Proceedings of the International Conference on Machine Vision and Machine Learning
Prague, Czech Republic, August 14-15, 2014
Paper No. 96

Hand Segmentation for Optical See-through HMD Based on Adaptive Skin Color Model Using 2D/3D Images

Gisu Heo, Dong-Woo Lee, Sungyong Shin, Gague Kim, Hyeongcheol Shin
Electronics and Telecommunications Research Institute
218 Gajeong-ro, Yuseong-gu, Daejeon, Republic of Korea
heogisu@etri.re.kr; hermes@etri.re.kr; seneru@etri.re.kr; ggkim@etri.re.kr; shin@etri.re.kr

Abstract - In this paper, we propose a robust hand segmentation method based on adaptive skin color histogram model using the fusion of 2D/3D images for bare hand interaction in the optical see-through head-mounted-display (OHMD). Initially, the hand area detection is performed with depth information from stereo vision. The adaptive skin color histogram models are created with a chromaticity-based constraint to select pixels in a scene for updating a dynamic skin color model under changing illumination or cluttered background conditions. Experiment results show that the proposed method can accurately segmented the hand under various environmental conditions in the OHMD.

Keywords**:** Hand segmentation, Adaptive skin color histogram model, OHMD stereo vision.

1. Introduction

An optical see-through head-mounted display (OHMD) is a wearable display that shows a user the real environment through a half-transparent mirror and the virtual environment reflected on the half-transparent mirror (Zhou et al., 2008). The hand gesture interaction is appropriate interaction techniques which allow end users to interact with virtual content in an intuitive way.

Segmentation is the first step of the image processing in the computer vision applications such as hand or gesture tracking and recognition. The purpose of hand segmentation is to detect the orientation and position of the hands before it can be further processed. Traditional vision-based hand segmentation methods are commonly based on color filtering. These methods are computationally easy task and can be done very efficiently but can be seriously affected by the appearance of skin color-like objects and by lighting conditions. Recently, many researchers has been attempting to combine color and depth information for the hand segmentation to overcome this drawback (Bergh et al., 2011). The most widespread method of acquiring 3D depth data is stereo vision. Stereo vision can work on a wide range of exposures while ToF has a very limited exposure under outdoor (shadow and sunlight) conditions. ToF sensors are generally more sensitive to sunlight than stereo vision (Kazmi et al., 2013).

In this paper we used the fusion of 2D color (HS and CbCr) information and 3D depth (stereo vision) information to the hand segmentation in the optical see-through head-mounted display.

2. Hand Segmentation

In the following, we describe our approach to segment the hand present in the image, which is divided in the three steps. In the first step, we find the initial hand area using depth information. The second step finds the coarse hand area using the fixed skin color model. The last part of the algorithm extracts the fine hand area using the adaptive skin color histogram models. Figure 1 shows the procedure of our proposed adaptive hand segmentation algorithm.

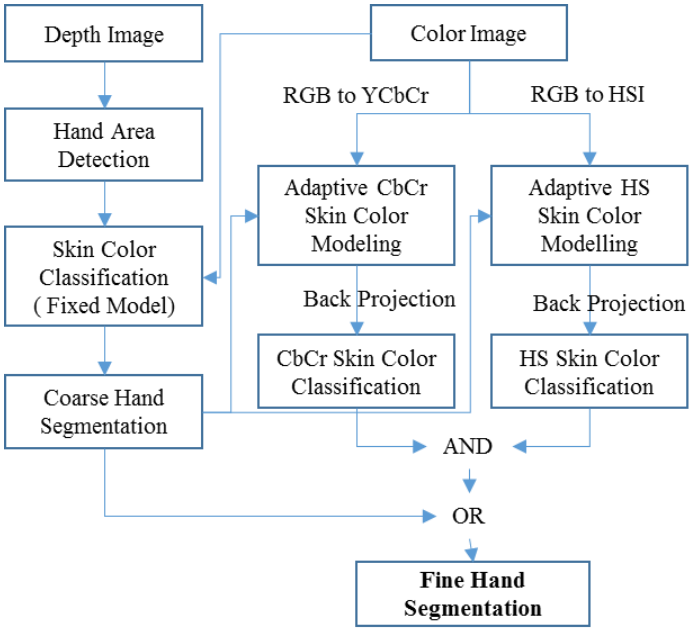

Fig. 1. The proposed Adaptive Hand Segmentation Algorithm.

2. 1. Hand Area Detection

The first step of the hands segmentation is that determine whether or not a hand is present within the frame. In order to accomplish this, we had used depth information that was created by stereo vision.

Generally a hands or fingers are used for interaction to the projected digital contents within the OHMD. Therefore, we have defined the distance within 20~40 cm from the wearable camera as the virtual user interaction area that user can naturally interact with the augmented objects in the OHMS using hands. The presence or absence of hands within this virtual interaction area can be detected using the depth information.

Fig. 2.Hand Area Detection.

2. 2. Coarse Hand Segmentation

The "Coarse Hand Segmentation" is the classification step of skin color pixels using the fixed range skin color model to the detected hand area by the depth information. In this step, the YCbCr color model is selected due to have a good performance under various illumination conditions. Among the YCbCr color components, we have opted fixed range skin color model from the Cb and Cr color components in order to find the hand area. The range of values for skin color used in this paper were according to experimental results of (Hu et al., 2003a) and (Hu et al., 2003b).

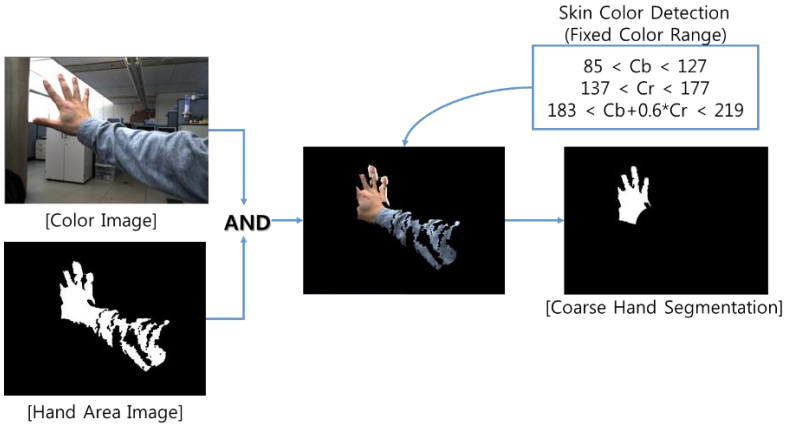

Fig. 3.Coasre Hand Segmentation.

2. 3. Fine Hand Segmentation

The "Fine Hand Segmentation" is the classification step of skin color pixels using the adaptive skin color models to the all pixels within the image. In this process an adaptive skin color histogram models are created from the color information of the coarse hand segmentation area, and a back projection algorithm is used to classify with skin color or non-skin color pixels for the all image pixels. As a result, the segmentation method using the fusion of color and depth information compared to using the only color information can be segmented into regions accurate hand.

2. 3.1. Generation of an Adaptive Skin Color Histogram Model

The skin color histogram model which a represent of the distribution of skin color in image is widely used method for skin color detection. The color histogram of an image is relatively invariant with translation and orientation about the viewing axis. In this paper, the adaptive skin color histogram is uses the H-S and Cb-Cr components in HIS and YCbCr color spaces, respectively, to model the human skin color in a way to minimize the effect of illumination variation by not considering luminance components in the color spaces. Since the proposed algorithm is based on the back projection of H-S and Cb-Cr color components, the histogram should be initialized under the current illumination condition of the hand. In this reason, to initialize and update the histogram, we had used the coarse hand segmentation area. To create the H-S and Cb-Cr color histograms, the histograms are quantized with the 30x32 for the H-S color components and the 30x30 for the Cb-Cr color components. After normalizing the value of histogram models, the values are set to the values in the associated quantized bin.

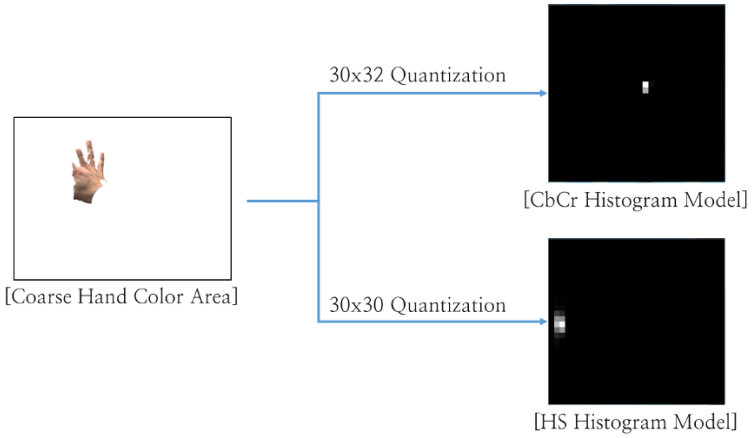

Fig. 4. Cb-Cr and H-S Skin Color Histogram Models.

2.3.2 Classification of Skin Color Using the Back Projection

The histogram back projection algorithm is used for the detection of know objects. Back projection algorithm is a way of recording how well the pixels fit the distribution of pixels in a histogram model. In other word, the probability that a pixel in an image is an area of the hand characterized by histogram can be calculated by the use of back projection method. Since the histogram is normalized, this value can be associated with a conditional probability value (Bradski et al., 2008).

Finally, a bitwise AND operation operates on each pair of the Cb-Cr and the H-S probability map to remove the noise, and a bitwise OR operation operates for add the information of the hand segmentation.

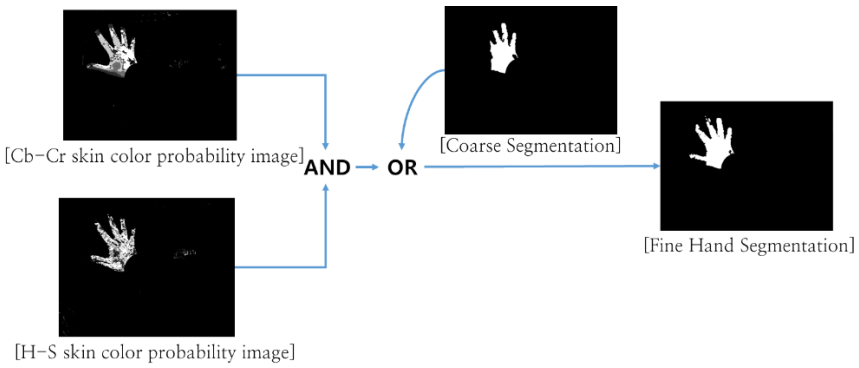

Fig. 5. Fine Hand Segmentation.

3. Conclusion

In this paper we have proposed a new approach to hand segmentation algorithm combining 2D color and 3D depth information provided by wearable cameras. We have taken advantage of the use of adaptive skin color histogram model based on real-time depth information to address the most conventional problem of video-based skin segmentation methods. The histogram of the detected hand area (only skin color) is dynamically updated by using the Cb-Cr and H-S color components respectively as adaptive skin color model. This adaptive skin-color histogram models are used for classification with skin-color or non-skin color pixels within the image by the back projection algorithm. The experimental results show that the proposed method is robust method for hand segmentation the use of the wearable OHMD devices under the various conditions such as changing illumination or cluttered background conditions.

Acknowledgements

This work was supported by the IT R&D program of MSIP/KEIT [10041656, Open Platform Development for See-Through Smart Glasses with Smart Vision].

References

Bergh M. V., Gool L. V. (2011). Combining RGB and ToF Cameras for Real-time 3D Hand Gesture Interface, WACA, pp. 66-72.

Bradski G., Kaeheler A. (2008). “Learning OpenCV: Computer Vision with the OpenCV Library”, O’Reilly, Cambridge. MA, USA, pp. 209-298.

Hu M., Worrall S., Sadka A.H., M.kondoz A. (2003), Face Feature Detection and Model Design For 2-D Scalable Model-Based Video Coding, “VIE”, Guildford- London.

Hu M., Worrall S., Sadka A.H., M.kondoz A. (2003), A Fast and Efficient Chin Detection Method for 2-D Scalable Face Model Design, “VIE”, Guildford- London.

Kazmi W., Foix S., Alenya G., Andersen H. J. (2013), Indoor and Outdoor Depth Imaging of Leaves with Time of Flight and stereo vision Sensors: Analysis and Comparison, “Journal of Photogrammetry and Remote Sensing”, pp. 128-146.

Zhou F., Duh H. B., Billinghurst M. (2008), Trends in Augmented Reality Tracking, Interaction and Display: A Review of Ten Years of ISMAR, “International Symposium on Mixed and Augmented Reality”, pp. 193-202.