Removing Shadow for Hand segmentation based on Background Subtraction

Rahmita Wirza Rahmat
Department of Multimedia
Faculty of Computer Science and Information Technology
UPM, 43400, Serdang, Selangor, Malaysia
rahmita@fsktm.upm.edu.my

Zaher Hamid Al-Tairi
Department of Multimedia
Faculty of Computer Science and Information Technology
UPM, 43400, Serdang, Selangor, Malaysia
zaher_tu@yahoo.com

M. Iqbal Saripan
Department of Computer and Communication System Engineering
Faculty of Engineering
UPM, 43400, Serdang, Selangor, Malaysia
iqbal@eng.upm.edu.my

Puteri Suhaiza Sulaiman
Department of Multimedia
Faculty of Computer Science and Information Technology
UPM, 43400, Serdang, Selangor, Malaysia
suhaiza@fsktm.upm.edu.my

***Abstract*—Hand segmentation is an important stage for accurate hand detection and background subtraction is one of the best solutions to detect the hand motion accurately; however the shadow is the critical problem in this technique which is not easy to separate the hand region from the shadow area. Removing shadow using an automatic threshold will be a good solution to detect the hand region where the variety of skin color and lighting condition affect the hand segmentation. The proposed approach involves three stages: First, we convert RGB color model to YUV space to get the benefit of separation the luminance channel (Y) from the chrominance channels (U, V) to reduce the effect of shadow, reflections and, etc. In the second stage; we applied background subtraction technique to the V channel to remove the unwanted background noise and to get the hand and shadow pixels. Finally, we used thresholding technique by considering a mean value of the pixels of foreground image (the hand and shadow pixels) as automatic threshold value and other tow static thresholds to distinguish the hand region from shadow pixels. After background subtraction, we used the famous morphology techniques (Erosion and Dilation) to enhance the accuracy of hand detection. We measure the accuracy for the results by compare the detect hand pixels to the actual hand pixels quantitatively. From the results, we noticed that our proposed approach is accurate and suitable for real time application systems.**

Keywords-Background subtraction, Automatic thresholding, Hand segmentation, Removing shadow.

I. INTRODUCTION

Hand detection and tracking are an interesting research for human-computer interaction applications such as sign language systems, fingerprint recognition, and hand gesture recognition that can be acted as a mouse to control some devices likes computer, cell phones, and TV. The basic goal of this work is to remove the shadow of hand for better hand segmentation. This study is related to the work mentioned in Wirza et al. (2011), where the user hand will interact with a collaborative learning environment for medical students. In this paper the literature presented the shadow problem in the hand interaction applications and the proposed solutions to enhance the process of hand segmentation.

Never the less, this paper introduced a hand detection method which used a background subtraction technique to extract the region of a person in HSV color model. This color space is sensitive for shadow noise when V (value) is small so this color is converted to xyz cone to reduce the effect of shadow, as explained by Tanibata et al. (2002). A hand gesture recognition method is presented by using motion sensing to extract hand data and classification process to recognize the hand gesture. Holden et al. (1999) discussed the shadow occlusion affects the fingers features extraction and made the process of fingers tracking is difficult and to cover this problem, they adopted a prediction algorithm to find the miss locations of fingers in an image. Malik et al. (2002) introduced a hand tracking method for interactive augmented realty system. It used image subtraction and image thresholding techniques for hand detection process and applied threshold technique to separate the shadow occlusion from the hand region. Winkler et al. (2007) presented a hand and finger interaction for mixed reality system. Their hand segmentation is based on image differencing approach, which extracts hand as foreground by comparing the successive frames. The shadow casting can merge with the detect finger and affect the accuracy of detection. Generally, a hand tracking method which based on image difference technique is introduced for gesture-recognition and human-computer interaction. To cover the problems of hand segmentation such as the changing in luminance and shadow noise, they

978-0-7695-4959-0/13 $25.00 © 2013 IEEE
DOI 10.1109/ACSAT.2012.71

used two images with difference processes; the first one between successive images and the other between a pre-sorted background image and the live captured images (Martin et al., 1998). A method for hand and fingertip tracking which based on background subtraction and skin color detection is proposed for human-computer interaction as in Malik (2003); however the proposed method fails sometimes to detect the accurate peak and valley due to the shadow noise. In the mean time Tsang and Pun (2005) introduced a finger tracking method for interacting with computer or TV system directly without extra devices. This method used skin-color segmentation for hand detection and applied a morphological technique such as dilation and erosion to reduce the effect of the shadow noises. Dmitry and Arjun (2006), proposed an approach for detecting and tracking hand and fingers for interacting with the piano keys. Their method adopted the background subtraction approach and skin color technique to improve the hand segmentation and to eliminate the light reflections and the shadows occlusion. Christian and François (2001), introduced a bare hand interaction system for human-computer interaction, which they used image differencing approach for hand segmentation. However the shadow can affect the accuracy of fingertip detection where it can recognize as a fingertip and it can close the area between fingertips. Anagnostopoulos and Pnevmatikakis (2008) introduced a hand and fingertips detection and tracking system for natural interaction between hand and the 3D object. It used skin color segmentation with a region growing algorithm to locate the hand area and to reduce the effect of shadow they approach adopted a low-pass filter. In Kofler et. al. (2007), a fingertip detection method is described for hand interaction with documents that captured from a live camera. It used background subtraction to get hand region and for reliable hand detection and robustness against noises such as shadows and lighting changes, it first detected shadow and eliminated it from the foreground image and then used Bayes classifier with skin probability map to detect the hand from other regions. Lew et al. (2002) introduced a method for hand segmentation, which was based on image subtraction and skin color molding. It investigated gray level image, RGB color image and normalized RGB color image for better hand segmentation and found that the normalized RGB color model is more robust for shadow casting and lighting condition. Alvarez et al. (2010) described a skin color and background subtraction techniques for robust hand segmentation in their paper to cover the problems such as changing lighting and background clutter. Their proposed method detects and removes the background first and then detects and eliminates the shadow and finally used skin color model to improve the results of hand segmentation. Fingers tracking method was introduced by Letessier and Bérard (2004) for interaction with the digital objects. Their proposed method adopted two techniques that were image differencing and the rejection filter for reliable fingers tracking and used a metric that called Chrominance Euclidian Distance to reduce the shadow effect.

As mentioned earlier this work is part of a research work for collaborative learning environment for medical students. We intend to develop a hand interaction for the learning environment. So the aim of this work is to remove the shadow occlusion for accurate hand segmenting, which later enhancing the accuracy of the hand interaction. The segmentation process is based on image differencing approach to get the foreground image that contains the hand and its shadow and image thresholding technique to eliminate the shadow noise.

II. METHODOLOGY

Basically, the proposed method consists of the following stages:

- capture the hand image live using a USB webcam
- Convert the RGB color model to YUV color space
- Split YUV color into Luminance Channel (Y) and the chrominance channels (U, V)
- Segment the V channel into background and foreground using image differencing process to get the hand with shadow
- Remove the shadow of hand using dynamic threshold
- Apply the morphology techniques on segmented image to enhance the hand region

For further clarification, refer to the proposed flowchart of these steps in Figure 1.

As described earlier, skin color segmentation is a popular technique in human-computer interaction applications due to its simplicity and the fast processing, but it suffers from noises such as shadow and light changing. There are many color models such as RGB, HSV, YUV, YCrCb and others. Those color spaces are differing from each other in the effect by noises. The YUV color is one of the less affected by shadow occlusion as explained by Zabulis (2009) where it can reduce shadow noise by separate the luminance channel (Y) from the chrominance channels (U, V). In our approach, we adopt this color model and to get this color space; we convert the captured image from RGB to YUV model based on the following equation as in [18].

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B \\ U &= -0.147R - 0.289G + 0.437B \\ V &= 0.615R - 0.515G - 0.100B \end{aligned} \quad (1)$$

Once we converted RGB Image to YUV image, we used V channel as input for background subtraction to extract the hand region and reduce the unwanted noises such as shadow and some other regions result to the change in lighting. To remove the occlusion and getting the hand region, we proposed a dynamic thresholding as following.

$$Th = \begin{cases} 10 & DTh > 10 \\ 5 & DTh < 5 \\ DTh & else \end{cases} \quad (2)$$

where Th is the proposed threshold and the DTh is the average pixels of the foreground image.
We managed to get the exact hand region by tested each pixel of the foreground image using the following equation where if the foreground pixel is greater than the threshold value the output image will take the value 1 unless the value will be 0.

$$\mathrm{H(i,j)} = \begin{cases} 1 & F(i,j) > Th \\ 0 & else \end{cases} \quad (3)$$

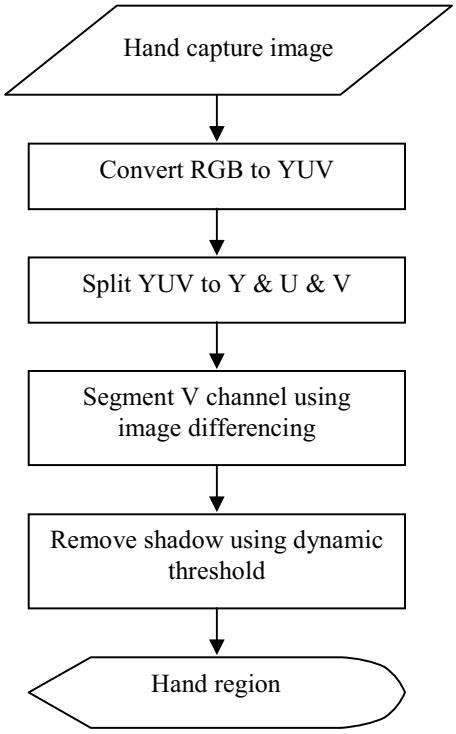
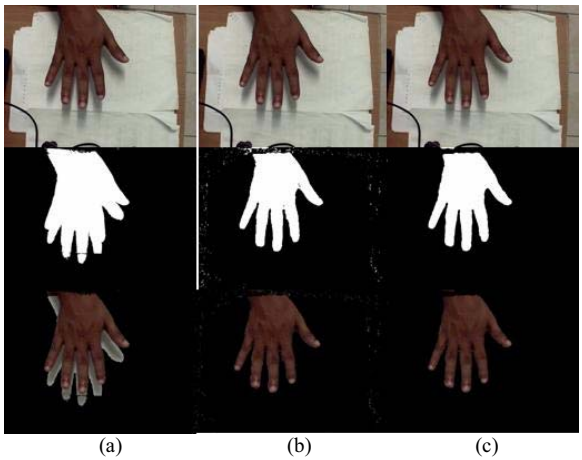


Figure 1. flowchart of the proposed method

III. RESULT

Some of the video data of hands used in this experiment were downloaded from the internet (please refer to appendix for the list of the address) and some were taken in the laboratory. The data is consists of 10 videos with a little different of simple background, light conditions and the data belongs to different persons, each one of these videos is converted into Xvid codec with 30 frames per second and the frame size is 640×480. The proposed method was coded in VC++ 2008 and OpenCV 2.1 under Microsoft Windows Professional, Intel Core 2 Duo Processor 1.83 GHz laptop with 3.0 GB memory. Figure 3 shows some of the final results of shadow removal for hand segmentation.



(a) (b) (c)

Figure 2. show background subtraction approach. (a) Hand images with background subtraction using OpenCV Gaussian mixture model and (b) Background subtraction using our proposed method without morphology technique and (c) are hand images with background subtraction using our proposed method with morphology technique.

(a) (b) (c)

Figure 3. show the result of background subtraction based on our proposed method and OpenCV Gaussian mixture model; (a) is the original images, (b) is background subtraction using our proposed method and (c) is the background subtraction using OpenCV Gaussian mixture model.

IV. DISCUSSION

To evaluate our proposed method we calculate the detection rate, false positive ratio, and false negative ratio based on the equation that introduced by Guoliang Yang et. al. (2010) as following:

$$\text{Detection Rate}(\%) = \frac{Ns}{Nf} \text{X100}$$

$$\text{False Positive Rate}(\%) = \frac{Nfp}{Nnf} \text{X100} \quad (4)$$

$$\text{False Negative Rate}(\%) = \frac{Nfn}{Nf} \text{X100}$$

Where,

N_F is the total number of shadow color pixels
N_S is the number of correctly detected shadow color pixels
N_{NF} is the total number of non-shadow color pixels
N_{FP} is the number of non-shadow pixels that are detected incorrectly as shadow color
N_{FN} is the number of shadow color pixels that are detected incorrectly as non-shadow color pixels

TABLE I. ACCURACY OF THE PROPOSED METHOD

DR	FPR	FNR
95.94	2.50	4.06

We calculate the Detection Rate (DR), False Positive Rate (FPR) and False Negative Rate (FNR) respectively as shown in Table I.

V. CONCLUSION AND FUTURE WORK

In this paper, we have described a simple and efficient background subtraction approach for hand segmentation. The used of YUV color model with the automatic threshold eliminates the shadow occlusion as shown in the results. The morphological operations: erosion and dilation is applied respectively to remove the unwanted noises and filling the holes of the hand region that happened during the process of removing shadow. From the results we can notice that the proposed method is reliable and acceptable. In our future work we will use both background subtraction with skin color segmentation to increase the accuracy and the efficiency.

REFERENCES

[1] A. Anagnostopoulos and A. Pnevmatikakis, "A realtime mixed reality system for seamless interaction between real and virtual objects," Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts (DIMEA 08). ACM, New York, NY, USA, 2008, pp. 199-204, doi:10.1145/1413634.1413673.

[2] C. Hardenberg and F. Bérard, "Bare-hand human-computer interaction," Proceedings of the 2001 workshop on Perceptive user interfaces, Nov. 2001, Orlando, Florida, doi:10.1145/971478.971513.

[3] D. O. Gorodnichy and A. Yogeswaran, "Detection and tracking of pianist hands and fingers," Proceedings of the The 3rd Canadian Conference on Computer and Robot Vision, p.63, June 2006 doi:10.1109/CRV.2006.26.

[4] G. Yang, H. Li, L. Zhang, and Y. Cao, "Research on a Skin Color Detection Algorithm Based on Self-adaptive Skin Color Model," Proceedings of the International Conference on Communications and Intelligence Information Security, 2010, pp. 266-270, doi:10.1109/ICCIIS.2010.67.

[5] E. J. Holden, R. Owens, and G. G. Roy, "3D Hand Tracker for Visual Sign Recognition," TR 99/2, University of Western Australia, 1999.
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.40.2235

[6] J. Letessier and F. Bérard, "Visual tracking of bare fingers for interactive surfaces," Proceedings of the 17th annual ACM symposium on User interface software and technology, Santa Fe, NM, USA, Oct. 2004, doi:10.1145/1029632.1029652.

[7] C. Kofler, D. Keysers, A. Koetsier, and T. M. Breuel, "Gestural interaction for an automatic document capture system," Proceedings of the Second International Workshop on Camera-Based Document Analysis and Recognition, Sep. 2007, pp. 161-167.

[8] Y. P. Lew, A. R. Ramli, S. Y. Koay, R. Ali, and V. Prakash, "A Hand Segmentation Scheme Using Clustering Technique in Homogeneous Background," Proceedings of the 2nd Student Conference on Research and Development (SCOReD 02), Grand Blue Wave Hotel, Shah Alam, 2002.

[9] J. Martin, V. Devin, and J. L. Crowley, "Active Hand Tracking," Proceedings of the 3rd. International Conference on Face & Gesture Recognition, April 1998, p.573.

[10] R. Wirza, N. P. Shan, A. A. Al-Surni, and M. Z. Dimon, "Utilization of Augmented Reality for Medical Students Collaborative Learning," Proceeding of International Conference on Engineering Education (ICEE 11), Madinah, Kingdom of Saudi Arabia, Dec. 2011, pp. 210-216.

[11] S. Malik , C. McDonald , and G. Roth, "Hand Tracking for Interactive Pattern-Based Augmented Reality," Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR 02), Sep. 30 - Oct. 01, 2002, pp.399-406.

[12] S. Malik, "Real-time Hand Tracking and Finger Tracking for Interaction", CSC2503F Project Report, December 18, 2003.

[13] S. Álvarez, D. F. Llorca, G. Lacey, and S. Ameling, "Spatial Hand Segmentation Using Skin Colour and Background Subtraction," Dublin , Trinity College Dublin's Computer Science Technical Report, November, 2010, (TCD-CS-2010-35)

[14] N. Tanibata, N. Shimada, and Y. Shirai, "Extraction of Hand Features for Recognition of Sign Language Words," Proceedings of International Conference on Vision Interface, 2002, pp. 391-398.

[15] W. M. Tsang and K. Pun, "A Finger-Tracking Virtual Mouse Realized in an Embedded System," Proceedings of the International Symposium on Intelligent Signal Processing and Communication Systems, Dec. 2005, pp. 781-784.

[16] S. Winkler, H. Yu, and Z. Y. Zhou, "Tangible mixed reality desktop for digital media management," In SPIE The Engineering Reality of Virtual Reality, vol. 6490B, 2007.

[17] X. Zabulis, H. Baltzakis, and A. Argyros, "Vision-based Hand Gesture Recognition for Human-Computer Interaction," In: The Universal Access Handbook. LEA, 2009.

[18] B. C. Ennehar, O. Brahim, and T. Hichan, "An Appropriate Color Space to Improve Human Skin Detection," INFOCOMP Journal of Computer Science, vol. 10, no. 1, 2011, pp. 18-27.

APPENDIX

- **2nd Whiteboard Stop Motion**
  http://www.youtube.com/watch?v=K8h9JUqfBbI
- **Stop Motion Story**
  http://www.youtube.com/watch?v=l-G7QtDk5gU
- **DNA origami add-ons - BIOMOD 2011 Team Slovenia**
  http://www.youtube.com/watch?v=zpa1YJXFAuk
- **Outside The Box Entertainment's First Experimental White Board Stop Motion Animation**
  http://www.youtube.com/watch?v=7SMjOnFrw8Q
- **Easy Cartoon Drawing : How to Draw a Cartoon Pig**
  http://www.youtube.com/watch?v=lBZEMXZwwic
- **Whiteboard Claymation - Stop Motion Animation**
  http://www.youtube.com/watch?v=7UJHUNCYHp0
- **stop motion**
  http://www.youtube.com/watch?v=z-pCSYNYCac&feature=related
- **whiteboard stop motion**
  http://www.youtube.com/watch?v=dIzRpP-ita8
- **Thinking Outside The Box**
  http://www.youtube.com/watch?v=dW2lr_J3pOk&feature=related