An Experiment of Three-Dimensional Point Clouds Using GoPro

Jong-hwa Kim, Mu-wook Pyeon, Yang-dam Eo, Ill-woong Jang

Abstract—Construction of geo-spatial information recently tends to develop as multi-dimensional geo-spatial information. People constructing spatial information is also expanding its area to the general public from some experts. As well as, studies are in progress using a variety of devices, with the aim of near real-time update. In this paper, getting the stereo images using GoPro device used widely also to the general public as well as experts. And correcting the distortion of the images, then by using SIFT, DLT, is acquired the point clouds. It presented a possibility that on the basis of this experiment, using a video device that is readily available in real life, to create a real-time digital map.

Keywords—GoPro, SIFT, DLT, Point Clouds.

I. INTRODUCTION

IN recent years, construction of geospatial information and service technology, are moving to multi-dimensional geospatial information services in three-dimensional [1]. In the past, spatial information was obtained and owned by experts only [2]. Users are able to use and join realistic three-dimensional geospatial information services and various Web-based contents recently [3]. When many people not only geospatial information available, it is possible to build in this way, HLM (Hyper-Live Map) service and building is determined by the fusion of a variety of information.

HLM uses various types of data as well as global information such as satellite imagery, to information on individual such as the location of people, the mobile phone photo, SNS (Social Networking Service) data [4]. An integral part in building of HML is the construction of real-time digital map. When creating a digital map of existing, a time interval exists between the data acquisition using airborne LiDAR (Light Detection and Ranging), satellite imagery and the mapping in the office. However, as HLM, a digital map in real time must be updated to reflect the current location of the user, the terrain change immediately and to connect to SNS. To make real-time digital map, the equipment is can be used such as CCTV videos, 3D Flash LiDAR, ToF (Time of Flight) Camera and Kinect, and a technique of the sensor and data fusion is essential [5].

II. VIDEO CAPTURE DEVICE

By the development of the video capture device, most people

Jonh-Hwa Kim and Yang-Dam Eo Kim are with the Department of the Advanced Technology Fusion, Konkuk University, South Korea (e-mail: fff307@ konkuk.ac.kr, eoandrew@konkuk.ac.kr, respectively).

Mu-Wook Pyeon and Ill-Woong Jang are with the Department of the Civil Engineering, Konkuk University, South Korea (e-mail: neptune@konkuk.ac.kr, jjang7551@konkuk.ac.kr, respectively).

carry a digital camera or a mobile phone with built-in camera. In addition, using a video camera, such as GoPro, to save and upload to the web site by shooting the video of their activity. It is used to being the most popular among a large number of cameras with similar size and function. Thus, GoPro was used that can be easily obtained and used widely as described above, to get the point cloud.

The model of GoPro is HERO 3+ Black Edition. Weight is 74g.

Fig. 1 Size of GoPro

III. EXPERIMENT AND RESULT

A. Video and Reference Points Acquisition

Using a 1080p SuperView as Shooting resolution, so the screen resolution is 1920×1080. The shooting location is engineering building, Konkuk University, in Seoul, South Korea, and a subject of the experiment is a part of C-dong, engineering building. Placing GoPro to the right, and shooting video for about 1 minute, then putting the camera to the left, and shooting video about one minute (see Fig. 2).

The reference points were obtained using a total station for the calculation of three-dimensional data at the DLT (Direct Linear Transformation) method.

Fig. 2 Stereo video acquisition using Gopro

TABLE I
VIDEO MODE SPECIFICATION

Resolution	4K	4K 17:9	2.7K	2.7K 17:9	1440p	1080p Super View	720p Super View
NTSC fps	15	12	30	24	48 30 24	48 30 24	100 60 48
PAL fps	12.5	12	25	24	48 25 24	48 25 24	100 50 48
FOV	Ultra Wide	Ultra Wide	Ultra Wide Medium	Ultra Wide Medium	Ultra Wide	Ultra Wide	Ultra Wide
Screen Resolution	3840x2160	4096x2160	2704x 1524	2704x 1440	1920x1440	1920x1080	1280x720
Screen Ratio	16:9	17:9	16:9	17:9	4:3	16:9	16:9

B. Fisheye Lens Correction

The GoPro, since using a fisheye lens, distortion occurs on the Peripheral part of the acquired video (see Fig. 3).

(a) Distortion of left video   (b) Distortion of right video

Fig. 3 Fisheye lens distortion occurs in stereo video

The distortion of fisheye lens was corrected using GoPro Studio that is a dedicated program for GoPro (see Fig. 4).

(a) Corrected left video   (b) Corrected right video

Fig. 4 Stereo video distortion has been corrected

It can be seen that the peripheral part of the video distortion has occurred severely is removed when comparing the video after correction to the uncorrected video.

C. Convert to Image

In order to perform the image processing, converting stereo video to stereo image and saving it (see Fig. 5).

(a) Converted left video   (b) Converted right video

Fig. 5 Converted stereo image

D. Enhancement of Image

Selecting the matching points of the two images using SIFT (Scale Invariant Feature Transform) method [9], [10], to build up a three-dimensional point clouds by using it. Therefore, it is important to extract the accurate and many matching points. In this case, in order to increase the number of matching points, left and right images were converted to gray images (see Fig. 6). Then, making the histogram equalization each image (see Fig. 7).


(a) Left gray image   (b) Right gray image

Fig. 6 Converted stereo grey image
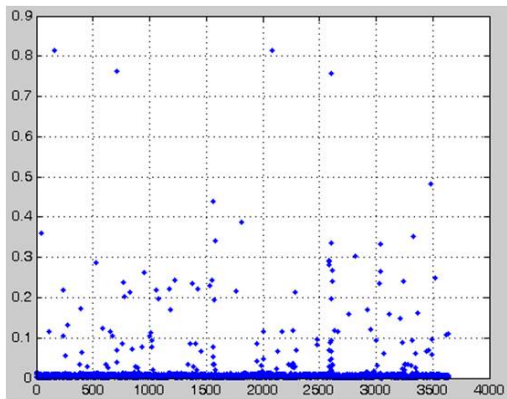

(a) Left image   (b) Right image

Fig. 7 Histogram equalization gray images

E. Matching of Image

Extracting the matching points from the stereo gray image and the Histogram equalization stereo gray image by using SIFT method obtained after the image enhancement. When obtaining a matching point from stereo gray image, the number of the matching point is 3609. When extracting from the Histogram equalization stereo gray image, the number of the matching point is 3641. However, there are incorrectly matched points that drop accuracy. Usually, these points are removed by using RANSAC (RANdom Sample Consensus) method [6], but removing easily by using the slope between the points are matched in this case. The buffer value of the slope can be displayed graphically (see Fig. 8), using only the buffer value of 0.01 or less in this case.
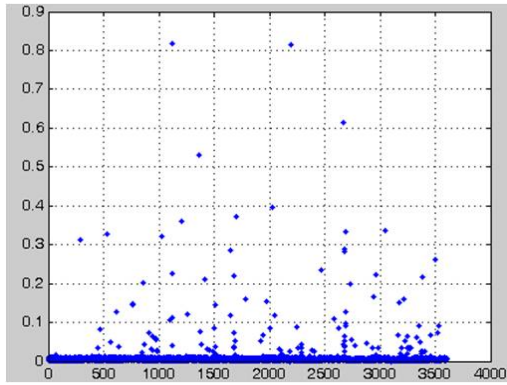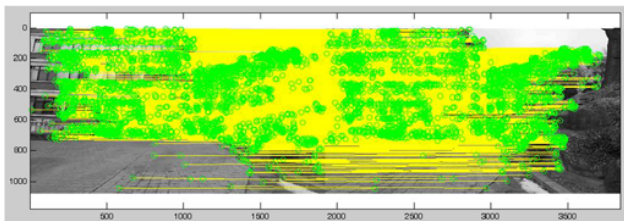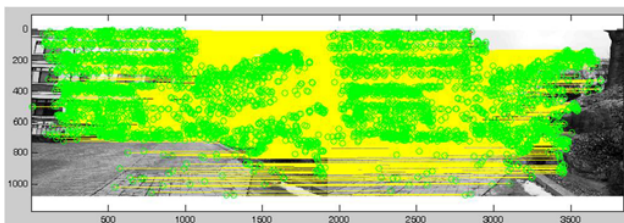
(a) Stereo gray image

(b) Histogram equalization gray images

Fig. 8 Buffer value of the slope between the points

After removing the points that are matched incorrectly, extracting the 3300, 3289 matching points from the stereo image of each. Matching points of each pair can be displayed by connecting a line (see Fig. 9)

(a) Stereo gray image

(b) Histogram equalization gray images

Fig. 9 Connecting a line between matching points

After extracting the matching points, adding the matching points for each to increase the number of matching points (see Fig. 10).
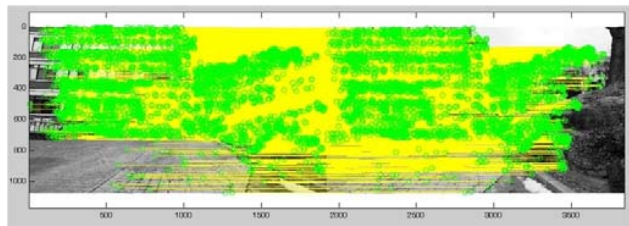

Fig. 10 Adding the matching points for each

F. Creating Point Clouds

DLT method can be divided into stages spatial intersection and spatial resection stage. In the spatial resection stage, the reference points and the matching points of the image are used to calculate the DLT factors and the image shooting location. In the spatial intersection stage, the calculated points of the image are applied to the DLT factors to produce three-dimensional data [7], [8].

In this case, the reference points are obtained using a total station. Fig. 11 represents the three-dimensional coordinate graph of the reference points. Fig. 12 represents 1:1 connections of the matching point and the reference point.
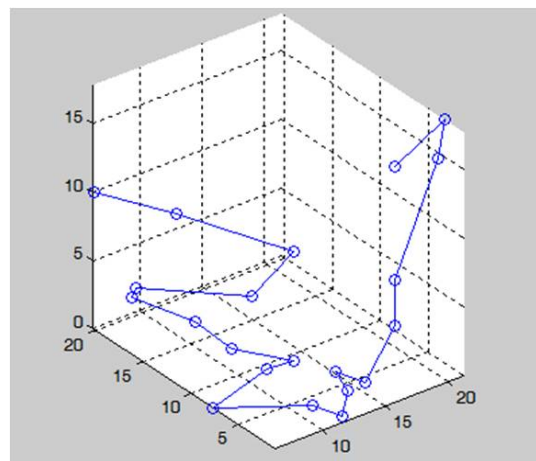

Fig. 11 Graph of the reference points

After calculating the DLT factors using the matching points connected to the reference points, to input the matching points extracted using the SIFT method. Then, the three-dimensional point clouds are created. Data created estimates coordinate in reverse, then comparing this coordinate and the coordinate of stereo image used. After this process, the error graph is plotted (see Fig. 13). Further advance the process of deleting the three-dimensional data which has been incorrectly calculated.

The data which is finally extracted is three-dimensional point clouds form of a total 4055 points as shown in Fig. 14, and every point has the pixel value.

(a) Left image

(a) Right image

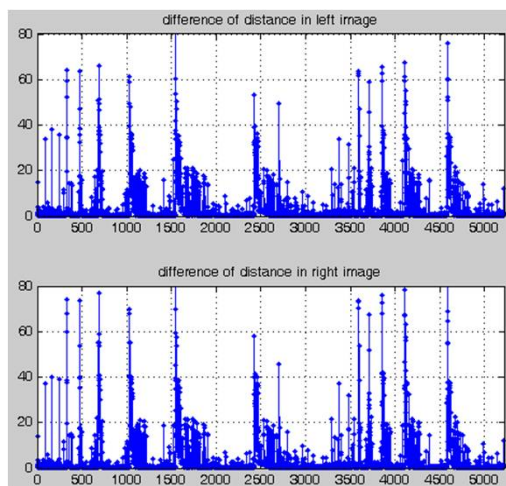Fig. 12 Connections of the matching and the reference point.
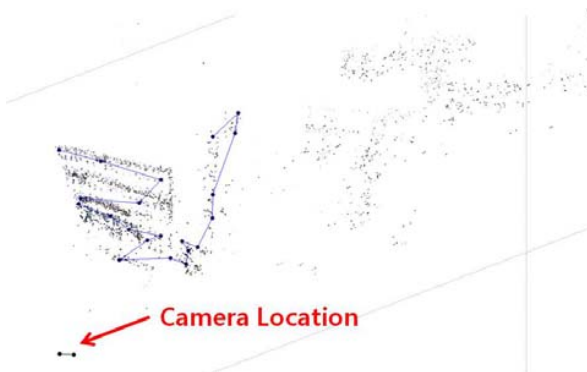

Fig. 13 Graph of the error

Fig. 14 Three-dimensional point clouds with the pixel value

IV. CONCLUSION

It was performed to configure the stereo camera by using widely used in public GoPro, and then to extract three-dimensional point clouds as a method for constructing a real-time digital map in this experiment. As a result, it was able to obtain the point cloud of total 4055 points. If it is determined in only the point clouds that have been acquired, it is true that the accuracy and the number of points acquired are not enough as compared to the terrestrial LiDAR and others surveying equipment. But if increasing the amount of point clouds by using the image of the same region obtained from multiple locations, and improving the accuracy, it can be used sufficiently as a method to construct the real-time digital map.

As future work, there is no shaking, while moving, to get the stereo video using the gimbal system that is one of the image capturing devices, and there are a number of studies and products current on it. In this way, getting a lot of data in a short period of time, and then making the mass data of large street and building by using the alignment method such as ICP on it.

REFERENCES

[1] Choi Seok Keun, Jo Ui Hwan, Lee Byung Yong, "Facilities Management Using Multi-Dimension Spatial Information," *Journal of The Korean Society for Geo-Spatial Information System*, 2010, vol. 18, no. 1, pp. 41–46.

[2] Han, S.H., "Geospatial Information System (GSIS) and LBS developments and trends in the study", *Journal of Korea Contents Association*, vol. 5, no. 1, 2007, pp. 45–53.

[3] Land & Maritime Knowledge Information Center, National Strategic Technology – Technology Development Trends in Major Advanced Countries.

[4] Jae-sun Park, Mu-wook Pyeon, Su-jung Moon, "Analysis of Cloud Computing Technology for Hyper-Live Map," *The 4th International Conference ON Internet,* Dec. 2012, pp. 5–8.

[5] Su-jung Moon, Mu-wook Pyeon, Yang-dam Eo, "Analysis of current technologies for real time multidimensional map service," *Information Science and Digital Content Technology*, vol. 3, Jun. 2012, pp. 665– 668.

[6] Wang Wei, Hong Jun, Tang Yiping, "Image Matching for Geomorphic Measurement Based on SIFT and RANSAC Methods," *2008 International Conference on Computer Science and Software Engineering.*

[7] Balletti, Caterina, and Francesco Guerra, "NM3digit: the DLT algorithm in a digital photogrammetrical software for the survey teaching," *The International Archives of Photogrammetry and Remote Sensing* 32, 1998.

[8] ABDEL-AZIZ, Y. I., "Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry," *Proc. the Symposium on Close-Range Photogrammetry, American Society of Photogrammetry*, 1971.

[9] Lowe, David G., "Object recognition from local scale-invariant features," *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, Vol. 2, 1999.

[10] Lowe, David G., "Distinctive image features from scale-invariant keypoints," *International journal of computer vision* , 2004, pp. 91–110.