The effect of angles and distance on image-based, three-dimensional reconstructions

Siyuan Chen, Debra F. Laefer, Jonathan Byrne & Atteyeh S. Natanzi
School of Civil Engineering Dublin, University College Dublin—Urban Modelling Group, Dublin, Ireland

ABSTRACT: This paper introduces a three-dimensional reconstruction experiment based on a physical laboratory-based experiment on a brick wall. Using controlled shooting distances and angles, different images sets were captured and processed with a structure from motion based technique, which can reconstruct 3D models based on multi-view, Two-Dimensional (2D) images. Those 2D geometries are shown to generate significant deformations within the resulting point cloud, especially where there were large angles (with respect the camera position and the wall's normal direction) and at close distances to the wall's surface. This paper demonstrates that by overlapping different flawed image sets, the deformation problem can be minimised.

Keywords: 3D reconstruction, Unmanned Aerial Vehicle (UAV), Structure From Motion (SFM), Aerial inspection, camera angle, camera distance

1 INTRODUCTION

Three-Dimensional (3D) reconstruction is widely used in urban modelling and infrastructure inspection. In the history of computer vision development, this topic has been intensively researched, and a variety of methods have been proposed. Generally, these methods can be classified as either active or passive. Active methods, such as laser scanning (Jarvis 1983), use active interference to detect the target and collect depth information. However, most of the complex detectors required for these methods are expensive and complex and result in restricting its scope of application. On the contrary, passive methods employ relatively inexpensive digital cameras to record light that already exists in a scene. These devices are less restrictive, but a challenge remains in determining the shape of a scene from images and videos from their Two-Dimensional (2D) information. Thus, further research has focused on abstracting the depth information from these 2D images. To achieve that, various methods have been proposed. The simplest way is imitating a human binocular visual system by using stereo cameras, which have two or more lenses with separate image sensors (Izadi et al. 2011). Like the active methods, this method relies on complex equipment and independently known locational information, which limits the ease of use. The ideal approach uses only one simple camera for the task to which image post-processing is applied. Several approaches involve shape identification through a shading method (Horn 1975) or some other means. Specifically by using shadow changes for calculation, the shape from texture method (Ikeuchi 1981) has

been used to find repeated patterns for measurement, the shape from contours method (Witkin 1981) used contour shape distortion for measurement estimation, the shape from focus method (Pentland 1987) using focus changes to compare distances. Alternatively the Structure From Motion method (SFM) (Ullman 1979) uses images taken from multiple viewpoints for triangulation. Among those methods, the SFM approach is the only method independent of specified conditions, such as shadows or surface patterns. Thus, it has been widely applied, especially as Unmanned Aerial Vehicles (UAVs) have become increasingly popular in photogrammetry applications such as bridges inspection (Hallermann 2014), landslide monitoring (Lucieer 2014), building damage detection (Sui 2014).

2 RELATED WORK

Building on the SFM approach, Longuet-Higgins (1981) introduced a means matrix to recover the camera position for better image match. Some researchers have used this with the RANdom SAmple Consensus (RANSAC) method as introduced by Fischler & Bolles (1981) to discount erroneous feature matches and with the work by Lowe (2004) in the form of the Scale Invariant Feature Transform (SIFT) method for feature extraction.

All of the aforementioned techniques focus on refining the processing procedure, either by increasing the speed or the accuracy. What is overlooked is the question of improved data acquisition strategies. As a passive reconstruction method, the

quality of SFM results is closely related to the data collection procedure. These relate to the camera angles, the offset distances, the light conditions, and the camera model. Among those factors, the angles and distances are most intrinsic to the survey process. To date, however, their impact on results and means to optimize their accuracy have yet to be comprehensively addressed. In UAV inspections, it is not clear whether a close distance or a better shooting angle will benefit the results. To begin to investigate this topic, a laboratory experiment was devised and conducted as described below.

3 EXPERIMENT SETTING

The laboratory experiment was conducted on a custom constructed of masonry wall (130×90 cm) (Figure 1). In an attempt to replicate field conditions, a lightweight digital camera or action camera commonly used with UAVs was selected: specifically, the Cannon IXUS 175 camera. The camera was tripod mounted to control the shooting angles (Figure 2). Image capture occurred at offsets from 1, 2, and 3 meters from the wall's front surface. Camera angles
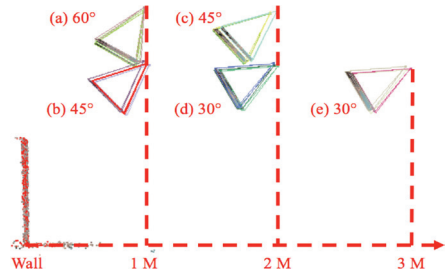

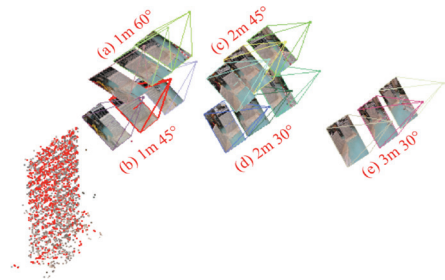Figure 3. Five positions for data capture from different angles and distances.

Figure 4. The 3 images from each of the 5 positions shown in Figure 3.
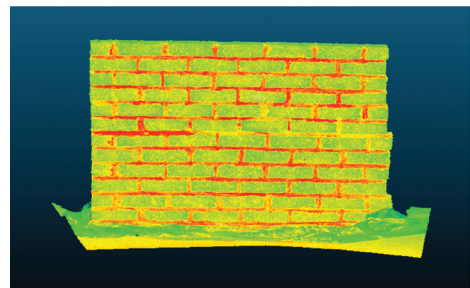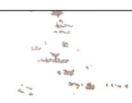

Figure 1. Camera setting.

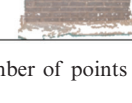Figure 2. Camera angle setting.

Figure 5. Laser scanner setting.

Figure 6. Laser scan based point cloud.

were set as 30, 45, and 60 degrees to the wall's normal direction, as shown in Figure 3. In total, 5 groups of data were collected; each group of data includes 3 images, as shown in Figure 4. As a reference dataset, the wall was scanned with a Leica ScanStation P20 terrestrial laser scanner, as shown in Figure 5. The scan resolution was 12.5 mm at 10 meters, 4698×833 points, yielding the results shown in Figure 6. For SFM reconstruction, the non-commercial platform VisualSfM developed by Wu (2013) was used. To analyse the results, CloudCompare and MATLAB software packages were used for point cloud comparison.

4 RECONSTRUCTED RESULT

Table 1 shows the SFM reconstructed result. Datasets 'a' to 'e' are based on single groups of images (3 for each time). Datasets 'f' and 'g' are based on processing two groups images (6 images for each time) in a single reconstruction. These observations, show that the datasets taken within 2 meters ('a' to 'd') have varying degrees of missing features, which intensify as the camera angle

Table 1. Point cloud result.

Dataset ID	Point cloud	Camera distance	Camera angle	Image numbers	Number of Points*
a		1 meter	60 degree	3	6,933
b		1 meter	45 degree	3	175,965
c		2 meters	45 degree	3	104,895
d		2 meters	30 degree	3	134,740
e		3 meters	30 degree	3	82.764
f		1 meter	45 degree + 60 degree	6	260,479
g		2 meters	45 degree + 30 degree	6	203,434

*The number of points only includes points in the wall area.

increases. As the camera distance increases, the situation improves. This defies common perceptions that closer data capture should be superior. A possible reason is that the longer shooting distance increased the scope of images with a fixed field of view, thus increasing the overlapping and matching between images, which is a significant part of the SFM process. Furthermore, by processing two sets of images together (f and g), reliability losses can be overcome.

5 DATA ANALYSIS

Evaluation goal was to assess both the geometric accuracy and point distribution of the reconstructions. To measure geometric accuracy, each dataset was aligned with the reference data (i.e. the scanned laser point cloud data), and five bricks were chosen as specific features (as shown in Figure 7). Then, the displacement of each feature was calculated by comparing it the SFM reconstruction to the corresponding feature in the reference data. As each feature was a group of points, the displacement was represented as mean distance and standard deviation, as shown in Table 2,
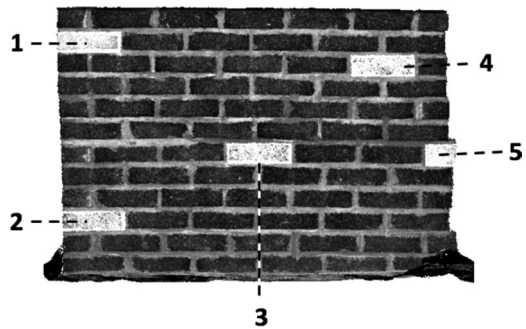

Figure 7. Selected features.

Table 2. Feature displacement.

Feature	Data set ID	a	b	c	d	e	f	g
	Camera pose	1 m 60d	1m 45d	2m 45d	2m 30d	3m 30d	1m 60d+45d	2m 45d+30d
1	Mean Distance	—	—	0.047587	0.020161	0.005395	0.001304	0.002597
	Std Deviation	—	—	0.044114	0.019675	0.006734	0.000938	0.001859
2	Mean Distance	—	—	0.005499	0.001881	0.002651	0.002539	0.00243
	Std Deviation	—	—	0.006594	0.002757	0.003003	0.00197	0.001215
3	Mean Distance	—	0.002051	0.001341	0.013858	0.001682	0.001045	0.001238
	Std Deviation	—	0.001511	0.000913	0.022349	0.001676	0.00062	0.000841
4	Mean Distance	—	—	0.002708	0.004282	0.003855	0.001244	0.001245
	Std Deviation	—	—	0.004678	0.006296	0.004886	0.000684	0.000855
5	Mean Distance	—	—	0.004184	0.040225	0.009473	0.001371	0.002007
	Std Deviation	—	—	0.004213	0.028035	0.011257	0.00078	0.002187

'—' indicates that no corresponding feature can be found.

which was calculated from the distance of each point in the SFM-feature to the nearest neighbor in the reference-features. Since datasets 'a' and 'b' were highly incomplete which resulted in indistinguishable features, the displacements could not be compared. For datasets 'c' to 'g', the results are plotted in Figure 8 showing that point clouds reconstructed from a single group of images have bigger errors while those that included a secondary image taken from a different angle had improved geometric accuracy. The datasets 'f' and 'g', which contain two groups of images, had nearly indistinguishably levels of high geometric accuracy for feature displacement.

All distances are shown in meters.

In addition to geometric accuracy, surface point distribution can be another important characteristics for SFM-based point cloud evaluation. Unlike laser scanning where a laser beam moves over the surface with a constant speed, SFM seeks common points that appear in overlapping images. The image overlap rate in different regions are different and result in a non-uniform point cloud. In this experiment, in order to compare the point distribution, a searching radius of 0.1 m was set to calculate the density variance on the surface. The results are shown in Figure 9.

In this comparison a greater distinction in the data quality in sets (f) and (g) can be seen. Over all, in the horizontal direction, the density in the center is much denser than the boundary. This mainly due to the overlapping is higher in the middle. In the vertical
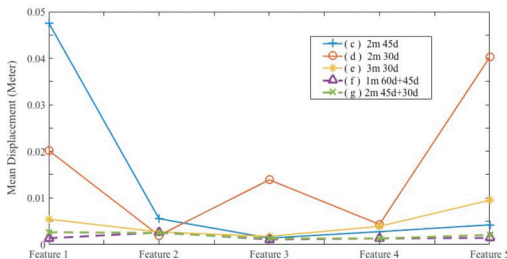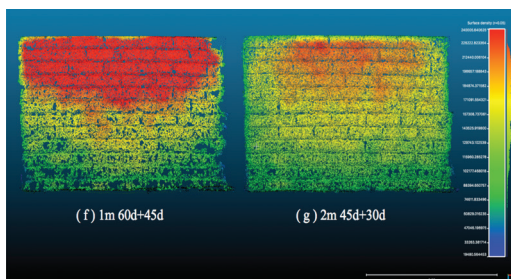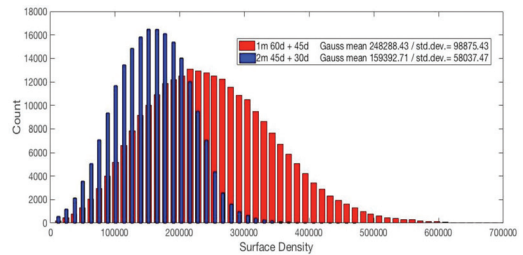

Figure 10.   Points statistical distribution.

direction, in case, the points are much denser on the upper side. In case (f) the data captured from 1 meter away at 60° and 45° shooting angles have a much greater variation. In the upper portion of both samples, there is a wide, high-density area, while the bottom portions of the wall have more empty spots. This is much worse in case (f). A possible reason is that the bigger shooting angle will increase the perspective effect. The upper part, which is closer to the camera, occupied most of the pixels in the images. Thus, the bottom part lacks sufficient data for uniform presentation. Furthermore, the shorter distance decreased the image scope, which results in poor overlapping. Thus, some features at the bottom cannot be matched from multi-viewpoints for triangulation. In contrast, dataset 'g' gives a better presentation of the bottom features and has a more uniform surface distribution.

To better interpret the situation, the statistical distribution of the points is shown in Figure 10. Although the average density of data taken from 1 m is higher than the data taken from 2 m, the standard deviation of the former is almost doubled. This means that even though the point cloud taken from 1 m contains more points, they are not well distributed. For example, some surface has extremely high density, even exceed what is needed, while some surface was not dense enough to describe the feature. Meanwhile, the smaller standard deviation number in the small angle dataset proved that the surface density is more uniform.

Figure 8.   Displacement of each feature.

Figure 9.   Density variance in the surface.

6   DISCUSSION

The experiment and analysis demonstrated that images obtained from closer locations or with larger shooting angles were less reliable, resulting in an incomplete point cloud. A good strategy to solve this problem is adding another group of non-colinear images to the current dataset. On the other hand, even when combining multiple groups of data, big shooting angles (i.e. 90 degrees) and short distances will lead to poorly distributed results in which most points are concentrated in only small areas and other details remain only partially documented.

The results provide meaningful insights for real-life UAV inspection strategies. When inspecting a slender object, such as road or coastline, scanning it twice with different shooting angles will improve the SFM results. Moreover, with a fixed lens camera, instead of getting too close to the target, keeping a relatively longer distance and a smaller shooting angle will be more appropriate when generating the flight path.

7 CONCLUSIONS

In this paper, camera position was tested in a lab experiment to consider the effects of varying camera angles and distances on SFM reconstruction accuracy when compared to a single terrestrial laser scan. Larger shooting angles were found to significantly degrade the data completeness level, geometric accuracy, and point uniformity. Overall, a relatively greater distance combined with a smaller shooting angle in UAV inspections is recommended. In addition to that, by adding one more group of offsetting images, the data reliability can be significant improved.

REFERENCES

Bay H, Tuytelaars T, Van Gool L. 2006 SU RF: Speeded Up Robust Features [C]. Proce edings of the 9th European Conference on Computer Vision. Austria, 2006: 404–417.

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

Hallermann, N. & Morgenthal, G. 2014. Visual inspection strategies for large bridges using Unmanned Aerial Vehicles (UAV). *Isbm 2014*, pp. 661–667.

Harris C, Stephens M 1988. A combined corner and edge Detector[C].Proceedings of the Fourth Alvey Vision Conference. [S.l.]: [s.n.], 1988: 147–151.

Horn, B.K.P. 1975. Obtaining shape from shading information. In Winston, P.H. (ed.), The Psychology of Computer Vision, pp. 115–155, McGraw-Hill, New York.

Ikeuchi, K. 1981. Shape from regular patterns. Artificial Intelligence, 22(1):49–75.

Izadi, S. et al., 2011. Kinect Fusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera. *Proceedings of the 24th annual ACM symposium on User interface software and technology—UIST '11*, p. 559. Available at: http://dl.acm.org/citation.cfm?id = 2047270%5Cnhttp://dl.acm.org/citation.cfm?doid = 2047196.2047270.

Jarvis RA 1983. A perspective on range finding techniques for computer vision. IEEE Transactions Pattern Analysis Machine Intelligence PAMI-5, 2:122–139.

Longuet-Higgins, H.C. 1981. A computer algorithm for reconstructing a scene from two projections. Nature, 293:133–135.

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

Lucieer, A., de Jong, S.M. & Turner, D. 2014. Mapping landslide displacements using Structure from Motion (SfM) and image correlation of multi-temporal UAV photography. *Progress in Physical Geograph*, 38(1), pp. 97–116.

Pentland, A.P. 1987. A new sense for depth of field. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-9(4):523–531.

Slama, C.C. (ed.). 1980. Manual of Photogrammetry. American Society of Photogrammetry, Falls Church, Virginia, fourth edition.

Sui, H. et al., 2014. A Novel 3D Building Damage Detection Method Using Multiple Overlapping UAV Images. *ISPRS—International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-7(October), pp. 173–179.

Ullman, S. 1979. The interpretation of structure from motion. Proceedings of the Royal Society of London, B-203:405–426.

Witkin, A. 1981. Recovering surface shape and orientation from texture. Artificial Intelligence, 17(1–3):17–45.

Wu, C., 2013. Towards linear-time incremental structure from motion, in: International Conference on 3-D Vision – 3-DV, Seattle, WA, USA, 127–134.