

HAND SEGMENTATION USED IN VIRTUAL ENVIRONMENT

¹Chen-Yu Kuo (郭振宇)

²Chiou-Shann Fuh (傅楸善)

¹ Department of Mechanical Engineering,

² Department of Computer Science and Information Engineering,
National Taiwan University, Taipei, Taiwan

ABSTRACT

In this paper, we propose a hand segmentation and the finger numbers detection method for interacting with virtual object. We used skin color detection to get the skin color region. Assume the area of the hand occupies the most pixels of the image. Then, we calculate the pixel in each region. The maximum pixel region will be segmented to remove the noise. We calculate the center of hand and draw a circle filled with black to get the finger segmentation image. Since we know the finger numbers, it can be applied to AR or VR applications.

Keywords: *Hand detection, Skin color detection, distance transform, YCbCr.*

1. INTRODUCTION

Human-Computer-Interface (HCI) is the study about how to improve the interaction between human and computer. People want to see real, so the application of Virtual Reality (VR) and Augmented Reality (AR) attracts people's attention. In the operation of VR and AR, the operating experience will be more immersive and intuitive by using bare hand control.

VR is a technology that uses a computer to create and control a virtual world in which people can feel as if they were in a real environment.

AR is a technology that combines virtual objects with real scene. A real scene is photographed by a camera. Then, it calculates coordinates using the computer vision methods and renders virtual objects by computer graphics. As the computer technology advances, there are more state-of-the-art and convenient AR development devices. Therefore, AR applications are common in our lives such as engineering simulation, skill training, and teleoperation. VR requires a Head-mounted display (HMD) to generate three-dimensional view for the users. It also can be used in AR to make the users feeling more realistic. Common HMDs are: Oculus Rift, HTC Vive, and Google Card Board. Oculus Rift and HTC Vive need a powerful computer to compute and render the scene.

Hand is the most important natural labor tools for human. If hands are used to operate in AR, the users will feel more immersive and intuitive than using the handheld controller. For tracking hand, there are two methods. One is using data glove and magnetic sensor. The other is using the vision-based techniques. In this paper, the latter will be used because it is less costly than the former.

There will be two main challenges in hand segmentation applied in VR or AR system. One is locating hand in a complex background. The other is computing speed. It must be real-time hand segmentation in VR or AR system to avoid user dizziness.

In this research, the hand gesture will be used to interact with the virtual environment.

2. RELATED WORK

Herumurti, Yuniarti et al. (2017) implements Mixed Reality (MR) to develop an application in virtual room arrangement. In that research, Google Card Board is used for HMD. They capture the real scene with the phone's camera and use fingertip as input pointer to interact with virtual environment.

Khattak, Cowan et al. (2014) used HMD, RGB-D camera, Leap Motion to build a AR system. In that research, Leap Motion is used to capture hand's motion and RGB-D camera is used to get depth map to handle occlusion.

Bruder, Steinicke et al. (2010) mask the room in green color that is suitable for color keying. The green color space will be replaced by the virtual environment and the hand and the marker will show in the scene.

Oliveira and Conci (2009) convert RGB image to HSV color space. The H channel is used to characterize the colors range for skin detection.

3. METHOD

In this chapter, we will introduce the techniques used in our system.

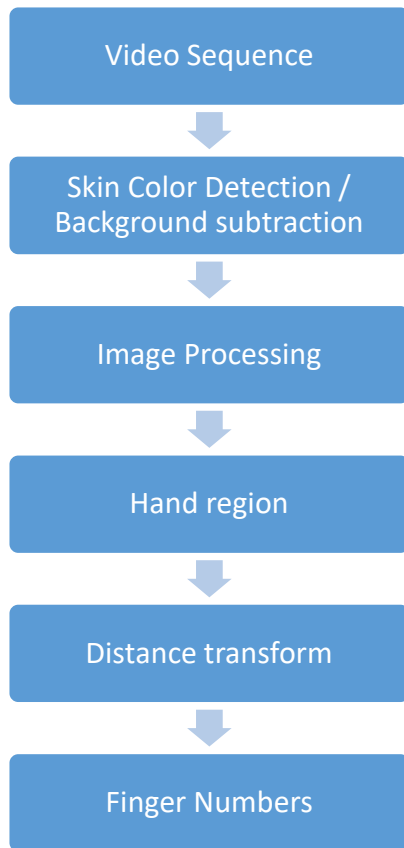


Fig. 1. Flow Chart of Hand Segmentation

3.1 Background subtraction

Background subtraction is widely used in detecting objects in videos from static scene. The background model should be recorded. The difference between the updated frame and background model can quickly segment the hand image.

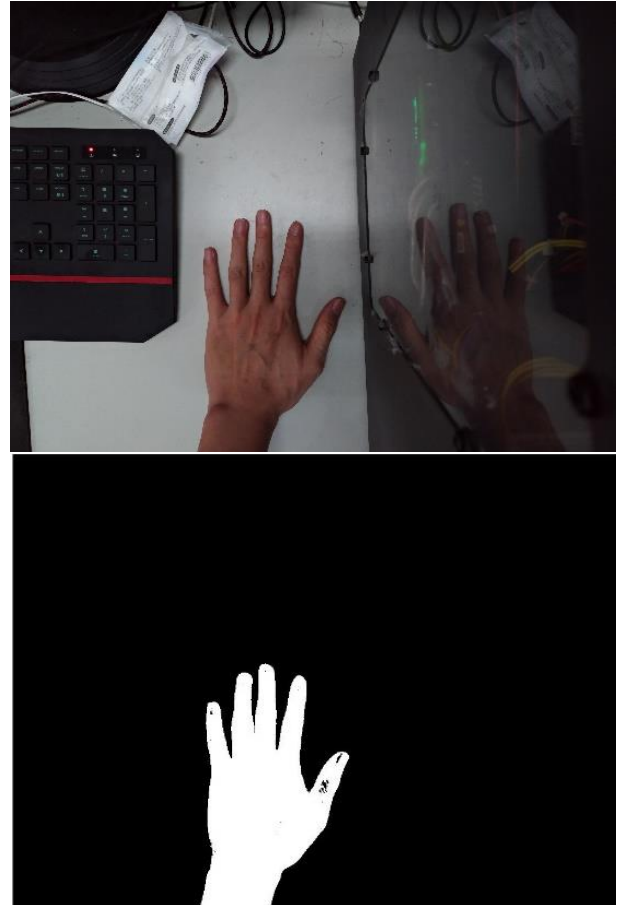
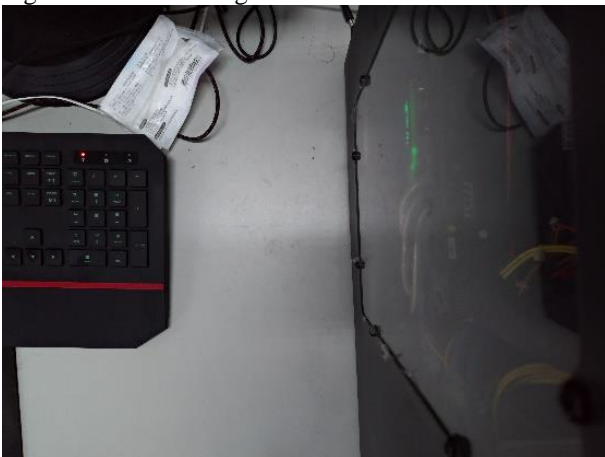


Fig. 2. Background subtraction.

3.2 Skin Detection

Skin detection is the process of finding pixels that are similar to skin color. It can be used for hand detection and face detection. In this paper, the YCbCr color space is used for hand detection. The YCbCr values were determined to be as follows:

- Y value selected for skin is between 70 and 255.
- Cb value selected for skin is between 140 and 180.
- Cr value selected for skin is between 80 and 135.

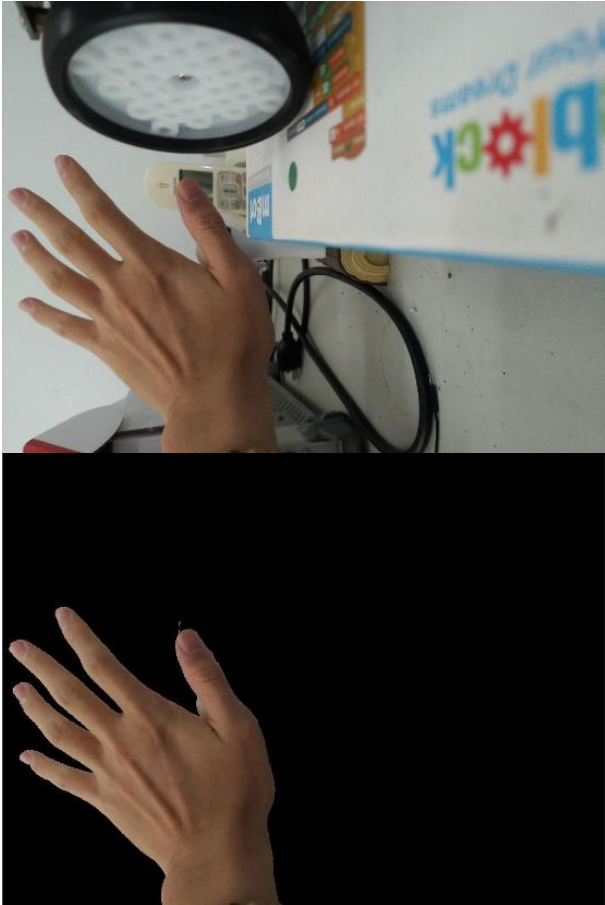


Fig. 3. Skin detection.

3.3 Noise Removal

In a complex background, there should be some noise in the skin color detection. We calculate the number of pixel in each area. We remove the noise by selecting the max area in the result of skin color detection. We fill the contour, so some holes will be fill in Figure 5.



Fig. 4. Skin color detection.

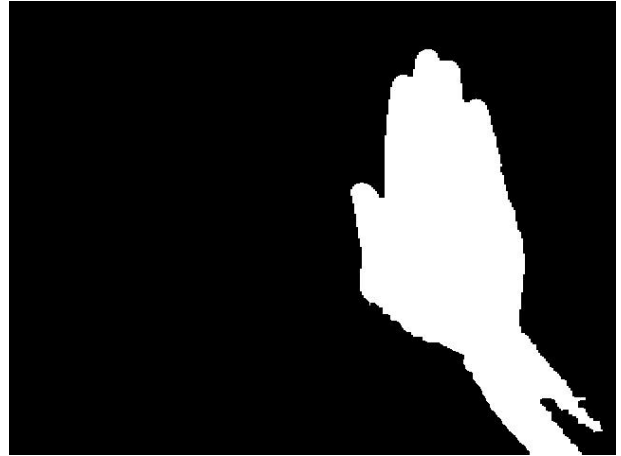


Fig. 5. Hand region.

3.4 Distance transform

Distance transform is used to find the center of hand (Figure 6). Draw a circle in the center of hand and subtract it from the hand image. We can get the fingers image (Figure 7).

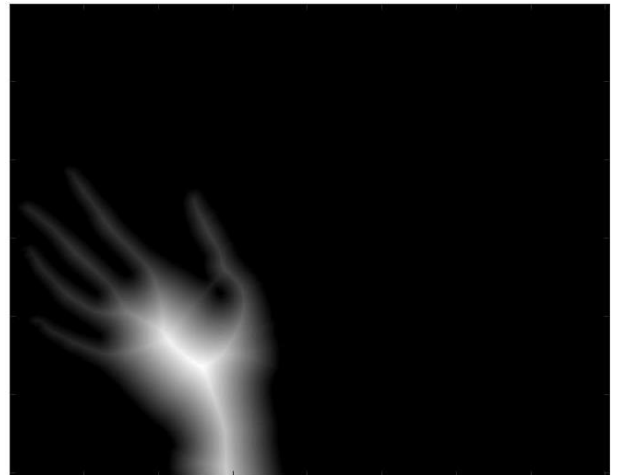


Fig. 6. Distance transform.

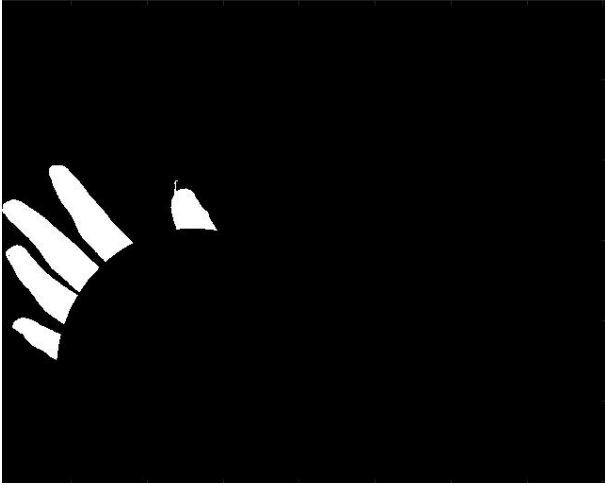


Fig. 7. Finger segmentation.

3.5 Position Estimation

Hand movement is continuous. If we know the position of hand at frame n and $n-1$, we can use the difference between frame n and $n-1$ to estimate the position of hand at $n+1$. This method can help us to reduce the computation and quickly segment hand image. The estimation function is as follows:

$$P_{n+1} = P_n + \Delta P \quad (1)$$

4. RESULT

In this section, we show our system result. We compare some method used in this paper with others to evaluate our system performance.

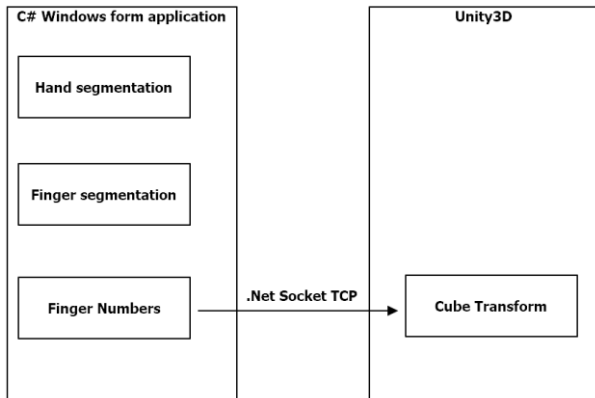


Fig. 8. System diagram.

4.1 Software and hardware

The system is developed by using Visual Studio 2017. The programming language C# is used because it can make a cross-platform project and is the main language used in Unity3D scripts. Unity3D is a cross-platform game engine. Most of HMD can be used on it, so we can build AR or VR Environment on it. The Emgu CV

library is used to process image. The .Net Socket is used to communicate between C# project and Unity3D project.

4.2 Skin color detection

We compare our method with HSV skin color detection. As Figures 8 and 9 show, HSV skin color detection has more unnecessary details and poor detection.



Fig. 9. HSV skin color detection.



Fig. 10. YCbCr skin color detection.

4.3 Noise removal

Our system can remove some noise like holes and other skin color object.

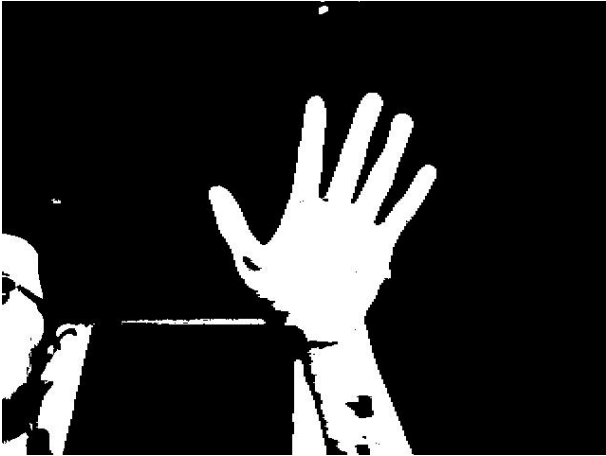


Fig. 11. Skin color segmentation result with some noises.



Fig. 12. After noise removal.

4.4 Finger Number

We used distance transform to get the finger segmentation and calculate the number of fingers. Then, we used bounding box to mark the finger as the figure shown below.

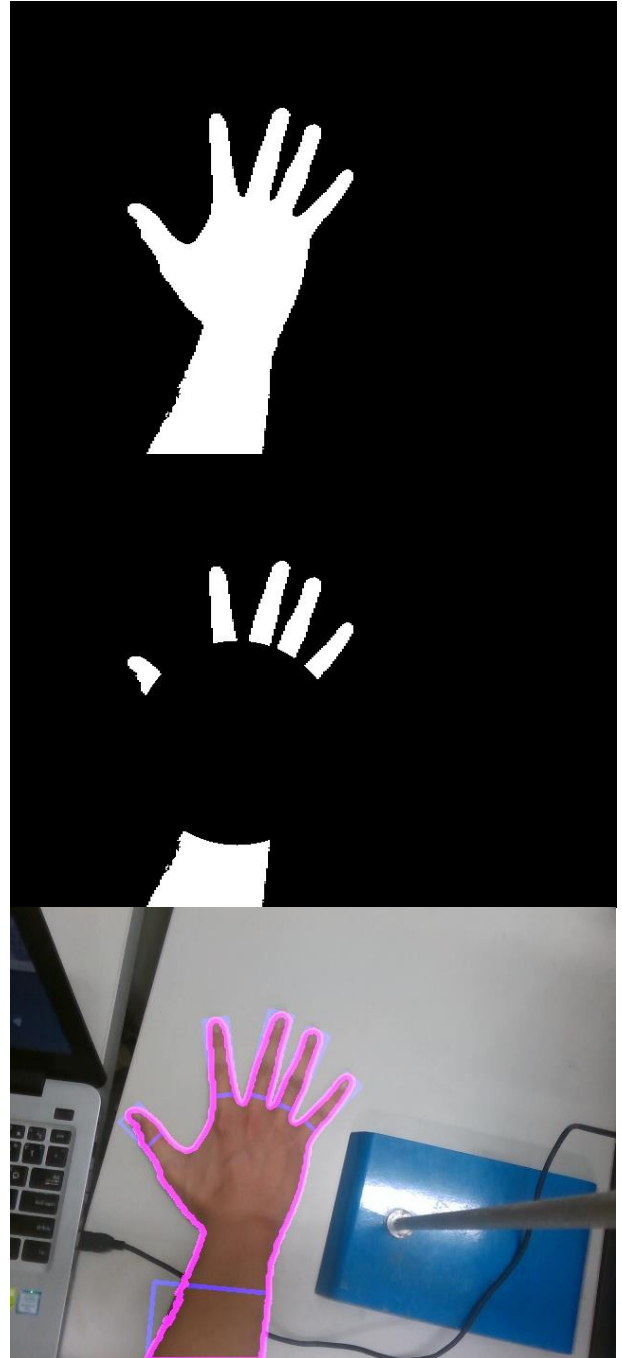


Fig. 13. Step of calculating fingers number.

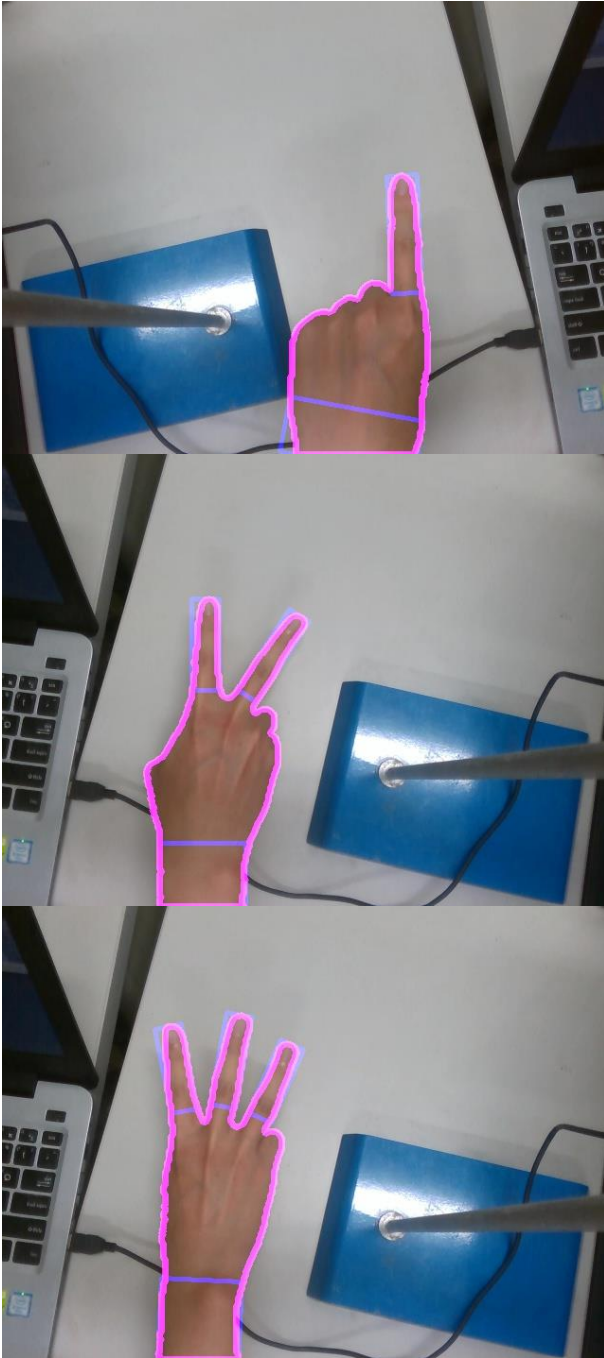


Fig. 14. Result of finger numbers 1, 2, and 3.

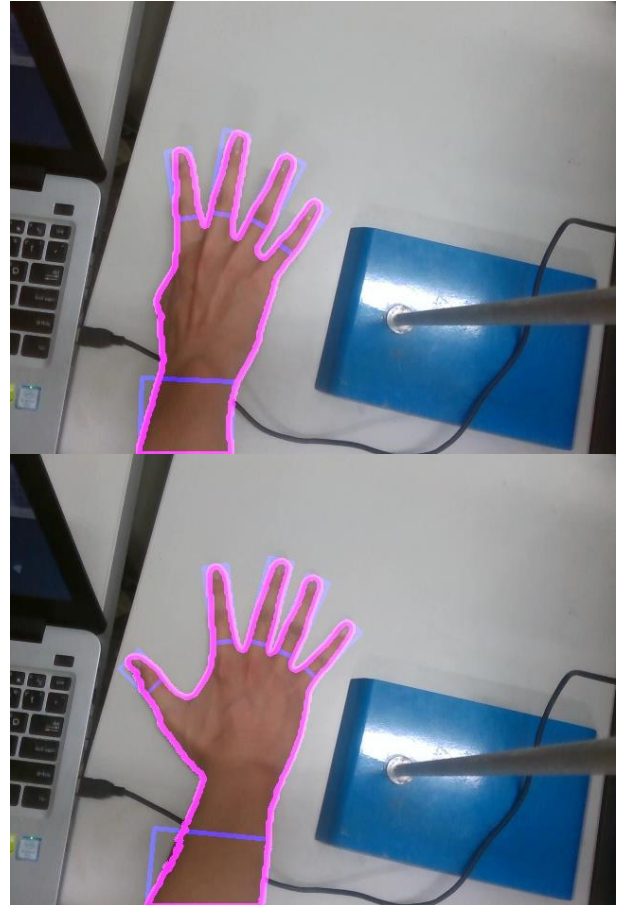


Fig. 15. Result of finger numbers 4 and 5.

4.5 Interaction with virtual object

In this section, we will show that how we use hand segmentation to interact with virtual object.

The C# project is used to do hand segmentation and set as the server side. Unity3D project is set as the client side. The C# project send the finger numbers to Unity3D. We define different uses of the finger numbers:

1. The transform of the cube will be reset.
2. The cube will rotate.
3. Do nothing.
4. The cube will be enlarged.
5. The cube will be zoomed out.

The result is shown in the below.

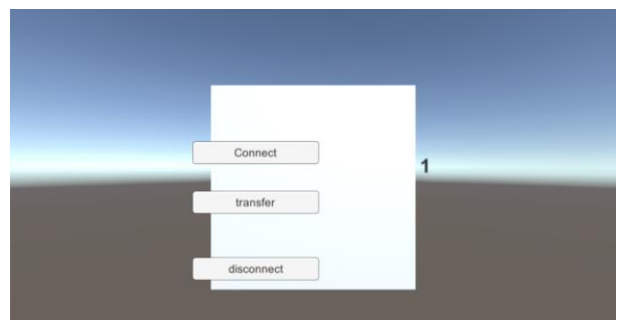


Fig. 16. Reset the cube transform.

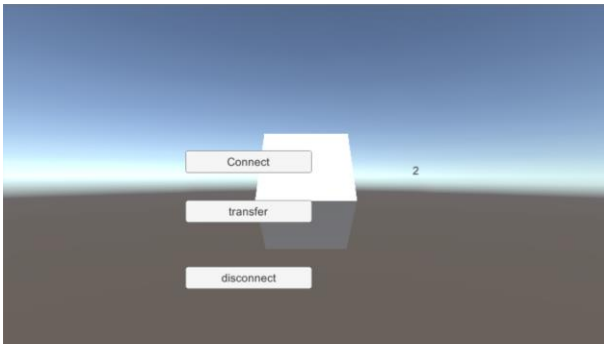


Fig. 17. Rotate the cube.

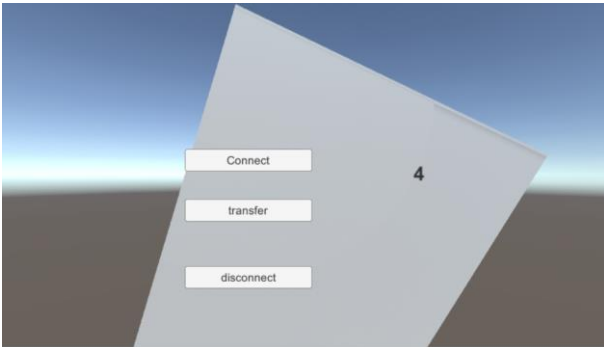


Fig. 18. Enlarge the cube.

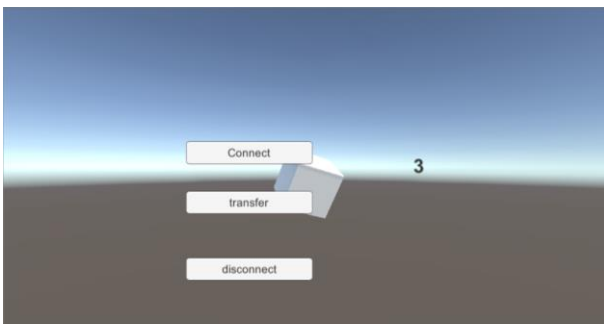


Fig. 19. Zoon out the cube

5. CONCLUSION

In this paper, a real-time hand detection and gestures recognition system was developed. The result of our system can detect the user's hand even if his face is viewed by the webcam.

Our project can be used in the VR or AR system to increase the sense of immersion. Our system can provide more intuitive control than using handheld controller. Our result may be fail on some condition, because most cameras automatically correct images according to the environment.

6. FUTURE WORK

In the future, this system can be an input of other software. Some message like finger number, center of hand, contour area of hand is useful for VR or AR system. Finger number can be defined as different work.

Center of hand can be a pointer. Contour area of hand can be defined as z-axis coordinate.

REFERENCES

- [1] Bruder, G., et al. (2010). Immersive virtual studio for architectural exploration. 2010 IEEE Symposium on 3D User Interfaces (3DUI).
- [2] Herumurti, D., et al. (2017). Mixed reality in the 3D virtual room arrangement. 2017 11th International Conference on Information & Communication Technology and System (ICTS). Surabaya, Indonesia: 303-306.
- [3] Khattak, S., et al. (2014). A real-time reconstructed 3D environment augmented with virtual objects rendered with correct occlusion. 2014 IEEE Games Media Entertainment.
- [4] Oliveira, V. and A. Conci (2009). Skin Detection using HSV color space. H. Pedrini, & J. Marques de Carvalho, Workshops of Sibgrapi, Citeseer.