International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.9, No.6 (2016), pp.355-368
http://dx.doi.org/10.14257/ijsip.2016.9.6.31

Dynamic Hand Gesture Segmentation Method Based on Improved Kalman Filter and Weighted Skin-Color Model

Zhang Qiu-yu, Lu Jun-chi, Wei Hui-yi, Zhang Mo-yi and Duan Hong-xiang

School of Computer and Communication, Lanzhou University of Technology, Lanzhou, 730050, China
zhangqylz@163.com, lujunchi9@163.com

Abstract

In order to improve the problems of segmentation accuracy and real-time existing in dynamic hand gesture under complex backgrounds, this paper presents a kind of dynamic hand gesture segmentation method based on improved Kalman filter and weighted skin-color model. Firstly, improved Kalman filter is utilized to process hand gesture image of hand gesture video sequences and get rough hand gesture results. Secondly, weighted skin-color model is applied to process rough results of hand gesture segmentation and segment hand gesture. Finally, morphological method is utilized to deal with gesture segmentation result, getting rid of the holes in the hand gesture's binary image to realize the segmentation of dynamic hand gesture. Experiments show that the proposed method can segment hand gesture from dynamic hand gesture video sequences with complex backgrounds effectively. And the accuracy of hand gesture segmentation is high.

***Keyword**: Human-Computer Interaction, Dynamic hand gesture segmentation, Improved Kalman filter, Weighted skin-color model, Morphological processing*

1. Introduction

As a significant research of computer vision, dynamic hand gesture segmentation plays an important role in the field of hand gesture recognition and has profound meaning. With the development of hand gesture recognition, hand gesture recognition technology is widely applied in the fields of Human-Computer Interaction, telecommunications, machine control and language interpretation. Hand gesture segmentation algorithms have also become the focus of research in the field of hand gesture recognition [1].

For now, there are no general hand gesture segmentation methods that can solve all problems existing in the procedure of hand gesture segmentation, such as illumination, similar skin color interference and poor real-time. A lot of researchers put efforts on proposing novel hand gesture segmentation algorithms, which can make novel algorithms applied in hand gesture recognition. Therefore, how to segment hand gesture from hand gesture video sequences efficiently is meaningful to follow-up procedure. There are always complex colors, similar skin color and human skin color in hand gesture backgrounds of natural environment, which has great effects on hand gesture segmentation and influences the recognition work. To improve real-time and accuracy of hand gesture segmentation and shorten running time of hand gesture segmentation algorithm, research on real-time dynamic hand gesture segmentation with complex backgrounds has theoretical and practical values.

To segment hand gesture from hand gesture video sequences accurately and efficiently, a kind of dynamic hand gesture segmentation method based on improved Kalman filter and weighted skin-color model is proposed in this paper. Kalman filter is a kind of efficient and autoregressive filter. Original Kalman filter is improved in this paper and the improved Kalman filter is used to track hand gesture images in dynamic video sequences.

ISSN: 2005-4254 IJSIP

The purpose is to solve the problem of hand gesture segmentation with complex backgrounds. RGB, YCbCr and HSV (Hue, Saturation, Value) color spaces are combined and a novel weighted skin-color model is proposed in this paper. Utilize weighted skin-color model to process hand gesture region that is processed by improved Kalman filter. Finally, process hand gesture segmentation results with morphological method and get rid of holes and cracks of hand gesture images, which make the procedure of hand gesture segmentation done.

The rest of this paper is organized as follows. Section 2 describes related work. In Section 3, the proposed dynamic hand gesture segmentation method is elaborated, and this section includes average background difference, improved Kalman filter and weighted skin-color model. Subsequently, Section 4 gives the experimental results and analysis as compared with different backgrounds, illumination environment and similar skin color interference. Finally, we conclude our paper in Section 5.

2. Related Work

For motion hand gesture segmentation, many researchers have adopt more conventional approaches, background difference method or frame difference method, and obtained motion hand gesture images of video sequences by making the difference between current fame image and background image or adjacent frame images [2-5]. H. Duan *et al*. [6], the authors proposed a method based on color and background difference. They detected skin color region of hand gesture video sequences with chrominance in YCbCr color space and detected motion hand gesture region with background difference. Final hand gesture results were obtained from motion region hand skin color region of dynamic hand gesture. Background difference and frame difference have low time complexity, little calculation. And they are easy to meet the demand of real-time for hand gesture recognition. However, hand gesture results are not good as people wish. Besides two methods mentioned above, optical flow is also widely used in the field of dynamic hand gesture tracking and segmentation [7-9]. Optical flow performs well. However, it has high time complexity and it is hard to implement real-time hand gesture system. Y. Fang *et al*. [10], the authors used Adaboost to train hand gesture features and detect hand gesture region. Detected hand gesture region was used to segment hand gesture in the procedure of hand gesture segmentation. S. Patidar *et al*. [11], the authors selected proper color space to segment hand gesture image by comparing several color spaces widely used in the field of hand gesture segmentation. The disadvantage of their proposed method is easy to be interfered by other colors. A. S. Ghotkar *et al*. [12], the authors used Lab color space to detect skin color region and segmented dynamic hand gesture marked by Camshift algorithm. P. Kakumanu *et al*. [13], the authors elaborated features of each color space and described skin color detection method in each color space. In hand gesture video sequences, there is no doubt that hand gesture is main motion objection. Therefore, many researchers used particle filter, Kalman filter, Meanshift algorithm and Camshift algorithm to mark motion hand gesture and segmented motion marking objection [14-16]. There have been a lot of different methods of hand gesture segmentation proposed. However, these methods cannot meet the demands for accuracy of hand gesture segmentation, complex backgrounds and the real-time of dynamic hand gesture.

In summary, existing hand gesture segmentation methods show a lot of disadvantages. Most of hand gesture segmentation methods cannot segment hand gesture from complex backgrounds and have low accuracy, high time complexity and poor real-time. Therefore, it will be the focus of research for hand gesture segmentation algorithm to improve accuracy and real-time of dynamic hand gesture segmentation.

3. Proposed Scheme

3.1. Hand Gesture Segmentation Procedure

Experimental steps of dynamic hand gesture segmentation method based on improved Kalman filter and weighted skin-color model are as follows. The flow chart of proposed method is shown in Figure 1.

Step 1: Acquire frame images of dynamic hand gesture video sequences.

Step 2: Utilize average background difference to detect motion objection and track motion objection with improved Kalman filter.

Step 3: Detect skin color and segment motion hand gesture with weighted skin-color model.

Step 4: Handle binary hand gesture images with morphological process and get rid of noise in image.

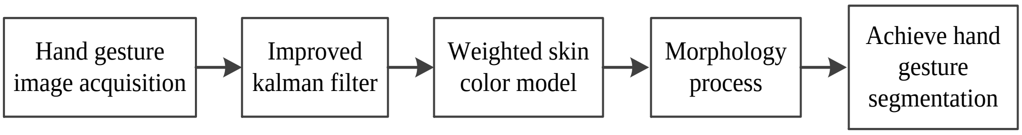

Figure 1. The Flow Chart of Proposed Method

3.2. The Improved Kalman Filter

Background difference is used to detect motion hand gesture of video sequences in this paper. Frame difference, background difference and optical flow are the most commonly used methods for motion detection in the field of hand gesture segmentation. Background difference is easy to implement and has high accuracy and low time complexity, which can meet the demand of real-time. On the basis of original background difference, average background difference is used in this paper. We average the initial motion background of background difference algorithm so that proper background of motion difference can be acquired and running time can be shortened. Formula of average background difference is as follows:

$$\mathrm{I}_{Background} = \left(\sum_{k=1}^{n} I_k\right) \Big/ n \tag{1}$$

where I_k is the k-th frame image of motion hand gesture video sequences, n is the maximum frame number of unobvious motion, $k \leq n$.

After detection of motion hand gesture, we utilize Kalman filter to estimate frame images of video sequences. Kalman filter is simple, robust and fast convergence. It also has low time complexity.

Priori estimated state of Kalman filter:

$$\bar{x}_i = A_i x_{i-1} + B_i \mu_i \tag{2}$$

Priori error covariance:

$$P_i^- = A_i P_{i-1} A_i^T + Q_i \tag{3}$$

where A_i and B_i have relationship with current stare estimated value x_{i-1} and procedure model vector μ_i. $\bar{x}_i$ is priori estimated value. P_{i-1} and P_i^- are current error covariance and priori error covariance. Q_i is system noise covariance. x_{i-1} and P_{i-1} can be initialized freely, but $P \neq 0$.

Kalman gain:

$$K_i = P_i^- H_i^T (H_i P_i^- H_i^T + R_i)^{-1} \tag{4}$$

Modify prediction value with z_i:

$$x_i = \overline{x}_i + K(z_i - H_i^- x_i) \tag{5}$$

Modify estimated error of prediction:

$$P_i = (i - K_i H_i) P_i^- \tag{6}$$

where K_i is the current Kalman gain, H_i relates the state vector to the measurement vector z_i, and R_i is the measurement error covariance.

As time progresses, it will produce some errors in parameters of Kalman filter or lose tracking target when original Kalman filter is utilized to track motion hand gesture of video sequences. On the basis of original Kalman filter, parameters acquisition procedure of Kalman filter is improved in this paper. Adding judgment section makes new parameters used to track real-time motion hand gesture in video sequences.

x_j and P_j are used to update prediction:

$$x_j = \overline{x}_i + K(z_{i-1} - H_{i-1}^- x_{i-1}) \tag{7}$$

$$P_j = (i - K_{i-1} H_{i-1}) P_i^- \tag{8}$$

where if $|x_i\text{-}x_j|>\varepsilon$, x_i equals to x_j. If $|P_i\text{-}P_j|>\delta$, P_i equals to P_j. New x_i and P_i are used to estimate parameters. ε and δ are judging parameters.

Steps of improved Kalman filter as follows:

Step 1: Calculate priori estimated state value of Kalman.
Step 2: Calculate priori error covariance.
Step 3: Calculate Kalman gain.
Step 4: Calculate prediction value x_i.
Step 5: Calculate the prediction value of estimated error.
Step 6: Modify x_i and P_i.
Step 7: Judge whether x_i and P_i need to be updated.

3.3. Weighted Skin-Color Model

For now, RGB, YCbCr and HSV color spaces are widely used in the field of hand skin color detection. However, using only one of several color spaces cannot obtain good segmentation results. To take advantage of features of color spaces and make up disadvantages of skin color segmentation of each color space, a kind of weighted skin-color model based on RGB, YCbCr and HSV color spaces is proposed in this paper. According to differences of different color spaces, hand gesture results of each color space are weighted. Utilize weighting method to eliminate interference of segmentation results of different color spaces and remove the effects of non-target region of the whole image so that hand gesture can be segmented from dynamic video sequences efficiently and accurately.

Utilize Otsu algorithm to segment hand images in RGB color space. Otsu algorithm is simple and has lower time complexity and smaller probability of misclassification. It will not be affected by the luminance of image and contrast either, and these characters are the basis for this paper [17].

Mark t as the segmentation threshold of hand gesture foreground and background, w_0 is the proportion that foregrounds hand gesture accounts for, u_0 is the average gray, w_1 is the proportion that background pixels accounts for, and u_1 is the average gray.

The u is the total average gray of the original image:

$$u = w_0 \times u_0 + w_1 \times u_1 \tag{9}$$

The g is the variance of hand gesture foreground and background image:

$$g = w_0 \times (u_0 - u) \times (u_0 - u) + w_1 \times (u_1 - u) \times (u_1 - u) = w_0 \times w_1 \times (u_0 - u_1)^2 \tag{10}$$

When variance g has the max value, difference between foreground and background is biggest and the gray t_{sb} is the optimal threshold:

$$t_{sb} = w_1 \times w_0 \times (u_1 - u_0) \times (u_0 - u_1) \tag{11}$$

Gray image of hand gesture is divided into foreground hand gesture and background through (10) and (11).

We utilize skin color ellipse model to segment hand gesture image in the YCbCr color space. Because chrominance values and luminance Y have relationship of functional dependency, image in YCbCr color space is transformed to the image in YCb'Cr' color space through nonlinear transformation. And skin color ellipse is used to detected skin color [18]. Pixels that satisfy the ellipse equation can be seen as skin-color pixels in this paper, and they can be called pixels of foreground hand gesture too. Compared with other skin color detection methods, skin-color ellipse model has better skin color segmentation results, meets the demand of real-time and is easy to implement.

It is a common skin color detection method that sets skin color thresholds. According to the feature that skin color in HSV color space is less sensitive to illumination, fixed skin color threshold method is used in this paper. Setting skin color threshold in HSV color space is easy to implement, which can segment hand gesture quickly. The idea of setting threshold is as follows: if hand gesture pixels in HSV color space satisfy conditions ($H \in [0.01, 0.43]$, $S \in [0.001, 0.278]$, $V \in [0.0784, 1]$) as the same time, pixels can be seen as skin color pixels and their values will be reset as 255. Else pixels will be seen as non-skin color pixels and their values will be reset as 0.

Assign different weights to hand gesture segmentation results that are from Otsu algorithm, ellipse skin color equation and setting threshold method in RGB, YCbCr and HSV color spaces. Eliminate effects of non-target hand gesture on hand gesture segmentation and add weighed segmentation results. Steps of weighted skin-color model are as follows:

Step 1: Utilize Otsu algorithm, skin-color ellipse model and setting threshold method to segment hand gesture image in RGB, YCbCr and HSV color space.

Step 2: Hand gesture segmentation result of YCbCr color space is multiplied by the weight β. Hand gesture segmentation result of HSV color space is multiplied by the weight α.

Step 3: Add two results with weight α and β. Regulate result with binary method.

Step 4: Hand gesture segmentation result of RGB color space is multiplied by the weight λ. Hand gesture segmentation result of Step3 color space is multiplied by the weight γ.

Step 5: Add two results with weight λ and γ. Regulate result with binary method. Get hand gesture segmentation results of weighted skin-color model.

$I_{HSV}(i, j)$, $I_{YCbCr}(i, j)$ and $I_{RGB}(i, j)$ are hand gesture segmentation results of HSV, YCbCr and RGB color spaces. Formula of weighted skin-color $I_{skin\text{-}color}(i, j)$ is as follows:

$$I_{skin-color}(i, j) = \gamma[\alpha I_{HSV}(i, j) + \beta I_{YCbCr}(i, j)] + \lambda I_{RGB}(i, j) \tag{12}$$

where α, β, γ and λ are weight parameters, $\alpha+\beta=1$ ($\beta<\alpha$), $\gamma + \lambda = 1$ ($\gamma < \lambda$).

Weighted skin-color model building process is shown in Figure 2.

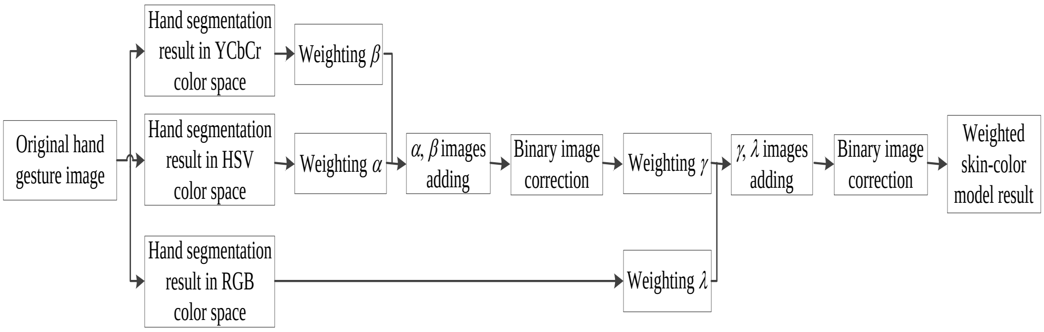

Figure 2. Building Process of Weighted Skin-Color Model

3.4. Morphological Process

Morphological open and close operations can process special image details that are smaller than structural elements, repair cracks and have better smoothing effects [19]. Based on effects and features of open and close operations, binary image processed by the proposed algorithm is regarded as object of morphology process in this paper. We carry on further process to get better results of hand gesture segmentation. The image that needs to be processed is obtained by logic operation of other binary images. Therefore, there will be many interference points, such as image holes, connection of image points, saw tooth and cracks of image edge. To remove all these interference factors exist in binary image, morphological open and close operations are used to process binary images in this paper. Handle holes and saw tooth in binary image with structural elements of morphology algorithm. Reduce impacts of these factors on the experimental results and get better binary image so that binary image can be used for hand gesture recognition perfectly.

Steps of morphological process:

Step 1: Process binary hand gesture image of weighted skin-color model with erode operation and eliminate holes effects on binary image.

Step 2: Process results of Step1 with dilate operation and recover basic contour of binary image eroded.

Step 3: Process results of Step2 with erode operation and improve edge effects of binary image. Finally, regulate binary image result.

4. Experiment Results and Analysis

Hand gesture with different gestures, different illumination environments and similar skin color are tested as experimental samples in experiment section in this paper. The size of each frame image is 640×480.

The experimental platform: Inter Core i3, 2450M, 2G, 2.27GHz, Win7 OS, and MATLAB 2012b.

In Experiment One, three different hand gestures of video sequences are segmented. Experimental illumination source is normal laboratory light. Setting threshold method in YCbCr color space is used to compare with the proposed method in this paper. In Experiment Two, hand gestures of video sequences with different illumination conditions are segmented. Illumination conditions of experiment two include faint light condition and dark condition. Setting threshold method in YCbCr color space is used to compare with the proposed method in this paper. In Experiment Three, hand gestures of video sequences with similar skin color interference are segmented. Experimental illumination source is normal laboratory light. Setting threshold method in HSV color space is used to compare with the proposed method in this paper.

Experiment One: Different hand gestures of video sequences under the condition of indoor fluorescent lamp, the results are shown in Figure 3-Figure 5.

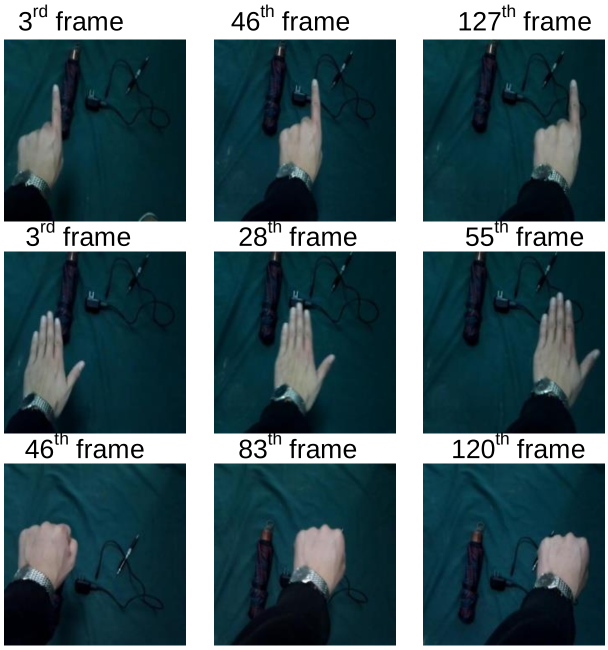

Figure 3. Hand Gestures of Different Video Sequences

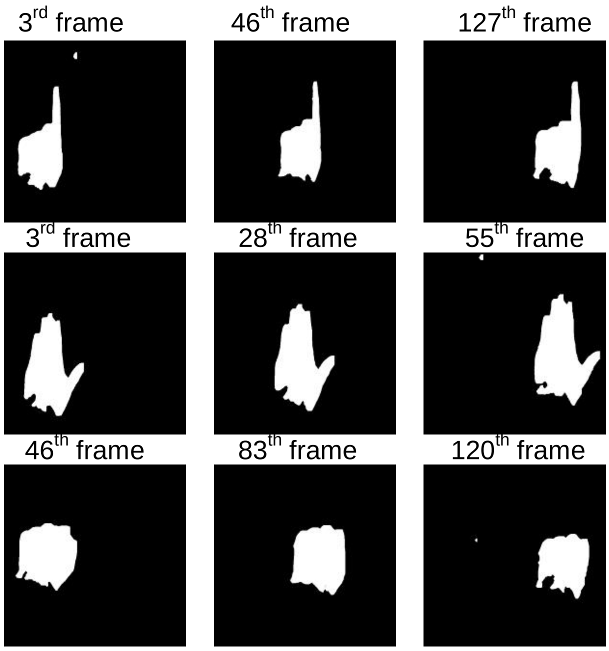

Figure 4. Segmentation Effects of Figure 3 with the Proposed Method

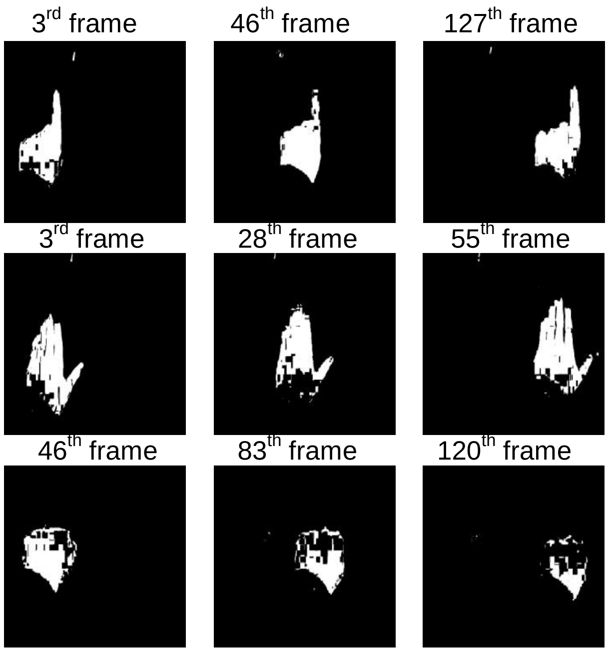

Figure 5. Segmentation Effects of Figure 3 with Threshold Method in YCbCr Color Space

Experiment Two: Hand gestures of video sequences under different illumination conditions, the results are shown in Figure 6-Figure 8.

Figure 6. Hand Gestures of Video Sequences with Various Illumination Conditions

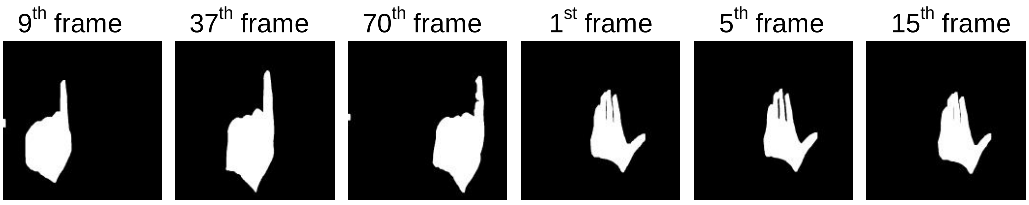

Figure 7. Segmentation Effects of Figure 6 with the Proposed Method

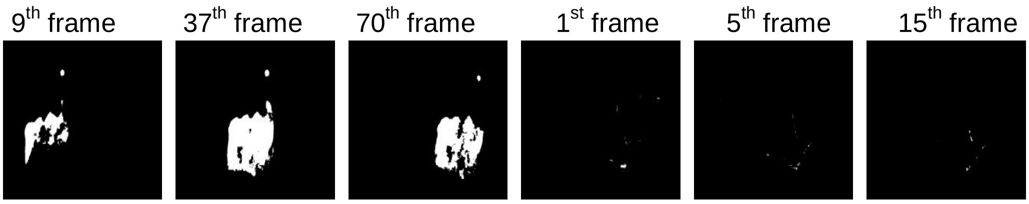

Figure 8. Segmentation Effects of Figure 6 with Threshold Method in YCbCr Color Space

Experiment Three: Hand gestures of video sequences with similar skin color interference under the condition of indoor fluorescent lamp, the results are shown in Figure 9-Figure 11.

Figure 9. Hand Gestures of Video Sequences with Similar Skin Color Interference

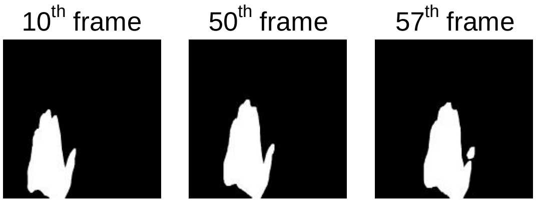

Figure 10. Segmentation Effects of Figure 9 with the Proposed Method

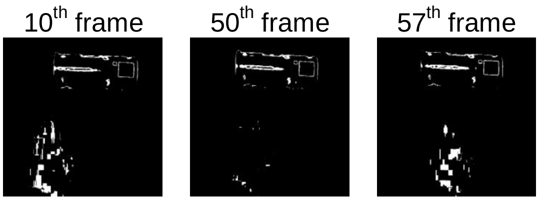

Figure 11. Segmentation Effects of Figure 9 with Threshold Method in HSV Color Space

According to the results of Experiment One, we can make a conclusion that using the proposed method can segment different hand gestures of video sequences effectively. To prove that the proposed method is robust to illumination, we design Experiment Two as comparing experiment. Comparing experiment results of Experiment One and Two, we can conclude that using the proposed method can segment hand gestures with different illumination conditions, which illustrates that the proposed method is robust to varying illumination. To prove that the proposed method can avoid interference of background colors, this paper designs Experiment Three, we can conclude that using the proposed method can get good results of hand gestures segmentation when similar skin color exists in experiment background. Comparing the proposed method's results with threshold method's results, we can conclude that only using threshold method in one color space cannot segment hand gesture from video sequences when background illumination changes or background contains similar skin-color interference. The analysis results of three experiments illustrate that the proposed method has improved the ability of skin color detection and effects of segmentation algorithm.

Time complexity of segmentation procedure of the proposed method:

$$O(mn+mn+mn) = O(3mn) = O(mn)$$

Time complexity of segmentation procedure is $O(mn)$ for RGB, YCbCr and HSV color spaces, where m is the length of each frame image, n is the width of each frame image.

Comparing the time complexity above, it can be concluded that running time of segmentation has relationship only with image size. With the same image resolution, time complexity of the proposed dynamic hand gesture segmentation method and threshold method is $O(mn)$. However, utilizing the proposed method can segment hand gesture better and be applied better to follow-up work. According to analysis of conclusions, the proposed method can segment hand gestures from video sequences efficiently and it is robust to various illumination conditions and similar skin color interference.

5. Conclusions

This paper proposes a kind of hand gesture segmentation method based on improved Kalman filter and weighed skin-color model. The proposed method improves original Kalman filter. The paper also combines existing skin color detection and segmentation methods, and proposes weighted skin-color model to detect skin-color region of hand gesture image. Experiment results illustrate that the proposed method can segment hand gestures from video sequences with complex backgrounds efficiently. It is also robust to illumination condition and has low time complexity, which can meet the demand of real-time and has practical values. How to segment hand gesture form backgrounds that have faces or other hands and improve the real-time of segmentation procedure will be the focus of follow-up work.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (No. 61363078), the Natural Science Foundation of Gansu Province of China (No. 1212RJZA006, No. 2014GS01729). The authors would like to thank the anonymous reviewers for their helpful comments and suggestions.

References

[1] H. Li and M. Greenspan, "Model-based segmentation and recognition of dynamic gestures in continuous video streams", Pattern Recognition, vol. 44, no. 8, **(2011)**, pp. 1614-1628.
[2] H. Xu, X. Hou and R. Su, "Real-time hand gesture recognition system based on Associative", Proceedings of the 2nd IEEE International Conference on Computer Science and Information Technology (ICCSIT 2009) ; Beijing, China., **(2009)** August 8-11.
[3] S. Alvarez, D .F. Llorca, G. Lacey and S. Ameling, "Spatial Hand Segmentation Using Skin Color and Background Subtraction", Trinity College Dublin's Computer Science Technical Report, Dublin (2010),
[4] H. Wang, "Research on Hand Gesture Tracking Based on Adaptive Active Contour Model", MS Thesis, Lanzhou University of Technology, Gansu, China, (2012).
[5] B. Toni and J. Darko, "A robust hand detection and tracking algorithm with application to natural user interface", Proceedings of the 35th IEEE International Convention on MIPRO, ; Opatija, Croatia, (2012) May 21-25
[6] H. Duan and Y. Luo, "A Method of Gesture Segmentation Based on Skin Color and Background Difference Method", Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering, Atlantis Press, ; Hangzhou, China, (2013) March 22-23.
[7] S. Zhang and H. Wang, "Human-Computer Interaction by a Robust Hand Tracking Method", Proceedings of the 2009 3rd International Conference on Teaching and Computational Science (WTCS 2009), Springer Berlin Heidelberg, Shenzhen, China, (2012) February 6.
[8] X. Shen, G. Hua and L. Williams, "Dynamic hand gesture recognition: An exemplar-based approach from motion divergence fields", Image and Vision Computing, vol. 30, no. 3, (2012), pp. 227-235.
[9] P. Premaratne, S. Ajaz and M. Premaratne, "Hand gesture tracking and recognition system for control of consumer electronics", In Advanced Intelligent Computing Theories and Applications, with Aspects of Artificial Intelligence, Springer Berlin Heidelberg, Berlin, (2012), pp. 588-593.
[10] Y. Fang, K. Wang and J. Cheng, "A real-time hand gesture recognition method", Proceedings of 2007 IEEE International Conference on Multimedia and Expo, Beijing, China, **(2007)** July 2-5.
[11] S. Patidar and C. S. Satsangi, "Hand Segmentation and Tracking Technique using Color Models", Hand, vol. 1, no. 2, (2013), pp. 18-22.
[12] A. S. Ghotkar and G. K. Kharate, "Hand segmentation techniques to hand gesture recognition for natural human computer interaction", International Journal of Human Computer Interaction (IJHCI), vol.3, no.1, (2012), pp. 15-25.
[13] P. Kakumanu, S. Makrogiannis and N. Bourbakis, "A survey of skin-color modeling and detection methods", Pattern recognition, vol. 40, no. 3, (2007), pp. 1106-1122.
[14] M. S. M. Asaari and S. A. Suandi, "Real Time Hand Tracking System Using Predictive Eigenhand Tracker", Australian Journal of Basic and Applied Sciences, vol. 6, no. 2, (2012), pp. 190-198.

[15] S. Mo, S. Cheng and X. Xing, "Hand gesture segmentation based on improved kalman filter and TSL skin color model", Proceedings of 2011 IEEE International Conference on Multimedia Technology (ICMT), ; Hangzhou, China, (2011) July 26-28.
[16] W. P. Tang, "Hand Gesture Recognizing and Tracking System Based on Camshift", MS Thesis, Xidian University, Shanxi, China, (2012).
[17] N. Otsu, "A threshold selection method from gray-level histograms", Automatica, vol.11, no.285-296, (1975), pp. 23-27.
[18] R. L. Hsu, M. Abdel-Mottaleb and A.K. Jain, "Face detection in color images", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 5, (2002), pp. 696-706.
[19] H. Zhao, "Research on Image Edge Detection Based on Mathematical Morphology", MS Thesis, Dalian University of Technology, Liaoning, China, (2010).