ToF와 스테레오 융합을 이용한 3차원 복원 데이터 정밀도 분석 기법

정석우* · 이연성 · 이경택

한국전자기술연구원

Analysis of 3D Reconstruction Accuracy by ToF-Stereo Fusion

Sukwoo Jung* · Youn-Sung Lee · KyungTaek Lee

Korea Electronics Technology Institute (KETI)

E-mail : swjung@keti.re.kr / yslee@keti.re.kr / ktechlee@keti.re.kr

요 약

3차원 복원은 AR, XR, 메타버스 등에서 활용되고 있는 중요한 주제입니다. 3차원 재구성을 하기 위해서는 스테레오 카메라, ToF 센서 등을 이용해 깊이 지도를 구해야 합니다. 우리는 두가지 센서를 모두 상호보완적으로 이용하여 3차원 정보를 정밀하게 구하는 방법을 고안하였습니다. 우선 두 카메라의 캘리브레이션을 적용하여 색상 정보와 깊이 정보를 일치시킵니다. 그리고 두 센서로부터의 깊이 지도는 3차원 정합과 재투사 방법을 통하여 융합하였습니다. 융합된 3차원 복원 데이터는 RTC360을 이용해 구한 정밀 데이터와 비교 분석하였습니다. 평균 거리 오차 분석을 위해 상용프로그램 Geomagic Wrap을 활용하였습니다. 제안하는 방법을 구현하고 실공간 데이터를 이용하여 실험을 진행했습니다.

ABSTRACT

3D reconstruction is important issue in many applications such as Augmented Reality (AR), eXtended Reality (XR), and Metaverse. For 3D reconstruction, depth map can be acquired by stereo camera and time-of-flight (ToF) sensor. We used both sensors complementarily to improve the accuracy of 3D information of the data. First, we applied general multi-camera calibration technique which uses both color and depth information. Next, the depth map of the two sensors are fused by 3D registration and reprojection approach. The fused data is compared with the ground truth data which is reconstructed using RTC360 sensor. We used Geomagic Wrap to analysis the average RMSE of the two data. The proposed procedure was implemented and tested with real-world data.

키워드

3D reconstruction, ToF, stereo camera, sensor fusion

Ⅰ. Introduction

Sensors using computer vision system usually provide both color and depth information. To achive accurate depth infromation, fusion of multiple sensors or expensive sensors are needed. The depth information can be applied to various computer vision solutions, e.g., Virtual Reality (VR), Extended Reality (XR), Metabus, and 3D map reconstruction.

There are many devices and algorithms for real-time depth map acquisition including active and passive camera systems. The active sensor includes structured light cameras and well kown Time-of-Flight (ToF) sensors while the passive camera includes stereo camera. ToF camera computes the depth by sending an electromagnetic wave signal and measuring the phase shift of the reflected signals. Even if the ToF sensor robustly estimates the 3D information robustly and has compact configuration, it has a high level of noise and a low performance on irreflective

* corresponding author

surfaces. Also, the ToF camera can not extract color information of the scene. On the other hand, stereo camera computes the depth by finding the two corresponding pixels in the two images acquired by the calibrated two cameras. Stereo vision system only uses simple camera setup and it is widely used for the high resolution range image. However, the stereo vision system is highly affected by the texture of the scene, and it has low accuracy relative to the active sensors[1].

This paper proposes a method to obtain accurate depth maps using both 3D information and 2D image acquired by a ToF sensor and stereo camera.

II. Proposed Method

The acquired two depth map from two sensors can be converted to point cloud data. As the two point cloud data are jointly calibrated in our procedure, the reconstruction procedure as follows.

Figure 1. ToF-stereo fusion result

First, the depth map acquired from the ToF sensor is projected to the referenced stereo camera view point. As the ToF captured low resolution image, it is necessary to interpolate the depth map to the high resolution depth map image aligned with the stereo camera lattice[2]. Next, A two point cloud is extracted from the two depth map, and they are aligned using ICP (Iterative Closest Point) algorithm[3]. The point cloud from the stereo camera is set as target, and the point cloud from the ToF camera is set as source point cloud. Then, The two point cloud are fused and reprojected to the color camera viewpoint. As the point cloud is reprojected to the different viewpoint of the ToF camera, the depth map includes sparse hole looking depth map as shown in Fig. 3. The hole in the depth map is filled with simple hole-filling algorithm. Finally, The fused depth map of ToF and stereo camera is colored with texture mapping of corresponding color image [4,5]. Fig.1 shows the example of the proposed method.

In the next section, various methods are tested to analyze the performance of 3D reconstruction procedure.

III. Experiments and Result

We used RTC360 (Leica, Germany) laser scanning device to obtain the ground truth of the real world indoor and outdoor scene data. We also used Artec Leo (Artec3D, Luxembourg) to obtain the ground truth data of the small object. The tested data is shown in Fig. 2.

Figure 2. Real-world image and the reconstructed ground truth data

The tested proposed method reconstruction results and the analysis are shown in Fig. 3.

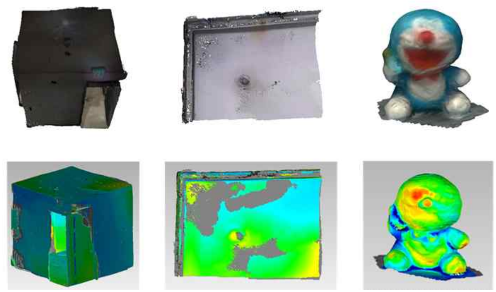

Figure 3. Tested result of the proposed procedure

III. Conclusion

In this paper, we proposed a 3D reconstruction method using both ToF and Stereo Camera. The proposed method is implemented and tested.

The limitation of the proposed method is the calibration error by the two camera. As the test configuration is not packed with rigid frame, the abrupt motion could cause the two camera to vibrate separately. The result of the calibration error cause the mismatch of the depth map and color image and cause error to the feature matching process. This limitation should be considered in the future research.

Acknowledgement

This research was supported by the Institute of I nformation & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea Go vernment (MSIT) (No.2022-0-00970, Development of multi-image-based 3D virtual space and dynamic obj ect reconstruction technology for manufacturing site support).

References

[1] Y. Kim, C. Theobalt, J. Kosecka, B. Miscusik, S. Thrun, “Multi-view image and ToF sensor fusion for dense 3D reconstruction,” *IEEE International Conference on Computer Vision Workshops*, pp. 1542-1549, 2009.

[2] S. Jung, S. Park, KT. Lee, “Pose Tracking of Moving Sensor using Monocular Camera and IMU Sensor,” *KSII Transactions on Internet & Information Systems*, vol 15, no 8, 2021.

[3] S. Jung, S. Song, M. Chang, S. Park, "Range Image Registration based on 2D Synthetic Images,” *Computer-Aided Design*, vol 94, pp.16-27, 2018.

[4] S. Jung, Y. Cho, D. Kim, M. Chang, "Moving Object Detection from Moving Camera Image Sequences Using an Inertial Measurement Unit Sensor,” *Applied Sciences*, vol 10, pp. 268, 2020.

[5] S. Jung, Y. Cho, KT. Lee, M. Chang, "Moving Object Detection with Single Moving Camera and IMU Sensor using Mask R-CNN Instance Segmentation,” *International Journal of Precision Engineering and Manufacturing*, vol 22, no 6, pp. 1049-1059, 2021.