A study on performance of three-dimensional imaging system for physical models

B.T. Le
Ho Chi Minh city University of Transport, Ho Chi Minh city, Vietnam, formerly City, University of London, London, UK

S. Nadimi
University of Leeds, Leeds, UK, formerly City, University of London, London, UK

R.J. Goodey & R.N. Taylor
City, University of London, London, UK

ABSTRACT: A study by Le et al. (2017) reported the application of computer vision techniques structure from motion (SfM) and multi-view stereo (MVS) to measure three-dimensional soil displacements at the surface of physical models. However, little information exists on the significance of the camera resolution and the number of images to the measurement performance. This study assesses the measurement performance of the SfM-MVS, provided by an open source software Micmac, with input images taken by two different types of camera including DSLR (18Mega-pixel) and mobile phone cameras (12Mega-pixel). Rigorous quantifications were carried out to examine the precision of the image analysis, in measuring vertical and horizontal displacements, over a region of interest of 420x200mm. The measurement precision, achieved by different numbers of images, ranged from 0.06mm to 0.03mm. The results from this paper can be useful for researchers to select appropriate camera that satisfies their measurement requirements.

1 INTRODUCTION

The three-dimensional imaging system in this research combines the computer vision technique “structure from motion” and “multi-view stereo” (SfM-MVS) delivered by the open source software Micmac (Galland et al. 2016) with 2D PIV (Stanier et al. 2015) to measure displacements in three dimensions (3D).

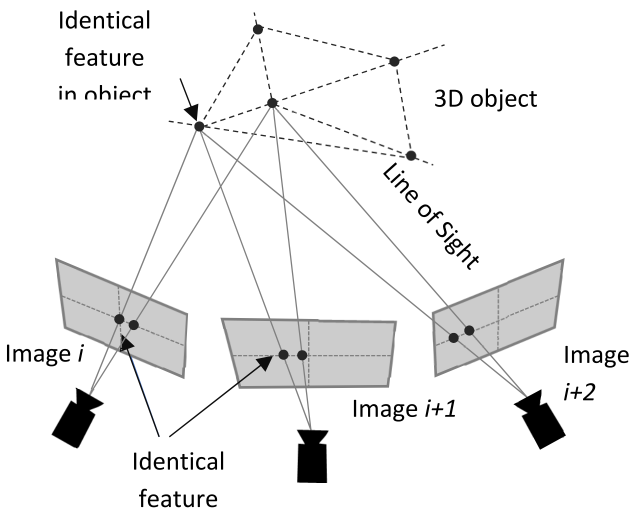


Figure 1. Structure from Motion (SfM) principle (Le et al. 2016).

SfM is a technique from computer vision and photogrammetry field that analyses input images to produce a high quality 3D point cloud (Ullman 1979). It has wide range of applications on both large and small scales. For large scale, Smith et al. (2015) reported SfM-MVS were used in 3D topographic surveys, monitoring glacier movements, observing and tracking lava movements and landslide displacements. For small scale such as experiments in laboratory, Galland et al. (2016) and Le et al. (2016) used SfM-MVS to measure three-dimensional displacements.

Simply put, SfM-MVS processes a minimum of two images to reproduce the 3D point cloud of the observed scene. The SfM-MVS algorithm is described in detail by Robertson & Cipolla (2009) and Le et al. (2016). The fundamental principles are illustrated in Figure 1 and described briefly below.

1.1 *Structure from Motion*

Firstly, the identical features known as keypoints in each image are detected and assigned with a unique identifier. An identical feature is a set of pixels that are invariant to changes in scale and orientation and can be detected in other images. The feature detection used in this research is the Scale Invariant Feature Transform (SIFT) algorithm (Lowe 2004).

The locations of the features in multiple images determined in the first step are used in a process named bundle adjustment to estimate the parameters of the scene including individual positions of the cameras, orientation of the cameras, intrinsic camera parameters and relative locations of the features in object space. Images taken from different positions add more data to the bundle adjustment process which improve the precision of the parameters estimated (Triggs et al. 1999) which is confirmed later in this paper.

1.2 *Multi-View Stereo Image Matching*

The 3D point cloud obtained from SfM has a coarse density as only identical features were included. Additional matching algorithm named Multi-View Stereo (MVS) is normally carried out after SfM which can increase the density of the 3D point cloud by at least two orders of magnitude (Furukawa & Ponce 2010, James & Robson 2012, Smith et al. 2015). During MVS analysis, most of noise data points (outliers) will be removed.

1.3 *Georeference*

After the MVS step, the obtained 3D point cloud is in image space (i.e. unit: pixel) and needs to be transformed to object space (e.g. unit: mm) by the Georeferencing process. Basically, georeferencing process uses the provided positions of the Ground Control Points, in image and physical space, to transform the 3D point cloud to physical space. The minimum number of Ground Control Points is three but more points provide better precision in the transformation process.

2 REVIEW ON SFM-MVS TECHNIQUE AND MICMAC SOFTWARE

There are several software that feature SfM-MVS technique such as commercial software Agisoft Photoscan and free software Bundler (Snavely et al. 2006), VisualSfM and Micmac. Smith et al. (2015) reported that MicMac, with sophisticated self-calibration camera models, outperformed Agisoft Photoscan and VisualSfM.

Galland et al. (2016) used Micmac to analyse four images taken by 24 Mega-Pixel cameras and the achieved measurement precision was 50µm. Similarly, Le et al. (2016) reported a precision of 50µm was achieved by using Micmac to analyse three images, taken by three 2Mega-pixel cameras.

Despite the technique SfM&MVS and the software Micmac were known to be able to produce high quality 3D point clouds, there was no guidance on the effects of the camera resolution and the number of images to the measurement performance. This paper aims to provide a clearer insight into these two factors.

3 EXPERIMENT AND RESULTS

Two cameras used in the following experiments are Canon EOS 700D DSLR Camera (18Mega-pixel sensor, 18-55mm lens) and Iphone 6s (12Mega-pixel sensor, 4mm focal length lens).

3.1 *Experiment setup to determine precision of vertical measurement*

Figure 2 illustrates the set up that includes a reference plate with 59 Ground Control Points (GCPs) and three blocks with known heights obtained from micrometer. A paper sheet with speckle texture was fixed to the observed flat objects to aid the SfM-MVS process. This is because SfM-MVS requires textures to detect identical features as plain surfaces can not be distinguished.

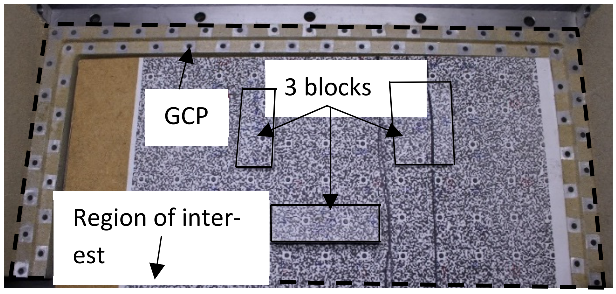


Figure 2. Experiment set up to determine precision of vertical measurement.

Five photos were sequentially taken by each camera for image processing purpose. For each camera, there were three different analyses using three, four and five photos to investigate the effect of number of images, in addition to the effect of the camera resolutions, to the performance of the measurements.

The precision of the measurement is determined by comparing the heights of the blocks in the 3D point clouds with the known heights as described by the value mean absolute error (MAE) (Equation 1);

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left(\left|H_i^{3D} - H_i^{M}\right|\right) \tag{1}$$

where

H_i^{3D} is the height of block *i* in the 3D point cloud,

H_i^{M} is the height of block *i* measured by a micrometer,

n is the number of the objects, for the experiment for vertical measurements n=3 blocks.

Tables 1 and 2 shows the performance of the measurement, in terms of number of data points and precision, using photos from the two different cameras. Over a ROI of 420x200mm, the number of data points obtained from the DSLR camera with a 18 Megapixel sensor was approximately 1.5 times more than that for the phone camera with a 12 Mega-pixel sensor (Table 1). This is thought to be analogous with the ratio of the number of pixels in the sensors of the two cameras. Interestingly, the number of the images does not have considerable impact to the number of data points.

The best precision was approximately 30μm achieved in analyses utilised five photos for both cameras (Table 2). For analysis that utilised only three or four photos, the precision was decreased to approximately 50-60μm for both cameras. All the experiments, apart from the abnormally high MAE in iPhone 6s with four photos, shows that more images yielded higher measurement precision. This is because more images are beneficial for the camera calibration process to determine more precise camera parameters.

The effect of camera resolution seems to be not significant to the measurement precision as IPhone 6s (12Mega-pixel) and DSLR Canon (18Mega-pixel) yielded similar precisions which are in line with the results reported by Galland et al. (2016) and Le et al. (2016).

It can be seen that camera resolution is not the only factor that governs the quality of images and hence the quality of 3D point clouds. Even though high resolution images enable more identical keypoints to be detected and corresponded in SfM&MVS process, the quality of lens is also important. If a low quality lens is used with a high resolution camera, then the obtained image may decrease the quality of the 3D point clouds (Furukawa & Hernández 2015). The lens also controls the depth of field and the sharpness of the image across the whole field of view

Other factors that also needs to be taken into account when considering the quality of the camera parameters and SfM-MVS process are the type of the camera sensor that affects the noise in the obtained image.

CCD sensors are known to be able to capture high quality images with low noise but they are more expensive than traditional CMOS sensors which are susceptible to noise. However, recent developments in imaging technology allow new CMOS sensors to capture images with comparable quality to those obtained by CCD sensors but at a more affordable price.

The EMVA (European Machine Vision Association) data of the sensors are normally available and is useful for comparison on the characteristics of the sensors. In addition to sensor types, pixel size and sensor size are also important factors. Larger sensor size and pixel size allow larger amount of light into the sensor hence better quality images.

Table 1. Number of data points

	Number of photos		
Camera	3	4	5
DSLR Canon (18Mega-pixel)	2,645,196	2,117,084	2,114,078
IPhone 6s (12Mega-pixel)	1,345,463	1,351,630	1,348,967

Table 2. Mean absolute error in vertical direction (Unit: μm)

	Number of photos		
Camera	3	4	5
DSLR Canon (18Mega-pixel)	57	55	30
IPhone 6s (12Mega-pixel)	44	61	32

3.2 *Experiment setup to determine precision of horizontal measurement*

Figure 3 presents the experiment set up that comprises Ground Control Point and a ROI which can be displaced in a precise manner using a sliding bed controlled by two micrometers.

Four experiments have been conducted using the controlled movement of a sliding bed to estimate the measurement precision in the horizontal directions. In each experiment, the sliding bed was moved by 1 mm in either X or Y directions. At each displacement, one image (test image) was captured for later analysis.

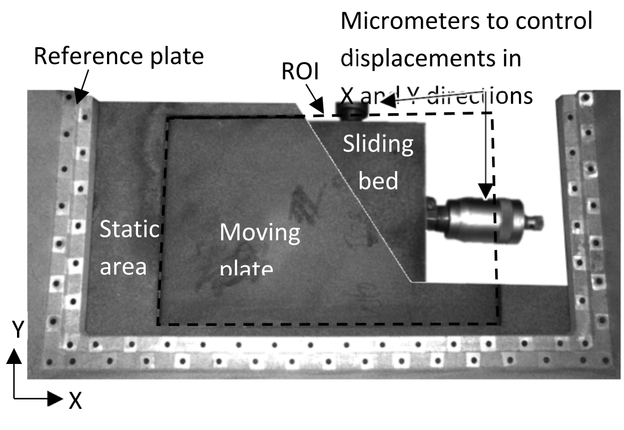


Figure 3. Experiment set up to quantify horizontal displacement measurement precision (after Le et al. 2016).

The test images were undistorted and unwarped before being analysed by conventional 2D PIV (Stanier et al. 2016). The detailed procedure is described by Le et al. (2016) and illustrated in Figure 4.

Apart from the test images, a separated set of calibration images containing a series of ring patterns were taken by each camera. These calibration images were used for camera calibration process to determine the camera parameters. A minimum of three calibration images are normally required. In this study, five images were used for calibration purpose (Figure 5). The ring patterns were used as their centres can be determined precisely in comparison with square and circular patterns. The centre of each ring is determined based on outer circle and is refined based on inner circle of the ring. This procedure is similar to the camera calibration procedure using the checkboard pattern described by Zhang (2000).

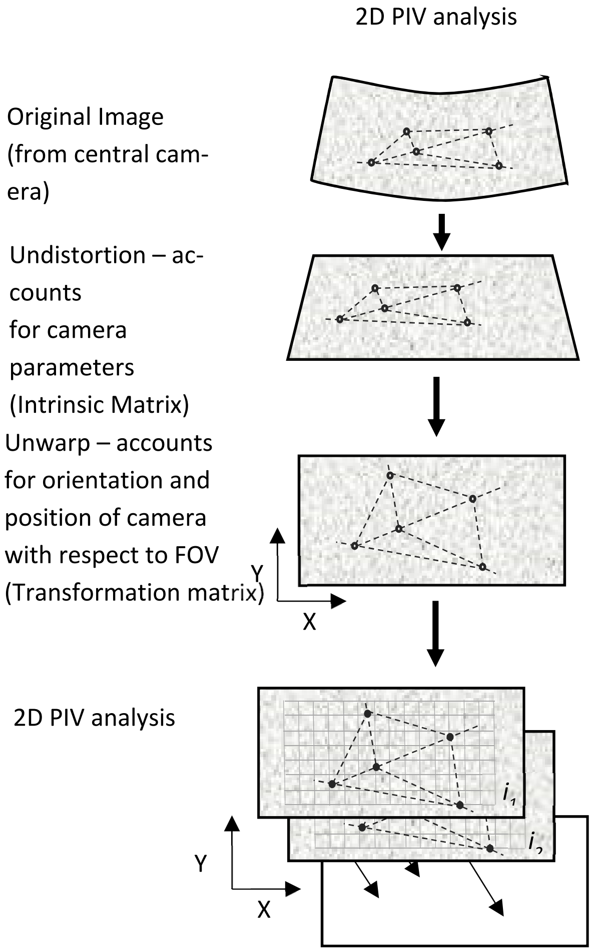
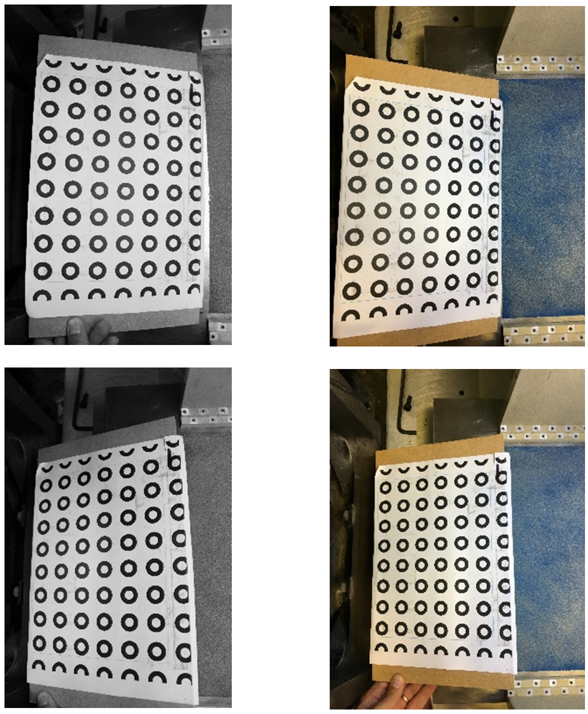


Figure 4. PIV analysis procedure to determine horizontal displacements (after Le at al. 2016)

The determined camera parameters were then used in an in-house Matlab code to remove the distortion in the test images.

The upwarp step performing on the undistorted test images requires the positions of the Ground Control Points in physical space and image space in order to determine the position and orientation of the camera, to correct the images. Finally, the PIV analysis is performed on the rectified images to determine horizontal displacements.



(a) (b)

Figure 5. Typical calibration images of the ring pattern of a) Canon camera and b) iPhone camera

The error is calculated by comparing the displacements determined from the PIV analysis and the movements caused by the micrometers. Table 3 presents the MAE in horizontal displacement measurements in from the four experiments. As can be seen from these experiments, higher camera resolution offers slight improvement on the measurement precision (lower MAE).

Table 3. Mean absolute error in the horizontal direction (Unit: μm)

Experiment	DSLR Canon (18Mega-pixel)	IPhone 6s (12Mega-pixel)
1	22	30
2	12	47
3	5	13
4	14	23

The error in this method accumulates from camera calibration, undistortion, unwarp and PIV analysis. Therefore, minimising the error in each step will reduce the error hence improve the measurement precision.

4 CONCLUSIONS

The paper presented simple setups that determined the measurement precision of a 3D imaging system, featuring SfM&MVS and 2D PIV analysis. The IPhone and DSLR cameras were chosen in this study because similar cameras are relatively widely available that allows researchers to quickly carry out simple experiments to ensure the technique is suitable to the intended experiments before purchasing expensive industrial cameras and lenses.

The results show that larger number of images yields higher measurement precision in vertical direction in SfM-MVS analysis. The camera resolution is an important factor that governs the number of data points in the obtained 3D point cloud and hence to the measurement resolution. The number of data points appears to be linear with the number of pixel in the camera sensor. The sensor resolution has a more significant impact to the precision of the proposed measurement system in both horizontal direction than that for vertical directions. Other factors that need to be considered to improve the precision of the measurements are the quality of the camera lens and type of sensor and the depth of field.

REFERENCES

Furukawa, Y. and Hernández, C. 2015. Multi-view stereo: A tutorial. Foundations and Trends® in Computer Graphics and Vision, 9(1-2), pp.1-148.

Furukawa, Y. & Ponce, J. 2010. Accurate, dense, and robust multiview stereopsis. IEEE Trans. Pattern Anal. Mach. Intell. 32, No. 8, 1362–1376.

Galland, O. Bertelsen, H. S. Guldstrand, F. Girod, L. Johannessen, R. F. Bjugger, F. Burchardt, S. & Mair, K. 2016. Application of open-source photogrammetric software MicMac for monitoring surface deformation in laboratory models. J. Geophys. Res. Solid Earth 121, 2852–2872.

James, M. R. & Robson, S. 2012. Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. Journal of geophysical research 117, F03017.

Le, B. Nadimi, S. Goodey, R. Taylor, R.N. 2016. System to measure three-dimensional movements in physical models. Géotechnique Letters 6, 256-262

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

Robertson, D.P. & Cipolla, R. 2009. Structure from motion. In: Varga, M. (Ed.), Practical Image Processing and Computer Vision. John Wiley and Sons Ltd. New York.

Smith, M.W. Carrivick, J. & Quincey, D. 2016. Structure from motion photogrammetry in physical geography. Progress in Physical Geography 40: 247-275.

Snavely, N, Seitz. S, Szeliski, R. Photo Tourism: Exploring image collections in 3D. ACM Transactions on Graphics (Proceedings of SIGGRAPH 2006), 2006.

Stanier, S. Dijkstra, J. Leśniewska, D. Hambleton, J. White, D. and Wood, D.M. 2016. Vermiculate artefacts in image analysis of granular materials. Computers and Geotechnics, 72, pp.100-113.

Ullman, S. 1979. The interpretation of structure from motion. Proc. R. Soc. Lond. B, Biol. Sci. 203, No. 1153, 405–426

Zhang, Z. “A Flexible New Technique for Camera Calibration.” IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 22, No. 11, 2000, pp. 1330–1334.