

Adaptive Threshold Gesture Segmentation Algorithm Based on Skin Color

Chengyuan Liu^{1, a}, Jingqiu Wang², Ting Zhang^{3, c} and Dongsheng Ding^{3, c}

¹ School of Automation, Beijing Institute of Technology, Beijing 100081, China

² Institute of Automation, Chinese Academy of Sciences (CASIA)

³ School of Automation, Beijing Institute of Technology, Beijing 100081, China

^achyuanliu@vip.qq.com, ^bzhangting2003@bit.edu.cn, ^c137995622@qq.com

Keywords: Gesture recognition; Gaussian model; Skin segmentation; Otsu algorithm

Abstract. In gesture recognition systems, segmenting gestures from complex background is the hardest and the most critical part. Gesture segmentation is the prerequisite of following image processing, and the result of segmentation has a direct influence on the result of gesture recognition. This paper proposed an algorithm of adaptive threshold gesture segmentation based on skin color. First of all, a whole image should be transformed from RGB color space to YCbCr color space. After eliminating luminance component Y, similarity graph of skin color will be obtained from the already established Gaussian model. Then Otsu adaptive threshold algorithm is used to carry out binary processing for the similarity graph of skin color. After the segmentation of skin color regions, the morphology method is used to process binary image for determining the location of hands. Experimental results show that the detailed segmentation of skin color using the dynamic-adaptive threshold can improve noise resistance and can produce better results.

1. Introduction

Gesture is a kind of human-interactive means that is natural, direct and easy to learn. Using hands as the input device of computers, users can control surrounding machines by gestures directly without the need of a communication medium^[1]. Since gesture segmentation is the first and the most important step in the whole processing of gesture recognition, its results will have direct effects on the following feature extraction and recognition precision^[2].

Gesture segmentation refers to extracting useful gesture regions from images of hand gestures. It can be divided into two kinds: segmentation based on skin color and segmentation based on movement. The advantages of skin color segmentation are not affected by the sizes, directions and postures of hands. And the disadvantage is requiring for retraining the background color and light conditions. In hand images, the skin color information is the most direct and effective feature, the main color pixel value. Although different people have different skin colors, studies show that skin color is clustered in certain regions of color space^[3], especially in the color space which has been transformed and the effect of light luminance changes are eliminated^[4].

This paper analyzes color spaces, models with Gaussian model in YCbCr color space which has a better result of clustering and proposes a skin color segmentation algorithm based on adaptive threshold.

2. The establishment of skin color model

2.1 Common skin color models

Skin color models can roughly be divided into regional model, histogram model and Gaussian model which can be further subdivided into single peak Gaussian model and mixture Gaussian model. Gaussian model is simple in mathematical form. It shows continuous data and information by calculating the probability value of pixels. After that, a similarity graph of skin color is established and skin color regions can be distinguished by numerical values. Since it is difficult to estimate the

parameters of mixture Gaussian model, and iteration is necessary while convergence is affected by the initial parameters^[5], mixture Gaussian model performs slowly, although the fall-out ratio of this model is very low. By considering accuracy and rapidity of skin color segmentation, in this paper, skin modeling was based on single peak Gaussian model.

2.2 Skin color similarity computation

First, transform color images from RGB color space to YCbCr color space, then calculate the probability of belonging to the skin color regions for each pixel, according to 2-dimensional Gaussian distribution of skin color that is in Cb-Cr chrominance space. And then color image is transformed to gray image^[6]. The formula is as follows, e.g.

$$P(Cb, Cr) = \exp\left[-0.5(x-m)^T C^{-1}(x-m)\right]. \quad (1)$$

$x = (Cb, Cr)^T$ is the Cb and Cr chrominance component value of skin color pixels. $m = E(x)$ is the mean of Cb and Cr chrominance component of skin color samples. $C = E[(x-m)(x-m)^T]$ is the covariance matrix of skin color components. By gathering experimental statistics of amounts of skin color samples' pixel values, the parameters of common single peak Gaussian model are obtained:

$$m = (118.5728 \quad 152.9346)$$

$$C = \begin{pmatrix} 174.6293 & 23.8105 \\ 23.8105 & 286.9407 \end{pmatrix}. \quad (2)$$

After the completion of computing each pixel in the image using the formula above, the maximum value of each pixel's skin color similarity $P(Cb, Cr)$ is found, and the values of each skin color similarity in the whole image is normalized: with the value of each pixel's skin color similarity $P(Cb, Cr)$ divided by the maximum value of skin color similarity $P(Cb, Cr)$ in the image^[7], the value of skin color similarity of each pixel which is obtained and normalized in [0,1]. It indicates that there is a greater possibility of skin color if the value is closer to 1, or non-skin color if closer to 0.

3. Otsu skin color segmentation algorithm

In order to conduct binarization processing of skin color similarity images, each normalized value of similarity should be multiplied by 255 and rounded, and an appropriate threshold value is selected to divide skin color and non-skin color regions. It will mistake non-skin color spot for skin color spot and increase the skin color region if the threshold value is set too small; or it will mistake skin color spot for background causing some skin color region undetected if the threshold value is set too large^[8]. This paper selects maximum inter-class variance method, proposed by Otsu, as the binarization processing method. Otsu algorithm is derived on the basis of the principle of least squares, its basic thoughts is to use the histogram to divide the image into two parts, the target and the background, with a gray value, then calculate the inter-class variance of the two parts. The greater the difference between target and background are, the greater the variance between the two parts. To segment image with the threshold value of maximum inter-class variance minimizes the classifying error^[9]. The specific process of calculating the threshold value is as follows.

Setting all possible gray levels in the image as L, the number of pixels of gray level i is n, then the total number N of pixels in the image and the probability p_i of each gray levels are:

$$N = \sum_{i=0}^{L-1} n_i, p_i = \frac{n_i}{N}, \sum_{i=0}^{L-1} p_i = 1. \text{ Selecting a threshold value } k \text{ to divide the gray levels into two}$$

groups, $C_0 = [0, 1, K, k-1]$ and $C_1 = [k, k+1, K, L-1]$, the probability ω_0 of C_0 and the probability ω_1 of C_1 are:

$$\omega_0 = \sum_{i=0}^{k-1} p_i = \frac{\sum_{i=0}^{k-1} n_i}{N}, \omega_1 = \sum_{i=k}^{L-1} p_i = \frac{\sum_{i=k}^{L-1} n_i}{N} = 1 - \omega_0. \quad (3)$$

The mean gray values of C_0 , C_1 and the whole image are μ_0 , μ_1 and μ respectively:

$$\mu_0 = \frac{\sum_{i=0}^{k-1} ip_i}{\omega_0}, \mu_1 = \frac{\sum_{i=k}^{L-1} ip_i}{\omega_1}, \mu = \sum_{i=0}^{L-1} ip_i = \omega_0\mu_0 + \omega_1\mu_1. \quad (4)$$

The average variance values σ_0 , σ_1 and σ^2 of C_0 , C_1 and the whole image are:

$$\sigma_0^2 = \frac{\sum_{i=0}^{k-1} (i - \mu_0)^2 p_i}{\omega_0}, \sigma_1^2 = \frac{\sum_{i=k}^{L-1} (i - \mu_1)^2 p_i}{\omega_1}, \sigma^2 = \sum_{i=0}^{L-1} (i - \mu)^2 p_i. \quad (5)$$

The inter-class variance σ_B^2 between C_0 and C_1 is defined as formula (6):

$$\sigma_B^2 = \omega_0\omega_1(\mu_1 - \mu_0)^2 = (\mu_1 - \mu)(\mu - \mu_0). \quad (6)$$

In order to be more adaptive, this paper makes improvement for Otsu algorithm mentioned above. The idea of traditional Otsu algorithm is segmenting images when the inter-class variance is the maximum. From the definition of inter-class variance, we see that the distance between the target and background's mean gray value and the whole image's mean gray value should be as far as possible in order to make σ_B^2 the maximum. Only in this way could promise the maximum weighted sum or product of the above two^[10]. This idea can be extended to replacing mean gray value with gray average variance. The new threshold calculating formula is as follows:

$$k^* = \text{Arg Max}_{0 \leq i \leq L-1} \left[(\sigma_0^2 - \sigma^2)^2 (\sigma_1^2 - \sigma^2)^2 \right]. \quad (7)$$

Experiments show that the method not only is feasible and has a very good effect, but also can resist the effect of the contrast and brightness changes of images. Small thorns and holes still remain in the binary images of segmentation, which requires further treatment based on the morphological transformation. Morphology operation can simplify the image data, remove the irrelevant structures and reduce the image noises while maintaining the original basic shape features^[11].

4. Analysis of experimental results

Software and hardware testing platforms used in this paper are: the PC with Intel Core i3 3.10GHz and a 4GB memory, Linux OS, QT/E4.7+OpenCV2.4.6, 1280 x 720 resolution camera with the frame rate of 30 fps. The experiments were carried out indoors under the normal condition of light; gesture images captured by the camera were real-time processed. In order to verify that the algorithm has better adaptability, amounts of hands with different skin colors were especially sampled for experiments. Figure 1 and figure 2 show the gesture segmentation results of yellow race and black race respectively. Figure 1b and figure 1b, in which the gesture segmentation images with fixed threshold were used, show that edge effects were not ideal, and in skin color regions there were many small black holes, which damaged some valuable features of hands and may have effects on following gesture recognition. Figure 1c and figure 2c show the gesture segmentation images processed by the improved Otsu algorithm introduced in this paper. It could be seen that the skin color region is continuous and smooth, showing a clear boundary; the effect of hand image is satisfying, with an accuracy rate of complete skin color segmentation above 96.3%.



Fig.1a. Original image



Fig.1b. Fixed threshold



Fig.1c. Adaptive threshold



Fig.2a. Original image

Fig.2b. Fixed threshold

Fig.2c. Adaptive threshold

5. Conclusions

This paper proposed a gesture segmentation algorithm based on adaptive threshold for people with different skin colors. By analyzing features of different color spaces, we decided to transform images to YCbCr color space, in which the chrominance and luminance are separated for segmentation. Therefore, single peak Gaussian model was established to calculate skin color similarity of each pixel. Then improved Otsu algorithm was used to acquire the best dynamic threshold of gesture segmentation, and binarization processing of skin color similarity images was performed. Next, holes and thorns in images were filled or eliminated by morphological operations. The algorithm has finally realized accurate segmentation of skin color regions, with good anti-noisy capability and robustness.

References

- [1] Jian-sheng P A N. The Application of Hand Gesture Recognition in Human Computer Interaction[J]. Computer Knowledge and Technology, 2011, 35: 083.
- [2] CAO X, ZHAO J, LI M. Monocular Vision Gesture Segmentation Based on Skin Color and Motion Detection[J]. Journal of Hunan University (Natural Sciences), 2011, 1: 017.
- [3] Yang J, Lu W, Waibel A. Skin-color modeling and adaptation[M]. Springer Berlin Heidelberg, 1997, 687-694.
- [4] Sun L, Zhang L, Guo C. Technologies of hand gesture recognition based on vision[J]. Computer Technology and Development, 2008, 18(10): 214-216.
- [5] Wang P, Cai Z, Liu W. EM estimation of PDF parameters for Gaussian mixture processes[J]. SHENGXUE JISHU, 2007, 26(3): 498.
- [6] Liu Z, Liu J. Research on Face Detection Algorithm Based on Complexional Segmentation [J]. Computer Engineering, 2007, 4: 061.
- [7] Wei S, Weigong Z, Xiaorui Z, et al. Adaptive Face Location for Detecting Fatigue Driving[J]. Modern Transportation Technology, 2009, 1: 024.
- [8] Zeng X, Zuo W, Shi Y, et al. A NOVEL SELF-ADAPTIVE THRESHOLD CHOSEN METHOD OF COMPLEXION SIM ILARITY IMAGE[J]. Computer Applications and Software, 2010, 27(9): 125-127.
- [9] Zhao M, Zhao Y. Skin color segmentation based on improved 2D Otsu and YCgCr[C]//Electrical and Control Engineering (ICECE), 2010 International Conference on. IEEE, 2010: 1954-1957.
- [10] Phung S L, Bouzerdoum Sr A, Chai Sr D. Skin segmentation using color pixel classification: analysis and comparison[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2005, 27(1): 148-154.
- [11] Xiao-qian C. An Edge Detection Operator Based On Mathematic Morphology[J]. Journal of Hebei Polytechnic University (Natural Science Edition), 2009, 2: 014.