PROCESSING SPEED IMPROVEMENT ON SFM USING PANORAMIC IMAGES BASED ON CAMERA DIRECTION CONSTRAINTS

Yuichiro Yamaguchi, Yoshiki Sasaki, Masafumi Nakagawa

Shibaura Institute of Technology, 3-7-5, Toyosu, Koto-ku, Tokyo 135-8548, Japan
Email: ah16076@shibaura-it.ac.jp; mnaka@shibaura-it.ac.jp

KEY WORDS: Structure from Motion; Omni-directional Camera; BIM

ABSTRACT: Point clouds are generated with structure from motion (SfM) and laser scanning for 3D modeling of objects and ground surfaces. Although SfM is a lower cost approach than laser scanning, SfM requires plenty of time for image acquisition from various positions and angles. When many images are used for dense point cloud generation, the image processing cost increase nonlinearly. Thus, SfM holds two technical issues such as image acquisition cost and image processing cost. Therefore, we firstly focused on the use of low-price omni-directional camera to improve image acquisition cost. We secondary focused on the efficiency improvement in image matching procedures in SfM processing based on camera direction constraints. We select stereo pairs to reduce processing time based on bundle constraints with image matching lists prepared in advance. We conducted experiments using panoramic images to evaluate our methodology. Through the experiments, we confirmed that our methodology can improve the efficiency of image acquisition cost and image matching processing cost.

1. INTRODUCTION

Recently, point clouds have been used for 3D modelling of structures and ground surfaces for building information modelling (BIM). BIM can improve the efficiency of construction projects with 3D models and information related to the planning, investigation, design, construction, and maintenance. Terrestrial laser scanners, mobile mapping systems, and unmanned aerial vehicles (UAVs) are generally used for 3D acquisition. Other methodology for the point cloud data acquisition is structure from motion (SfM) (Tomasi and Kanade, 1992) and multi-view stereo (MVS) (Seitz et al., 2006) with ground-based image acquisition (ground-based SfM). The ground-based SfM can achieve lower cost than laser scanners and UAVs, because inexpensive cameras can be used for image acquisition. However, the ground-based SfM generally requires more images than SfM using images taken from the UAV. When the number of input images is increased, the SfM processing time drastically increases. Based on these issues, we firstly focused on the use of low-price omni-directional camera to improve image acquisition cost. Although the low-price omni-directional camera can acquire panoramic images, the panoramic images are not available for point cloud generation with conventional SfM software products. Therefore, we proposed a methodology to generate multi-directional centric projection images from each panoramic image for the use of conventional SfM software products. We secondary focused on the efficiency improvement in image matching procedures in SfM processing based on camera direction constraints. In our methodology, we prepared image matching lists to describe stereo pairs selected from multi-directional centric projection images based on bundle constraints to reduce SfM processing time. We conducted experiments to evaluate our methodology with panoramic images. We selected a low-cost omni-directional camera. Through the experiments, we confirmed that our methodology can improve the efficiency of image acquisition cost and image matching processing cost.

2. METHODOLOGY

Our methodology consists of five steps, as shown in Figure 1. First, the panoramic images are acquired at multiple positions with an omni-directional camera. Two images are acquired at each camera position with the omni-directional camera mounted on a tripod. When the second image is acquired, the photographer should move to a different position.

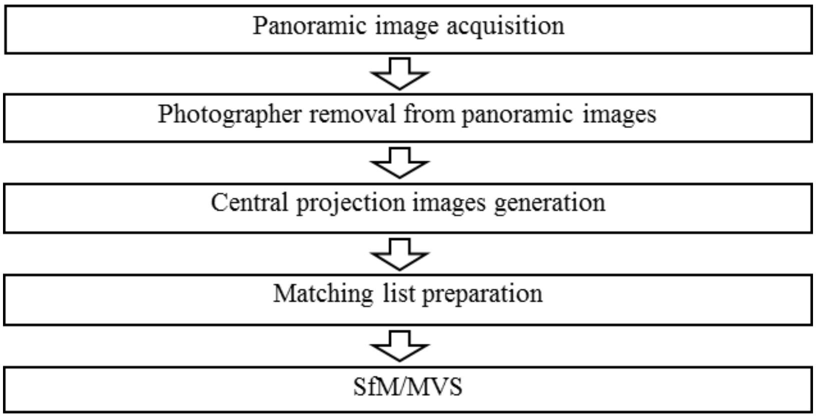


Figure 1. Proposed methodology

Second, the photographer captured in panoramic images is removed. When a camera can be controlled with wireless communication, a photographer can control the camera from a hidden position behind a wall in each image acquisition. However, the efficiency of image acquisition would decrease. Thus, we applied image mosaicking approach to remove the photographer from images.


Figure 2. Photographer removal from panoramic images

Third, 18 central projected images are generated from each panoramic image with 60 degrees angle of view. Pixel-based rendering is applied to the image projection. In the pixel-based rendering, missing parts occur after the image projection. Thus, the missing parts are interpolated with bicubic interpolation.

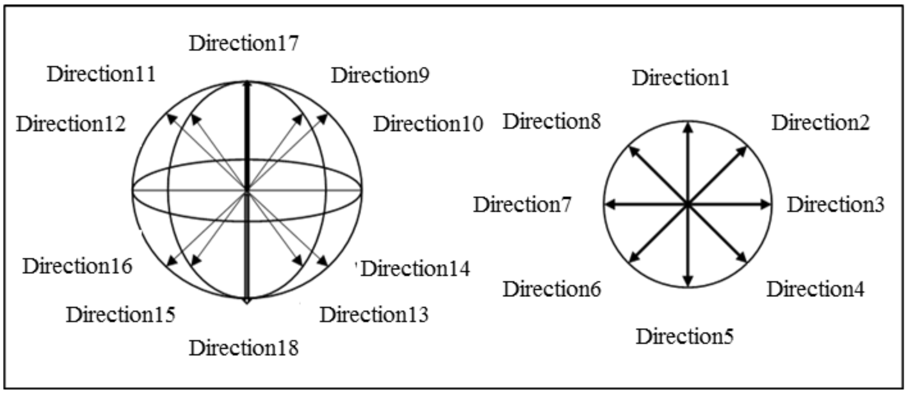


Figure 3. Directions in image projection from panoramic image to central projection images

Fourth, a matching list is prepared based on camera direction constraints. Generally, conventional SfM tries to find corresponded points from all images. Therefore, when the number of input images increases, processing time is required for image matching in SfM drastically. However, image matching candidates can be predicted with camera direction constraints. We select stereo pairs to reduce processing time based on bundle constraints with image matching lists prepared in advance.
Finally, point clouds are generated with SfM/MVS using the central projection images. In image matching, feature points and corresponded points are extracted from the images with scale-invariant feature transform (SIFT) (David, 2004), random sample consensus (RANSAC) (Fischler and Bolls, 1981), and sparse bundle adjustment (SBA) (Lourakis and Argyros, 2009). Sparse point clouds are generated with SfM, and dense point clouds are generated with MVS.

3. EXPERIMENTS

3.1. Experiments Environments

We conducted an experiment at a road bridge to verify our methodology and used THETA S (RICOH) (Table 1) mounted on a tripod. Panorama images were acquired on three baselines along the bridge length direction (Figure 4). The camera position interval was set to 50cm (Ichii et al., 2015) along baselines, and camera heights and directions were fixed. We acquired 60 images, and 30 images were obtained after photographer removal processing.

Table 1. Camera specification

Product	Pixels	Image sensor	F Value
RICOH THETA S	Approximately 14million pixels	1/2.3 CMOS	F2.0

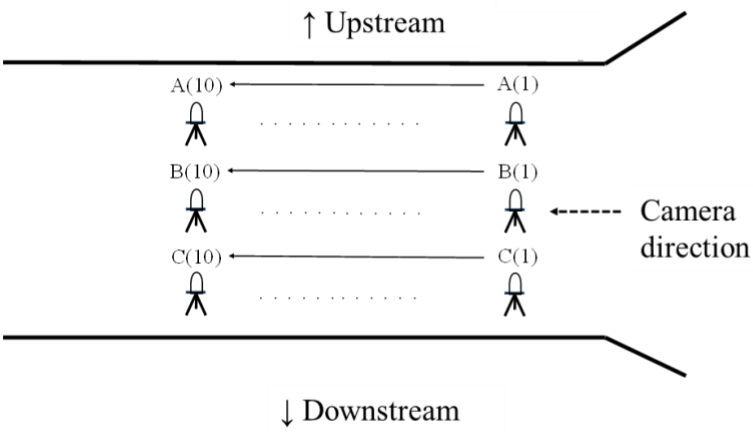


Figure 4. Image acquisition

3.2. Matching Lists

We prepared four types of matching lists to verify our methodology. We used VisualSFM for SfM/MVS and the matching lists were imported to VisualSFM. Details of matching lists are described, as shown in Figure 5. In the first constraint, central projection images were classified into four groups, such as front direction, rear direction, right direction, and left direction. In the second constraint, image matching candidates were filtered with camera directions based on bundle constraints for image matching. In the third constraint, we focused on image overlaps. We selected stereo pairs with 40% overlap or more. In the fourth constraint, we focus on easy stereo pair rejection. In the constraint, we rejected only opposite direction image pairs.

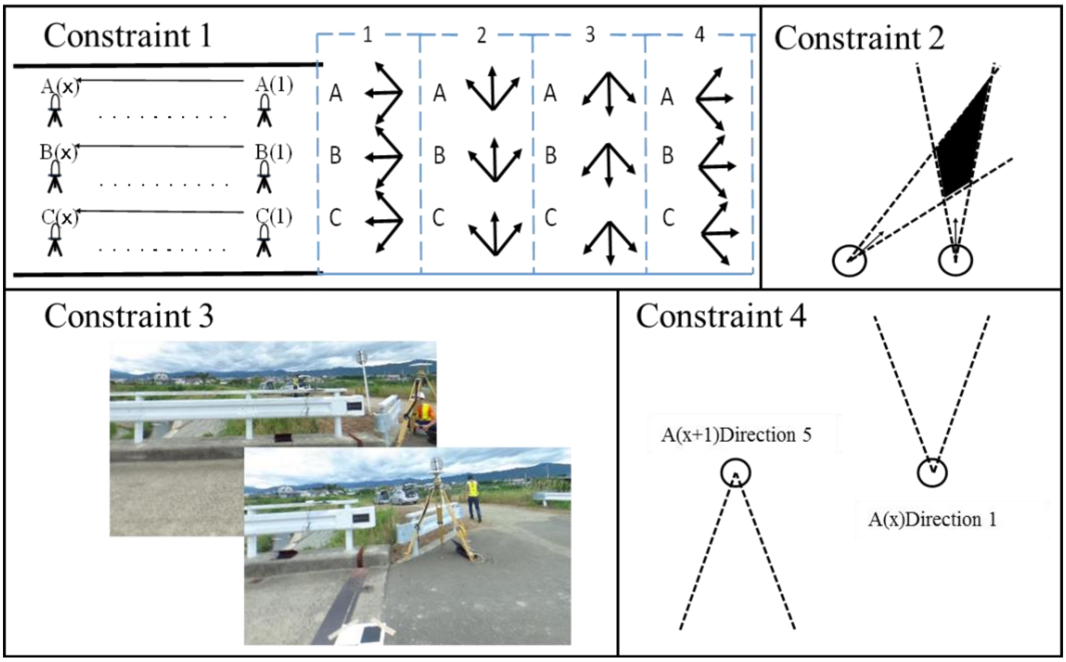


Figure 5. Four Matching Lists

4. RESULTS

4.1. Image Acquisition Cost and Image Projection

Panoramic image acquisition requires approximately 2 minutes at each position, including wireless communication and camera installations. The number of acquired images and projected images are shown in Table 2. We excluded images in perpendicular directions (upward and downward) and used 480 projection images for SfM.

Table 2. Image projection results

The number of panoramic images	The number of projected images	Processing time [s]
30	540	834

4.2. Matching Lists

The point cloud acquisition results with matching lists are shown in Table 3.

Table 3. Point cloud generation results with matching lists

Matching lists	The number of images	The number of image matching	The number of matching points	Processing time [s]	Results of point clouds
Constraint 1	480	15,648	6,960	389	×
Constraint 2	480	15,249	10,232	473	×
Constraint 3	480	43,537	28,560	1,526	△
Constraint 4	480	47,200	43,540	1,646	○
Without constraints	480	-	114,960	4,549	○

When we used constraints 1, 2, and 3, point cloud generation was almost failed because image matching was insufficient. When we used constraint 4, point cloud generation was done successfully. Generated point clouds were 90,573 points, while point clouds generated without the matching list was 101,314 points.

5. Discussion

5.1. Image Acquisition and Image Matching Processing Cost.

Although we have improved image acquisition cost with multiple central projection image generation from panoramic images, overall image acquisition time was not improved in our experiment, because tripod installation and image acquisition required 2 minutes. In order to reduce the image acquisition cost, our idea is to use an omni-directional camera mounted on a cart instead of a tripod. When we use a general digital camera, the instability of overlap management remained. On the other hand, an omni-directional camera captures a spherical space with a single shot, thus we can reduce missing areas and can manage overlap rate after image acquisition.
In image matching processing cost, the total number of image combinations without the matching list was 114,960, and the number of image combinations with the fourth constraint was 43,540. Our proposed approach provided, 38.5 % image combinations of the result estimated from the conventional approach. Although the number of output point clouds was almost same, our proposed approach was 2.7 times higher than a conventional approach. Experimental results confirmed that our approach can improve the efficiency of SfM/MVS. Moreover, our approach has the possibility to improve the scalability for SfM/MVS processing to generate massive point clouds.

5.2. Matching list

In the first constraint, a matching list was created based on camera direction clustering, such as front, rear, right, and left. However, clustering results were separated by baselines. As a result, point cloud generation was failed. In the second constraint, we confirmed that the stereo pair design was too severe to generate point clouds. Therefore, point cloud generation was failed because many image combinations were extremely omitted. In the third constraint, several combinations were checked visually, we selected image combinations with 40 % overlap or more. However, the point cloud generation failed because the number of stereo pairs was insufficient to apply triplet matching. In the fourth constraint, although the number of image matching was decreased, experimental results confirmed that the efficiency of SfM/MVS was improved and point clouds were generated successfully, as shown in Figure 6.

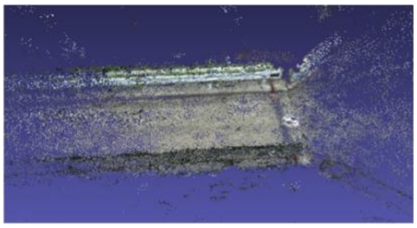
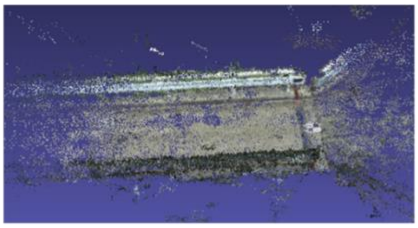
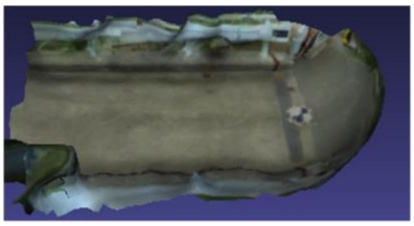
	Without Matching Lists	With Constraint 4
Point clouds		
3D models		

Figure 6. Generated point clouds and 3D models

In our experiment, the tripod was installed at 50 cm intervals. Although point clouds were obtained, many lacked areas occurred. Thus, we would examine several tripod installation patterns to improve both of point cloud generation quality and image acquisition cost.

6. CONCLUSION

In this study, we focused on the use of an omni-directional camera to improve the efficiency of image acquisition and image matching processing in SfM processing at the same time. We prepared four types of filters based on bundle constraints with image matching lists used for SfM processing. We conducted experiments with an omni-directional camera to evaluate our methodology. Through our experiments, although several filters were failed to generate point clouds, we confirmed that our methodology can improve the efficiency of image acquisition cost and image matching processing cost.

REFERENCES

David, G. L., 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2), pp. 91–110.

Fischler, M. and Bolles, R. 1981. Random sample consensus: a paradigm for model fitting with application to image analysis and automated cartography. Communications of the ACM, 24(6), pp. 381-395.

Ichii, K., Tamaki, T. Fujii, T. 2015. Detection of the deformation and condition changes of structures by 3D points clouds obtained by photographs. Report Detection of the deformation and condition changes of structures by 3D point clouds obtained by photographs, Hiroshima University.

Lourakis, M. and Argyros, A. 2009. SBA: A software package for generic sparse bundle adjustment. ACM Transactions on Mathematical Software (TOMS), 36(1), article No.2.

Seitz, S. M., Curless, B., Diebel, J., Scharstein, D. and Szeliski, R., 2006. A comparison and evaluation of multi-view stereo reconstruction algorithms. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1, pp. 519-528.

Tomasi, C., and Kanade, T. 1992. Shape and motion from image streams under orthography: a factorization method. International Journal of Computer Vision, 9(2), pp. 137-154.