3D Virtual Image Composition System based on Depth Map

Hye-Mi Lee
Dept. of Computer Science
Sunchon National University,
Sunchon, the Republic of Korea
e-mail:lhrooh@sunchon.ac.kr

Nam-Hoon Ryu
WithusBiz Co., Ltd.
Gwangyang, the Republic of Korea
e-mail: nhryu@sunchon.ac.kr

Eung-Kon Kim
Dept. of Computer Engineering
Sunchon National University,
Sunchon, the Republic of Korea
Corresponding Author
e-mail: kek@sunchon.ac.kr

***Abstract*— To complete a film, it needs to go through the process to capture the actual actor's motion and compose it with virtual environment. Due to the excessive cost for production or lack of post-processing technology to make the film, however, it is mostly conducted by manual labor. The actor plays his role depending on his own imagination at the virtual chromakey studio, and at that time, he has to move and consider the possible collision with or reaction to an object that does not really exist. And in the process of composition applying Computer Graphics(CG), when the actor's motion does not go with the virtual environment, the original image may have to be discarded and it is necessary to remake the film. This paper presents and realizes depth-based real-time 3D virtual image composition system to reduce the ratio of remaking the film, shorten the production time, and lower the production cost. As it is possible to figure out the mutual collision or reaction by composing the virtual background, 3D model, and the actual actor in real time at the site of filming, the actor's wrong position or action can be corrected right there instantly.**

Keywords-3D Image; Composition; Chromakey; Depthmap

I. INTRODUCTION

It is common that the digital technology is used to produce the video in the broadcasting media including movie. The computer graphic technology, which can realize the creative idea, removed the limitations in the production space and allowed the production of diverse contents by composing the actual image with the virtual environment. In case of the video that the live action is hard, the background image is composed with the actual actor using the computer graphic technology, for which composing the virtual background image or objects with the actual actor should be made seamlessly and the sense of difference in the images should be minimized. Since the realistic and high quality video invested with high cost raised the demand of the audiences for the quality, it is now that the development of low cost and high efficient video production technology is needed to meet such demand of the audiences.

In this article, the system, which can monitor the screen at the same time as filming is made by composing the virtual environment applied with 2-D background and 3-D virtual model with the actor, is intended to be designed and realized. Since the conflicts and the reaction between the actor and 3D virtual objects can be identified during the filming, the filming can be completed correcting the wrong action and location of the actor. In Section II, the depth information-based real-time virtual image composite system will be explained. And in Section III, the results of its realization will be examined, in Section Ⅳ, the conclusion and the expected effects will be presented.

II. REAL-TIME VIRTUAL VIDEO COMPOSITING SYSTEM

In this section, the video composing system, which replaces the 2D background by converging the depth information and the actual object and the 3D object can respond each other in real-time, is designed excluding the fragmentary composing by chromakey. Since the purpose of this system is to produce the video according to the story or scenario, all the works like the animation of the virtual environment and 3D object should be performed in advance. This system is divided into the 2-layer composing module, 3D virtual space generation module and the 2D-3D virtual video composing engine.

In case of the foreground and background separation and the background superimposing by the existing chromakeying, the depth buffer is not generated. Since to link with the 3D model, 3D space data should be generated, the foreground and background are separated through the depth-keying method.

As the hardware to obtain precise depth data is expensive equipment, to obtain the depth image and color image at the low cost simultaneously, the RGB image and the depth stream are entered using Kinect. Kinect is the device that generates the 3D space data and provides diverse user experiences by measuring the depth of the video entered through the infrared pattern recognition [1][2]. Figure 1 shows the pipeline of the 2-layer composing module.

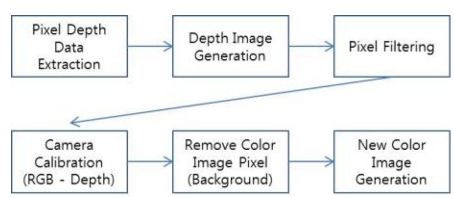


Figure 1. Pipeline of the 2-Layer Composition Module

To separate the foreground from the depth image generated, the filtering should be performed. Kinect returns the depth value from 0 to 4095 and each value is responded

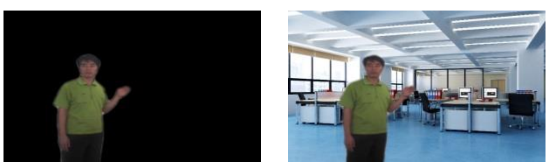
 ISBN: 978-1-61208-400-8

to the length unit in mm. In 2-layer composing module, the threshold value for the foreground position is set to between 800 ~ 4000 mm, and the pixels in the region, whose depth threshold exceeds 4000, are processed to background through filtering. Figure 2(a) shows the color video entered through RGB camera, and Figure 2(b) shows the depth image, which visualizes the depth of space.

(a) (b)

Figure 2. The process of Separating the Foreground from the Depth Image

The pixel of the depth image that entered through the depth camera has a problem that it is not converted 1 vs 1 with the color stream entered through the RGB camera. The reason for this is because the positions of two cameras are different and therefore, they convert the position of the depth image pixel position into the pixel position of the color image through the camera calibration [3][4]. And then our system removes the pixels, which are not the foreground, generates the new color image and draws it to the area corresponded to the background. Figure 3(a) shows the RGB image of the foreground after filtering the background and Figure 3(b) shows the image that the background of actual filming image entered through the device was converted and composed with the foreground.



(a) (b)

Figure 3. Synthesis of the Foreground and Background Image based on the Depth Information

At this moment, since the noise is generated around the boundary of the foreground, unless the smoothing process is not performed, the pixels around the foreground remove the noise by compensating the edge.

III. RESULTS OF SYSTEM REALIZATION

Since this system is the technology to reduce the refilming rate when filming the virtual video, the purpose of this system is to prevent the mismatching between the position and the eye line of the actual actor according to the virtual objects that appear in video. Figure 4 shows that the virtual background, actual actor, 3D virtual model are continuously moving. Figure 4(a) shows the results realized as if the 3D virtual model is located closer to the camera than the actor and Figure 4(b) shows the results realized as if the actor is located far away from the camera than 3D virtual model. Each object is moving continuously and is rendering on the screen in real-time.

(a) (b)

Figure 4. Screen of the Final Implementation Results

The actor acts in the virtual studio where the 3D objects to be arranged in final outcome do not exist, but in the video, the virtual environment and 3D objects appear together. Therefore, the movement and the flow of the eye line of the actual actor matched with the virtual environment can be observed.

IV. CONCLUSIONS

In the dangerous scene or in the production of video using the computer graphic technology for the SF movie where the live action is difficult, there are many problems such as technical problem, work time, and cost. To complete the high quality video, the work should be done manually but if the action of the actor is not delicately matched with the 3D virtual environment, the entire work done by the actor might have to be redone. About 10% of footage would need to be re-shot, which results in additional time and cost.

Since the depth-based real-time 3D virtual image composition system can verify the position where the virtual environment and the 3D model are inserted in the video in advance and monitor the outcome that the action of the actor and the virtual environment are composited through the real-time screen in the filming site, it can reduce the error that might occur in the post production process. If the number of additional takes can be reduced through this system, the production period of high quality video can be reduced and significant amount of production cost can be saved. Therefore the vitalization of diverse and experimental video production will be expected.

ACKNOWLEDGMENT

This work was supported by the Human Resource Training Program for Regional Innovation through the Ministry of Education and National Research Foundation of Korea (NRF-2013H1B8A2032217).

REFERENCES

[1] C.-K. Lee and G. Park, “A Study on Comparison of background chroma studio for Virtual Studio,” Journal of KOSST. vol. 7, no. 2, pp. 36–41, 2012.

[2] J. Park and J. Park, "Upper Body Exercise Game using a Depth Camera," Journal of Korean Society For Computer Game, vol. 25, no. 1, pp. 61–66, 2012.

[3] J. Webb and J. Ashley, Beginning Kinect Programing with the Microsoft Kinect SDK. California: O'Reilly, 2012.

[4] K.-I. Kim, S.-H. Han, and J.-S. Park, "A Distortion Correction Method of Wide-Angle Camera Images through the Estimation and Validation of a Camera Model," Journal of the Korea institude of Electronic Communication Sciences, vol. 8, no. 12 , pp. 1923–1932, 2013.

[5] S.-C. Kwon, W.-Y. Kang, and Y.-H. Jeong, "Stereoscopic Video Composition with a DSLR and Depth Information by Kinect," Journal of. Korea Inst. of Commun. Inform Sci. (KICS), vol. 38C, no. 10, pp. 920–927, 2013.