International Journal of Latest Research in Science and Technology **ISSN (Online):2278-5299**
Volume 3, Issue 4: Page No.65-71. July-August 2014
https://www.mnkpublication.com/journal/ijlrst/index.php

GESTURE SEGMENTATION USING AN ADAPTIVE THRESHOLD ALGORITHM

[1]Mei Wang, [2*]Jzau-Sheng Lin, [1]Zhou Xing Fu, [1]Guo Qing Meng
[1]Xi'an University of Science and Technology, College of Electrical and Control Engineering, Xi'an, China
[2] Department of Computer Science and Information Engineering, National Chin-Yi University of Technology, Taichung, Taiwan
[*]Corresponding Author, e-mail: jslin@ncut.edu.tw, Phone:+886-4-23924505 ext 8713
No.57, Sec. 2, Zhongshan Rd., Taiping Dist., Taichung 41170, Taiwan

Abstract*- *Hand gesture segmentation is a key step for gesture recognition. Based on the construction of a new color space of skin model, a new dynamic-thresholding segmentation approach named Adaptive Threshold Segmentation Algorithm (ATSA) was further developed and segmentation effect evaluation was conducted. Some images of hand gesture were processed by using ATSA and the Fixed Threshold Segmentation (FTS) algorithm as well as the Similarity algorithm of Skin Color (SSC). Comparing with FTS and SSC algorithms, The ATSA is experimentally demonstrated that, the segmentation results have a less brightness impact, a lower redundancy rate, a lower rate of false alarm and missing, and a higher integrity rate.

Keywords - Image Segmentation; Adaptive Threshold, Hand Gesture

I. INTRODUCTION

Hand gesture segmentation is a key step for hand gesture recognition on computer vision. As a new human-computer interaction interface technology, hand gesture recognition is a high-profile research field and application technology [1, 2]. In recent years, vision-based gesture recognition on computer vision abstracts the hand shape, hand pose, hand position, motion and other information from the gestures to explain the actions of human in the forms of text, virtual human and the other forms. It is widely used in clinical medicine, human-computer interaction, remote communication, language translation and other fields. For example, the remote control robot acts according to the movements of commander's hands. In these applications, hand gesture segmentation is the first and the most crucial step of the hand gesture recognition based on computer vision. The quality of hand gesture segmentation directly impact on the results of hand tracking, feature extraction and hand gesture recognition [3, 4].

Early gesture recognition mainly used sensor gloves or marked special signals on the hand. The human-computer interactions were implemented by the signal processing such as the instrument gloves to recognize the hand gesture [5]. The advantages of this approach are a few input data needed, high recognition rate, and good real-time performance respectively. But it required expensive equipment and unnatural interactive actions.

The gesture segmentation approaches were implemented usually by setting the threshold in the RGB color space so that the pixels in the hand gesture region are separated from the background on an image. But these approaches were seriously impacted by the lighting condition. To solve this problem, Rafael et al. [6] used HSV color space to segment image. Later Ankit and Sayjay [7] adopted the skin color segmentation method by using the ratio of R/G. Also, Dawod et al. [8] segmented the skin color in the YCbCr color space. Ma and Liu [9] as well as Gupta and Ma [10] used the skin color to segment and recognize hand gesture.

When the color information is used for the segmentation, the segmentation accuracy of hand gesture is closely related to the background complexity and the similarity between the background color and the skin color. Furthermore, the segmentation threshold of color model should satisfy a certain condition. Once the lighting condition was changed, it will have an impact on the segmentation results [4, 11].

Currently, the methods for hand gesture segmentation included mainly 3 types [12]. The first type of hand gesture segmentation is detection and segmentation based on the color information. This method has high efficiency and robustness. This is because the skin color is relatively special when it is compared with the other objects and the skin color would not be changed too much when the size of gesture image and the rotation of gesture were changed.

The second type of the hand gesture segmentation is to detect motion information based on the segmentation. The background is often relatively stationary to the human hand in an image. The gesture and background could be separated by establishing and comparing the background. A dynamic 3D human posture modeling was built by using the optical flow [13]. This method is able to compensate for some deficiencies of the color-information-based detection. However, the gesture action should be fast and the background should be relatively static.

The third type is the clustering algorithms of image pixel similarity. Such as, the probability, Bayesian decision rule, and mean-shift algorithm were used for image segmentation [14, 15]. Although such segmentation methods were intelligent, it should be adjusted the parameters when the environment was changed owing to the foreground and background being utilized [16, 17].

Publication History
Manuscript Received : *15 August 2014*
Manuscript Accepted : *26 August 2014*
Revision Received : *28 August 2014*
Manuscript Published : *31 August 2014*

Taking account of the impact by the environment change, the ATSA was developed for gesture segmentation based on a new skin color space model. This paper is organized as follows. The related theories of HSV color space, conversion of color space, binary segmentation, and HSV color segmentation model are introduced in Section 2. The proposed ATSA is developed in Section 3 which includes the definitions of image color distance, the construction of a new color space model, adaptive threshold rule, and segmentation evaluation respectively. Experimental results by using the proposed ATSA, the FTS, and the SSC strategies are conducted in Section 4. Finally, Section 5 demonstrates the conclusions.

II. RELATED WORKS

In order to segment hand gesture on an image effectively, a suitable color space is needed. The color space should keep the human skin for extracting gesture. Besides, the space should not be sensitive to the change of illumination. In image processing, RGB color space model is the most popular. But it has seriously flawed in the applications of machine vision because it is difficult to adjust the digital detail. The hue, brightness and saturation are mixed together and difficult to be separated from each other. The HSV color space model is designed for digital color processing. It is a non-linear color space created with hue (H), saturation (S), vision (V) brightness. It is also based on the color of visual characteristics. In HSV color space, H, S, and V are defined as

$0^\circ < H < 360^\circ$, $0 < S < 1$, and $0 < V < 1$.

According to the reference [18], HSV space could be obtained by the conversion from RGB space to HSV space which is shown as below from Eq. (1) to Eq. (4).

$$H = \begin{cases} H_1 / 360 \ , B \le G \\ 360 - H_1 \ , otherwise \end{cases} \tag{1}$$

$$S = \frac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B)} \tag{2}$$

$$V = \frac{\max(R,G,B)}{255} \tag{3}$$

where

$$H_1 = \cos^{-1} \frac{0.5[(R-G)+(R-B)]}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \tag{4}$$

and R, G, B represent the intensity of the three channels of red, green, and blue in RGB space, respectively.

Usually, the hue and the saturation of the object are determined by the characteristics of light absorption and reflection of the material as well as the brightness is affected apparently by the light and the vision. Therefore, image segmentation is more reliable according to the hue and saturation. The segmentation formula is shown in Eq. (5) if F is the binary image after segmentation [19].

$$F = \begin{cases} a_0, \ if \ H_l < H < H_u \quad \text{and} \quad S > S_u \\ a_1, \quad otherwise \end{cases} \tag{5}$$

where H is the hue, H_l and H_u are the lower and the upper boundaries of H, respectively; S is the saturation, and S_u is the upper boundaries of S; a_0 and $a1$ are two constants. a_0=0 and a_1=1 that results in a binary image F with black and white. Human skins are mainly concentrated in the red zone. This characteristic could be represented by $R > G > B$. Even if the color of skin or the ambient light is changed, the feature is quite stable. Inserting Eq. (3) into Eq. (1) and Eq. (2), the following relation could be obtained.

$$H = \frac{(B-G) * \pi/3}{R-B}, \quad S = \frac{R-B}{R}, \quad V = R \tag{6}$$

then

$$H = \frac{\pi(B-G)}{3(R-B)}, \quad S' = S * V = R - B, \quad V = R \tag{7}$$

where H, S', and V are used as a segmented space model in the FTS algorithm. This is because that H, S' and V are more concentrated for human skin.

The centroid of skin region is usually located in the skin region. For a hand skin region, the centroid is located in the palm. The coordinate (x_c, y_c) of centroid can be calculated as

$$x_c = \frac{m_{10}}{m_{00}}, \quad y_c = \frac{m_{01}}{m_{00}} \tag{8}$$

where

$$m_{10} = \sum_{x=1}^{w} \sum_{y=1}^{h} x f(x,y), \quad m_{01} = \sum_{x=1}^{w} \sum_{y=1}^{h} y f(x,y), \quad \text{and}$$

$$m_{00} = \sum_{x=1}^{w} \sum_{y=1}^{h} f(x,y). \tag{9}$$

where $f(x, y)$ is the intensity value at the coordinate (x, y).

FTS method is a traditional image segmentation method which is the most commonly used because of its simplicity, small amount of calculation and more stable performance. Let the original image is $f(x, y)$, the fixed threshold value T could be found according to a certain criteria, then the image could be transformed into a binary image $g(x, y)$ obtained by the following formula.

$$g(x,y) = \begin{cases} b_0 & if \quad f(x,y) < T \\ b_1 & if \quad f(x,y) \ge T \end{cases} \tag{10}$$

where T is a fixed threshold value; b_0 and b_1 are respectively two intensity values of the image after segmentation.

Besides, the extended centroid method was taken to compute the centroid of skin color region and get the hand contour because the color information of this method is adequate and computation amount is acceptable. The Hand

Contour Extraction (HCE) method is directed as follows from the modified extended centroid method [20].

Step1: Set the hand region with the intensity of 255 and 0 on other regions.

Step2: Calculate the centroid $M=(x_c, y_c)$ using Eq. (8).

Step3: Starting with the centroid M, the rectangle can be extended by a length of n in up, down, left and right directions.

Step4: The extended processes are continued until contour edges are reached or exceeded by all of four sides of the extended rectangle. Then, record locations for these four sides.

Step5: Extract the skin color from the rectangular area of the test image with the extended centroid method.

III.ADAPTIVE THRESHOLD ALGORITHM

Normally, we select a fixed threshold for gesture segmentation when image background is fixed or has little change in brightness. But the result of segmentation is not good when the background or the brightness changes obviously. To solve this problem, the presented ATSA was developed in which the threshold could be automatically updated for the variant brightness.

It is easy to identify the difference of the colors on an image. But it is not quite obvious to identify the difference between two images with same backgrounds. So the image color distance must be defined in advance. Assume there are two images, $f(x, y)$ and $g(x, y)$, and their mean values are

$$Aver_f = \frac{\sum_x \sum_y f(x,y)}{M_1 * N_1} \tag{11}$$

$$Aver_g = \frac{\sum_x \sum_y g(x,y)}{M_2 * N_2} \tag{12}$$

where $M_1 * N_1$ is the number of total pixels on image $f(x, y)$, and $M_2 * N_2$ is the number of total pixels on image $g(x, y)$. After the mean-value filtering, the image color distance d between image f and image g is defined as

$$d = \left| Aver_f - Aver_g \right| \tag{13}$$

where $Aver_f$ and $Aver_g$ are mean values of images f and g.

The segmented results on an image were changed when the environmental light is changed. That is the white light or the colour intensity being variant. Normally, the values of R, G, and B in the white light are basically equal to each other, and the increment quantities of $\Delta R, \Delta G, \Delta B$ are almost considered the same. But, there are some differences among them in fact. The values of R, G, B are not seriously equal in color intensity, and $\Delta R, \Delta G, \Delta B$ are significantly different when the brightness is changed. Therefore, we have the following relationship.

$$\begin{cases} S*V = (R-B) + (\Delta R - \Delta B) \\ V = (R + \Delta R) \\ H = \dfrac{[(B-G) + (\Delta B - \Delta G)] * \pi / 3}{(R-B) + (\Delta R - \Delta B)} \end{cases} \text{ then} \tag{14}$$

$$\begin{cases} S*V = (R-B) + (\Delta R - \Delta B) \\ V = (R + \Delta R) \\ H*S*V = [(B-G) + (\Delta B - \Delta G)] * \pi / 3 \end{cases} \tag{15}$$

When the image brightness is changed, we still believe that the image satisfies $R + \Delta R > G + \Delta G > B + \Delta B$. Analyzing Eq. (15), we can know that the threshold of $S*V$ is changed only by the value of $\Delta R - \Delta B$, and it has nothing to do with ΔG. V is affected only by ΔR. $H*S*V$ is influenced by $\Delta B - \Delta G$ and it has nothing to do with ΔR. So we selected $H*S*V$, $S*V$ and V as the segmentation space model. After the above analysis, we develop the threshold adjustment method as follows.

$$T_{S'}(K+1) = T_{S'}(K) + T_0[\Delta R(K+1) - \Delta B(K+1)] + R_0 \tag{16}$$

$$T_{HSV}(K+1) = T_{HSV}(K) + R_1 + \frac{\pi T_1[\Delta B(K+1) - \Delta G(K+1)]}{3 * T_{S'}(K+1)} \tag{17}$$

$$T_V(K+1) = T_V(K) + T_2 \Delta R(K+1) + R_2 \tag{18}$$

where $T_{S'}$, T_{HSV}, and T_V are respectively the threshold values of the S', $H*S*V$, and V in the space model on the current image. ΔR, ΔB and ΔG are respectively the variant value of each channel on the image; T_0, T_1, and T_2 are the adjustment parameters; R_0, R_1, and R_2 are the current coefficients of relaxation and they can be constants or can be adjusted automatically according to the changes of ΔR, ΔB, and ΔG.

In order to avoid the inconvenient threshold selection process owing to large value and a broad range of $H*S*V$, H is normalized firstly, and then $H*S*V$ is calculated to obtain the model H'.

$$H' = \frac{H}{360} * S * V = \frac{\pi}{3} * [(B-G) + (\Delta B - \Delta G)] \Big/ 360 \tag{19}$$

The threshold adjustment rule should be designed to determine the state in which the threshold should be adjusted. That is what conditions are satisfied and then the threshold adjustment should be started. For this purpose, we designed the threshold adjustment rule shown in Fig. 1.

In Fig.1, N and $N+n$ are the serial numbers of image frames; P_U and P_L are the upper and the lower protection limits. If the thresholds of n frames skip beyond the two protection limits, the threshold should be maintained. This case is shown with the red curve. This rule could decrease the disturbances of the threshold adjustments from the wrong gesture regions, and could not lead to the malfunctions because of the hand gesture motions too fast.

D_U and D_L are the upper and the lower threshold adjustment limits respectively. The effective threshold adjustment regions are respectively constructed by P_U, D_U and P_L, D_L. If the threshold values of the n frames reach the effective regions, then the threshold would be adjusted. This case is shown with the blue curve. If the threshold values of the n frames change slowly and do not reach D_U and D_L, then the new threshold is equal to the mean value of the n frames.

In this section, we divide the adaptive threshold adjustment into 4 kinds of situation. The first is the brightening case in which the threshold is changed into the effective threshold adjustment range of Range1. The second is the dimming case in which the threshold is changed into the effective threshold adjustment range of Range2. The third is the disturbance case in which the threshold is changed between the effective threshold adjustment ranges of Range1 and Range2. The fourth is the malfunction case in which the threshold is changed beyond the effective threshold adjustment ranges of Range1 and Range2.

Before we developed the ATSA algorithm, the contour extraction algorithm should be firstly described because it plays an important role in the process of ATSA. Hand contour extraction algorithm is designed as follows.

Step 1: Extract the outer contour of image and save it sequentially.

Step 2: Set the parameters a_0, a_1 of contour judgment criterion, and the parameters S_0, S_1 of statistical judgment criterion.

Step 3: If the coordinate of a pixel $f(x, y)$ on a contour satisfies one of the following conditions, then remove the contour.

$$x < a_0 \text{ or } x < i_w - a_0 \tag{20}$$

$$y < a_1 \text{ or } y < i_h - a_1 \tag{21}$$

$$\sum f(x,y) < S_0 \text{ or } \sum f(x,y) > S_1 \tag{22}$$

where i_w and i_h are the width and the height of an image, respectively.

Step 4: If no hand is found after traveling all over the contours, go back to step (2), and then the parameters would be adjusted. Otherwise, hand contour is found and the contour extraction process is ended.

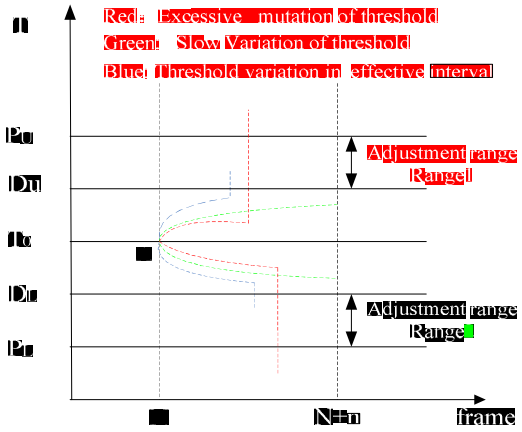


Fig. 1 Adaptive threshold rule

For each contour, conduct the above process in the same way. The parameters of a_0, a_1, S_0 and S_1 are set according to the image size and the distance from the camera. After the contour extraction, the ATSA is developed as follows.

Step1: Set the initial values of segmentation threshold $T_S(0)$, $T_H(0)$ and $T_V(0)$, skin color mean value S_M, the upper and the lower limits D_{MU} and D_{ML} of the mean value difference, the upper and the lower skin color limits S_U and S_L, the upper and the lower limits difference D_{SU} and D_{SL} of skin color, and the threshold protection limit ΔP.

Step 2: Input a gesture image and project the gesture image into HSV space, then conduct the preprocessing of mean-shifted filtering.

Step 3: Decompose the image into HSV space, and calculate the skin color segmentation model H', S' and V.

Step 4: Calculate the thresholds of H', S' and V for color segmentation model, and obtain the binary image *Img_H*, *Img_S* and *Img_V* after the threshold processing. 3

Step 5: Fuse image *Img_H*, *Img_S* and *Img_V* by the operation "and", then obtain the fused segmentation image *Img_and*.

Step 6: Get the hand region by the morphological processing, remove the small regions, extract the hand contour by using the above hand contour extraction algorithm, and detect the hand location.

Step 7: Obtain the hand gesture on the original image.

Step 8: Take the skin color region from the original gesture image, and calculate the mean value $S_{M'}$ and its histogram. The upper and the lower skin color limits $S_{U'}$ and $S_{L'}$ are determined by the histogram.

Step9: If one of the following constraints is satisfied, go back to step 2. Otherwise, stop.

$$abs(S_{M'} - S_M) < D_{ML} + \Delta P \tag{23}$$

$$abs(S_{M'} - S_M) > D_{MU} + \Delta P \tag{24}$$

$$abs(S_{U'} - S_U) < D_{SL} + \Delta P \tag{25}$$

$$abs(S_{U'} - S_U) > D_{SU} + \Delta P \tag{26}$$

$$abs(S_{L'} - S_L) < D_{SL} + \Delta P \tag{27}$$

For the effectiveness evaluation of the hand gesture segmentation, there is almost no the uniform standard. Therefore, we define four kinds of evaluation indexes as follows.

Definition of Redundancy Degree R_e: R_e is defined as an absolute value of the difference between the amount pixels A of the detected region and the amount pixels B of the hand region.

$$R_e = |A - B| \tag{28}$$

The smaller R_e , the better segmentation result is.

Definition of Integrity Rate *IR*: *IR* is a ratio in which the numerator is the absolute value of the difference between the pixel number N_a of actual target and the pixel number N_i of ideal target. The denominator is N_i.

$$I = \left| N_a - N_i \right| / N_i \tag{29}$$

The smaller *IR*, the better the segmented result is.

Definition of the Rate R_{fm} of False Detection or Missing Detection (FDMD): For the detection of multi-frame images, the ratio of the number of the false detected images and the missing detected images to the total number of the detected images is defined as the rate R_{fm} of false detection or missing detection.

$$R_{fm} = \left| N_1 + N_2 \right| / N \tag{30}$$

where N_1 is the number of false detected images, N_2 is the number of missing detected images, and N is the total number of detected images.

Definition of Brightness Impact (BI): Observe the segmentation results meanwhile change the brightness. If the object disappears after segmentation, the BI is considered strong. If the object exists but some information is missed after segmentation, the BI is considered moderate. If the object information is still integrated after segmentation, the BI is considered small. The smaller BI, the better segmentation result is.

IV. EXPERIMENTAL RESULTS

In this section, we segment the hand gesture by using the ATSA to compare with FTA and SSC strategies. The processing platform is C++6.0 and Opencv 1.0 library. The light source is the incandescent light. The images are collected from the camera of the laptop. The size of the experimental images is 320∗240 mm^2.

Fig. 2 shows the hand gesture images and segmentations using ATSA algorithm and FTS algorithm. For the original hand gesture images in Fig. 2 (a) ~ (b), the hand gesture regions are firstly detected in the red rectangles, and the small skin regions are secondly extracted from the hand gesture regions in the white rectangles. Finally the experimental data of *H*, *S′* , and *V* are obtained and shown in Table 1 with *S′*=*S***V*. Meanwhile, the experimental data of *H′*, *S′*, and *V* can be also obtained. In order to simplify calculation, the actual data of color model H'S'V are converted into the range of 0-255. For Gesture1 and Gesture2, the relative data of H′, S′ and V are also shown in Table 1. MeanBC and RangeBC represent respectively the mean value and the range before conversion while MeanAC and RangeAC indicate the mean value and the range after the conversion, respectively.

To illustrate the effect of the constructed color space H'S'V, Gesture1 and Gesture2 were segmented by using the FTS algorithm with HS′V model and the ATSA algorithm with the constructed H′S′V model. Figs. 2(c) and 2(d) are the segmented results of Gesture1 by the FTS algorithm with HS′V model and by the ATSA algorithm with H′S′V model respectively while Figs. 2(e) and 2(f) are the segmented results of Gesture2 by the FTS algorithm with HS′V model and by the ATSA algorithm with H′S′V model individually.

From Fig. 2, it can be seen that the developed ATSA with the constructed space model H′S′V results in a better segmentation performance than the FTS algorithm with the space model *HS′V*.

After the above process, the data and their difference between Gesture1 and Gesture2 were obtained in Table 2. Then $\Delta R - \Delta B = -7.272$, $\Delta B - \Delta G = 6.46$, and $\Delta R = -74.56$. Let $T_0 = T_1 = T_2 = [0.5, 0.8]$, $R_0 = R_1 = R_2 = 0$, the updated thresholds would be calculated according to Eq. (16), Eq. (17), Eq. (18) and Eq. (19) as follows.

$$T_{S'}(1) = \begin{cases} 1 + 0.5 * (-7.272) = -2.636 \,(\text{minus} \Rightarrow 0) \\ 22 + 0.8 * (-7.272) = 16.1824 \end{cases}$$

$$T_{H'}(1) = \begin{cases} 1 + 0.5 * \dfrac{\pi * (6.46)}{3 * THR\_S'(1)} = 1.2089 \\ 30 + 0.8 * \dfrac{\pi * (6.46)}{3 * THR\_S'(1)} = 30.345 \end{cases}$$

$$T_V(1) = \begin{cases} 110 + 0.5 * (-74.56) = 72.72 \\ 220 + 0.8 * (-74.56) = 160.35 \end{cases}$$

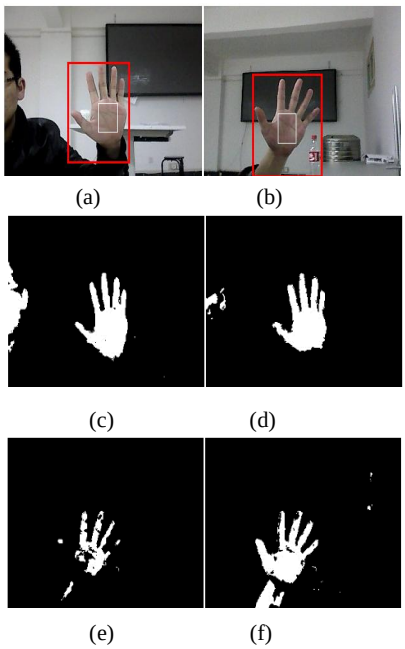


Fig. 2 Hand gesture images and segmentations using ATSA algorithm and FTS algorithm: (a) Gesture1, (b) Gesture2, (c) Segmented result of Gesture1 by FTS algorithm with HS′V model, (d) Segmented result of Gesture1 by ATSA algorithm with H′S′V model, (e) Segmented result of Gesture2 by FTS algorithm with HS′V model, and (f) Segmented result of Gesture2 by ATSA algorithm with H′S′V model.

Table I. H', S' and V before and after the conversions of Gesture1 and Gesture2

Item	MeanBC	RangeBC	MeanAC	RangeAC
H′ of Gesture1	0.006	0.006-0.007	15.702	1-30
S′ of Gesture1	0.083	0.05-0.10	21.133	1-24
V of Gesture1	0.716	0.65-0.76	182.497	110-220
H′ of Gesture2	0.012	0.001-0.009	21.005	0-26
S′ of Gesture2	0.112	0.008-1.13	14.271	0-17
V of Gesture2	0.451	0.35-0.53	115.067	70-150

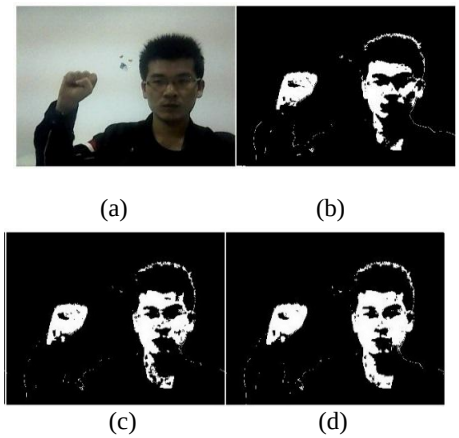
Table 2. Data and difference between Gesture1 and Gesture2

Means of Gesture1			Means of Gesture2			Difference		
$\overline{R}$	$\overline{G}$	$\overline{B}$	$\overline{R}$	$\overline{G}$	$\overline{B}$	$\overline{R}$	$\overline{G}$	$\overline{B}$
161.33	166.40	182.44	86.77	92.65	115.16	74.56	73.75	67.28

Experimental results show that the segmentation result of Gesture2 has the multiple fractures in hand region. The result shows that the proposed ATSA algorithm and composed H′S′V model improved the effect of segmentation at a certain extent. Furthermore, the integrity rates ***IR*** of segmentation results can be calculated according to Eq. (29). The integrity rates ***IR*** are 0.2 and 0.7 by the ATSA and the FTS algorithms respectively that implies the performance of ATSA is better that the FTS.

Fig. 3 shows the segmentations by using the ATSA and the SSC algorithms. Fig. 3(a) is the original image. Fig. 3(b) is segmentation result using the ATSA algorithm in H′S′V space. Figs. 3(c) and 3(d) are the segmentation results using the SSC algorithm with similarity levels L=11 and L=30 respectively. In a simple background case, the ATSA and the SSC algorithms reach the almost same segmentation result. But in complicated cases of different backgrounds, the ATSA algorithm will obtain much better segmentation results than the SSC.

To illustrate the influence of brightness for the segmentation, we took 1 image every 10 images from a hand gesture video, and the 56th, 66th, 76th and 86th frames are totally taken and shown in Figs. 4(a) ~ (d), respectively. The brightness of every image is different from each other because the foregrounds and backgrounds are not in same states. Fig. 4(a) shows gesture with background light; the foreground and background light are both added in Fig. 4(b); Fig. 4(c) displays gesture with foreground light; and the background light is put in Fig. 4(d).



(a) (b)

(c) (d)

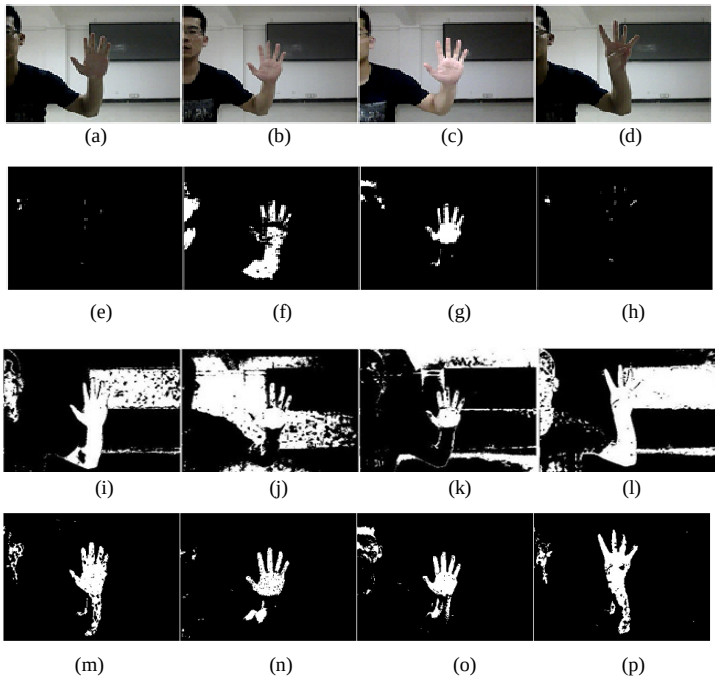
Fig. 3 Segmentations using ATSA algorithm and SSC algorithm: (a)original image, (b) ATSA algorithm in H′S′V space, (c)SSC algorithm with similarity level L=11, and (d) SSC algorithm with similarity level L=30.

Segmentation results using FTS, SSC and the proposed ATSA algorithms are shown from Fig. 4(e) to Fig. 4(p). Segmentation results using the FTS algorithm in HS′V space for the 56th, 66th, 76th, and 86th frames are shown in Figs. 4(e) ~ 4(h), respectively. Figs. 4(i) ~ 4(l) show the segmentation results by using of the SSC algorithm for the 56th, 66th, 76th, and 86th frames, individually. Segmentation results by using of the ATSA algorithm in H′S′V space for the 56th 66th, 76th, and 86th frames are displayed in Figs. 4(m) ~ 4(p). Table 3 shows relative data for these 4 images. From Fig. 4, it can be seen that the changes of brightness have obvious impact on the segmented results. This is because the threshold value is impacted by the change of brightness. Clearly, Fig. 4 displays worse segmented results by using of FTS and SSC in complex background.

For the purpose of objective viewpoint, the properties of Redundancy Degree, Integrity Rate, Rate of FDMD, and Brightness Impact for the segmented results of different algorithms are compared in Table 4. From Table 4, we can indicate that the proposed ATSA can obtain better segmentation results than those got by FTS in HS′V and SSC space, respectively. The ATSA is applied to hand gesture recognition in a test platform. The recognition processes are shown in Fig. 5. In the recognition platform, it can obtain better segmentation results in fingertips detection and gesture recognition.

Table 3 Data of 4 images from a hand gesture video

Item	56th	66th	76th	86th
Mean of R	55.727	165.325	108.442	48.774
Mean of G	65.285	167.191	115.598	59.215
Mean of B	77.627	193.915	137.474	72.982
ΔR	108.432	109.598	-56.883	-59.668
ΔG	100.385	101.400	-51.593	-56.383
ΔB	110.206	116.288	-56.441	-64.492
Threshold $T_{S'}$	0-15	0-9.648	0-9.648	0-14
Threshold $T_{H'}$	0-35	0-35	0-35	0-35
Threshold $T_{V'}$	60-150	118-253	89-207	59.7-159



(a) (b) (c) (d)

(e) (f) (g) (h)

(i) (j) (k) (l)

(m) (n) (o) (p)

Fig. 4 Different brightness and segmentation results using FTS, SSC and ATSA: (a) image of 56th, (b) image of 66th, (c) image of 76th, and (d) image of 86th; (e) ~ (h) Segmented by the FTS from (a) to (d); (i) ~ (l) Segmented by the SSC

from (a) to (d); and (m) ~ (p) Segmented by the ATSA from (a) to (d).

Table. 4 Property comparison of segmentations

Property	SSC	FTS in HS′V	ATSA in H′S′V
Redundancy Degree	0.9	0.6	0.4
Integrity Rate	0.5	0.7	0.2
Rate of FDMD	0.2	0.5	0.2
Brightness impact	Middle	Large	Small

V. CONCLUSIONS

This paper presents a gesture segmentation method named ATSA algorithm based on the new constructed color space model HSV. The proposed ATSA algorithm can reduce the brightness impact of background environment change, improve the redundancy degree of gesture segmentation, increase the integrity rate of the segmentation, and reduce the rate of false detections or missing detections.

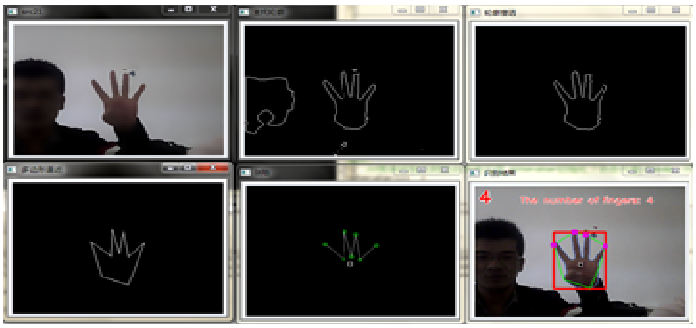


Fig. 5 The application of ATSA in a hand gesture recognition platform

ACKNOWLEDGMENT

This paper is supported by Scientific Research Foundation for Returned Scholars, Ministry of Education of China ([2011]508), Natural Science Foundation of Shanxi Province of China (2011JM8005), and sponsored by the National Science Council of Taiwan under the Grant NSC101-2221-E-167-040.

REFERENCES

[1] Guan, R. and Xu, X. M., "A computer vision-based gesture detection and recognition technique," Computer Applications and Soft-ware. 30, pp. 155-164, 2013.

[2] Choi, J., Park, J., Park, H., and Park, J, "iHand: an interactive bare-hand-based augmented reality interface on commercial mobile phones," Opt. Eng. 52, 027206 [doi: 10.1117/1.OE. 52.2.027206], 2013.

[3] Liu, J. M. and Yuan Q. Q., "A new method of hand gesture segmentation with complex backgrounds," Journal of Beijing Electronic Science and Technology Institute. 14, pp. 23-27, 2006.

[4] Nolker, C. and Ritter, H., "Visual recognition of continuous hand postures," IEEE Trans. on Neural Networks. 13, pp. 983-994, 2002.

[5] Pradhan, G., Prabhakaran, B., and Li, C. J., "Hand-gesture computing for the hearing and speech impaired," IEEE Trans. on MultiMedia. 15, pp. 20-27, 2008.

[6] Rafael, B., Miguel, S. D, and Miguel, S., "Skin color profile capture for scale and rotation invariant hand gesture recognition," Lectures Notes in Computer Science. 5085, pp. 81-92, 2009.

[7] Ankit, A. B. and Sanjay, N. T., "Vision-based Authenticated Robotic Control using Face and Hand Gesture Recognition," Proceeding on the 3rd International Conf. of Electronics Computer Technology, pp. 64-68, 2011.

[8] Dawod, A. Y., Abdullah, J., and Alam, M. J., " Adaptive skin color model for hand segmentation," Int. Conf. on Computer Applications and Industrial Electronics, pp. 486-489, 2010.

[9] Ma, G. Y. and Liu, X. Y., " Operational gesture segmentation and recognition," Tsinghua Science and Technology,. 8, pp. 169-173, 2003.

[10] Gupta, L. and Ma, S. W., "Gesture-based interaction and communication: automated classification of hand gesture contours," IEEE Trans. on Systems, Man, and Cybernetics, Part C: Applications and Reviews. 31, pp. 114-120, 2001.

[11] Li, C. M., Huang, R., Ding, Z. H., Gatenby, J. C., Metaxas, D. N., and Gore , J. C., "A level set method for image segmentation in the presence of intensity in homogeneities with application to MRI,". IEEE Trans. on Image Processing. 20, pp. 2007 – 2016, 2011.

[12] Yang, D. G., Wan, H. J., and Yang, M., "Natural image segmentation based on contour detection," Journal of Huazhong Normal University (Natural Sciences).46, pp. 18-21, 2012.

[13] Schwarz, L. A., Mkhitaryan, and A., Mateus, D., "Estimating human 3D pose from time-of-flight images based on geodesic distances and optical flow," Proceeding on IEEE Int. Conf. of Automatic Face & Gesture Recognition and Workshops, pp. 700-706, 2011.

[14] Unnikrishnan, R., Pantofaru, C., and Hebert, M., "Toward objective evaluation of image segmentation algorithms," IEEE Trans. on Pattern Analysis and Machine Intelligence, 29, pp. 929-94, 2007.

[15] Park, J. G. and Lee, C., "Bayesian rule-based complex background modeling and foreground detection," Optical Eng. 49, [doi: 10.1117/1.3319820], 2010.

[16] Dardas, N., Georganas, H., and Nicolas, D., "Real-time hand gesture detection and recognition using bag-of-features and support vector machine techniques," IEEE Trans. on Instrumentation and Measurement. 60, pp. 3592-3607, 2011.

[17] Suppatoomsin, C., "Hybrid method for hand segmentation," Proc. SPIE 7546. [doi: 10.1117/12.856290], 2010.

[18] Lu, K., Li, X. J., and Zhou, J. X., "Hand signal recognition based on skin color and edge outline examination," J. of North China University of Technology," Beijing China. 18, pp. 12-15,2006.

[19] Guo, S., Gu, G. C., and Cai, Z. S., "Skin segmentation using similarity of skin color and dynamic threshold," Computer Engineering and Applications, 46, pp. 1-3, 2010.

[20] Zheng, Y. W., Feng, Z. Q., Yang, B., and Xu, T., "The research of gesture feature extraction algorithm based on the vector edge detection," China academic journal electronic publishing house, http://www.cnki.net, 2010.