

## A COMPARISON OF LOW-COST CAMERAS APPLIED TO FIXED MULTI-IMAGE MONITORING SYSTEMS

N. Bruno<sup>1\*</sup>, K. Thoeni<sup>2</sup>, F. Diotri<sup>1</sup>, M. Santise<sup>2</sup>, R. Roncella<sup>1</sup>, A. Giacomini<sup>2</sup>

<sup>1</sup> Dept. of Engineering and Architecture, University of Parma, Parco Area delle Scienze, 181/A, 43124 Parma – (nazarena.bruno, riccardo.roncella)@unipr.it; fabrizio.diotri@gmail.com

<sup>2</sup> Centre for Geotechnical Science and Engineering, The University of Newcastle, 2308 Callaghan, Australia – (klaus.thoeni, marina.santise, anna.giacomini)@newcastle.edu.au

### Commission II, WG II/10

**KEY WORDS:** Low-Cost, Digital Surface Model, Accuracy Assessment, Monitoring, 3D Reconstruction

### ABSTRACT:

Photogrammetry is becoming a widely used technique for slope monitoring and rock fall data collection. Its scalability, simplicity of components and low costs for hardware and operations makes its use constantly increasing for both civil and mining applications. Recent on site permanent installation of cameras resulted particularly viable for the monitoring of extended surfaces at very reasonable costs. The current work investigates the performances of a customised Raspberry Pi camera module V2 system and three additional low-cost camera systems including an ELP-USB8MP02G camera module, a compact digital camera (Nikon S3100) and a DSLR (Nikon D3). All system, except the Nikon D3, are available at comparable price. The comparison was conducted by collecting images of rock surfaces, one located in Australia and three located in Italy, from distances between 55 and 110 m. Results are presented in terms of image quality and three dimensional reconstruction error. Thereby, the multi-view reconstructions are compared to a reference model acquired with a terrestrial laser scanner.

### 1. INTRODUCTION

Over the last decade, the use of photogrammetry for repeated 3D mapping and monitoring has received an increased attention by specialists, researchers and practitioners. While some limitations were initially identified in the use of photogrammetry, if compared to other techniques (i.e. its dependency to daylight and good weather conditions, harder requirements to ensure an invariant datum definition etc.), its high scalability, simplicity of components and lower costs for hardware and operations made its use constantly increasing. Unmanned Aerial Vehicle (UAV) or terrestrial photogrammetry have been extensively used to monitor landslides, to assess the evolution of landforms, to support landscape monitoring and geohazard assessment practices (Scaioni et al., 2014; Tannant, 2015; Thoeni et al., 2018). In addition, on-site permanent installations of cameras for monitoring purposes has received an increased attention in the scientific literature (Eltner et al., 2017; Giacomini et al., 2019; James & Robson, 2014; Kromer et al., 2019; Mallalieu et al., 2017; Motta et al., 2013; Parente et al., 2019; Roncella et al., 2014; Travelletti et al., 2010). The use of low-cost camera systems results particularly viable for limited budget monitoring activities or when a significant number of cameras are required to capture the morphology of extended objects (or surfaces) while maintaining reasonable costs.

It is generally well known that low-cost systems with small sensors and unstable optics produce noisy images and less accurate results if compared to digital single-lens reflex (DSLR) cameras. Nevertheless, low-cost solutions are continuously evolving and improving their features and their capabilities (Thoeni et al., 2014). Indeed, Santise et al. (2017) and Blanch et al. (2020) showed that cheap credit card sized single board microcomputers such as the Raspberry Pi enable the development of automated monitoring systems with promising applications. Such systems have very low power consumption and provide fast processing ability at low costs. In particular, the study by Santise

et al. (2017) looked into the investigation of novel low-cost monitoring solutions for extended rock surfaces located in surface mining environments. The current work further extends this work to assess quality of acquisition and resolution of different low-cost camera solutions for the monitoring of large vertical rock faces located at significant distance (between 80 and 110 m) from the camera system. Indeed, such requirement is an important feature in mining applications where a “no go zone” of several meters from the object (rock wall) is generally imposed by safety regulations, and the potential installation of a fixed photogrammetric system needs to account for machineries and mining activities in continuous evolution.

The paper describes the application of four camera systems for the monitoring of extended vertical rock surfaces. The four systems were specifically chosen for their similarity in costs and accessibility on the market. Firstly, a modification of the Raspberry Pi2 camera system presented by Santise et al. (2017) was applied to monitor a vertical rock face in Australia (Hunter Valley, NSW) from a distance of around 110 m. Then, the use of three alternative off-the-shelf cameras has been considered to investigate the best system performances. Three test areas located in Italy (Valle d’Aosta region) consisting of sub-vertical rock slopes at a distance of 55 to 90 m from the cameras were considered. Details of the four camera systems are provided followed by a discussion on the accuracy and reliability of the generated Digital Surface Models (DSM). The overall aim was to assess the performance of different camera systems compared to their costs.

### 2. METHODOLOGY

#### 2.1. Image sensors

Four different image sensors were considered (Table 1). The first sensor corresponds to an adaptation of the one used by Santise et al. (2017) for bigger object distances. It is a Raspberry Pi camera module V2 (Rpi) with a resolution of 8 MPixel. The sensor type

\* Corresponding author

Camera	Resolution	Sensor Type	Sensor size [mm x mm]	Pixel size [µm]	Focal Length [mm]	Cost [€]
Raspberry Pi camera module V2 with custom lens (RPi)	8 MPixel (3280x2646)	CMOS Sony IMX219	3.68 x 2.76	1.12	15.3 (150*)	45-50
ELP-USB8MP02G (USB)	8 MPixel (3264 x 2448)	CMOS SONY IMX179	4.54 × 3.42	1.4	10.1 (80*)	50-55
Nikon Coolpix S3100 (S3100)	14 MPixel (4320 x 3240)	CCD	6.17 x 4.56	1.42	13.5 (79*)	40-80 (used)
Nikon D3 (D3)	12.1 MPixel (4256 x 2832)	CMOS	36 x 23.9	8.46	85	400-500 (used)

\*35mm equivalent focal length

Table 1. Technical specifications of the four tested camera systems.

is a CMOS Sony IMX219 (pixel size of ca. 1.12 µm). The original lenses used by Santise et al. (2017) (29 mm focal length equivalent to 35 mm format) shipped with the camera were removed and substituted with a longer focal length optic of 15.3 mm (150 mm focal length equivalent to 35 mm format). In addition, a custom 3D printed support was designed and built to centre the lens with respect to the sensor and to fix it in position. The total cost, including camera module, 150 mm optic and 3D printed support stage, is around 45 €. Initial tests of the system suggested that even the slightest decentring and off-axis positioning of the lenses (e.g. due to an inaccurate 3D printing process or assembly of the optical system) could potentially deteriorate the quality of the images (i.e. defocusing, vignetting, chromatic aberration etc.) due to the big focal length. Hence, three additional image sensors have been tested in parallel in Italy (note that the RPi has been designed and tested in Australia). These additional three camera sensors consist of off-the-shelf solutions. The sensors considered are a low-cost ELP-USB8MP02G (USB) camera module with variable focal lenses (2.8-12 mm corresponding to 21-95 mm 35 mm equivalent focal length), which was considered having similar features as the RPi, a Nikon S3100 (S3100) compact digital camera and a higher cost DSLR camera, a Nikon D3 (D3). Latter was used as reference for a higher cost system.

The USB, S3100 and D3 have significant different sensors and geometric features. The D3 has a full frame (FX) CMOS sensor with large pixel size (ca. 8.5 µm), a resolution of 4256x2832 pixel (12.1 MPixel) and mounted fixed focal length optics (85 mm). When it was put on the market in 2007, it was considered the Nikon's flagship model, and still today's its picture quality outperforms many low to mid-level DSLR cameras. Every single camera acquisition parameter can be easily controlled remotely (e.g. using gPhoto, Nikon capture or Dragonframe). The use of fixed focal length optics assures greater stability of interior orientation (IO) parameters if periodic acquisitions have to be performed. At the time of writing the Nikon D3 can be found on the used market for around 400-500 € (in Italy).

The S3100, although discontinued and with a higher resolution than the other systems, has been chosen for its similar price range. Even if probably not the best compact camera around in terms of image quality, it is one of the most popular Nikon point-and-shoots cameras and can be easily found and acquired for a price between 40 and 80 € (being discontinued it can only be found on the used market). It has a 14 MPixel 1/2.3" sized CCD sensor. The pixel size is ca. 1.42 µm. It features a 4.45-22.3 mm (26-130 mm focal length equivalent to 35 mm format) F3.2-6.5 zoom lens. The S3100 can also be controlled remotely using the same libraries as for the D3. However, as with many compact zoom lens cameras, zoom control is not supported and the zoom is reset (e.g. to 4.45 mm) every time the camera is powered on. In some camera models the zoom can be changed by sending specific software commands which incrementally modifies the zoom. However, it is nearly impossible to set a specific fixed

focal length. This represents the most critical drawback of the implementation of digital compact cameras in fixed photogrammetric monitoring systems. Not being able to precisely reset the zoom level of the camera after restart implies a new estimation of all the IO parameters, which, in many cases, can only be performed with a self-calibration procedure based on a usually not enough redundant, rigid and complete image block configuration.

The USB is a Plug-&Play (UVC compliant) USB2.0 camera widely used in many different machine vision applications and provides good image quality and sensitivity performance if equipped with good lenses. The sensor, a CMOS Sony IMX179, has a resolution of 8 MPixel and a pixel size of about 1.4 µm. The module can be equipped with a wide range of different optics (also fixed focal length lenses): in our experiments we opted for low-cost zoom lenses and fixed, in all the experiment, the focal length to 10.1 mm (80 mm focal length equivalent to 35 mm format). The overall quality of the lens is not very high but the overall cost of the system (USB including lens ca. 55 €) was comparable with the other two low-cost sensors. The USB camera system including lens is extremely compact (38x38x80mm) which makes it the ideal module to be implemented in low-cost mini case fixed photogrammetric systems.

## 2.2. Test sites and testing program

A total of four test sites were considered in the comparisons. Test Site A is a vertical rock face in a coalmine in Australia (Hunter Valley, NSW). The wall was formed by presplit blasting. It is highly fractured with lots of loose material. It comprises horizontal layers of predominantly fine to medium sandstones and siltstone, with minor claystone, and a layer of coal, as shown in Figure 1. The considered area is about 25 x 17 m<sup>2</sup>.



Figure 1. Test Site A.

Test Sites B, C and D are located in Italy, in the municipalities of Pollein (Site B and C) and Arnad (Site D) in Aosta Valley. They consist of sub-vertical rock-surfaces with similar geological characteristics. Sites B (Figure 2) and C (Figure 3) are two portions of the same rock wall. The rock wall is hanging over the Comboé river, a typical alpine stream that has excavated a deep rock canyon. Both rock surfaces, of about  $34 \times 22 \text{ m}^2$  and  $26 \times 18 \text{ m}^2$  respectively, are particularly suitable for the reconstruction of a good quality 3D model, as they are unvegetated surfaces with minor occlusions. Test Site D instead, has been excavated by a glacier and extends for about  $33 \times 32 \text{ m}^2$  and presents a more regular but also more vegetated surface, with evidence of outcropping trees and bushes (Figure 4).



Figure 2. Test Site B.



Figure 3. Test Site C.

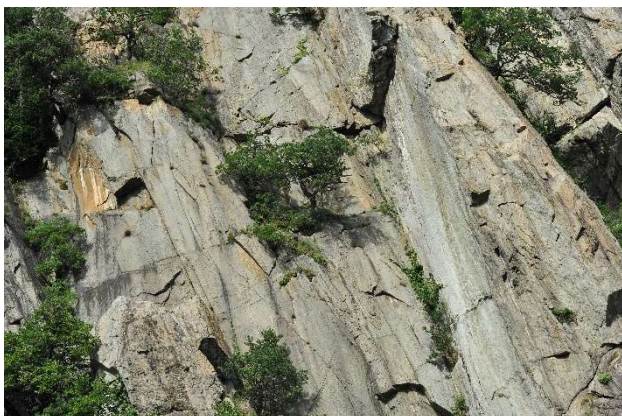


Figure 4. Test Site D.

The RPi sensors were installed at Site A as a fixed monitoring system. At the same time, the other three camera sensors were tested in Italy at Site B, C and D with the aim to find a valuable alternative low-cost solution to install at Site A. Unfortunately, due to mining production activities, Site A became soon inaccessible for further trials. Hence, data for Site A was acquired with the RPi only while data for Site B, C and D was acquired with the USB, S3100 and D3.

For each site, image blocks with the same geometry configuration were considered. In particular, five images were taken with a regular base-length along a straight line parallel to the rock wall. The images had a small convergent pose to grant a higher overlap and strengthen the camera network, resulting in the reduction of systematic DEM errors. The camera positions and the total base length (distance between the first and last camera station) to distance ratio (on average about 1:3) were similar for all sensors. All the acquisitions were performed in similar light conditions. In addition, a Total Station (TS) survey was carried out to determine Ground Control Points (GCP). Target installation was prevented by accessibility limitation, therefore, well distributed natural features over the rock face were surveyed and used as GCP. In addition to the TS survey, the rock faces were also scanned using a terrestrial laser scanner (TLS) set up at a mid-point position of each photogrammetric camera network. The point clouds obtained with the TLS were processed and used as reference models for the evaluation of the accuracy of the photogrammetric reconstruction.

During the test conducted at Site A with the RPi, images were taken over a period of five consecutive days. The first four acquisitions (A1-A4) were performed at 12.30pm whereas the last acquisition was performed on 6.10pm. The idea behind this was to test the system under different lighting conditions. The five camera systems were set up at a fixed position of about 102 m in front of the rock face with a base length between each other of about 7 m. Each DSM generated from the acquired images was compared to the TLS reference model in order to assess the accuracy of the system. In addition, a repeatability study was conducted where the first photogrammetric model (A1) was compared to the following four photogrammetric models (A2-A5) in order to assess the reliability of the system. It should be noted that the positions of the five RPi sensors were also measured using the TS since the system was set up as a preliminary fixed monitoring system.

The tests conducted at Site B, C and D primarily focused on the accuracy assessment of the three different systems. Hence, differently from Site A, all image acquisitions were made on the same day using one sensor (instead of five fixed sensors). The D3 and S3100 were hand-held, while the USB was installed on a tripod, therefore, even for the same test site, camera poses are slightly different. Between sites, camera poses vary more, especially when addressing object distance, due to morphological characteristics of each site. The average distance from the object was selected as far as possible considering the accessibility constraints and safety implications, so that the Ground Sampling Distance (GSD, i.e. the corresponding pixel size on the object) was similar to the one considered at Site A. In particular, object distances were of 55 m, 69 m and 88 m for Site B, C and D respectively, while the base lengths varied in a range of 5.5 to 6 m.

The D3 was equipped with a fixed focal length optic (85 mm) resulting in a GSD in the range of 5.6 to 8.7 mm/pixel. The zoom lenses of the S3100 and USB were set to approximately cover the same area as the D3 image block. Consequently, the focal lengths were 10.1 mm (80 mm equivalent) for the USB camera and 13.5 mm (79 mm equivalent) for the S3100. Since the USB camera has a significant lower resolution compared to the S3100 and D3, framing the same object area implies that the GSD of the USB

tests are higher compared to the one of the other camera models. For this reason, subsampling of the higher resolution images was also considered during image processing, i.e. the images of the S3100 and D3 were processed twice (full resolution and subsampled lower resolution). Although using subsampled images to simulate a lower resolution sensor is questionable, it might still give some additional insights. When using the subsampled images, the GSD is consequently higher and comparable to the one of the USB camera. Table 2 summarises the GSD for the different sites and camera models. The subsampled GSD is reported in parenthesis.

	RPi	USB	S3100 (Subsamp.)	D3 (Subsamp.)
Site A	7.1	-	-	-
Site B	-	10.5	7.2 (9.6)	6.9 (9.0)
Site C	-	8.3	5.7 (7.6)	5.6 (7.3)
Site D	-	13.4	9.6 (12.8)	8.7 (11.4)

Table 2. Average GSD of the acquired images (in mm).

Assuming the image block orientation is performed correctly and without systematic effects, the theoretical precision of each single multi-image convergent block can be computed using the following equation (Fraser, 1984):

$$\sigma_p = \frac{m_b \cdot q \cdot \sigma_i}{\sqrt{k}} \quad (1)$$

where  $m_b$  represents the image scale ( $Z/c$  where  $Z$  is the distance from the object and  $c$  is the principal distance of the optical system),  $q$  is a quality factor describing the camera configuration,  $\sigma_i$  represents the expected precision of the image coordinates, which can be assumed to be  $\pm 1$  pixel (see e.g. Luhman et al., 2013) and  $k$  is the average number of photos taken from each station (in our case  $k = 1$ ). For the quality factor  $q$ , some authors suggest to use a value in the range of 0.5-1.2, with the lowest values associated with very redundant, rigid and convergent image block geometry. Others, with weaker configurations, suggest to use much higher  $q$  factors around 3.0-3.5 (see e.g. Passoni et al., 2018).

Table 3 shows the computed theoretical precision according to Eq. 1 for every single combination of test site and camera model. Since the block geometry is not very redundant and the cameras are only slightly convergent, a  $q$  factor equal to 3.5 was considered.

	RPi	USB	S3100 (Subsamp.)	D3 (Subsamp.)
Site A	24.9	-	-	-
Site B	-	36.8	25.2 (33.6)	24.2 (31.5)
Site C	-	29.1	20.0 (26.6)	19.6 (25.6)
Site D	-	46.9	33.6 (44.8)	30.5 (39.9)

Table 3. Estimated precision of image blocks (in mm) based on Eq. 1.

In fact, Eq. 1 usually provides affordable estimation of the level of precision in strongly convergent and controlled image blocks where the base length to object distance ratio of consecutive images is quite high. Conservatively we also opted for the calculation of the theoretical precision of the image block using the equation for stereo restitution in the normal case (Kraus, 2011) considering the first and last image of each single block only:

$$\sigma_z = \frac{m_b \cdot Z \cdot \sigma_{p_x}}{B} \quad (2)$$

where  $Z$  is the distance from the object,  $B$  is the base length and  $\sigma_{p_x}$  represents the expected precision of the parallax between two homologous points (the precision of single image point measures was considered to be  $\pm 1$  pixel as before and parallax precision was estimated using the error propagation law). As can be seen in Table 4, in this case the estimated precision is significantly lower compared to the one estimated using Eq. 1.

	RPi	USB	S3100 (Subsamp.)	D3 (Subsamp.)
Site A	50.7	-	-	-
Site B	-	43.7	30.0 (40.0)	28.7 (37.5)
Site C	-	27.5	18.9 (25.2)	18.6 (24.2)
Site D	-	68.8	49.2 (65.8)	44.7 (58.5)

Table 4. Estimated precision of image blocks (in mm) based on Eq. 2.

### 2.3. Image processing and DSM reconstruction

Image processing followed a pipeline quite similar for all the test sites using the commercial software package Agisoft Metashape Professional. Image subsampling for the data acquired with the S3100 and D3 was performed using a bilinear interpolation. Since the three camera sensors used at Site B, C and D have slightly different aspect ratios, the subsampled resolution was set based on the number of pixels along the horizontal direction of the sensor.

Images were initially oriented using all the GCP surveyed by the TS. The use of a relatively large number of GCP makes the solution of the orientation more stable and reliable, especially when using natural features, whose collimation both with the TS and on images is usually difficult and less precise. To assess the accuracy of the orientation, the residuals on check points (also natural features) have been analysed. Only four of the available control points (the ones located at the corner of the rock wall) were constrained as GCP, while the remaining points were used as check points (CP).

Since the camera poses of the RPi sensors were fixed and known from the topographic survey, image orientation of the acquisitions of Site A was performed by constraining the Bundle Block Adjustment (BBA) with the positions (X, Y, Z) of the camera centres and 11 GCP. Camera rotations were instead estimated during the orientation process. For the estimation of the IO parameters, pre-calibrated values were used for lens distortion while a self-calibration procedure was performed for the focal estimation because the distances between the rock face and the five installed cameras was greater than the one available during the laboratory calibration stage, leading to different focal estimates during calibration and operation. The collinearity equation residuals are consistent over the five acquisition periods, with an average root mean square reprojection error of 1.5 pixel. The root mean square error (RMSE) differences on the CP instead increase progressively over time. They are similar during the first two acquisitions A1 and A2 (27 and 34 mm) but much higher during the consequent acquisitions A3, A4 and A5 (75, 66 and 72 mm). The RMSE on the CP is on average 55 mm which is very close to the expected precision reported in Table 4. For Sites B, C and D, the external orientation (EO) parameters were computed during the BBA using the GCP as constraints. The IO parameters were estimated applying an on-the-job self-calibration procedure. Even though this procedure is mainly recommended for unstable optics (such as the S3100 and USB

cameras used in these tests), to ensure uniformity of the results, the self-calibration was performed also for the D3. The number of GCP for these three sites is larger compared to Site A with an average of 22, 23 and 26 GCP for Site A, B and C respectively. The RMSE are quite low, especially for the D3 and S3100, with an average of 0.18 and 0.36 pixel respectively over the three sites. On the contrary, the USB sensor gives a much higher average RMSE of 1.1 pixel, with a maximum of 1.74 pixel for Site D. This is similar to the RMSE of the RPi of Site A. The accuracies of the orientations vary according to sensor size but they show a similar behaviour for the three sites. Table 5 summarises the average residuals per sensors and per sites. In particular, comparing the data in Table 5 with the expected precisions highlighted in Table 4, it should be noted that the results of the USB sensor are always worse than expected, while the other two camera models provide (sometimes significantly) better results than expected. The residuals obtained for the D3 and S3100, using full resolution images, are totally comparable, with slightly better results for the S3100 probably related to the higher resolution of the sensor. On the contrary, using subsampled images at 8 MPixel, the D3 gives the best results (23 % better than S3100 and 55 % better than USB), highlighting the prevalence of sensor size and quality over resolution.

Camera	Site B (Subsamp.)	Site C (Subsamp.)	Site D (Subsamp.)
USB	85.9	34.0	69.8
S3100	20.8 (37.6)	16.4 (22.5)	14.1 (49.0)
D3	23.5 (33.3)	13.2 (15.9)	18.7 (34.4)

Table 5. Average residuals on the check points (in mm) for Sites A, B and C.

Finally, for each site and for each camera model, dense matching and DSM generation were performed using the best quality settings provided by the software. Interpolation was disabled during DSM generation to avoid incorrect surface reconstruction around holes. The number of 3D reconstructed points for the different sites and camera models varies according to the GSD, ranging from about 15000 points/m<sup>2</sup> (Site C, Nikon D3 and S3100 full resolution) to about 3000 points/ m<sup>2</sup> (Site D, USB). For Sites B, C and D, due to the different sensor aspect ratio and the difficulties in framing exactly the same portion of the rock wall, the DSM obtained had slightly different extensions. To assure comparable results between the different sensors, for each site the DSM were cut so that the same identical portion of the wall was considered in the final comparisons. The DSM were then compared with the reference datasets for the accuracy assessment. The reference TLS point clouds were initially co-registered with the photogrammetric DSM using the GCP, then an Iterative Closest Point (ICP) algorithm was applied to optimise registration and eliminate systematic errors. The comparisons evaluated the signed distances between two models and were performed in CloudCompare considering a maximum distance of 0.2 m, meaning that distances greater than 20 cm are not considered in the comparisons (gross-errors).

### 3. RESULTS

#### 3.1. Test Site A

The comparisons with the TLS point cloud showed that the DSM generated from the images of the RPi are generally complete but noisy. The level of noise of the models can be attributed to different issues. The quality of the lens and the excessively long focal length make it difficult to accurately determine the IO parameters during the calibration procedure. Moreover, the low-quality and sensible optic make it very difficult to focus the

sensor. Consequently, some of the images are partly blurred. The light conditions are also really important with this type of sensor, as the presence of numerous small recesses on the monitored wall can generate shadows which are more or less evident depending on the time of image acquisition.

Figure 5 shows the colour map of distances between the first photogrammetric model A1 and the reference point cloud of the TLS. Two things can be noted. Firstly, the photogrammetric model is less accurate towards the edges since the images tended to be blurred at the edges due to the issue with the optic. In addition, the image overlap is lower towards the edges. Secondly, there is an almost horizontal white strip from left to right. This strip represents the black coal layer in the rock face (cf. Figure 1), which was not entirely picked up by the laser scanner. When looking at the repeatability study (Figure 6) it can be seen that these issues disappears. The colour maps of the other comparisons are very similar and, hence, not shown for brevity. Instead, a summary of all comparisons in terms of RMSE is provided. Figure 7a summarises the RMSE of the distances between the reference TLS point cloud and the RPi models whereas Figure 7b summarises the RMSE of the distances between the first RPi model A1 and the consecutive four RPi models (A2 to A5).

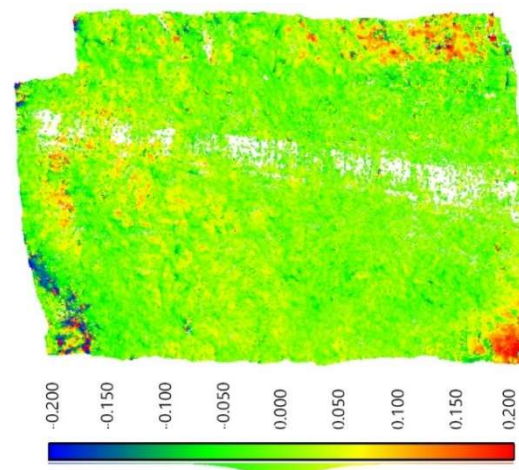


Figure 5. Colour map of distances between the photogrammetric model A1 and the TLS data at the maximum distance of 0.2 m.

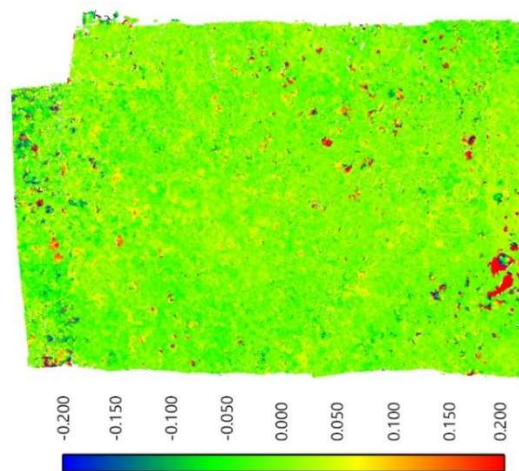


Figure 6. Colour map of distances between the photogrammetric models A1 and A2 at the maximum distance of 0.2 m.

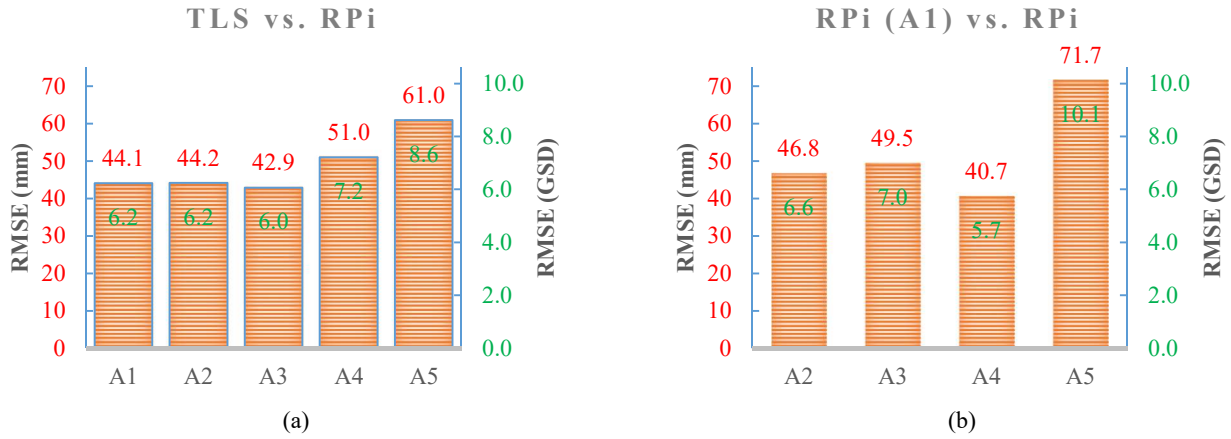


Figure 7. Summary of the RMSE of the distances between (a) the reference TLS point cloud and the RPi models and (b) the first RPi model (A1) and the consecutive four RPi models (A2 to A5).

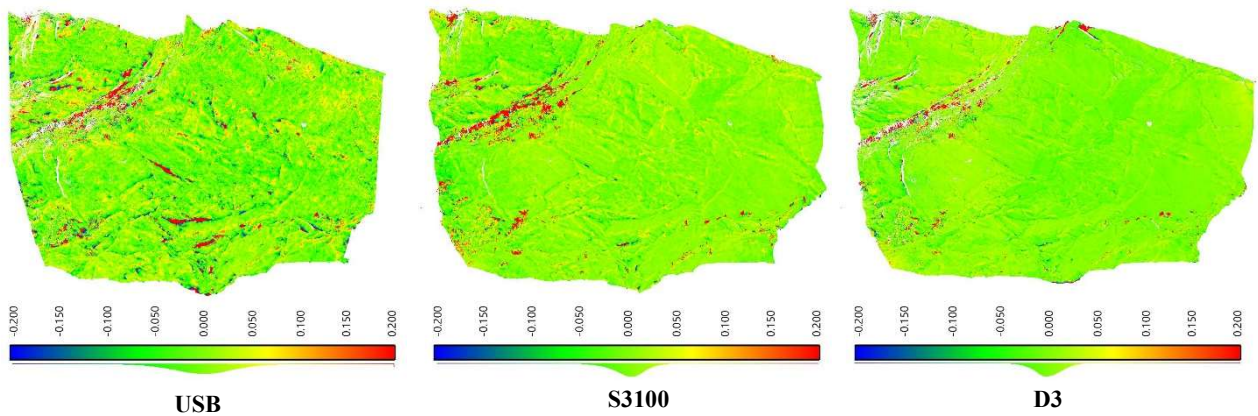


Figure 8. Example colour maps of distances between the photogrammetric models of the three sensors and the TLS point cloud for Site B at the maximum distance of 0.2 m.

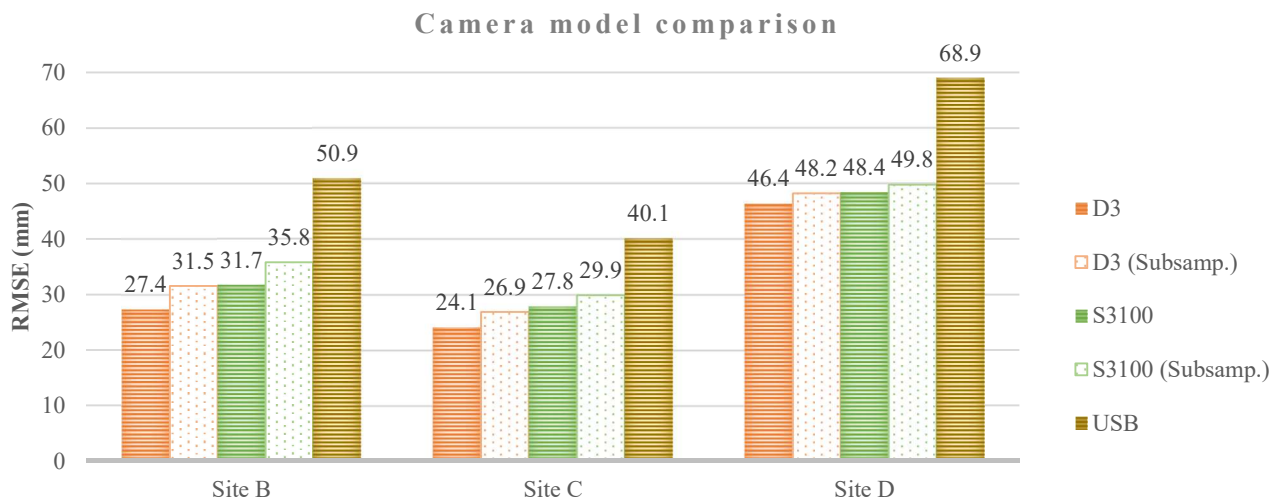


Figure 9. RMSE of the distances (TLS vs. photogrammetry) obtained for Sites B, C and D.

### 3.2. Test Sites B, C and D

The DSM obtained for Sites B, C and D are less complete compared to the one obtained for Site A. This is due to the presence of vegetation. The DSM of Site D, in particular, has large holes and reconstruction uncertainties due to trees. As far as the noise is concerned, the models produced by the USB are as noisy as the ones of the RPi, while the S3100 and D3 seem producing DSM free of apparent noise, even when using subsampled images. As an example, the colour maps of distances for Site B are shown in Figure 8. Figure 9 summarises the RMSE for the three sensors and the three sites.

## 4. DISCUSSION

The comparisons performed on Site A using the RPi show that the overall agreement between image based DSM and reference TLS point cloud is in the range of 44 to 61 mm, which corresponds to 6.2 to 8.6 times the GSD (Figure 7a). The results agree with the reported average residuals on the CP (the total RMSE considering all the acquisition is around 55 mm). However, it should be noted that while in the first two acquisitions (A1 and A2) the CP residuals are on average 31 mm, the subsequent acquisitions (A3, A4 and A5) have much higher residuals. This is most likely due to an increasing difficulty by the operator to identify the correct location of the object points on the images due to changes in the lighting condition on the wall (e.g. A3 and A4, differently from A1 and A2, were acquired during cloudy days, and A5 was acquired at a different time of the day). According to Eq. 2 (cf. Table 4) and an estimated TLS accuracy of 5 to 10 mm, the expected RMSE of the distances should be around 52 mm. The results are therefore consistent with the expected accuracies. It should be noted, however, that the DSM of the last acquisition A5 (6.10pm) provides significantly worse results with a total RMSE of about 61 mm. In this case, the quality of the images is inferior, due to worse lighting conditions and strong presence of shadows on the whole rock wall. This produced a much noisier and, therefore, less accurate reconstruction. In the repeatability analysis, the comparison data comes from the same source (image-based reconstruction) and the same precision for both the reference and the compared DSM should be considered. The variance propagation law would imply that, according to the precision reported in Table 4, an RMSE of about 72 mm should be expected. However, in this case, most of the possible systematic effects that might affect the photogrammetric reconstruction (e.g.

image block and DSM deformation due to inaccurate image orientation and/or image calibration) are the same for all the dataset and nullify reciprocally in the comparison. This is particularly interesting as far as the system aims to monitor the evolution of an object over time (e.g. rock fall or rock/landslide detection) where the final user is probably much more interested in identifying and quantifying specific events rather than having a very accurate global object reconstruction. The results reported in Figure 7b clearly show that when comparing photogrammetry with photogrammetry the RMSE is significantly lower than the expected 72 mm, except for the comparison with the DSM of A5 where images were taken under different lighting conditions. The second set of comparisons was performed on data acquired at Sites B, C and D using the D3, S3100 and USB camera models. The results show that overall the accuracy of the USB camera is significantly lower compared to the other two cameras. From this point of view, the RPi, even if equipped with custom lenses, seems to provide better results. It should be noted that both camera models, the RPi and USB, mount very cheap optics, whose quality is far lower than the one used, for instance, by the S3100. Also from a qualitative visual inspection it is quite evident that the images acquired with these two cameras are worse than the other ones and the operator experienced much more difficulties in accurately locating homologous points during manual identification of GCP and CP. This issue is confirmed when looking at the reprojection residuals of the BBA and the RMSE of the CP (Section 2.3), in particular for the USB. The S3100, on the contrary, provides good accuracies even if compared with the more expensive D3. The reported RMSE are just 15% higher for Sites B and C and 4% higher for Site D. The good performance of the S3100 can be partly justified by its bigger resolution (14 MPixel compared to 12 and 8 MPixel for the D3 and USB respectively). Hence it was decided to run the same comparison using the DSM generated from subsampled images. When looking at the RMSE of the subsampled data it can be seen that the values are only slightly higher compared to the one obtained using the original images. The difference is generally between 3 and 10%. This means that the results are still considerably better compared to the one of the USB even when using the subsampled images. Figure 10 confirms this when comparing the RMSE to the expected precision according to Eq. 2. It is also interesting to note that the subsampled data for Sites B and D perform better than expected. For Site C, all sensors perform a little bit worse than expected. Nevertheless, overall the results are in line with the expected precision.

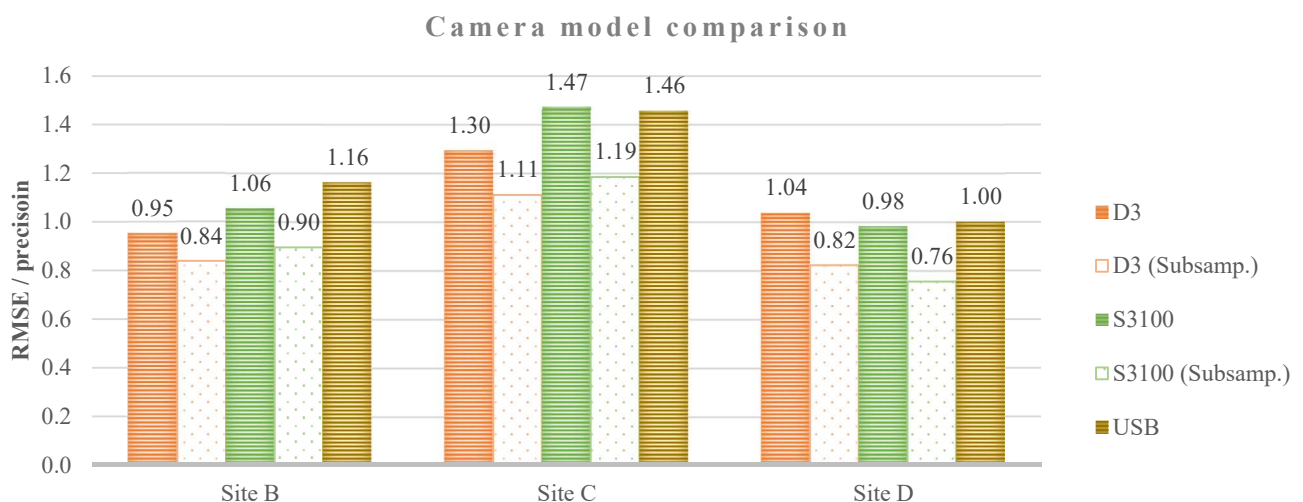


Figure 10. The ratio between RMSE of the distances and expected precision according to Table 4 for Sites B, C and D.

## 5. CONCLUSIONS

This work investigated the performances of four different low-cost camera sensors for the monitoring of sub-vertical rock faces from distances of 50 to 100 m. The reconstructed DSM were compared to reference TLS point clouds and overall the error was within the expected precision. However, it should be noted that the precision was estimated using the equation for the stereo restitution in normal case considering the first and last image in each block.

From the practical point of view, it was observed that the customised RPi and the USB camera, although very practical and compact, produce lower quality images and consequently less accurate 3D reconstructions. The lower quality of the images, and hence the less accurate reconstruction, can mainly be attributed to the very unstable optics. The S3100 compact camera provided much better results for the same price. However, using such a camera in a monitoring system requires setting up additional control units (Raspberry PI + battery) and setting acquisition parameters is not straightforward (e.g. external control of zoom lenses is often not possible). As expected, the D3 DSLR had the best performance but the S3100 came very close. This might be due to the fact that its resolution is slightly higher. Overall, it can be concluded that low-cost sensors can be applied to monitoring purposes but special attention should be given to the optics.

## ACKNOWLEDGEMENTS

The financial support of the Australian Research Council (LP160100370) provided to the Newcastle authors is greatly acknowledged.

## REFERENCES

- Agisoft Methashape Professional <http://www.agisoft.com> (Accessed on 24 April 2020).
- Blanch, X., Abellan, A., Guinau, M., 2020. Point Cloud Stacking: A Workflow to Enhance 3D Monitoring Capabilities Using Time-Lapse Cameras, *Remote Sensing* 12(8), 1240.
- CloudCompare (version 2.10.2) (GPL software) 2020. <http://www.cloudcompare.org/> (Accessed on 24 April 2020).
- Eltner, A., Kaiser, A., Abellan, A., Schindewolf, M., 2017. Time lapse structure-from-motion photogrammetry for continuous geomorphic monitoring. *Earth Surface Processes and Landforms*, 42(14): 2240–2253.
- Fraser, C. S., 1984. Network design considerations for non-topographic photogrammetry. *Photogrammetric Engineering and Remote Sensing*, 50(8), 1115–1126.
- Giacomini, A., Thoeni, K., Santise, M., Fityus, S., Roncella, R., 2019. A low-cost terrestrial photogrammetric system for rockfall monitoring and hazard assessment in open-pit mines, *Rock Mechanics for Natural Resources and Infrastructure Development: Proceedings of the 14th International Congress on Rock Mechanics and Rock Engineering*, 1524–1531.
- James, M. R., Robson, S., 2014. Sequential digital elevation models of active lava flows from ground-based stereo time-lapse imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 97, 160–170.
- Kraus, K., 2011. *Photogrammetry: geometry from images and laser scans*. Walter de Gruyter.
- Kromer, R., Walton, G., Gray, B., Lato, M., 2019. Development and optimization of an automated fixed-location time lapse photogrammetric rock slope monitoring system. *Remote Sensing*, 11(16), 1890.
- Luhmann T, Robson S, Kyle S, Boehm J., 2013. *Close-range photogrammetry and 3D imaging*. Walter de Gruyter.
- Mallalieu, J., Carrivick, J. L., Quincey, D. J., Smith, M. W. and James, W. H. M., 2017. An integrated structure-from-motion and time-lapse technique for quantifying ice-margin dynamics. *Journal of Glaciology*, 63(242): 937–949.
- Motta, M., Gabrieli, F., Corsini, A., Manzi, V., Ronchetti, F., Cola, S., 2013. Landslide Displacement Monitoring from Multi-Temporal Terrestrial Digital Images: Case of the Valoria Landslide Site. In C. Margottini, P. Canuti, K. Sassa (Eds), *Landslide Science and Practice*, Vol. 2, pp. 73–78.
- Parente, L., Chandler, J. H., Dixon, N., 2019. Optimising the quality of an SfM-MVS slope monitoring system using fixed cameras. *The Photogrammetric Record*, 34(168), 408–427.
- Passoni, D., Federici, B., Ferrando, I., Gagliolo, S., & Sguerso, D. (2018). The estimation of precisions in the planning of UAS photogrammetric surveys. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 42(2).
- Roncella, R., Forlani, G., Fornari, M., & Diotri, F., 2014. Landslide monitoring by fixed-base terrestrial stereo-photogrammetry. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2(5), 297.
- Santise, M., Thoeni, K., Roncella, R., Sloan, S.W., Giacomini, A., 2017. Preliminary tests of a new low-cost photogrammetric system. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, 229.
- Scaioni, M., Longoni, L., Melillo, V., Papini, M., 2014. Remote sensing for landslide investigations: an overview of recent achievements and perspectives. *Remote Sensing*, 6(10):9600–9652.
- Tannant, D. D., 2015. Review of Photogrammetry-Based Techniques for Characterization and Hazard Assessment of Rock Faces. *International Journal of Georesources and Environment*, 1(2):76–87.
- Thoeni, K., Giacomini, A., Murtagh, R., Kniest, E., 2014. A comparison of multi-view 3D reconstruction of a rock wall using several cameras and a laser scanner. *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-5, 573–580.
- Thoeni, K., Santise, M., Guccione, D.E., Fityus, S., Roncella, R., Giacomini, A., 2018. Use of low-cost terrestrial and aerial imaging sensors for geotechnical applications, *Australian Geomechanics Journal* 53(3), 101–122.
- Travalletti, J., Malet, J., Schmittbuhl, J., Toussaint, R., Bastard, M., Delacourt, C., Allemand, P., van Dam, D., 2010. Multi-temporal terrestrial photogrammetry for landslide monitoring. *Proceedings of Mountain Risks International Conference*, pp. 183–191. Florence (Italy).