Automation System and Device based on Kinect for Implementing 3D Models

Hack-Joong Joo and Yeong-Seok Seo
Department of Computer Engineering
Gachon University, Korea
{gis8282 & ddlsck}@naver.com

Jung-II Jung and Jin-Soo Cho*
Department of Computer Engineering
PCT CO., Ltd., Gachon University
Seongnam, Korea
ceo@powerct.kr, jsocho@gachon.ac.kr

Abstract—Creating a 3D model is a difficult process that requires significant amounts of time and money. This paper proposes a system and a device for effectively implementing on-site 3D models of people or objects. The proposed method uses the depth sensor of Microsoft Kinect, which is capable of measuring and acquiring the depth data of a person or an object based on the infrared structure network. We also devised an automated video capturing device to reduce errors that occur in the process.

I. INTRODUCTION

Three-dimensional (3D) modeling is a process for creating a mathematical model that can be implemented using computer graphics in a virtual 3D space. 3D models are widely utilized in various fields, including computer games, film, and animation. Most 3D models are created using authoring tools such as 3ds Max and Maya, but using these tools requires significant amounts of time and money. Furthermore, it is difficult to express original shapes and characteristics when implementing 3D models.

To overcome such difficulties, 3D modeling can be performed by using multiple DSLR cameras or a 3D scanner. In the case of the former, precise data can be extracted for the 3D model, whereas the problem of requiring excessive time and cost is not resolved. As for the latter, although precise modeling is possible, a substantial amount of manual labor is required to capture images. If the model is large, there can be lost tracking data or errors caused by incorrect registration, and numerous repetitions are required to capture proper images.

In order to resolve these problems, this paper proposes a system and an automation device for effective on-site 3D modeling of people or objects utilizing Kinect and Kinect SDK from Microsoft, which is capable of measuring image’s depth data, as well as Kinect Fusion from Point Cloud Library.

II. DISCUSSION

A. Acquiring Kinect depth data

Microsoft Kinect is a device that senses motion by acquiring depth data through infrared (IR) pattern recognition [1]. As shown in Fig. 1, Kinect consists of a color camera, depth sensor, and microphone array. The primary object of the color camera is to acquire RGB color video. Consisting of an IR projector and IR camera, the depth sensor acquires depth data by transmitting IR patterns on to the IR projector and measure the time it takes for the patterns to return to the IR camera. The microphone array identifies and uses sound for motion sensing.

The angle of view of Kinect v2 is shown in Fig. 2. The RGB video acquired by Kinect has a resolution of 1920 x 1080 at 30 frames per second, and the depth video has a resolution of 512 x 424 at 30 frames per second.

B. 3D modeling based on depth data

Kinect Fusion is an application designed to cover 30–50 cm of indoor range. It uses the sensor on Kinect for Windows to detect 3D objects and to create a 3D model. Interacting with
the GPU, Kinect Fusion can be executed non-interactively on various hardware platforms [2].

Figure 3 displays the workflow of Kinect Fusion and the applications for implementing Kinect SDK and Kinect Fusion. Figure 3(a) depicts the workflow process of tracking and registering a 3D model by creating a depth map based on the depth data acquired from the depth sensor. According to this workflow, an application for creating a 3D model was implemented using the depth threshold and the volume option, as shown in Fig. 3(b).

**Figure 3.** 3D modeling using Kinect Fusion

(a) Workflow of Kinect Fusion (b) 3D modeling application

C. Device for automated video capturing

Figure 4 displays a miniature prototype of the device we implemented for automated scanning of 3D models. It was built by determining the minimum size of the device according to depth sensor’s angle of view.

The device consists of vertical and lateral movement units. If scanning is properly performed within depth sensor’s angle of view, the vertical movement unit moves vertically to register the overall size of the model. The lateral movement unit was designed to rotate continuously so that the model can be registered from all directions.

The device communicates with the application shown in Fig. 3(b) via Bluetooth. When tracking data is lost or integration error occurs during model registration, an error flag is sent, device motion is reversed to identify the point of the error, and the model is re-registered.

**Figure 4.** Device for automated video capturing

III. CONCLUSION

This study proposed a system and a device for effectively implementing on-site 3D modeling of people or objects. In addition to addressing the issues of requiring excessive amounts of time and money, we were also able to resolve the difficulty of having to hold a 3D scanner for a prolonged period. Furthermore, we implemented an automation device to reduce errors that frequently occur when scanning a large object, alleviating the fatigue factor that arises from numerous repetitions of image capturing.

Nonetheless, because the system proposed in this study tracks a person or an object through images captured by Kinect, we were not able to resolve the errors caused by the surrounding environment (such as light and material) and those caused by incorrect integration. Therefore, we intend to conduct further research to eliminate such errors and to enhance system functionality.

REFERENCES


