimated as:

$$H_m = \frac{1}{n} \sum_{x,y} H_{xy} \quad (4)$$

$$S_m = \frac{1}{n} \sum_{x,y} S_{xy} \quad (5)$$

Dispersion of hue values is calculated applying the Manhattan distance as follows:

$$H_d = \frac{1}{n} \sum_{x,y} |H_m - H_{xy}| \quad (6)$$

H_{xy} and S_{xy} are the hue value and the saturation value of

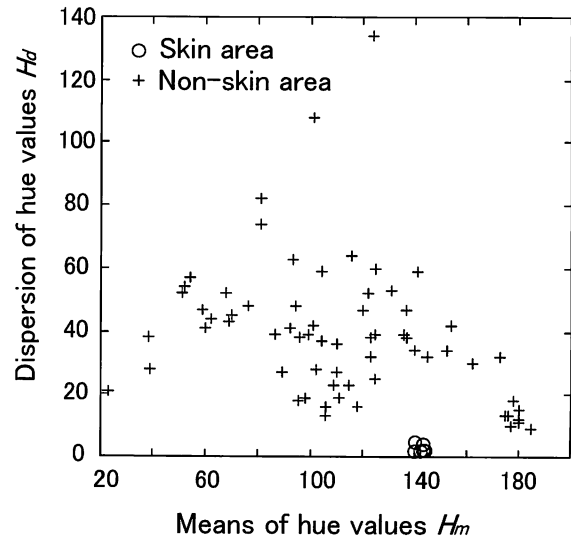


Fig. 4 Means and dispersion of the hue values in the skin area and the non-skin area

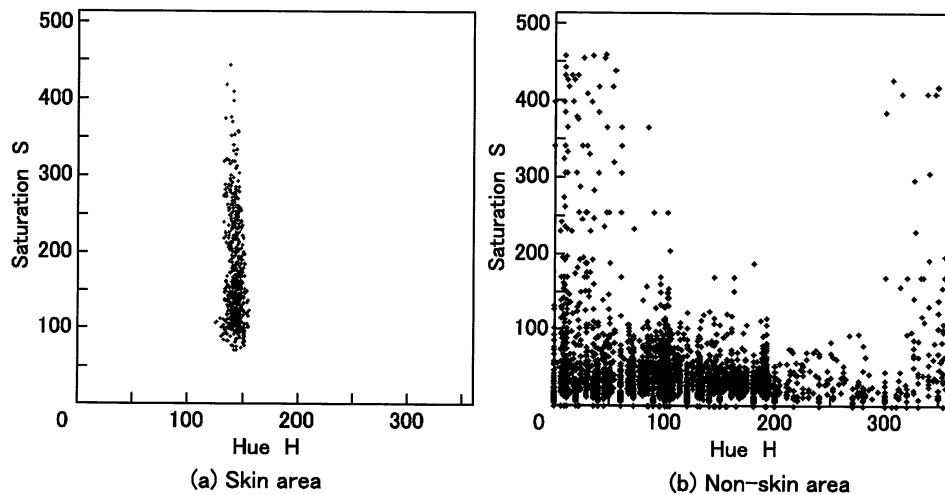


Fig. 3 Hue and saturation values for the skin area **A** and the non-skin area **B**

the pixel of the coordinate x, y and n is the number of pixels in the relevant square area. Fig. 4 shows the mean, H_m , and dispersion, H_d , of the hue values calculated using pixels in 10×10 square areas which are defined by sectioning close the skin area **A** and the non-skin area **B**. Fig. 4 reveals that a skin area can be distinguished from a non-skin area by using both H_m and H_d : the relevant area of a image can be regarded as a skin area when H_m is in a narrow range and H_d is very small. In what follows, an area having a H_m value between 120 and 160, a H_d value less than 10 and a S_m value greater than 50 is taken as a skin area, i.e., an area on a hand.

3. Hand contour recognition

A hand contour is recognized from a captured image in two steps: (1) find a square area on a hand; (2) trace the contour moving square skin areas that are bounded by non-skin background area.

At the first, a square area of 20×20 pixels at the center of the image is inspected by using the aforementioned criteria based on the mean of hue values, H_m , the dispersion of hue values, H_d , and the mean of saturation values,

S_m . If the area at the center is a skin area, then, a 20×20 square area neighboring to a non-skin background is searched while the square area is moved from the center to left direction. If not, a raster scan operation is used to find the boundary square skin area. Fig. 5 illustrates the first case: a center area **A** is judged as a skin area and a boundary square area **B** is found while the area is moved left from **A**.

A contour of the hand is traced by using square areas of 10×10 pixels starting at the square area **B** of Fig. 5. Eight 10×10 neighbor square areas (horizontal, vertical and diagonal direct neighbor areas of the current 10×10 boundary square area) are investigated based on the H_m , H_d and S_m to find next boundary square area. Then the current boundary area is updated. These operations are iterated until the current boundary area comes back to the square area **B** of Fig. 5. Fig. 6 illustrates the resultant boundary 10×10 square areas for the same image as Fig. 5. Almost the correct boundary is detected.

4. Hand sign recognition

In order to identify shape of the hand, especially a number of fingers stuck out, two parameters are investigated. One is a tangential direction, T , of the contour of the

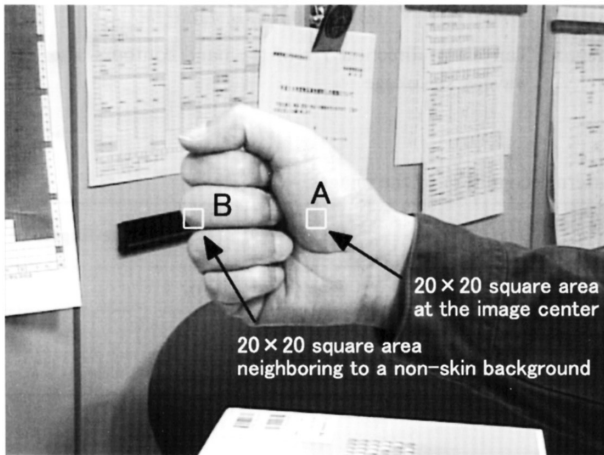


Fig. 5 Square area at center recognized as skin and a square area neighboring a background

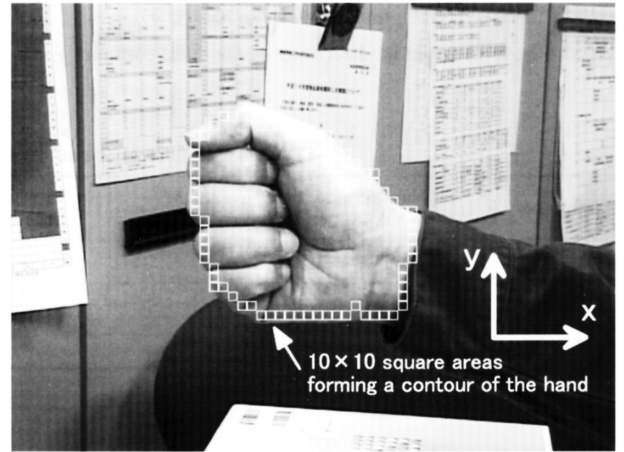
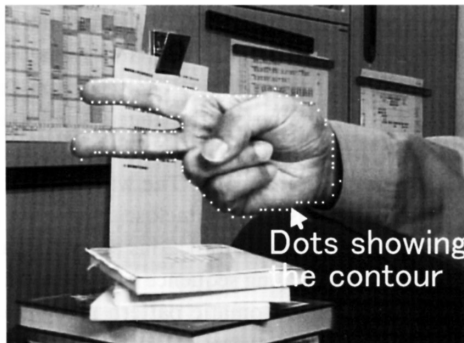
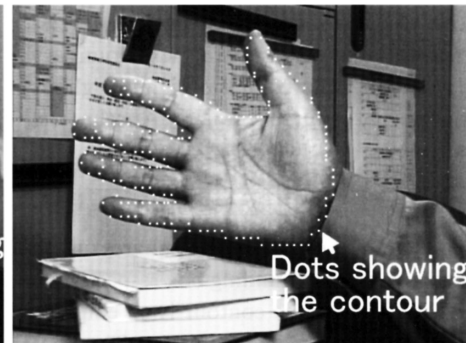


Fig. 6 Boundary 10×10 square areas of the hand



(a)



(b)

Fig. 7 Examples of hand images

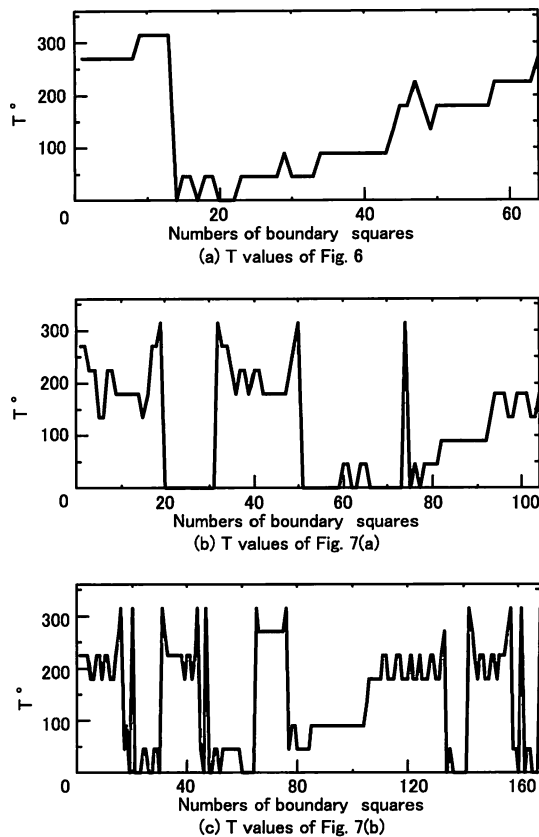


Fig. 8 Tangential directions of the hand contours

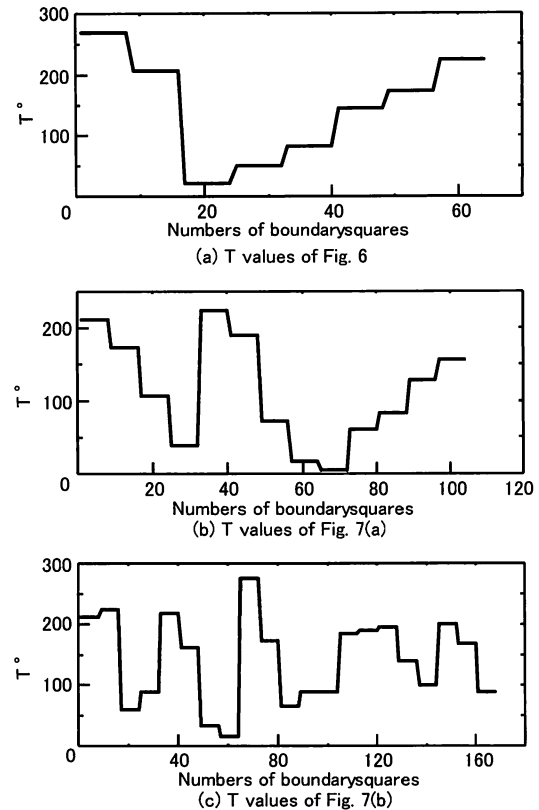
Fig. 9 Tangential direction T_s of the hand contours after the wavelet shrinkage

Table 1 Results of the automatic pattern recognition of the images

Images	Numbers of the convexes M in the tangential directions	Ratio r : the contour against the bounding box	Recognized number of the fingers
Fig. 6	1	0.78	0 or 1
Fig. 7(a)	2	1.03	2
Fig. 7(b)	5	1.33	5

hand. The tangential direction, T , is one of the eight (two horizontal, two vertical and four diagonal) directions of the current boundary square: $T=0^\circ, 90^\circ, 180^\circ$ and 270° mean $x, -y, -x$ and y direction of the coordinate shown in Fig. 6, respectively. T is decided by the positional relation among the former boundary square, next boundary square and the non-skin square in the eight neighbor squares around the current boundary square. Another is a ratio r defined as follows:

$$n = \frac{n_C}{n_B} \quad (7)$$

where n_C and n_B are numbers of the 10×10 squares on the contour of the hand and on the bounding box respectively. The ratio r is less than 1 when the contour is an ellipse.

Fig. 7 shows two examples of hand images. The contours shown, only for illustrations, in Fig. 7 by dots are determined based on the positional relations between the boundary square areas and the non-skin background

areas.

Fig. 8 shows the tangential directions, T_s , of the contours of these hands in Fig. 6, Fig. 7(a) and Fig. 7(b). A change of tangential direction, T , over 90° means a convex or a concave. The changes of the tangential directions vary by the outward form of the hands, so using the tangential directions of the contours can recognize the shapes of the hands: the number of the fingers stuck out. The changes shown in Fig. 8, however, exhibit variations tough to recognize automatically. The wavelet shrinkage using the Haar wavelet is used to unclothe the feature patterns. Fig. 9 shows the tangential directions after the wavelet shrinkage by using the scaling sequence of level 3. Table 1 shows the results of the automatic pattern recognition; that is, the number of the convexes, M , of Fig. 9 and the ratio, r . One convex in Fig. 9 means two changes of the tangential directions: the recognized hand is regarded as having a shape like ellipse including a shape with one finger stuck out.

Two convexes in Fig. 9 means a shape with two fingers.

5. Conclusion

A method to recognize hand shape, i.e., numbers of fingers, using a color image is presented in this paper. Skin areas in an image are distinguished from a non-skin area by valuating three features: a mean and dispersion of hue values and a mean of saturation values, applying the HSL color model, in a small square area of an image are calculated for skin-recognition. Moving the small square area on a captured image and valuating the three features coordinates and tangential directions of a contour of a hand are estimated. Changes of the tangential directions of a contour reflect a shape of the hand, that is, a number of fingers stuck out. Although only a few images evaluated, numbers of the fingers stuck out were recognized automat-

ically.

References

- 1) J. Isaacs and S. Foo, Optimized Wavelet Hand Pose Estimation for American Sign Language Recognition, Proceedings of the 2004 Congress on Evolutionary Computation Vol.1, 797–802.
- 2) N. Kawarazaki et al., Cooperative Work System between Human and Manipulator using Hand Gesture Instructions (in Japanese), Journal of the Robotics Society of Japan, Vol.23, No.6 (2005), 121–126.
- 3) M.A. Akbari and M. Nakajima, A New Skin Detector for Different Color Spaces, Technical report of IEICE, IE2004–143, 111–115.
- 4) M. TAKAGI, H. SHIMODA (Eds.), Handbook of Image Analysis (in Japanese), University of Tokyo Press, (2004), 1191–1193.