Monitoring coastal morphology: the potential of low-cost fixed array action cameras for 3D reconstruction

Samantha Godfrey,1* James Cooper,1 Frederic Bezombes2 and Andrew Plater1
1 School of Environmental Sciences, University of Liverpool, Liverpool, UK
2 General Engineering Research Institute, Liverpool John Moores University, Liverpool, UK

Received 14 July 2019; Revised 23 April 2020; Accepted 26 April 2020

*Correspondence to: Samantha Godfrey, School of Environmental Sciences, Roxby Building, University of Liverpool, Liverpool, L69 7ZT, UK.
E-mail: s.godfrey@liverpool.ac.uk
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

ABSTRACT: The combination of structure-from-motion with multi-view stereo (SfM-MVS) photogrammetry has become an increasingly popular method for the monitoring and three-dimensional (3D) reconstruction of coastal environments. Climate change is driving the potential for increased coastal landward retreat meaning geomorphological monitoring using methods such as SfM-MVS has become essential for detecting and tracking impacts. SfM-MVS has been well-researched with a variety of platforms and spatial and temporal resolutions using mainly rectilinear digital cameras in coastal settings. However, there has been no assessment of the potential of fixed multi-camera arrays to monitor landward retreat or on the significance of camera placement in relation to the scene. This study presents an innovative method of image acquisition using a purpose-built camera grid and GoPro® action camera to evaluate the combined effects of camera height, obliqueness and overlap at a site of known landward retreat. This approach examines the effect of camera placement on scene reconstruction to aid the design of a multi-camera array. SfM-MVS dense point clouds display millimetre accuracy when compared to equivalent terrestrial laser scans and strong image network geometry with internal precision estimates of < 3 mm. Comparable point cloud reconstruction can be achieved with a small number of images stationed in appropriate positions. Initial results show as few as five images positioned at a cliff to camera ratio of 3:4.18 and camera obliqueness of 40° can provide reconstruction in the range of millimetres (mean error of 4.79 mm). These findings illustrate the importance of camera placement when using multiple cameras and aid the design of a low-cost, fixed multi-camera array for use at sites of small-scale landward retreat. © 2020 The Authors. Earth Surface Processes and Landforms published by John Wiley & Sons Ltd

KEYWORDS: action camera; coastal erosion; coastal monitoring; image optimization; structure-from-motion photogrammetry; 3D reconstruction

Introduction

Over the last decade advances in remote sensing and three-dimensional (3D) image reconstruction techniques have made it easier to monitor dynamic and rugged coastal environments (Maiti and Bhattacharya, 2009; Earlie et al., 2013; Mancini et al., 2013; Earlie et al., 2015; Conlin et al., 2018; Westoby et al., 2018) Technological advancements have produced a range of surveying methods to monitor patterns of morphological change in various coastal settings. For example, large-scale spatial coverage can be achieved with airborne light detection and ranging (LiDAR) (Earlie et al., 2013; Dudzińska-Nowak and Wężyk, 2014), video monitoring with Argus systems (Holman and Stanley, 2007) and satellite imagery (Maiti and Bhattacharya, 2009). Smaller scale topographic coverage is achieved with a terrestrial laser scanner (TLS), capable of acquiring fine-resolution spatial data (in millimetres) at short stand-off distances (Anthony et al., 2006; Calligaro et al., 2014; Westoby et al., 2018) or roving Global Navigation Satellite System (GNSS) with high spatial accuracy (in centimetres) (Young, 2015). However, many of these surveying methods are expensive to acquire and operate meaning they are most commonly used in developed nations and, even in these settings, infrequently. The use of manual surveying methods such as Abney levels, the Emery Method or optical levels, provide a cost-effective option but offer very limited spatial resolution. There is, therefore, a case for alternative methods of monitoring coastal change that have high spatial resolution, are cost effective and can be used routinely. It would be essential that any such methods were of comparable accuracy to those that have become established as the ‘industry standard’ (Westoby et al., 2018).

Photogrammetric based methods of 3D reconstruction for topographic surveys have become increasingly popular due to their lower-cost and flexibility. Based on traditional photogrammetric principles, two-dimensional (2D) overlapping images are used to reconstruct 3D scene geometry. Advancements in
computer vision (Bemis et al., 2014) have allowed the underlying mathematical calculations of this to be developed into structure-from-motion with multi-view stereo (SfM-MVS), the fundamentals of which have been described in Westoby et al. (2012) and James and Robson (2012). The incorporation of this process into automated commercial software packages and other open-source alternatives has made it accessible to both professional and ‘non-expert’ users.

The flexibility of SfM-MVS has provided opportunities for a wide range of geographic applications, including coastal monitoring. The majority of research has been undertaken with a single rectilinear (pin-hole) digital camera deployed from a variety of platforms including poles (Conlin et al., 2018), drones (Mancini et al., 2013), blimps (Fonstad et al., 2013) and kites (Duffy et al., 2018). Comparative studies between TLS and SfM-MVS outputs have shown SfM-MVS compares well to TLS with centimetre accuracy (Wilkinson et al., 2016; Westoby et al., 2018). In addition, the impracticalities of TLS equipment such as reduced accuracy outside a specified temperature range and long surveying periods (James and Robson, 2012) during short tidal windows, lead some to favour SfM-MVS.

The quality of SfM-MVS output is often considered to be positively associated with the number of images used: the more images the better to optimize the number of keypoints present (Westoby et al., 2012). This has prompted recent studies (for example, Eltner et al., 2017) to experiment with multiple cameras. Although this work has produced many encouraging findings it has also identified the necessity for further research on the significance of camera position and setup. With this in mind, and the need for quick surveying during short tidal windows, the ability to simultaneously acquire images from a fixed array of multiple cameras would be advantageous – ideally with cameras set in positions to maximize efficacy. The potential approach of using a multiple fixed camera array – contrasts with the previously cited studies where single cameras were used for image acquisition. However, to truly make this approach a possibility it is first necessary to understand the significance of camera placement in relation to the scene. Optimal camera placement would result in a simplification of image capture geometry and would entail fuller scrutiny of the combined effect of some positional parameters (beyond simply number of images) that effect image suitability: overlap, obliqueness and convergence (Eltner et al., 2016).

Another critically important consideration to multiple camera use is choice of camera. Action cameras offer an accessible, easily operable, manoeuvrable and rugged alternative to digital single lens reflex (DSLR) cameras. In addition, GoPro action cameras offer the option for wireless multi-camera synchronization – a significant advantage for a camera array. Previously, action cameras were considered inappropriate for accurate 3D reconstructions due to radial distortion created by the wide angle of view of action cameras (AOV) or fisheye lens. AOV is a function of sensor size and lens type and describes the angular extent of a camera’s view. The short focal length of action cameras allows a greater field of view (FOV) or measurable ‘footprint’ but creates image distortion – particularly on the extremities of the image frame (Thoeni et al., 2014; Phillips and Eliasson, 2018). The distortion previously rendered fisheye lenses unsuitable for creating 3D models with accurate metric integrity (Perfetti et al., 2017). However, the advancement in commercial photogrammetric software has meant fisheye remapping functions are available to correct some of these previously unsuitable radial distortions. Research outside geomorphology has successfully used fisheye lenses with SfM-MVS (Ballarin et al., 2015; Hastedt et al., 2016). Increasing use of action cameras with SfM-MVS means that understanding potential 3D reconstruction quality is now essential.

The use of commercially popular and accessible action cameras, such as the GoPro Hero® range, and a simplification of image acquisition with action would allow regular close-range, high resolution surveys to be achieved at low-cost, rugged equipment. The question of how camera position affects fisheye image capture quality for SfM-MVS – and the degree to which attention to camera position may diminish the number of images required for a multi-camera setup – has received little academic attention. In this contribution the aim is to provide evidence on how the combined effect of camera height and obliqueness (the inclination of the optical camera axis towards the ground – see Figure B3, Appendix B) impacts the overall reconstruction of the subsequent dense point cloud within a set of practical limitations for field deployment.

This article presents an innovative method of image acquisition using a purpose-built camera grid and GoPro® action camera to evaluate the combined effects of camera height, obliqueness and overlap. This evaluation was conducted to inform the design of a fixed multi-camera array which, once thoroughly tested, could be deployed at a wide range of tidal and landward retreat. The grid was designed to allow controlled camera movement without changes to the optical axes (X, Y and Z) caused through human error or environmental conditions. A trial using the camera grid was conducted at a typical site of a small-scale landward retreat at Crosby, northwest England and compared against results obtained using an ‘industry standard’ TLS (assumed ground-truth).

Study Site

Crosby is located on the Sefton Coast in northwest England, UK (Figure 1a), situated north of the Mersey Estuary in Liverpool Bay. The Sefton coastline extends for ~36 km and is influenced by the processes occurring in the Irish Sea and the adjacent estuary (Dissanayake et al., 2014). The coastline is susceptible to some of the highest storm surge conditions in the UK owing to the shallow nature of the north-eastern Irish Sea. The site is located in a macro-tidal environment with a mean spring tidal range of ~ 8 m (Gladstone Dock tide gauge). Local waves are generated by dominant west and northwesterly winds (Plater and Grenville, 2010).

The test location is a ~7 m long section of Crosby/Hightown coast (Figure 1b) surveyed in February 2018. The ~1.5 m high cliff is a combination of unconsolidated material and rubble which has become smoothed and sorted due to wave action. The rubble provides a level of protection but, deprived of sediment deposition, the wave and storm action can cause landward retreat during severe storms (Figure 1b). The underlying stratigraphy is glacial till variably overlain by peat and dune sand (Plater et al., 2010).

Methods

The research used systematically acquired images that were processed with SfM-MVS for comparison with TLS data. Point clouds were evaluated using a two-stage process of assessment. The first established optimal camera positions based on combined height and camera obliqueness within a camera grid. The second used that result to establish a minimal image capture network for a fixed multi-camera array and an estimate of precision for the final output. This section provides an outline of camera grid design, initial site representation and fieldwork procedure which is followed by details on point cloud generation and the process of performance assessment.
Camera grid design

The camera grid was designed to test for ideal camera positions in a multi-camera setup using a GoPro Hero 4 Black action camera. This camera has a 1/2.3-inch (6.2 mm × 4.65 mm), 4:3 CMOS sensor with 1.55 μm pixels. The 'fisheye' lens has a fixed focal length (prime lens) of ~3 mm (17.2 mm equivalent). As with other GoPro cameras, the Hero4 Black has an ultra-wide AOV with differing image capture modes - 'Wide' was used. Still image resolution is 12 Megapixels (4000 × 3000 pixels). The GoPro has a small size, 80 mm × 80 mm × 38 mm with waterproof casing and stand, and a low weight, 152 g.

The camera grid was constructed from timber and wire to separate grid squares (Figure 2a). The specifications of the grid frame and mounting can be found in Appendix A. The grid had 15 rows and 9 columns. The spacing of the grid squares created an image overlap of ~99% and a distance between adjacent cameras (baseline) of 0.11 m. The grid had a 2 m stand-off distance from the cliff which was held constant throughout image acquisition; distance to the scene has a known impact on image resolution and so it was important to maintain this parameter to ensure that any changes were the result of other tested variables (height, obliqueness and overlap). The distance of 2 m would be part of any systematic image acquisition procedure using the subsequent multi-camera setup. Though it is understood that convergent imagery may improve 3D reconstruction, for fixed multi-camera arrays this is not possible as neighbouring cameras may intrude into another camera’s FOV. Convergent imagery, therefore, lies outside the scope of this research.

The ‘traveller’ (Figure 2b) (see Figure A2 in Appendix A for technical specifications) on the front of the grid moved the
camera into each grid square and maintained the X, Y and Z optical camera axes with no deviation in the camera orientation or angle (unless intentional). Further, the traveller allowed movement of the camera between columns and transfer of the camera to the next row. The overall camera grid and base platform was manually levelled using bricks as packers. A levelled platform was essential for the testing procedure to ensure all point clouds were the outcome of the tested inputs (changes to height, obliqueness and overlap) and not incorrect camera position. The grid was mounted on two tripods with props to maintain the Z-axis (Appendix A – Figure A3). The practical design of the camera grid meant it had two heights: ‘Height 1’ in which the top of grid Row A was at a height of 1.64 m from the levelled base; ‘Height 2’ in which the top of grid Row A was at a height of 2.52 m. The result was that Row A of ‘Height 1’ and Row I of ‘Height 2’ were set at an equivalent height.

Initial site representation
A scaled 3D representation of the site and equipment was created to inform practical camera placement prior to fieldwork (Appendix B – Figures B1, B2). The camera grid equipment was reconstructed using SketchUp 2017 and the GoPro FOV estimated to provide prior knowledge on hypothetical scene capture before deployment. The 3D model depicted scene coverage at different positions on the camera grid. The aim was to ensure correct scene capture at different positions without the unintentional inclusion of equipment and to guarantee the scene was central in the image frame (see Appendix B).

Visual analysis of the modelled setup revealed that varying camera angles would be required to maintain a viewsheet of the target surface and prevent encroachment of equipment. This analysis revealed the potential change of camera obliqueness between Rows D and E at ‘Height 2’ where the angle moved from 40° declination from the Z-axis to 30° (Appendix B – Figures B2, B3). Another potential obliqueness change was needed at ‘Height 1’ Row A from 30° declination to 0° (Appendix B – Figures B2, B3). Therefore, Heights 1 and 2 showed varied incidence angles to the scene.

Data acquisition
Fieldwork was undertaken at spring low tide over a nine-hour period and covered a ~7 m section of coastal frontage. Meteorological conditions were suitably diffuse with overcast cloud (James and Robson, 2012).

First, camera grid ‘Height 1’ (Figure 3) was set-up with 0° camera angle as defined in the SketchUp Model. This camera angle captured images perpendicular to the cliff front. Image acquisition began from Row A, Column 1 and proceeded through the grid to Row I, maintaining camera angle. Second, the camera grid was increased to ‘Height 2’, reaching 2.52 m from the base to the top of Row A. Oblique imagery was acquired. The camera obliqueness was adjusted to 40° for Rows A to D. The obliqueness was then adjusted to 30° for Rows E to I to ensure appropriate viewsheet for the target surface as interpreted from the SketchUp representation and confirmed in the field.

Images from Row A at ‘Height 1’ and Row I at ‘Height 2’ were at equivalent heights but different degrees of camera declination (0° for Row A and 30° for Row I). Therefore, they could be used to explore the impact of obliqueness on the dense point cloud. Overall, 270 out of 288 images were used in processing; rejected images were those used for the purpose of identifying row change.

Three converging TLS scans were captured using a ‘Faro 3D Focus 330’ to provide thorough scene reconstruction. Overall scan time for the TLS was ~30 minutes (see Appendix C for further details). The three scans were processed in Faro SCENE 3D (version 7.1). The average point error was 3.2 mm.

Prior to image acquisition, ground control points (GCPs), in the form of three 0.15 m² checkerboards were placed in the scene approximately 1 m apart and georeferenced with Trimble RTK-GPS R6 (see Appendix C for details).

SfM-MVS point cloud generation
Point clouds were generated through a two-stage process of assessment. The first established optimal camera positions and the second used the first result to investigate the impact of decreased redundancy which is essential for establishing a fixed multi-camera array.

Initial SfM-MVS reconstruction for ideal camera positions was undertaken with eight alternate images from each row of the camera grid at both ‘Height 1’ and ‘Height 2’ (Figure 4). The eight images from each row were uploaded into the software Agisoft Photoscan Professional Edition (version 1.3.2.4205) and a dense point cloud was produced. Eight images from each row allowed a balance between computational speed and maintaining an equal baseline between images for initial assessment of dense point cloud reconstruction.

The software provides a workflow in which 3D reconstruction of the scene is established. As a fisheye lens was used for image capture, the camera model was changed to ‘fisheye’ to match the specifications of the GoPro Hero 4 Black. Initially, photographs are aligned (setting was ‘Highest’) through keypoints identified and tracked across the uploaded images. A bundle adjustment then solves external and internal camera parameters. This adjustment results in the creation of a sparse point cloud which is optimized with GCPs. The GCPs (British National Grid/OSNT02) were identified in the images and the Trimble data uploaded into the software to be referenced with the known GCP positions. The addition of coordinates meant the model had real-world reference, scale and orientation which would improve comparison with the TLS dense point cloud. A densification process (setting was ‘Ultra High’) builds the dense point cloud based on the image set and estimated camera positions. This was exported as LAZ files to maintain the coordinate system when uploaded for performance analysis. A more detailed description of the fundamentals of the SfM-MVS workflow can be found in Nouwakpo et al. (2016).
The procedure of generating dense point clouds was repeated for the second stage of assessment – establishing the impact of reduced redundancy. Images were systematically reduced from the image set established in the first stage of assessment.

Performance assessment

To evaluate the impact of the different camera positions on point cloud reconstruction a systematic method of assessment was established. The goal at this stage was to identify the most suitably reconstructed dense point cloud by SfM-MVS compared to a TLS. This is a necessary first step to inform the design of a fixed multi-camera array.

The two-stage process of assessment first established the row with greatest overall performance. The second used the images from that row to investigate the impact of decreased redundancy and establish a minimal image capture network. The process of assessment followed three comparative tests using TLS as the benchmark: two of the tests evaluated positional point accuracy (deviation analysis and GCP analysis) and one, point cloud completeness (completeness analysis). After each stage an aggregated weighted average of the three tests was used to assess optimal camera position and image redundancy. The chosen point cloud was then assessed independently using precision estimates.

Stage One: positional camera parameters

Earlier studies have discussed the need for greater scrutiny of the positional parameters that affect image suitability and interaction (Eltner et al., 2016). However, direct control over camera movement can be difficult due to environmental conditions and human error. Here, the reconstructed dense points clouds from a rigid camera grid with combined variations in camera height and obliqueness were evaluated with the aim to define an optimum set of positional parameters that could be used in a fixed multi-camera set-up.

A dense point cloud was created for each row of the camera grid based on eight alternate images (Figure 4). The comparative metrics are set out below:

i. Deviation metric (B): cloud-to-cloud (C2C) closest point distance calculation is a direct method for 3D point cloud comparison (Lague et al., 2013). The C2C distance is calculated using ‘Nearest Neighbour’ analysis in CloudCompare V2.9 and is based on the point cloud generated by the TLS and those created from SfM-MVS. The method uses two aligned point clouds and defines each point’s nearest neighbour in the reference point cloud with those in the compared point cloud (Ruggles et al., 2016). This test was used because the TLS and SfM-MVS point cloud were of a similar point distribution and density, and C2C offered a direct comparison to the TLS point cloud. The distance (combined X, Y and Z) between the two points is calculated and the mean of these values is termed the mean C2C distance. The resulting mean C2C distance (j) was expressed relative to a 100 mm scale in the form of a deviation metric (B):

\[ B = \lim_{j \to 100} 1 - \left( \frac{j}{100} \right) \]  

A 100 mm scale was chosen as it offers a meaningful range of values against which C2C values could be understood (see Appendix D for further details).

ii. Completeness metric (C): The estimation of ‘holes’ (areas of missing points) in point clouds is an important step to understanding a truly representative 3D reconstruction. Therefore, Python programming language was used to develop an estimation of ‘holes’ present within each point cloud based on 2D JPEGs with nadir views produced in CloudCompare (see Appendix D for details). This estimate was used to produce a completeness metric based on the ratio of filtered pixels in the SfM-MVS images to those within the TLS images.

iii. GCP metric (G): The inclusion of GCPs of known dimensions in the scene allowed for a comparative test of the relative reconstruction accuracies of SfM-MVS and the TLS. The reconstructed GCPs were scaled by the inclusion of the GNSS data and X and Y of each GCP measured using a two-point measurement in CloudCompare. The degree to which SfM-MVS and TLS were able to accurately measure these 0.15 m² squares provided the basis for this test – set out in Equations (D6)–(D9) in Appendix D.

vi. Aggregated test of SfM-MVS performance: Once three comparative tests were completed for each row on the camera grid an aggregated weighted average (A) was calculated for each row.

\[ A = 0.25(B) + 0.5(C) + 0.25(G) \]  

Deviation (B) and GCP (G) metrics evaluated positional point accuracy, and C analysed point cloud Completeness. Both positional metrics are essential for a truly representative reconstruction and so a 50% weighting was given to positional point accuracy (each of the two tests given a weighting of 25%) and 50% given to point cloud completeness – the calculation of which is shown in Equation 2. A score of 1 would imply that SfM-MVS had produced results that were (in aggregate across the three tests) of equivalent quality to those generated by the TLS. Similarly, a score of above 1 would imply that SfM-MVS had been more effective than its comparator in some regard. The row with the highest value from the aggregated test of SfM-MVS performance was deemed to be the optimal camera position and used in ‘Stage Two’ analysis.

Stage Two: minimal image capture parameters

The aim of Stage Two was to assess the impact of the number of images on dense point cloud reconstruction to create a minimal image network. The row with the most appropriate combined camera height and obliqueness, and therefore the best overall average, was established in Stage One. The images from this row were used to create dense point clouds from varying image combinations. Figure 5 displays these combinations and percentage overlap between neighbouring images. The
maximum number of images used was 15 and the minimum was three.

The point clouds created from various image combinations (Figure 5) were assessed using the same three comparative tests (B, C, and G) described earlier. For each combination the aggregated test of SfM-MVS performance (A) was calculated to establish suitable image redundancy in comparison to the TLS. Subsequently, the cloud with a suitable image redundancy was taken through a final precision assessment using precision maps (James et al., 2017). These maps were used to highlight the influence of image geometry on the overall point cloud quality, independent of the TLS. In order to have a practical system of image acquisition for a fixed multi-camera array, image combinations from different rows were not explored.

Results

Stage One: positional camera parameters

Stage One analysis, which was to establish optimal camera position based on combined height and obliqueness, produced 18 point clouds. Nine from the camera grid positioned at ‘Height 1’ (maximum height = 1.64 m) and nine from ‘Height 2’ (maximum height = 2.52 m). The comparative tests provided TLS results with benchmark score of 1 from which the equivalent SfM-MVS result were compared.

Deviation analysis

The mean C2C was in the range of millimetres for all rows from ‘Height 1’ and ‘Height 2’. Camera grid ‘Height 2’ showed better replication and accuracy with lower mean differences overall in the order of 4 to 6 mm (Figure 6). Figure 6 also shows ‘Height 2’ provides better precision with generally lower standard deviation than ‘Height 1’. Greater discrepancy in accuracy was present within ‘Height 1’ with a range of 4 to 10 mm.

The highest mean C2C value was Row A from ‘Height 1’ and the lowest, Row D from ‘Height 1’ (Figure 6). Differences between each SfM-MVS point cloud and the TLS were illustrated with a colour scale of difference. Greatest deviation is seen on the peripheries of the point clouds where there is less overlap of images, reflected in the long tail of the histograms of C2C (Figure 7).

A change in camera obliqueness has an impact on the C2C result; Row I (‘Height 2’) and Row A (‘Height 1’), at equivalent heights but different degrees of declination (Row I at 0° and Row A at 30°), show an increased C2C deviation from 5.30 to 9.69 mm (Figure 6). The increase occurs because of inadequate scene coverage in Row I, which subsequently impacts the ability of the SfM-MVS algorithms to locate and track keypoints within the image set, and therefore
FIGURE 7. C2C scalar field for (a) ‘Height 1’, Row A (highest mean C2C value – 9.69 mm) dense point cloud. (b) ‘Height 1’, Row D (lowest mean C2C value – 4.12 mm) dense point cloud. [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE 8. Perspective view of point clouds for the best and worst completeness results with a source image from the image set. (a) Point cloud from Height 2 Row D had a value of 1.046 – above that of the TLS. (b) Point cloud from Height 1 Row I which had the worst value for completeness at 0.795. [Colour figure can be viewed at wileyonlinelibrary.com]
reconstruct scene geometry. This finding highlights the importance of correct camera positioning in the design of a fixed multi-camera array.

Completeness analysis
Completeness results from each row were compared to the ground-truth set by the three TLS scans (given a representative value of 1). Completeness varied greatly by row and some rows offered better results than the TLS. Overall, rows from ‘Height 2’ displayed consistently higher completeness values than those from ‘Height 1’. Figure 9(a, b) reveals Rows A-D on ‘Height 2’ offered 4% (Figure 8a) more coverage than the reference TLS and the lowest coverage achieved was Row I from ‘Height 1’ (Figure 8b).

As with the previous C2C result, a change in camera obliqueness displayed an impact on the resultant dense point cloud. There was a coverage loss of 19.2% from ‘Height 2’ Row I to ‘Height 1’ Row A through which a 30° angle change was made.

FIGURE 9. The results from the three comparative metrics (Deviation, Completeness, GCP) compared to the TLS reconstruction benchmark for (a) Height 1 and (b) Height 2. (c) The overall SfM-MVS point cloud performance for each row of the camera grid at ‘Height 1’ and ‘Height 2’ compared to the equivalent TLS point cloud. Row D from Height 2 provided the best overall balance for camera positional parameters at a height of 2.13 m and an angle of declination of 40°.

The increased obliqueness in ‘Height 2’, Rows A–D, improved overall completeness through reducing in the impact of shadowing from rock, debris and improving the overall camera FOV.

GCP analysis
Figure 9(a, b) show the relative reconstruction accuracies of SfM-MVS and the TLS for each row of the camera grid at ‘Height 1’ and ‘Height 2’. A result of 1 would imply that SfM-MVS and the TLS were equivalently accurate in surveying the GCPs.

Overall, rows in ‘Height 2’ provided higher GCP accuracy and reduced error than rows in ‘Height 1’, with all results above 0.99. Row D (‘Height 2’) provided the highest GCP reconstructions with an error range of 0.2 to 1.2 mm. The error range produced here is in line with a calculated theoretical error of 0.25 mm (James and Robson, 2012; Eltner et al., 2016). The calculation of the theoretical error is based on the use of parallel-axis imagery captured under ideal conditions. The oblique camera angles used in ‘Height 2’ is likely to have produced a reduction in occlusions and subsequent shadowing effect, improving detail in the images and the keypoint matching process.

Aggregated test of SfM-MVS performance
The calculation of an aggregate weighted average for the three tests provided each row with an overall score relative to the benchmark score of 1 for the TLS (Figure 9c). Overall, the rows from ‘Height 2’ represent the greatest level of performance; Rows A, B, C and D produced results higher than the TLS based on the three comparative tests. Row D at a height of 2.13 m from the base platform and an angle of 40° provided the highest score with 1.015; images taken from this row using SfM-MVS produced a point cloud with a 1.5% greater overall performance than that produced by the TLS, offering the best balance between point positional accuracy and point cloud completeness, both of which are vital for 3D reconstruction. The images from Row D (‘Height 2’) were used in the second stage of analysis to evaluate the impact of image redundancy.

Stage Two: minimal image capture
Stage Two of the analysis examined six point clouds based on a combination of the 15 images captured along Row D at ‘Height 2’. The maximum number of images used was 15 and the minimum was three.

Deviation analysis
Mean C2C distance was in a range of 4.4 to 89.1 mm (Figure 10). The major difference in mean C2C and standard deviation was between three and four images resulting in an 84.7 mm decrease in accuracy. Above four images, there are only slight inconsistent changes in accuracy and precision with the number of images. Figure 11(b) displays the smallest value was the result of six images (4.43 mm) which had an image overlap ~97%; the greatest C2C value was the result of three images with an overlap ~94% (Figure 11a). This latter C2C value may also have resulted from the sensitivity of the C2C test to the larger gaps created by poorer image overlap in the three-image point cloud.

The spatial error distribution of the three image point cloud (Figure 11a) appeared to show a severe deformation, similar in shape to that of a ‘pincushion’ lens distortion whereby the centre of the image bends inwardly – near opposite to that of a ‘doming’ effect (James and Robson, 2014). The distortion present within fisheye lenses means that the image has a high
resolution near the centre or principal point where deformation is negligible. However, resolution decreases non-linearly towards the peripheries of the image and is at its most severe near the corners (Phillips and Eliasson, 2018) which could exacerbate errors. This feature is less prevalent in DSLR cameras but Agisoft Photoscan has proven effective at the modelling and removal of radial distortion for wide-angle lens cameras (Nouwakpo et al., 2014). However, during the self-calibration procedure within Agisoft Photoscan determination of key parameters, such as principal point coordinates, is vital for 3D reconstruction. Increasing the number of images used limits the negative impacts of potentially reduced accuracy of the self-calibrated parameters (Boufama and Habed, 2004; Nouwakpo et al., 2014). Bearing this in mind, the decrease in images to three displays a severe deformation of the point cloud, suggesting the limit of image redundancy has been reached, where the decrease in the image set has potentially impacted the self-calibration process and the accurate determination of key parameters such as the principal point. The reduction has also decreased the area of image overlap to areas of lower resolution and potentially more distorted portions of the image. This causes a reduction in image observations that can be tracked across the image set, and those that are tracked are present within the more highly distorted regions.

Completeness analysis
Completeness results from each image combination were compared to the ground-truth set by the three TLS scans (given a representative value of 1). Figure 12(a) contains the completeness results; all image combinations above three provided a point cloud with completeness greater than or equivalent to the TLS. Similar to the deviation results earlier, there is a dramatic drop-off of 17.7% in completeness at the transition from three (Figure 12a) to four images. Fifteen images provided the

FIGURE 10. Mean C2C distance for point clouds created from image combination when compared to the equivalent TLS point cloud.

FIGURE 11. C2C scalar field for (a) three images (highest deviation value – 89.08 mm) dense point cloud, (b) six images (lowest deviation value – 4.43 mm) dense point cloud. [Colour figure can be viewed at wileyonlinelibrary.com]
highest completeness, 5.5% greater than the TLS (Appendix E – Figure E1a). This result is likely owing to the increased number of images used during the densification process.

GCP analysis
Greater than four images used in the SfM-MVS point cloud reconstruction was able to provide 0.82 to 1.4% higher reproduction accuracy than the TLS (Figure 12a). Fifteen images produced the highest accuracy with an error range of 0.01 to 0.19 mm. The second highest accuracy was produced by six images with a range of 0.03 to 0.13 mm.

Aggregated test of SfM-MVS performance and Precision Maps
Figure 12(b) displays the results of the aggregate weighted average calculation for Stage Two analysis. All point clouds with greater than five images showed results equivalent or better than the TLS, whereas those created from three and four images showed poorer results. Three images, ~94% overlap, resulted in a 45% reduction in performance compared to the dense point cloud construction with four. The use of 15 images, the maximum number available, did not improve the reconstructed point cloud proportionately compared to the point cloud created with eight images. Indeed, just five or six images were required to produce a similar, if not better, performance score than the TLS and a smaller mean C2C value than the point cloud created with 15 images.

Therefore, the six-image point cloud was used for assessment of precision (Figure 13).

The precision maps show results similar, both in scale and spatial distribution, to the C2C deviation results. All the precision estimates for the six-image point cloud produced from Row D (Height 2), are generally the same order of magnitude as C2C deviations from the TLS data. Point precision estimates have been separated into those associated with the external coordinate system and those associated with relative ‘internal’ precision (James et al., 2017). The mean values derived for overall survey precision (Figure 13a), which includes georeferencing error, display an offset from the internal precision of the point cloud (Figure 13b). The internal mean precision (i.e. relative measurable distances in the cloud) are <3 mm for x, y and z components. The good internal precision
suggests strong photogrammetry through high quality tie points and a strong network geometry created by the six images (James et al., 2017). The 'Surface Shape' error does not appear to show signs of systematic deformation but instead a reduced precision on the borders of the point cloud, equivalent to the C2C results. This lack of surface doming and good internal precision values (Figure 13b) would suggest impact from a reduction in image observations on the peripheries from reduced image overlap.

In comparison, overall survey precision appears to be limited by the distribution and, potentially, precision of GCPs. Overall survey georeferencing was calculated as < 3.7 mm in all three translational components (x, y, and z) suggesting a good measurement precision. The strong internal precision (Figure 13b) means that relative measurable distances will be precise across the point cloud. When combined with a strong external precision the measurement of GCP dimensions (such as in the GCP metric) will potentially have sub-millimetre precision. However, despite good values for translational components the spatial distribution of overall survey precision (external) displays radial degradation (Figure 13a). The degradation shows georeferencing certainty reduces away from the GCPs, where the georeferencing datum is initially defined in the bundle adjustment stage (James et al., 2017), subsequently affecting overall external precision. Although translational precision is good, uncertainty may be the result of fewer GCPs distributed on the edges of the point cloud. This same uncertainty is not seen in the internal precision due to the strong image network and tie-points though oblique imagery and high image overlap.

Discussion

Overall, the results support the use of GoPro Hero 4 Black action cameras with Agisoft Photoscan to provide accurate photogrammetric results when acquiring topographic data at a small-scale site of landward retreat. SIM-MVS with GoPro is a low-cost alternative to TLS on the condition that images are captured from optimal camera positions. This result contrasts with those of Thoeni et al. (2014) who found that images from GoPro cameras provided poor 3D reconstruction capabilities. However, this contrast may be due to the use of optimal camera positions in this research, the updated versions of Agisoft allowing the calibration and rectification of fisheye lens distortion and the improvement in GoPro cameras to a 12 Megapixel sensor.

Camera height and obliqueness proved to be dominant factors in reconstruction performance. Overall, reconstructions from camera grid ‘Height 2’ evidenced superior replication those than in ‘Height 1’. The keypoint matching algorithms used in SIM-MVS software rely on unobscured features to be visible in the scene. The presence of oblique imagery and improved viewshed of the camera created by the increased height reduced such surface occlusions and allowed previously shadowed areas from images in ‘Height 1’ to become visible in ‘Height 2’. James and Robson (2014) and Nouwakpo et al. (2016) documented similar improvements in 3D reconstruction when an off-nadir or oblique image acquisition strategy was used.

All the comparative tests displayed a dependence on camera obliqueness. A change in angle from 0° to 30° (height remained consistent) produced a decrease in mean elevation difference between SIM-MVS and TLS of 4.39 mm. Point cloud completeness likewise responded with a 19.2% reduction following the same angle change. The change in angle from 0° to 30° resulted in oblique imagery which is likely to have strengthened image geometry and improved overall reconstruction.

Height, within the set parameters of the camera grid, displayed a more marginal impact on the results of the three comparative tests. The results displayed a general inverse trend; deviation results generally improved with a reduction in height (variation of approximately 1 mm) and point cloud completeness generally decreased with a reduction in height (approximately 1–2% variation). The increased deviation performance may be related to a somewhat improved resolution of images and ground sampling distance (variation of ~ 0.01 mm per row) as the camera moved closer to the scene (Eltner et al., 2017).

For a fixed multi-camera system an optimum combination of both height and obliqueness parameters is essential. This combination (Row D, ‘Height 2’) was established through the use of an aggregate weighted average metric. With a camera obliqueness of 40° (declination) and a height of 2.13 m relative to the cliff, Row D provided the best balance between the three comparative test results. The characteristics of Row D produced a height ratio of approximately 3:4:18 between cliff and camera height allowing the camera obliqueness to remain at 40° to benefit from more oblique imagery and improved viewshed. The use of this ratio would help to account for a degree of natural variability in cliff height at sites of small-scale landward retreat. There is always a compromise between operational practicalities and improving the image geometry for a fixed multi-camera array. However, optimal camera position and the inclusion of off-nadir or oblique imagery cannot only reduce the impacts of shadowing but may also aid the reduction of systematic error present within SIM-MVS processing (James and Robson, 2014).

The point clouds created from different combinations of image overlap were compared relative to the TLS. As few as five well positioned images provided a point cloud similar in accuracy and completeness to three TLS scans. In contrast, the use of three images produced a severe deformation of the point cloud and reduction in overall aggregated performance. Six images provided the least deviation in elevations between SIM-MVS and TLS. The internal precision estimates for the six-image point cloud were < 3 mm (x, y, and z) on average suggesting a strong image network and high-quality tie point estimates. An increase in images from 6 to 15 produced a decrease in deviation accuracy by 0.73 mm. Though only a small change, capturing six images rather than 15 could reduce the cost of hardware and surveying and processing time for a fixed multi-camera array and potentially produce an improved performance. Michelelli et al. (2015) discuss a similar result for non-fisheye lenses, where, when a strong image geometry is present within the image set, a large number of images is not always necessary for an accurate reconstruction. Consequently, these results question the idea that more images always result in a vastly improved dense point cloud (Westoby et al., 2012).

Previous work has shown that the position and overlap of cameras has a considerable impact on the subsequent point cloud (James and Robson, 2012; Smith and Vericat, 2015; Eltner et al., 2016). However, the positional parameters for a multi-camera system and cameras with fisheye lens is less well documented. This information may aid other research projects in the organization and implementation of ground-based image acquisition. This research has provided an adaptable and systematic method of image acquisition which will prove useful to other SIM-MVS projects. Additionally, the research has evidenced that small changes in camera parameters can improve the overall quality of dense point clouds.

The findings of this research point to the potential of SIM-MVS with an array of GoPros to play an important role in the future of low-cost coastal monitoring but also to the further development and use of SIM-MVS applications. The
THE POTENTIAL OF FIXED ARRAY ACTION CAMERAS FOR 3D RECONSTRUCTION

importance of this article does not simply lie in its results but in
the reported methodology and ideas around the use of con-
trolled camera movement, where handheld cameras would
have natural variability. Similarly, the work is innovative in ex-
ploring the possibilities for multiple camera systems (e.g. for
time-lapse and rig setups) and also in its technical image acqis-
tion that could be adapted and transferred to other research
settings. However, there is further research that could build
upon this study:

a. The setup of cameras was designed for reconstructing the
cliff and base of typical small-scale site of landward retreat
which are common areas of environmental interest. How-
ever, exploring these optimal parameters in other landscapes
is required to examine the potential of fixed-array SIM-MVS
further.

b. The distance from sensor to the surface of reconstruction
was set at 2 m and would be held constant for an image ac-
quision procedure using a multi-camera setup. Therefore,
holding this variable constant was an important aspect of
the research. However, distance is likely to impact the re-
construction as each pixel covers a larger area with in-
creased distance from the surface (James and Robson, 2012; Fonstad et al., 2013; Smith and Vericat, 2015; Mosbrucker et al., 2017). Consequently, the impact of dis-
tance for other multi-camera scenarios remains an avenue
for further research.

c. Further scrutiny of other parameters that effect reconstruc-
tion such as lighting, complexity of object, the number and
distribution of GCPs could be investigated in a laboratory
setting.

d. Reconstruction at a small-scale was essential for initial test-
ing of camera capabilities and maintaining X, Y and Z opti-
cal axes. However, extending the size of the reconstruction
along the cliff front and with small variations in cliff height
would aid the development of the technique for larger scale
coastal monitoring.

e. The use of convergent imagery was not suitable for this in-
vestigation due to the nature of multi-camera setups. How-
ever, further scrutiny of the specific impacts of convergent
imagery in a similar systematic format may advance
SIM-MVS research.

f. Further comparisons could also be made with other
SIM-MVS image acquisition procedures (e.g. a single DSLR)
to provide a further detailed analysis of the accuracy of
fixed-camera arrays.

Conclusions

This article illustrates the viability of SIM-MVS with GoPros to
inform the design of a fixed multi-camera array for correctly
reconstructing sites of coastal landward retreat. The results of
which provide a readily available alternative to TLS at a frac-
tion of the financial investment. The performance tests under-
taken illustrate that, when the crucially important positional
variables are taken into account, a small number of well-sited
GoPro images can produce a dense point cloud of equivalent
and, on some measures, superior performance to three TLS
scans. Moreover, the findings show that five images at a height
ratio of 3:4:18 and obliqueness of 40° can produce a point
cloud of sufficient reconstruction quality with an average error
of 4.79 mm to the TLS. Generally, it is true that a larger number of
images will achieve a higher quality output, but with a con-
sidered approach to image acquisition, this article shows that it
is possible to reduce the number of images for a site of this
scale which could potentially shorten survey and processing
time. The implications of these findings point to the potential
of creating an optimized fixed multi-camera array that mini-
mizes the number of cameras needed for image acquisition
through good camera placement. This optimization is particu-
larly relevant to coastal zones at greatest risk in low- and
middle-income countries where frequent monitoring is corre-
spondingly most necessary and access to the repeated use of
expensive equipment can be limited.

Acknowledgements—This research was supported through the Euro-
pean Regional Development Fund (22R15P00045) Low Carbon
Eco-innovatory with industrial partner Marlan Maritime Technologies.

Data Availability Statement

The data that support the findings of this study are available
from the corresponding author upon reasonable request.

Conflict of Interest

The authors declare no conflict of interest.

References

Anthony EJ, Vanhee S, Ruiz M-H. 2006. Short-term beach–dune sand
budgets on the north sea coast of France: Sand supply from shoreface to
dunes, and the role of wind and fetch. Geomorphology 81(3–4):
316–329. https://doi.org/10.1016/j.geomorph.2006.04.022

Ballarin, M., Balletti, C., & Guerra, F. (2015). Action cameras and
low-cost aerial vehicles in archaeology, Proceedings of SPIE Volume
9528, Videometrics, Range Imaging, and Applications XIII, F.
Remondino and M. R. Shortis (eds), 21 June. SPIE Publications: Belling-
ham, WA. https://doi.org/10.1117/12.2184692

Bemis SP, Micklethwaite S, Turner D, James MR, Akciz S, Thiele ST,
Boufama B, Habed A. 2004. Three-dimensional structure calculation:
Achieving accuracy without calibration. Image and Vision Compu-
ting 22(12): 1019–1049. https://doi.org/10.1016/j.
imavis.2004.03.015

Calligaro S, Sofia G, Prosdocimi M, Dalla Fontana G, Tarolfi P. 2014.
Terrestrial laser scanner data to support coastal erosion analysis:
The Conero case study. ISPRS – International Archives of the Photo-
grammetry, Remote Sensing and Spatial Information Sciences XL–5/
W3: 125–129. https://doi.org/10.5194/isprsarchives-XL-5-W3-125-
2013

Conlin M, Cohn N, Ruggiero P. 2018. A quantitative comparison of
low-cost structure from motion (SIM) data collection platforms on
beaches and dunes. Journal of Coastal Research 34(6): 1341–1357.
https://doi.org/10.2112/SI70-011.1

Dissanayake P, Brown J, Karunarathna H. 2014. Modelling
storm-induced beach/dune evolution: Sefton coast, Liverpool Bay,
UK. Marine Geology 357: 225–242. https://doi.org/10.1016/j.
margeo.2014.07.013

Dudzińska-Nowak J, Wężyk P. 2014. Volumetric changes of a soft cliff
coast 2008–2012 based on DTM from airborne laser scanning
(Wolin Island, southern Baltic Sea). Journal of Coastal Research
70: 59–64. https://doi.org/10.2112/COASTRES-D-17-00160.1

Duffy J, Shutter J, Witt M, DeBell L, Anderson K, Duffy JP, Anderson K.
2018. Tracking fine-scale structural changes in coastal dune mor-
phology using kite aerial photography and uncertainty-assessed
structure-from-motion photogrammetry. Remote Sensing 10(9):
1494–1515. https://doi.org/10.3390/rs10091494

Earlie C, Masselink G, Russell P, Shall R. 2013. Sensitivity analysis of
the methodology for quantifying cliff erosion using airborne LiDAR
– examples from Cornwall, UK. Journal of Coastal Research
65: 470–475. https://doi.org/10.2112/SI65-080.1

© 2020 The Authors. Earth Surface Processes and Landforms published by John Wiley & Sons Ltd
Earth Surf. Process. Landforms, Vol. 45, 2478–2494 (2020)
Earlie CS, Young AP, Masselink G, Russell PE. 2015. Coastal cliff ground motions and response to extreme storm waves. Geophysical Research Letters 42(3): 847–854. https://doi.org/10.1002/2014GL062534

Eltnar A, Kaiser C, Castillo C, Rock G, Neugirt F, Abellán A. 2016. Image-based surface reconstruction in geomorphometry – merits, limits and developments. Earth Surface Dynamics 4(2): 359–389. https://doi.org/10.5194/esurf-4-359-2016

Eltnar A, Kaiser A, Abellán A, Schindewolf M. 2017. Time lapse structure-from-motion photogrammetry for continuous geomorphic monitoring. Earth Surface Processes and Landforms 42(14): 2240–2253. https://doi.org/10.1002/esp.4178

Fonstad MA, Dietrich JT, Courville BC, Jensen JL, Carbonneau PE. 2013. Topographic structure from motion: a new development in photogrammetric measurement. Earth Surface Processes and Landforms 38(4): 421–430. https://doi.org/10.1002/esp.3366

Hastedt H, Ekkel T, Luhmann T. 2016. Evaluation of the quality of action cameras with wide-angle lenses in UAV photogrammetry. ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXI-B1: 851–859. https://doi.org/10.5194/isprs-archives-XLI-B1-851-2016

Holman RA, Stanley J. 2007. The history and technical capabilities of Argus. Coastal Engineering 54(6–7): 477–491. https://doi.org/10.1016/j.coastaleng.2007.01.003

James MR, Robson S. 2012. Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. Journal of Geophysical Research - Earth Surface 117(F3): n/a-n/a. https://doi.org/10.1029/2011JF002289

James MR, Robson S. 2014. Mitigating systematic error in topographic models derived from UAV and ground-based image networks. Earth Surface Processes and Landforms 39(10): 1413–1420. https://doi.org/10.1002/esp.3609

James MR, Robson S, Smith MW. 2017. 3-D uncertainty-based topographic change detection with structure-from-motion photogrammetry: precision maps for ground control and directly georeferenced surveys. Earth Surface Processes and Landforms 42(12): 1769–1788. https://doi.org/10.1002/esp.4125

Lague D, Brod N, Leroux J. 2013. Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (N-Z). ISPRS – Journal of Photogrammetry and Remote Sensing 82: 10–26. https://doi.org/10.1016/j.isprsjprs.2013.04.009

Maiti S, Bhattacharya AK. 2009. Shoreline change analysis and its application to prediction: A remote sensing and statistics based approach. Marine Geology 257(1–4): 11–23. https://doi.org/10.1016/j.margeo.2008.10.006

Mancini F, Dubbini M,Gattelli M, Stecchi F, Fabbri S, Gabbianelli G, Gabbianelli G. 2013. Using unmanned aerial vehicles (UAV) for high-resolution reconstruction of topography: The structure from motion approach on coastal environments. Remote Sensing 5(12): 6890–6898. https://doi.org/10.3390/rs5126880

Micheletti, N., Chandler, J. H., & Lane, S. N. (2015). Structure from motion (SfM) photogrammetry. In British Society for Geomorphology Geomorphology Techniques (Vol. 2). Retrieved from www.photosynth.net

Mosbrucker AR, Major JJ, Spicer KR, Pitlick J. 2017. Camera system considerations for geomorphic applications of SFM photogrammetry. Earth Surface Processes and Landforms 42(6): 969–986. https://doi.org/10.1002/esp.4066

Nouwakpo SK, James MR, Weltz MA, Huang C-H, Chagas J, Lima L. 2014. Evaluation of structure from motion for soil microtopography measurement. The Photogrammetric Record 29(147): 297–316. https://doi.org/10.1111/jphor.12072

Nouwakpo SK, Weltz MA, McGwire K. 2016. Assessing the performance of structure-from-motion photogrammetry and terrestrial LiDAR for reconstructing soil surface microtopography of naturally vegetated plots. Earth Surface Processes and Landforms 41(3): 308–322. https://doi.org/10.1002/esp.3787

Perletti L, Polari C, Fassi F. 2017. Fish eye photogrammetry: Tests and methodologies for the survey of narrow spaces. ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXII-2/W3: 573–580. https://doi.org/10.5194/isprs-archives-XLII-2-W3-573-2017

Phillips, J. B., & Ellasson, H. (2018). Camera Image Quality Benchmarking. Wiley: Chichester. https://doi.org/10.1002/97811191905404

Plater AJ, Grenville J. 2010. In Liverpool Bay: Linking the eastern Irish Sea to the Sefton Coast, in: Sefton’s Dynamic Coast, Worsley AT, Lymberry G, Holden VJC, Newton M (eds). Coastal Defence: Sefton MBC Technical Services Department: Southport; 28–54. ISBN:978-0-9566350-0-6 2010.

Plater, A. J., Hodgson, D., Newton, M., & Lymberry, G. (2010). Sefton South Shore: Understanding coastal evolution from past changes and present dynamics. Sefton’s Dynamic Coast, Proceeding of the Conference on Coastal and Geomorphology, Biogeography and Management. 83-106.

Ruggles S, Clark J, Franke KW, Wolfe D, Reimschissel MR, Martin RA, Hedengren JD. 2016. Comparison of SIM computer vision point clouds of a landslide derived from multiple small UAV platforms and sensors to a TLS-based model. Journal of Unmanned Vehicle Systems 4(4): 246–265. https://doi.org/10.1119/juvs-2015-0043

Smith MW, Vericat D. 2015. From experimental plots to experimental landscapes: topography, erosion and deposition in sub-humid badlands from Structure-from-Motion photogrammetry. Earth Surface Processes and Landforms 40(12): 1656–1671. https://doi.org/10.1002/esp.3747

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. https://doi.org/10.5194/isprsarchives-XL-5-573-2014

Westoby MJ, Brasington J, Glasser NF, Hambrey MJ, Reynolds JM. 2012. ‘Structure-from-Motion’ photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology 179: 300–314. https://doi.org/10.1016/j.geomorph.2012.08.021

Westoby MJ, Lim M, Hogg M, Pound MJ, Dunlop L, Woodward J. 2018. Cost-effective erosion monitoring of coastal cliffs. Coastal Engineering 138: 152–164. https://doi.org/10.1016/j.coastaleng.2018.04.008

Wilkinson MW, Jones RR, Woods CE, Gilment SR, McCaffrey KJW, Kokkalis S, Long J. 2016. A comparison of terrestrial laser scanning and structure-from-motion photogrammetry as methods for digital outcrop acquisition. Geosphere 12(6): 1865–1880. https://doi.org/10.1130/GES01342.1

Young AP. 2015. Recent deep-seated coastal landsliding at San Onofre State Beach, California. Geomorphology 228: 200–212. https://doi.org/10.1016/j.geomorph.2014.08.005
Appendix A

Camera grid design

**FIGURE A1.** Technical drawing of the Camera Grid. The internal dimensions of the grid frame were 0.99 m in height and 1.65 m in length. The frame was constructed with fifteen, 0.11 m horizontal grid squares (columns) and nine 0.11 m vertical grid squares (rows). The grid size of 0.11 m × 0.11 m was chosen to accommodate the camera’s dimensions.

**FIGURE A2.** Technical drawing of the ‘traveller’ used for camera movement across the grid.
Appendix B

Initial site representation

The SketchUp modelled camera grid was positioned parallel to the cliff at a distance of 2 m. Thus the modelled FOV (Figure B1) in SketchUp was calculated a distance of 3 m distance to show where the camera FOV would ‘touch’ the cliff front and base (Figures B1, B2). This estimation is of particular importance for fixed multi-camera arrays as the cameras are stationary and understanding how the images cover the scene is essential.

Hypothetical vertical \( (H_v) \) and horizontal \( (H_h) \) FOVs were estimated based on the GoPro AOV specifications: 122.6° horizontal (AOV\(_h\)) and 94.4° vertical (AOV\(_v\)). These values were used in Equation B1 to estimate potential FOV or ‘footprint’ of the camera, where \( D \) is the desired FOV distance (in metres):

\[
D = \frac{H_v \times H_h}{AOV_h 	imes AOV_v}
\]
Appendix C

Data acquisition

TLS scan details

The Faro has an integrated digital camera which automatically captures images of the scene. These images were used to colourize the dense point clouds and 3D models (Thoeni et al., 2014). The three scans were processed in Faro SCENE 3D (version 7.1) and edited to remove noise, errors, duplicated points and crop areas not relevant to the survey. The three scans were registered together into a coordinate system using GCPs to provide correct orientation and scale.

GCP details

Checkerboards were chosen as they are recognized in both the Faro Scene and Agisoft Photoscan software. One of the GCPs was a cube at 0.15 m³. The cube was levelled and could be used as an axes reference for model orientation. The checkerboards were georeferenced using a Trimble RTK-DGPS R6 which optimized orientation and alignment for the dense point cloud. The horizontal coordinates for the reference points were set to the British National Grid (OSTN02) while the vertical coordinates were referenced to the mean sea level using the geoid model OSGM02.

Appendix D

Performance Assessment

Deviation metric further details (B): Prior to calculation, the reference point cloud (TLS) and the SfM-MVS point cloud were aligned using GCPs as point locations. An Iterative Closest Point (ICP) (Nouwakpo et al., 2014; Westoby et al., 2018) was used for a finer registration then the point clouds were cleaned and cropped to equivalent sizes. The final product was a mean C2C value, standard deviation and scalar field that was coloured to represent areas with greater deviation. The result of this test is used in Equation 1 in the main text, where the C2C mean value ($\bar{j}$) is measured relative to a 100 mm scale.

Completeness metric further details (C): Initially, the point clouds were equally segmented along the base of the cliff and separate nadir images produced for cliff and floor for testing. The background of the image was changed to an artificial colour which would allow for the identification of ‘holes’ and a standardized nadir viewshed of the point cloud established for the cliff and the floor. This process was repeated for the TLS point cloud segmentation images as a benchmark for comparison with equivalent SfM-MVS images (Figure D1).

Each image was filtered using Python OpenCV Image analysis functions to determine the number of green pixels present. Using these numbers, the ratio of filtered pixels (i.e. holes) within an image was estimated:

\[
H_c = 2 \cdot \frac{\text{AOV} \cdot \text{Row} \cdot \text{Depth}}{2}
\]

Equation (1)
where $R_s$ and $R_t$ are the ratio of filtered pixels, $n_t$ is the total number of image pixels, and $n_{fs}$ and $n_{ft}$ are the number of filtered pixels in the SfM-MVS and TLS images, respectively. The following was used to compare the ratio of filtered pixels in the SfM-MVS images to those within the TLS images:

$$M = \frac{R_s}{R_t}$$ (D4)

This was done separately for floor ($M_f$) and cliff ($M_c$). The $M_f$ and $M_c$ values were each given a weighting of 50% to reflect their equal importance:

$$C = 0.5(M_f) + 0.5(M_c)$$ (D5)

GCP metric further details ($Q_i$): Equations (D6) and (D7) describe the test of accuracy for both TLS and SfM-MVS ($P_s$ refers to the accuracy of SfM, $P_t$ refers to the performance of the TLS). Firstly, under- and over-measurement of the GCPs had to be treated equitably. The conditional statement ("if, then") denoted by the logical operator $\rightarrow$ occupying the numerator space between the parentheses in Equations (D6) and (D7) describes this process. The same set of steps was undertaken for the TLS summarized in Equation (D7).

Following this logical process the result of the under- or over-representation from the method of reconstruction was then divided by the GCP known value ($R$) to obtain a ratio of each method of reconstruction’s error relative to reality. Subtracting this result from 1 provided a measure of how accurate the method of reconstruction had been at recreating the known dimensions of the GCP.

$$P_s = 1 - \left\{ \frac{[(S > R) \rightarrow (S - R)] \lor [(S < R) \rightarrow (R - S)]}{R} \right\}$$ (D6)

$$P_t = 1 - \left\{ \frac{[(T > R) \rightarrow (T - R)] \lor [(T < R) \rightarrow (R - T)]}{R} \right\}$$ (D7)

$$Q = \frac{P_s}{P_t}$$ (D8)

$$G = 0.167(Q_1) + 0.167(Q_2) + 0.167(Q_3) + 0.167(Q_4) + 0.167(Q_5) + 0.167(Q_6)$$ (D9)

Equation (D8) describes the ratio of the results of Equations (D6) and (D7) where the capacity of SfM-MVS to accurately recreate the GCPs is understood relative to the same test applied to the TLS. If SfM-MVS proved more accurate than the TLS a value for $Q$ of greater than 1 would be returned for each of the GCPs ($Q_1$–$Q_6$). This test was applied to the X and Y axes of the three GCPs (six measurements in total). Each of these six measurements was then given a weighting of 0.167 in the overall test set out in Equation (D9).

**Appendix E**

**Results**

Minimal image capture Completeness results.

![Fig E1](wileyonlinelibrary.com)

**FIGURE E1.** Perspective view of point clouds for the best and worst completeness results for minimal image network. (a) Point cloud created from 15 images had a value of 1.055 – above that of the TLS. (b) Point cloud created from three images which had the worst value for completeness at 0.831. [Colour figure can be viewed at wileyonlinelibrary.com]

**Supporting Information**

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Supporting Information S1