

Investigating the geomorphological potential of freely available and accessible Structure-from-Motion photogrammetry using a smartphone

N. Micheletti^a, J. H. Chandler^b, S. N. Lane^a

^a*Institute of Earth Surface Dynamics, University of Lausanne, Switzerland*

^b*School of Civil and Building Engineering, Loughborough University, LE11 3TU, UK*

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/esp.3648

Abstract

We test the acquisition of high-resolution topographic and terrain data using hand-held smartphone technology, where the acquired images can be processed using technology freely available to the research community. This is achieved by evaluating the quality of digital terrain models (DTM) of a river bank and an Alpine alluvial fan generated with a fully automated, free to use, Structure-from-Motion package and a smartphone integrated camera (5 megapixels) with terrestrial laser scanning (TLS) data used to provide a benchmark. To evaluate this approach a 16.2 megapixel digital camera and an established, commercial, close-range and semi-automated software are also employed, and the product of the four combination of the two types of cameras and software are compared. Results for the river bank survey demonstrate that centimeter precision DTMs can be achieved at close range (10 m or less), using a smartphone camera and a fully automated package. Results improve to sub-centimeter precision with either higher resolution images or by applying specific post-processing techniques to the smartphone DTMs. Application to an entire Alpine alluvial fan system shows that the degradation of precision scales linearly with image scale, but that (1) the expected level of precision remains and (2) difficulties in separating vegetation and sediment cover within the results are similar to those typically found when using other photo-based techniques and laser scanning systems.

Keywords:

Structure-from-Motion (SfM), Close-range photogrammetry, Terrestrial laser scanning (TLS), Digital Terrain Model (DTM), Smartphone technology

1. Introduction

The last two decades has seen a revolution in topographic data measurement for geomorphic research, with both a substantial increase in the rate at which it is possible to acquire precise, three-dimensional terrain data and the ease with which associated methods can be applied. Initially, these developments focused upon constructing digital elevation models (DEMs) or digital terrain models (DTMs) using both photogrammetric (e.g. Lane et al., 1994; Barker et al., 1997; Lane et al., 1999; Lane, 2000) and differential global positioning system (dGPS) (e.g. Fix and Burt, 1995; Brasington et al., 2000) data. Whilst these approaches allowed users to generate DEM data themselves and so to control the data acquisition process (e.g. through ground-based surveys), they remained highly dependent upon both expensive equipment and expertise to manage and improve data quality (Lane et al., 2004). Over the last decade, terrestrial laser scanners proved capable of generating very high quality DEM data (e.g. Heritage and Hetherington, 2007; Alho et al., 2009; Hodge et al., 2009a,b; Schaefer and Inkpen, 2010) and have almost become routine in some DEM collection strategies. However, they remain relatively expensive items of technology and have only recently become truly portable. Thus, much interest remains in acquiring DEM data using much less expensive technologies and the last few years has seen a series of innovative adaptation of imaging systems for geomorphic research, including range imaging (Nitsche et al., 2013) and applications of the Kinect sensor

(Mankoff and Russo, 2013). Such methods have proved capable of measuring topographic surfaces with a precision in the mm to cm range as the basis of DEM construction.

Photogrammetric developments have been fundamental in allowing the development of DEM-based methods in geomorphology. approaches to photogrammetric DEM collection required access to expensive hardware and software, even when the traditional constraints imposed by aerial (or analogue) photogrammetric methods were relaxed through use of analytical methods (Chandler and Moore, 1989). They also required correct handling of geometrical distortions associated with image acquisition, notably when using non-metric cameras (e.g. Chandler et al., 1990; Butler et al., 1998, 2002; Brasington and Smart, 2003). Most recently, many of these constraints have been surpassed or have become more automated through the development of Structure from Motion (SfM) methods (e.g. James and Robson, 2012; Westoby et al., 2012; Fonstad et al., 2013). SfM has its origins in the machine vision community, particularly the tracking of points across sequences of images occupied from different positions. It has been developed and adapted for generating DEM data using potentially many images in a sequence Fonstad et al. (2013). In traditional photogrammetry, only two images of the same surface are required. After an estimation of image orientation, stereo-matching is then used, as it is in SfM approaches, to identify conjugate point pairs in the images and these can be used to determine 3D co-ordinates of points in the images, provided ground control data are available to determine

sensor position and orientation, and sensor internal geometrical distortions are known. SfM differs by starting with the image matching process, of multiple images of unknown position, orientation and distortion, to produce many matched points. By doing this for multiple images of the same surface, it becomes possible through iteration to determine relative sensor position, orientation and distortion. The resulting model can be transformed into a 3D object system after processing if ground control points are available (e.g. Westoby et al., 2012; Fonstad et al., 2013). As with conventional close range photogrammetry and laser scanning techniques, SfM is a truly 3D measurement technique in that by measuring the same surface from multiple perspectives, the result is a 3D point cloud rather than a 2.5 D surface (i.e. a set of x, y locations each with a single z value).

SfM-based techniques have been recently employed for a number of geosciences applications, including coastal erosion (James and Robson, 2012), fluvial environments (Fonstad et al., 2013; Woodget et al., in press), gully headcut retreat (Gomez-Gutierrez et al., 2014), etc. The aim of this paper is not to repeat a demonstration of the potential of the SfM approach for geoscience applications (see James et al., 2013; Westoby et al., 2012; Chandler and Fryer, 2013; Fonstad et al., 2013). Rather, it is to assess two developments. First, because of the ease with which sensor distortion can be modelled, the range of potential sensors has increased. Even ubiquitous imaging technologies, such as smartphone sensors, might be used to acquire valuable geomorphic data. Second, alongside these new SfM approaches has been the

development of low cost, sometimes free, internet-based processing systems: images can be uploaded, processed and the derived 3D data downloaded, sometimes only a few minutes after submission. Taken together, these two developments offer the possibility of very fast, fully automated and low cost acquisition of 3D data, based upon the analysis of smartphone-acquired images submitted to a processing service. Thus, the aim of this paper is to assess the quality that can be achieved using a freely available and instantly accessible SfM resource with smartphone imagery, a combination that can enable exceptionally low cost, rapid and easy 3D object capture and DTM acquisition in geomorphology. There is a range of options currently available for the acquisition of high resolution topographic data using a hand-held camera (Table 1), and 1a: smartphone and internet-based SfM system represents the lowest approach in terms of cost, expert supervision needs and possibly processing time. To assess the quality of such approach, we take a highly constrained photogrammetric approach, using a high quality sensor and subscription-based software capable of generating DEM's using two images (Table 1 approach 2c). We then compare this with three progressively less stringent elements of data acquisition: (1) the use of smartphone sensors (Table 1 approach 1c); (2) the use of a non-subscription internet processing service (Table 1 approach 2a); and (3) processing involving almost no ground control data. Two developments are not tested here: traditional, non-internet based processing SfM-MVS (multi-view stereo) (Table 1 image processing approach b) and the SfM-MVS approach integrated into subscrip-

tion base photogrammetry software (Table 1 image processing approach d, notably PhotoModeler release 2014).

[Table 1 about here.]

In all cases, our benchmark comparison is with data acquired using terrestrial laser scanning (TLS) technologies and we focus upon one close range measurement application (a 10 m long by 1.20 m high river bank) and one intermediate measurement scale application (an Alpine alluvial fan of approximately 87,000 m²). In addition, we perform an specific laboratory experiment to investigate the relation between the number of frames and quality of the derived DTMs.

2. Methodology

2.1. Close range measurement scale

Figure 1 shows the workflow adopted in this research.

[Figure 1 about here.]

2.1.1. Data acquisition

The close range case study chosen for the experiment is a 10 m long by 1.20 m high length of riverbank, in the Borgne d’Arolla, Val d’Hérens, Switzerland (Figure 2). The riverbank texture is heterogeneous, with grains ranging in size from a few centimetres to half a meter set in a coarse sand and fine gravel matrix. Ten photogrammetric “coded” targets were distributed

Accepted Article
along the riverbank and used as tie points for image orientation and for the computation of an appropriate transformation between the co-ordinate systems associated with the differing processing approaches examined. However, co-ordinates of the control points were not established using a conventional control survey involving a total station. Instead, TLS co-ordinates of the same coded targets were extracted and used for the co-registration purposes.

[Figure 2 about here.]

To assess the accuracies of the generated DTMs, terrestrial laser scanning (TLS) data were used as a benchmark. The riverbank was scanned from a distance of approximately 13 m using an Optech ILRIS 3-D scanner, with a point spacing of approximately 2 millimetres. In the absence of an absolute co-ordinate system, TLS co-ordinates are used to provide a reference. 3-D co-ordinates of targets in the TLS system were identified manually using the point cloud data management software Cloud Compare (EDF R&D, 2012), freely available at www.danielgm.net/cc.

The first device used for image acquisition was a Nikon D7000 model: a 16.2 megapixel digital single-lens reflex camera. 13 photographs were acquired from short distance (8-12 m) with a fixed focus 35 mm lens, attempting to obtain a uniform coverage of the feature and to maximise overlap. The same procedure was adopted using a smartphone, an Apple Iphone 4 device, equipped with a 5 megapixels camera also used to collect 13 photographs.

2.1.2. 3D scene reconstruction using PhotoModeler and 123D Catch

Photographs collected using Nikon D7000 and Iphone 4 devices were processed using both the Eos Systems Inc. PhotoModeler software (Version 2012) and Autodesk 123D Catch. A fundamental difference exists between the two software packages used. PhotoModeler provides more comprehensive facilities and control to the user at the expense of requiring greater understanding, making it more challenging for non-photogrammetrists. Effectively, it requires the user to know at least the basic concept of photogrammetry and technical terms, and to be able to correctly sustain the software in performing interior and exterior orientation and to understand the controls on output quality. In contrast, 123D Catch is a fully-automated, black-box tool and expert supervision is unnecessary. The potential expert input offered by PhotoModeler should generate better results, but 123D Catch has the distinct value of being freely available to all, particularly non specialized users.

PhotoModeler is a subscription software which provides tools for image analysis to generate 3D clouds of points. Its cost varies depending on options selected and whether the usage is commercial or academic. Prices and details about PhotoModeler are available at www.photomodeler.com.

To generate accurate spatial data using conventional photogrammetry it is essential to determine the geometrical characteristics of the imaging sensor (Chandler, 1999). This can be achieved using a targeted field and well established camera calibration methods involving a self-calibrating bun-

Accepted Article
dle adjustment. With modern software and appropriated coded targets this has become a routine and fully automated procedure (Sanz-Ablanedo et al., 2012). Thus, the two devices were used to capture different photographs of similar targets to those used on the riverbank but located on a planar surface (a wall). These images were used in PhotoModeler to estimate camera calibration models for the two sensors. These interior orientation models were then used for the PhotoModeler analysis of the riverbank imagery.

A bundle adjustment was used to estimate camera positions and orientations and to extract final point clouds, with a user control on point density. TLS targets co-ordinates were imported into PhotoModeler and associated with their corresponding targets in the images. The use of these tie points (TPs) allowed transformation of point clouds in PhotoModeler directly to the TLS co-ordinate system.

The software 123D Catch, developed by Autodesk implements a Structure-from-Motion based approach. It is freely available at www.123dapp.com/catch as PC software, smartphone app or web app, all providing the same services in generating the 3D models and differing only in the post-processing options. It requires the user to supply a minimum of three images of objects, to generate 3-D meshes automatically. Images are resampled down to a 3 MP resolution currently, so a high resolution sensor is not required. Moreover, images do not need to be from the same distance or have the same scale. The software does not pose restrictions upon camera type or focus settings; a camera calibration model is derived directly and automatically from

each photograph provided by the user. Nevertheless, 123D Catch provides a output file (in FBX format) where some basic calibration data are provided, as estimated focal length, exterior orientation parameters, sensor size and scale factors. All photographs obtained using the Nikon D7000 and the Iphone 4 devices were uploaded to the Autodesk server in distinct projects, and automatically processed to generate 3-D meshes. The maximum mesh density option was selected in 123D Catch. These meshes were exported in LAS-format files to allow further analysis in the Cloud Compare software. For practical reasons, the PC interface of 123D Catch was employed.

In both software systems, image-covered polygonal meshes were generated for visualisation purposes, by connecting the derived points in 3D space by line segments (Figure 3). The point clouds (mesh vertices) were used for the analysis.

[Figure 3 about here.]

2.1.3. Post-processing and comparison with terrestrial laser scanning

The Nikon D7000 and Iphone 4 photographs processed using the PhotoModeler and Autodesk 123D Catch softwares resulted in 4 point clouds. However, the 123D Catch models required a further post-processing operation after the DTMs had been generated to transform the DTMs to the TLS system. In our case, we used a Horn's absolute orientation algorithm (Horn, 1987) implemented in Matlab. A manual identification of targets across multiple images is performed in 123D Catch, and the corresponding co-ordinates

are extracted from the FBX output file. Afterwards, a rotation, translation and scaling transformation is used to compute 123D Catch co-ordinates in the TLS system. The accuracy of this transformation is demonstrated by the root mean square of the residuals shown in Table 2.

[Table 2 about here.]

To simplify the comparison process, points not in the area of interest were eliminated in all point clouds. The resulting points clouds include a number of points varying from 200,000 to 600,000 (Figure 4). The TLS data includes more than 3 million points.

[Figure 4 about here.]

The comparison between Structure-from-Motion models and TLS data is performed using the distance tool in Cloud Compare. The tool uses a Chamfer Matching Algorithm (Barrow et al., 1977) to compute the Chamfer Distance, that is, a value of dissimilarity between two point clouds. This value is computed by associating each point in the compared dataset to its closest point in the reference dataset, and calculating the three-dimensional distance. These are the reported error values, referred to as “errors” or “distance errors”. To improve visualisation of the results, a maximum distance of 0.1 m was set to isolate outliers.

A further comparison was made after the application of an Iterative Closest Point (ICP) algorithm. The algorithm iteratively revises a transformation solution in order to minimize the spatial difference between the two point clouds (Zhang, 1992). It optimizes a rotation matrix and a translation vector to fit the compared point cloud to the reference one. ICP has previously been used in both Structure-from-Motion (James and Robson, 2012) and close range photogrammetry (Eos Systems Inc., 2012) quality assessment to ensure that the co-ordinate system alignment is as tight as possible. This is also a logical step as the aim is to compare two methods used to generate data in the same co-ordinate system requiring any alignment problems to be removed.

2.2. Intermediate measurement scale

2.2.1. Data acquisition and 3D scene reconstruction

An Alpine alluvial fan located near the village of Satarma, Val d'Hérens, Switzerland (Figure 5) provided an intermediate range case study. The area of the alluvial fan is approximately 87000 m². As in the previous case study, TLS data was used to provide a benchmark for comparison. The TLS data have a point density of approximately 15 centimeters in the middle-upper part of the fan, where vegetation is absent and the surface is grass and sediment covered, and are registered in the Swisstopo LV95 co-ordinate system using dGPS measured targets. The focus of the intermediate range study is on the performance of the use of smartphone sensors and internet processing services over a larger area. Hence, the analysis was performed using an Autodesk 123D Catch model (see Section 2.1) obtained from 13 photographs captured with an Apple Iphone 4 device only. Images were taken upfan, from c. 350-450 meters from the channel turn visible in the middle of the alluvial fan in Figure 5.

[Figure 5 about here.]

This case study is more challenging compared to the close range one and includes a number of complications. First, and as expected, the acquisition of an appropriate set of images for the analysis is more difficult given the size of the object of study. In particular, very well defined structures on the fan might not have a sufficient representation from different angles and available

positions from the imaging survey. Second, the alluvial fan is partially covered by vegetation, and this represents a major problem for multiple reasons. At this scale, the occlusion caused by vegetation cannot be recovered by the acquisition of images from different angles, as data collection from the ground allows limited positions for the camera. Further, comparison with TLS data would be problematic because of the well-known characteristics of laser surveys in vegetated zones (i.e. points on leaves versus points passing through the vegetation and hitting the ground). Finally, in this case, the increased distance between the camera and the alluvial fan is not optimal for obtaining the necessary image texture, causing the 3-D reconstruction to be more difficult and of varying quality.

2.2.2. Post-processing and comparison with terrestrial laser scanning

Similar to the close range experiment (see Section 2.1), a Horn's absolute orientation algorithm (Horn, 1987) was used to transform the alluvial fan 123D Catch model co-ordinates into the official Swiss co-ordinate system LV95, used for the TLS data. Since the scale of this experiment is too large for the use of coded targets, 8 well-defined points were manually identified and used as tie points between the two sources of data to perform an approximate transformation. A Monte Carlo simulation to re-fine the transformation has then been executed. The resulting 123D Catch model is presented in Figure 6. Due to the presence of vegetation and related shading effects and the point cloud density differences between the 123D Catch model and the TLS

Accepted Article

data, the sampling of points for the comparison process proved challenging. To overcome this issue and to obtain reliable results, the point clouds were clipped so that the comparison is performed on the exposed main fan channel zone, where vegetation effects can be neglected (red perimeter in Figure 6). The calculation of error is performed as in the close range case study, by using the Chamfer algorithm to calculate distance errors and the ICP algorithm to reduce co-registration errors (see 2.1.3).

[Figure 6 about here.]

2.3. Laboratory experiment using a flat surface

One of the potential main controls on the successful application of SfM methods in geoscience is likely to be the number of images. It is reasonable to suppose that an increase in the number of images can generate a more dense mesh, potentially enhancing the quality of derived models. Very large datasets of more than one hundred images were employed in similar studies (James and Robson, 2012; Westoby et al., 2012; Fonstad et al., 2013), but these authors show how this lengthens processing times due to the computationally demanding nature of keypoint descriptor extraction, matching and reconstruction algorithms. The Iphone 123D Catch software currently allows a maximum upload of 70 images, while the PC interface allows more (note: generated meshes can be transferred between both interfaces).

The optimal or sufficient number of images required to maximize the quality of the derived models is likely to vary according to the complexity

of the surface being measured as well as the exact sensor (and hence frame coverage) being used. In order to investigate the impact of the number and distribution of images on the quality of the computed models in a generic way, we have undertaken an additional experiment using a simple methodology the authors had used previously (Chandler et al., 2005). We took the hypothesis that the number and distribution of the frames should be fundamental for controlling the quality and specifically the accuracy of the model. When a reduced number of frames are employed, the generated model could be deformed and/or successful only for a part of the scene. This suggests the need to consider (1) an appropriate distribution of photo locations to capture the scene, and (2) significant overlap between as many frames as possible. In a laboratory experiment, a simple A3 (29.7 cm x 42 cm) high quality print of a highly textured image (a pebble surface) was pasted onto a flat wooden board which can be assumed to be planar. This enables a very large number of checkpoints to be determined in the object space and allows the accuracy of the measurement system to be quantified (Figure 7). A total of 53 convergent and oblique frames were collected using the Iphone 4 device. Given the reduced scale and simplicity of this extreme test setting, it was possible to acquire all frames with significant overlapping areas between them, whilst maintaining a good coverage of possible image capturing positions. We then focus on the 123D Catch software, reflecting this paper's emphasis on low cost processing options. A number of Autodesk 123D Catch projects were generated, by varying the number of images and view angles.

[Figure 7 about here.]

The relation between the number of frames and the quality of derived DTMs was tested with the following method: after generating a 123D Catch model using all images available, groups of frames were progressively removed whilst maintaining a good scene covering and overlapping area. Hence, models are generated using 53, 40, 30, 20, 13, 10, 8 and 7 frames. Although 123D Catch can be used with only three images, projects with fewer than seven frames failed to provide satisfying results, probably because the iterative process for the exterior orientation could not converge with few images in this case and need more redundancy to be successful. To allow direct comparison, all meshes were scaled directly in 123D Catch using a measured distance.

Since the object surface is flat (topography is completely absent), model accuracy can be determined without the need for co-ordinate manipulation, which provides a major benefit of this particular methodology. A best fit plane is computed by minimizing the perpendicular distances between the plane and the points in the point clouds. Then, the distance between the plane and each 3D point is determined to estimate model accuracy.

3. Results

3.1. Close range measurement scale results

The error analysis for the Structure-from-Motion data and TLS for both direct registration and Iterative Closest Point algorithm for the initial river-

bank test is presented in Table 3. The histograms of distance error (Figure 8) have log-normal distributions, with a positive skew and hence we focus on the median rather than the mean error in the results analysis. Optimum results with direct registration were obtained using the Nikon D7000 frames in the PhotoModeler software, with a median error of 0.0038 m. In contrast, the application of Autodesk 123D Catch to Nikon D7000 photographs has a median error of 0.0044 m compared to TLS data. As expected, Iphone 4 images using PhotoModeler degrade to a median error of 0.0053 m. The application of Iphone 4 data to Autodesk 123D Catch degrades further to a 0.0148 m median error. For every approach, the error median is lower than the mean error, because of the positive skew of error distributions. Generally, all models provide very low median values, with the exception of the Iphone 4 123D Catch approach.

[Table 3 about here.]

With the exception of Iphone 4 123D Catch, all experiments achieve 95% of distance errors below 0.03 m. There is a high level of similarity between the two Nikon D7000 distributions (Figures 8(a) and 8(b)), which differ only in the number of outliers (more important for 123D Catch, as confirmed by the 99th centile in Table 3), and slightly in the proportion of errors between 0.01 and 0.02 m. The Iphone 4 Photomodeler approach (Figure 8(c)) has a similar distribution to Nikon D7000 approaches, but with a slightly greater proportion of points with a 0.02 to 0.03 m error (confirmed by higher quantiles

values). Finally, whilst Iphone 4 images processed in 123D Catch presented in Figure 8(d) have an error distribution which peaks at a similar error to other approaches (i.e. below 0.005 mm, with a slightly higher mode of 0.0044 m), its right tail is much more important, featuring higher errors than the others methods as demonstrated by the 75th centile, respectively 0.013, 0.0188 and 0.0204 m greater and more importantly by the 0.0615 value of the 95th centile, greater by more than 0.03 m.

[Figure 8 about here.]

The ICP algorithm was applied to help separate out registration and random errors. The application of the ICP algorithm did not improve significantly the Nikon D7000 PhotoModeler result; the median error decreases by only 0.0007 m, while the mean error drops only by 0.0005 m. On the other hand, the Iphone 4 PhotoModeler and the Nikon 123D Catch models benefit greatly from the refined co-ordinate system alignment by ICP. After the ICP application, mean errors between these models and the TLS data are approximately 0.005 m, with median values of 0.0032 and 0.0034 m, a performance that could be considered comparable to the Nikon D7000 PhotoModeler results. The Iphone 4 123D Catch result also improves considerably; the mean error decreases by one half and is now approximately 0.01 m. The median also decreases by one half (from 0.0148 m to 0.0079 m).

Figure 9 shows the distribution of errors after the ICP application. These distributions are consistent with the error statistics observed in Table 3: both

Nikon D7000 applications and Iphone 4 PhotoModeler results are comparable (Figures 9(a), 9(b) and 9(c)), achieving very satisfying performances. On the other hand, whilst Iphone 4 123D Catch errors reduce considerably, its error distribution demonstrates how performances are still inferior to the three aforementioned approaches (Figure 9(d)). Nonetheless, the model benefits greatly from the correction, as demonstrated also by important reduction in errors centiles.

[Figure 9 about here.]

For an additional visualization of the benefits of the ICP algorithms, 9 DEMs on a regular grid with 0.005 m spacing were generated (4 direct registration models, 4 post-ICP models, and 1 using the TLS data). Given the high density of points, a simple triangle-based linear interpolation was sufficient to generate DEMs at the chosen resolution. For visualization, the grid was generated on the X-Z co-ordinate axis, with elevation provided by the Y axis, and on a rectangular zone of the riverbank. DEMs of Difference (DoD) were generated by subtracting the 8 SfM and PhotoModeler surfaces from the TLS surface, which is used again as benchmark (the smaller the value of the DoD, the more accurate the model). Figure 10 illustrates the direct registration DoD compared with the ICP one, for the four SfM and PhotoModeler approaches. Results are broadly as expected with differences reduced following use of the ICP algorithm. The Nikon sensor appears to generate the smallest differences but optimum accuracies appear to be achieved

with 123D Catch. Table 4 illustrates the root mean squared error (RMSE) of each DEM compared to the TLS one. These values confirm the visual observations, as PhotoModeler models benefit less from ICP (from 0.0381 m to 0.0197 m using the Nikon sensor and from 0.0213 m to 0.0174 m using the Iphone 4 sensor) than the 123D Catch outputs (0.0647 m to 0.0168 m for the Nikon-PM approach). We attribute the high differences still remaining in the Iphone 4-123D Catch model after the ICP application to a systematic error introduced by the tilting clearly visible in Figure 10. The improvement inducted by the ICP for this model is c. 0.02 m but could be greater if the model were not tilted.

[Figure 10 about here.]

[Table 4 about here.]

Figure 11 shows the distance errors between the four Structure-from-Motion and PhotoModeler models and the reference TLS data as point clouds. Generally, the largest errors are concentrated around the edges of the riverbank, either on the top of it or on the blocks at its base. Secondly, a difference in point density between the PhotoModeler and 123D Catch models is visible; PhotoModeler derived point clouds have a point spacing of c. 0.007 m, while the distance between two neighbours points in 123D Catch is of c. 0.02 m. Third, it is possible to identify individual stones or others structures where errors are larger. Effectively, major errors are located in shaded spots, surrounding stones or in small crevices. Finally, Iphone models

have larger point densities on the right edge of the riverbank, although these zones correspond to higher errors. We attribute the larger point density to the fact that more images seem to contribute to matched points in this zone. Nevertheless, because of photo position and orientation, the baseline does not seem optimal for depth precision.

[Figure 11 about here.]

3.2. Intermediate measurement scale results

Distance errors between the Iphone 4, 123D Catch model and TLS data are shown in Table 5 for the alluvial fan. A median error of approximately 0.600 m is obtained. The error distribution for the alluvial fan - TLS comparison (Figure 12(a)) is log-normal, similar to the close range experiment distributions (Figure 5). The frequency peak corresponds to an error of approximately 0.5 m (with a 0.44 m mode), and more than 75% of distances are less than 1 meter (the 75th quantile is 0.9493 m).

[Figure 12 about here.]

As expected, the application of the ICP algorithm improves the results significantly: the aligned cloud points have a median error of 0.4226 m, with a mean of 0.5263 m. The error distribution histogram in Figure 12(b) and Table 5 show this improvement, with now more than 86% of errors less than 1 meter.

[Table 5 about here.]

Figure 13 visualises these errors. The largest distance errors are concentrated on the flanks of the main channel and discrepancies between the cloud points seem to increase upstream in that area. This may be explained by an insufficient coverage of the channel flanks in the pictures for a solid reconstruction of their shape. The camera positions are too few and too low to allow an optimal orientation to reconstruct channel flanks. In addition, the

increase in pixel size with distance up the fan translates into an increasing distance between the sensor and the observed surface and this also could be responsible for less accurate reconstructions. This also, reduces the density of derived samples in the point cloud and, as a consequence, the precision of the model. To verify this assumption, a subset of points has been extracted representing a narrow strip in the direction of slope. A scatterplot of the discrepancy of the sampled points between 123D Catch and TLS data, as a function of distance to one camera position, is presented in Figure 14 (a). This demonstrates a trend of increasing error with enhancement of distance between camera and object, with the channel flanks errors superimposed on this. Figure 14 also illustrate the error vector components for the same narrow strip (b,c,d). The error does not feature a particular trend in the X and Y dimensions. On the other hand, the elevation dimension Z error features a positive trend. This could because of (1) the strict correlation between elevation and distance from camera position given the positions from where images were collected (i.e. around the alluvial fan, but always at its front), or (2) uncertainties in the co-registration of data in this dimension (plausible given the steepness of the fan).

[Figure 13 about here.]

[Figure 14 about here.]

3.3. Laboratory experiment results

Table 6 presents the results of the flat surface laboratory experiment. The median distance error for a seven frames model is 0.171 mm, while for the use of all frames it is almost one half (0.091 mm). Thus, the accuracy of the results directly depends on the number of frames provided to the software that is, using more images translates into more accurate models. However, this pattern is non-linear. The addition of one frame considerably increases the model quality when using a reduced number of images, as demonstrated by the median distances errors of models with seven and eight frames (0.171 mm and 0.114 mm respectively). However, changing from eight to ten frames leads to almost identical median errors (0.114-0.115 mm). However, it would appear that outliers affect these results (e.g. the quantile .95 error reduces from 0.528 mm to 0.498 mm, 6). This tendency is also observed when moving from ten to thirteen to twenty images, where improvements are evident for outliers only (error reductions of 0.498 mm to 0.423 mm to 0.394 mm).

[Table 6 about here.]

4. Discussion

At close range, the application of SfM approaches provided very satisfying results in comparison to benchmark TLS data, as has been shown in previous studies (e.g. James and Robson (2012); Westoby et al. (2012); Fonstad et al. (2013)). Our results show that this approach can be extended to

acquire centimeter precision DTMs using smartphone imagery captured from a distance of up to 10 meters and a fully automated online analytical service. This precision is improved to a sub-centimeter scale by either: (1) the use of higher quality sensor or possibly more importantly higher resolution images; or (2) the use of more specific, commercial close range photogrammetry software for data generation. The approach here has three advantages over traditional photogrammetry and laser scanning techniques. First, and as with other SfM applications, images can be acquired with a reduction in the sensor, position and geometry constraints as compared with traditional photogrammetric methods. The result is more rapid data acquisition than both traditional photogrammetry and laser scanning methods, and the possibility of using a wider range of devices, including smartphones. The second is represented by the rapid data processing and very fast point cloud generation, using internet communication systems and fully automated services. This implies that digital elevation data as well as 3D images (e.g. Figure 3) can be generated in only a few minutes, without the need for post-processing of laser measurements or expert supervision. Finally, financial costs are close to zero, while laser scanning devices remain expensive pieces of equipment.

There is a co-registration uncertainty linked to the rotation, translation and scaling transformation (estimated using targets positions in Table 2) used to transform 123D Catch data into the TLS co-ordinate system. This was reflected in the change in distance errors after application of the ICP algorithm and is comparable to findings in similar studies (e.g. James and Robson

(2012); Westoby et al. (2012); Fonstad et al. (2013)), where it is indicated as an important source of uncertainty. In particular, Westoby et al. (2012) indicate how errors in the co-registration procedure can be linked to the manual identification of ground control points in both point cloud datasets and consequently impact upon the accuracy of the transformation matrix applied to the data. Unfortunately, contrariwise to other approaches, to compare 123D Catch and TLS data this manual identification is a necessity. Accordingly, the error estimated between the 123D Catch point clouds and the TLS data can be reduced with more precise co-ordinate transformation. The use of the Iterative Closest Point (ICP) algorithm (as demonstrated by James and Robson (2012)) confirmed this assumption. The ICP algorithm is expected to minimize the transformation error (but without changing the scale), so that distances between SfM and TSL data can be interpreted as approach-dependent errors. In that sense, the ICP is used to determine the source of the errors, for example registration issues versus random errors, by minimizing the effect of the former. Nevertheless, ICP can only be effective if there is an even distribution of data points. In fact, a surface with large voids would itself introduce errors in the surface matching and well-defined tiepoints can be useful. In the absence of a second dataset for comparison the ICP cannot be applied, but if the model needs to be geo-referenced anyway the use of GCP's (and targets) is advantageous and can provide a sufficiently precise registration for many geoscience applications. As mentioned previously, the ICP proved very efficient in the present study. However, there is a formal

distinction between SfM approaches. In PhotoModeler, TLS targets were used as ground control points, so that the registration errors are expected to be dominantly dependent on the bundle adjustment quality. Hence, 123D Catch surfaces are expected to benefit most from the ICP application. This is visually confirmed by the DEMs of difference in Figure 10: as expected, 123D Catch models benefit greatly from the ICP application, confirming that a non-negligible part of error could be linked to registration uncertainty possibly arising from the lower image resolution. ICP induced improvements are lower for PhotoModeler models. These results are consistent with values in Table 3, particularly when observing mean distance errors.

A fundamental control over DTM quality is exerted by the texture of the imagery. Areas of low image texture will yield lower density point clouds as well as more uncertain point qualities, because the image matching algorithm relies strictly on image texture. However, image texture is not only dependent upon surface characteristics but also image quality and resolution, particularly image scale. Autodesk currently limit image sizes to 3 MP, which can mean either increasing the image scale or number of images to achieve an acceptable model resolution. The practicality and speed of data generation associated with using a smartphone-123D Catch approach comes with the price of slightly poorer precision in the resulting surface data. However, the Autodesk 123D Catch smartphone “app” offers the possibility of near instantaneous acquisition of 3D data, based upon the analysis of smartphone-acquired images submitted to the free internet-based processing

Accepted Article

systems using wireless communication. This implies that a first indication on results can be made in near real time whilst still in the field. This is hugely beneficial as additional images can be acquired rapidly if either coverage or point density is inappropriate. However, this could only be done if good data network coverage is available, which could not always be the case in field applications.

Extension to an Alpine alluvial fan system demonstrated a linear degradation of precision with image scale, which is logical for a photograph-based approach and also observed in other studies (e.g. James and Robson (2012)). As explained by Fonstad et al. (2013), this is once again due to image texture; as viewing distances increase, the textural features required to calculate a point in that area could become insufficient. For that reason, and as with any photograph based technique, Fonstad et al. (2013) affirms that SfM is most effective in small study areas, where moderate-quality cameras have sufficient resolution to capture detailed texture. However, this is strictly dependent on the scale that provides the dominant texture in the images, as when capturing larger surface features, point matching could be more successful from longer distances than at a closer range. Nevertheless, we found that if the median errors for the riverbank and alluvial fan studies are scaled by the mean distance between generated points and the sensor, accuracies are approximately 1:625. Thus, despite precision degradation, intermediate scale models maintain the expected level of precision for their scale, as defined by image resolution. Improvements in alluvial fan models could be

achieved by taking imagery at a range of distances. Such a “multiscale” imaging solution suggests great promise. Fonstad et al. (2013) suggest that this procedure would reduce systematic distortions over large distances and still produce dense point clouds. However, this has not been verified and would require a more sophisticated strategy for data collection, involving advanced technology such as camera-equipped drones or kites for the alluvial fan case.

The flat surface experiment contributed to understanding the impact of varying the number of frames used by Autodesk 123D Catch. Additional frames generally increase model accuracy compared to a reduced number of images. Nevertheless, their benefit seem to affect outliers only when the number of images is already sufficient for a good representation of the surface of interest. This suggests that adding groups of images reduces the number of poorer quality points, a reasonable hypothesis given that point matching precision is dependent on image location. It appears that although mesh density increases and errors stabilize with increasing number of frames, denser meshes are not necessarily of better quality in terms of accuracies. Smaller image sets can still provide acceptable accuracies, provided their spatial distribution is adequate. The experiment featured an idealized case where (i) a frame can cover 100% of area of interest, and (ii) topography is completely absent. Nevertheless, the analysis proved useful in showing that unlike other traditional forms of photogrammetry, the number of images in an SfM application is absolutely critical. Nonetheless, as anticipated, the relation between

Accepted Article
results quality and number of frames is likely to differ from case to case.

SfM models are influenced equally by the difficulties of occlusion and separating vegetation and sediment cover typical of laser scanning systems. In SfM approaches, vegetation is important not only for occlusion effects, but also for degradation in the quality of surface reconstruction. This has been largely discussed by Westoby et al. (2012), with evidence of errors introduced by the presence of vegetation.

The investigation presented in this paper discussed a number of technical aspects. The mentioned approaches have advantages and disadvantages and their use depends on the specific application. Table 7 summarises these technical approaches. The list is not exhaustive, other approaches and solutions exist.

[Table 7 about here.]

Since 123D Catch is a process that automatically finds and matches common features between images, it is critical to capture imagery with the right characteristics. Autodesk provides useful guidelines and tutorials for potential users that can be found at <http://www.123dapp.com/>. Our recommendations would involve acquiring a multiscale image set which initially captures the whole site with a few frames at low resolution before obtaining a greater number of images at closer range to capture the required detail. It is important to collect images from varying directions, and for each 3D point to be identifiable in at least two or preferably three frames. The optimal

angle between each image will depend on the subject of interest. Since the features are extracted from the whole photographs, it is fundamental that (1) the whole scene is static, (2) light between the frames is consistent (no flash), and (3) images are not overexposed or underexposed as this can often confuse the feature matching process. If occlusions are present, increasing the number of photos is essential to ensure adequate image coverage. Moreover, increasing the number of photos may produce denser meshes and can improve model accuracies. The adequate number of images requires is likely to vary between case studies and cameras used and it is encouraged to investigate whether or not the derived data are sensitive to the number of images acquired and ensure that redundant frames are sufficient. It is advisable not to use zoom lenses, as it could be difficult to get accurate reconstruction with them. Finally, measuring transparent, reflective or homogeneous surfaces should be avoided because the features in this type of subjects are very difficult to match by the automatic feature matching process. If the generated model needs to be computed in real-world co-ordinates or co-registered with existing or future datasets, the use of coded and referenced targets can be advantageous. The marked objects need to be clearly identifiable and well represented on the images. Since 123D Catch operates with a fully automatic procedure, it is not possible to insert GCPs to directly compute the mesh in a real-world co-ordinate system. Hence, in order to minimize error introduced by the registration procedure, the manual identification of the targets needs to be eased by ensure a high density of points in target zones. If the size of

the object is so large to prevent the use of targets, easily identifiable natural features can be used. If a second dataset is available, after an approximative registration using common points or targets the application of the ICP algorithm can be employed in the attempt to minimize registration errors. If the aim is to monitor changes in an area or an object through time, it is suggested to isolate stable zones for the ICP application and further apply the transformation to the whole dataset. Finally, if a simply scaled model is desired, a known distance can be inserted as a reference in 123D Catch to scale the mesh. To summarise:

1. Plan camera survey, registration or method to introduce scale in advance
2. Capture the whole subject first, then the details
3. Ensure coverage is appropriate. Basic principle: ensure that every point on the object appears on at least three images, which are acquired from spatially different locations.
4. Static scene
5. Consistent light
6. Avoid overexposed and underexposed images
7. Consider occlusions (see 3, above)
8. Avoid transparent, reflective or homogeneous surfaces

The uploading of the images to Autodesk servers depends on their number and on the internet connection, but can be as rapid as a few minutes. The

Accepted Article

processing time is dependent not only on the number of images, but also on their texture and varies from 5 to 30 minutes for an initial result. It is then possible to generate a maximum-density mesh, which increases the processing time slightly. Overall, the generation of 3D data from raw data is considerably fast. Autodesk 123D Catch therefore provides an invaluable tool for geomorphologists, particularly in terms of cost, speed and ease of use.

5. Conclusion

In this contribution, the potential of a straightforward SfM approach using basic smartphone imaging technology and partial and fully automated data processing resources has been investigated. Two case studies have been chosen for the analysis, namely a riverbank as a close range example, and an Alpine alluvial fan as an intermediate distance case. With the use of TLS data used to produce a benchmark, the quality of DTMs generated using different sensors and processing methods has been assessed. Efforts were made to reduce the co-registration uncertainty in order to verify model accuracy. The quality of SfM results has proven to be related to image quality and photo scale. Nevertheless, the use of smartphone sensors allows practical and rapid geomorphic data collection and processing for all. Also, fully automated processing systems proved to be a valid resource for generating a spatial record of objects and DTM generation, especially at close range. These characteristics make fully automated SfM processing systems an ap-

pealing and promising approach for geoscience applications, deserving further efforts to investigate system performance in a wide range of environmental settings and terrain types.

Acknowledgements

This study was partially funded by the Herbette Foundation of the University of Lausanne, Switzerland and by the Swiss Geomorphological Society (www.geomorphology.ch). The paper benefitted substantially from critical but constructive comments by Mike James and a second anonymous reviewer, as well as from an Associate Editor.

References

- Alho, P., Kukko, A., Hyyppä, H., Kaartinen, H., Hyyppä, J., Jaakkola, A., 2009. Application of boat-based laser scanning for river survey. *Earth Surface Processes and Landforms* 34, 1831–1838.
- Barker, R., Dixon, L., Hooke, J., 1997. Use of terrestrial photogrammetry for monitoring and measuring bank erosion. *Earth Surface Processes and Landforms* 22, 1217–1227.
- Barrow, D. H., Tenenbaum, J. M., Bolles, R. C., Wolf, H. C., 1977. Parametric correspondence and chamfer matching: two new techniques for image matching. In: *Proc. 5th Int. Joint Conf. Artificial Intelligence*, Cambridge, MA, 659-663.

Brasington, J., Rumsby, B. T., McVey, R. A., 2000. Monitoring and modelling morphological change in a braided gravel-bed river using high resolution GPS-based survey. *Earth Surface Processes and Landforms* 25, 973–990.

Brasington, J., Smart, R. M. A., 2003. Close range digital photogrammetric analysis of experimental drainage basin evolution. *Earth Surface Processes and Landforms* 28, 231–247.

Butler, J., Lane, S. N., Chandler, J. H., Porfiri, E., 2002. Through-water close range digital photogrammetry in flume and field environments. *Photogrammetric Record* 17, 419–439.

Butler, J. B., Lane, S. N., Chandler, J. H., 1998. Assessment of DEM quality for characterizing surface roughness using close range digital photogrammetry. *Photogrammetric Record* 16, 271–291.

Chandler, J. H., 1999. Effective application of automated digital photogrammetry for geomorphological research. *Earth Surface Processes and Landforms* 24, 51–63.

Chandler, J. H., Cooper, M. A. R., Robson, S., 1990. Analytical aspects of small format surveys using oblique aerial surveys. *Journal of Photographic Science* 37, 235–240.

Chandler, J. H., Fryer, J., 2013. Autodesk 123d catch: how accurate it is? *Geomatics World* 2:21, 28–30.

Chandler, J. H., Fryer, J., Jack, A., 2005. Metric capabilities of low-cost digital cameras for close range surface measurement. *The Photogrammetric Record* 20:109, 12–26.

Chandler, J. H., Moore, R., 1989. Analytical photogrammetry: a method for monitoring slope instability. *Quarterly Journal of Engineering Geology and Hydrogeology* 22, 97–110.

EDF R&D, 2012. Cloudcompare (version 2.4) [gpl software]. Tech. rep., Telecom ParisTech.

Eos Systems Inc., 2012. Quantifying the accuracy of dense surface modeling within PhotoModeler Scanner. Tech. rep.

Fix, R. E., Burt, T. P., 1995. Global positioning system: an effective way to map a small area or catchment. *Earth Surface Processes and Landforms* 20, 817–827.

Fonstad, M. A., Dietrich, J. T., Courville, B. C., Carbonneau, P. E., 2013. Topographic structure from motion: a new development in photogrammetric measurements. *Earth Surface Processes and Landforms* Forthcoming.

Gomez-Gutierrez, A., Schnabel, S., Berenguer-Sempere, F., Lavado-Contador, F., Rubio-Delgado, J., 2014. Using 3D photo-reconstruction methods to estimate gully headcut erosion. *Catena* 120, 91–101.

Heritage, G., Hetherington, D., 2007. Towards a protocol for laser scanning in fluvial geomorphology. *Earth Surface Processes and Landforms* 32, 66–74.

Hodge, R., Brasington, J., Richards, K. S., 2009a. Analysing laser-scanned digital terrain models of gravel bed surfaces: linking morphology to sediment transport processes and hydraulics. *Sedimentology* 56, 2024–2043.

Hodge, R., Brasington, J., Richards, K. S., 2009b. In situ characterization of grain-scale fluvial morphology using terrestrial laser scanning. *Earth Surface Processes and Landforms* 34, 954–968.

Horn, B. K. P., 1987. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America* 4, 629–642.

James, M. R., Ilic, S., Ruzic, I., 2013. Measuring 3D coastal change with a digital camera. In: *Proceedings of Coastal Dynamics*.

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

Lane, S. N., 2000. The measurement of river channel morphology using digital photogrammetry. *Photogrammetric Record* 16, 937–961.

Lane, S. N., Hicks, D. M., Westaway, R. M., Duncan, M. J., 1999. Development of digital photogrammetry for automated monitoring of large gravel-bed rivers. *Image and Vision of Computing New Zealand*.

Lane, S. N., Reid, S. C., Westaway, R. M., Hicks, D. M., 2004. *Spatial Modelling of the Terrestrial Environment*. Wiley, Chichester, Ch. Remotely

- sensed topographic data for river channel research: the identification, explanation and management of error, pp. 157–174.
- Lane, S. N., Richards, K. S., Chandler, J. H., 1994. Developments in monitoring and modelling small-scale river bed topography. *Earth Surface Processes and Landforms* 19, 349–368.
- Mankoff, K. D., Russo, T. A., 2013. The Kinect: a low-cost, high-resolution, short-range 3D camera. *Earth Surface Processes and Landforms* Forthcoming.
- Nitsche, M., Turowski, M. J., Badoux, A., Rickenmann, D., Kohoutek, T. K., Pauli, M., Kirchner, J. W., 2013. Range imaging: a new method for high-resolution topographic measurements in small- and medium-scale field sites. *Earth Surface Processes and Landforms* Forthcoming.
- Sanz-Ablanedo, E., Chandler, J. H., Wackrow, R., 2012. Parametrising internal camera geometry with focusing distance. *Photogrammetric Record* 27(138), 210–226.
- Schaefer, M., Inkpen, R., 2010. Towards a protocol for laser scanning of rock surfaces. *Earth Surface Processes and Landforms* 35, 417–423.
- Westoby, M., Brasington, J., Glasser, N. F., Hambrey, M. J., Reynolds, M. J., 2012. 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* 179, 300–314.

Woodget, A. S., Carbonneau, P. E., F., V., Maddock, I., in press. Quantifying submerged fluvial topography using hyperspatial resolution UAS imagery and structure from motion photogrammetry. *Earth Surface Processes and Landforms* DOI: 10.1002/esp.3613.

Zhang, Z., 1992. Iterative point matching for registration of free-form curves. *International Journal of Computer Vision* 13, 119–152.

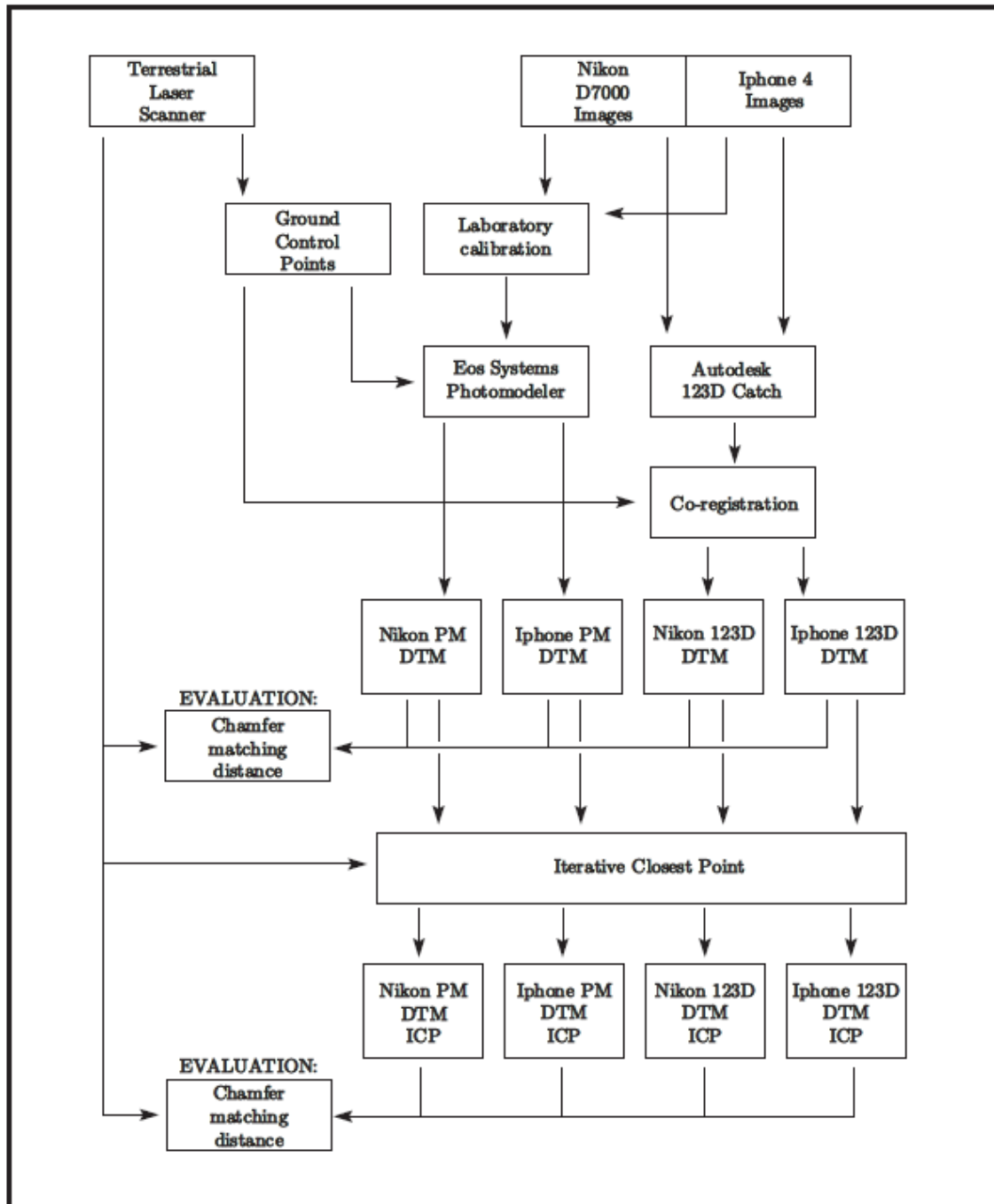


Figure 1: Methodology scheme: the comparison with benchmark (TLS) data is performed using the Chamfer distance twice (prior and after the application of the ICP algorithm).

ACC



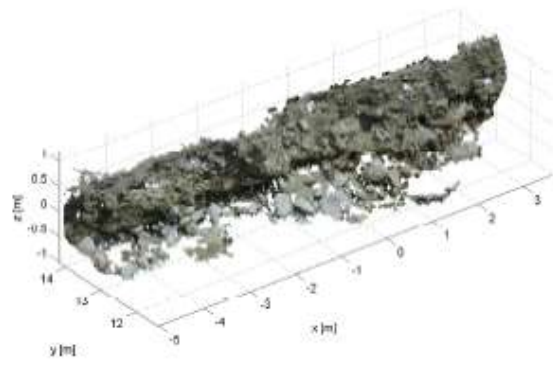
Figure 2: Riverbank of the Borgne d'Arolla, Valais, Switzerland. 10 targets printed on regular paper are disposed along it for referencing purposes.

Accepted

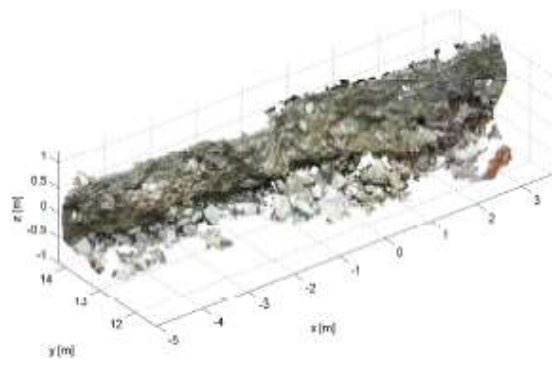


Figure 3: Examples of image-covered polygonal mesh (above) and its wireframe (below), generated using Iphone 4 images and 123D Catch.

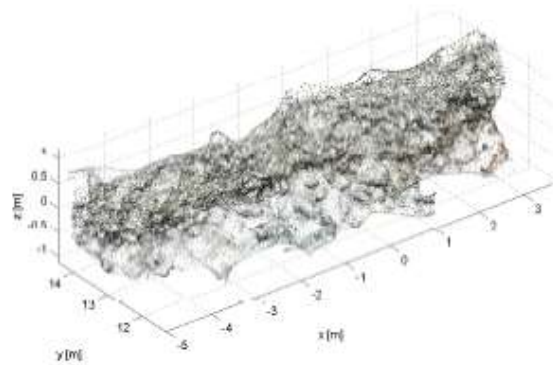
Accepted



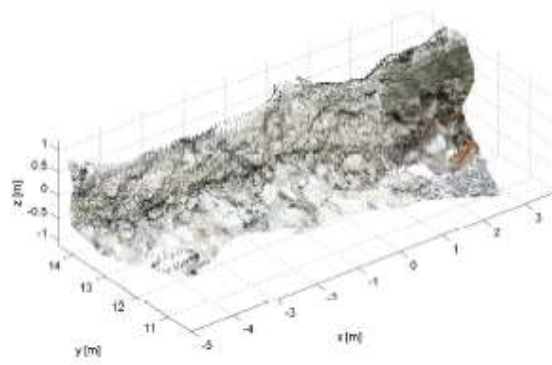
(a) Nikon D7000 Photomodeler



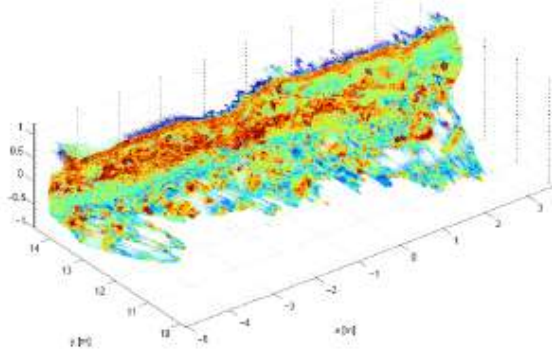
(c) Iphone 4 Photomodeler



(b) Nikon D7000 123D Catch



(d) Iphone 4 123D Catch



(e) Terrestrial Laser Scanner

Figure 4: Riverbank point clouds generated using different devices and softwares. Figure 4(d) presents a heterogeneous point density because of better image coverage in some zone, as discussed further in Section 3. Figure 4(e) is presented with laser reflectance as colours for visualization purposes.



Figure 5: Satarma alluvial fan, Valais, Switzerland.

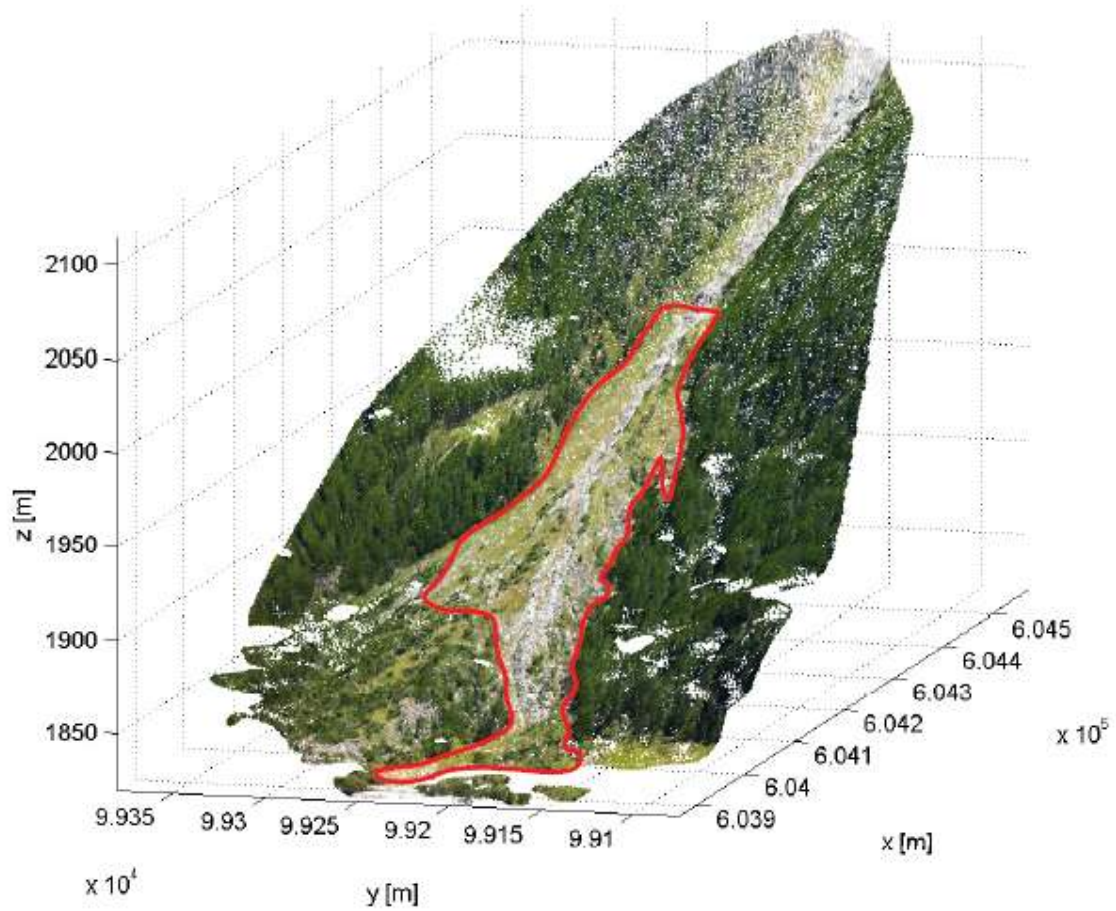


Figure 6: Satarma alluvial fan point cloud generated using an Iphone 4 and Autodesk 123D Catch. The highlighted zone is where the quantitative analysis is performed.

Accepted



Figure 7: Set up for the experiment: high quality print pasted on a wooded plank to be as at as possible.

Accept

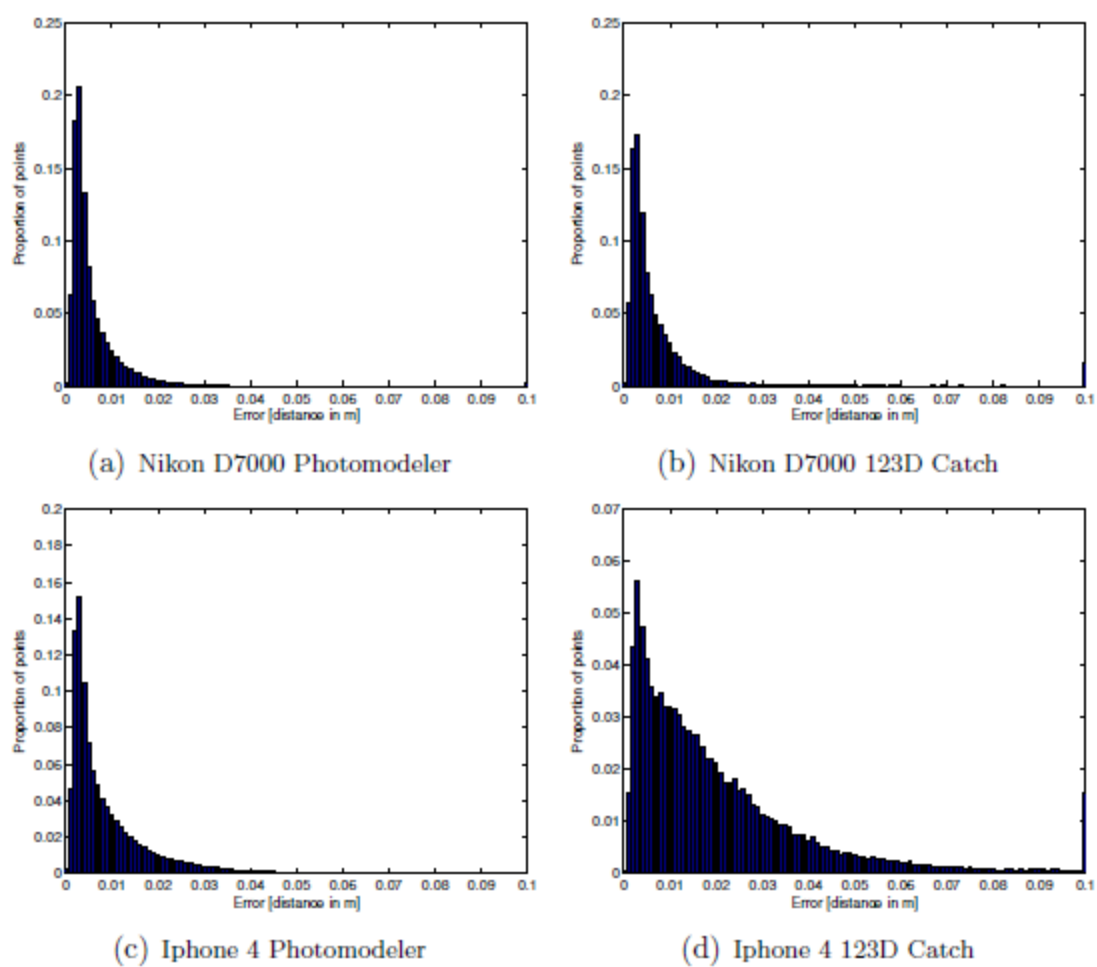
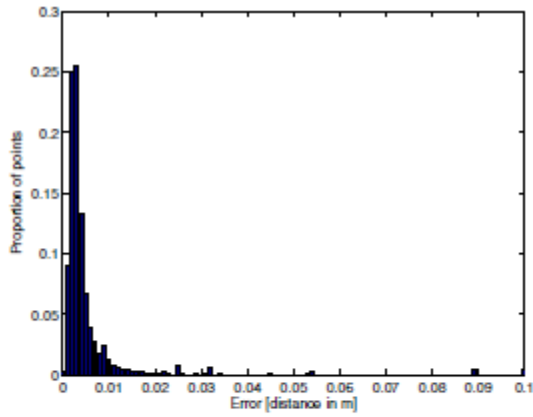
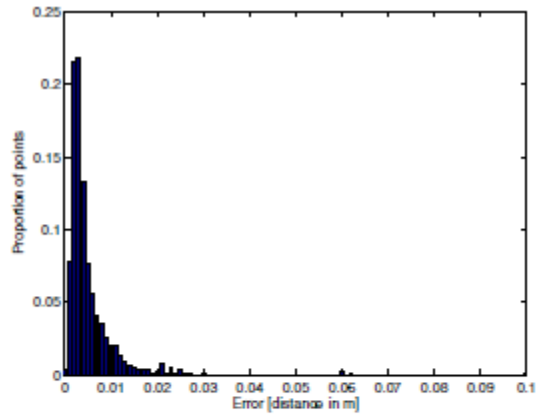


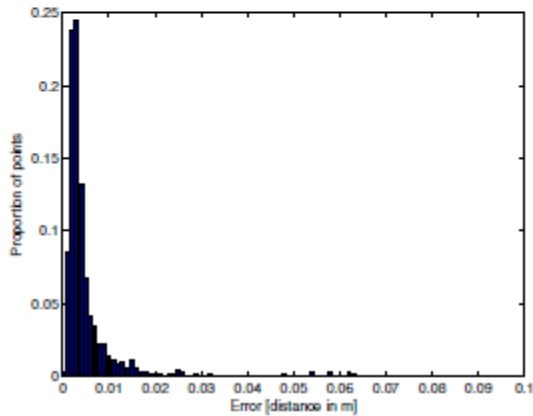
Figure 8: Histograms of distance errors distributions for the Borgne d'Arolla riverbank for each approach compared to TLS data



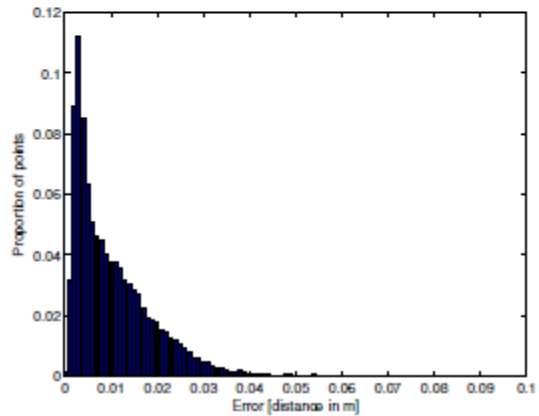
(a) Nikon D7000 Photomodeler - ICP



(b) Nikon D7000 123D Catch - ICP



(c) Iphone 4 Photomodeler - ICP



(d) Iphone 4 123D Catch - ICP

Figure 9: Histograms of distance errors distributions the Borgne d'Arolla riverbank for each approach compared to TSL data after the application of the ICP algorithm.

Accepted

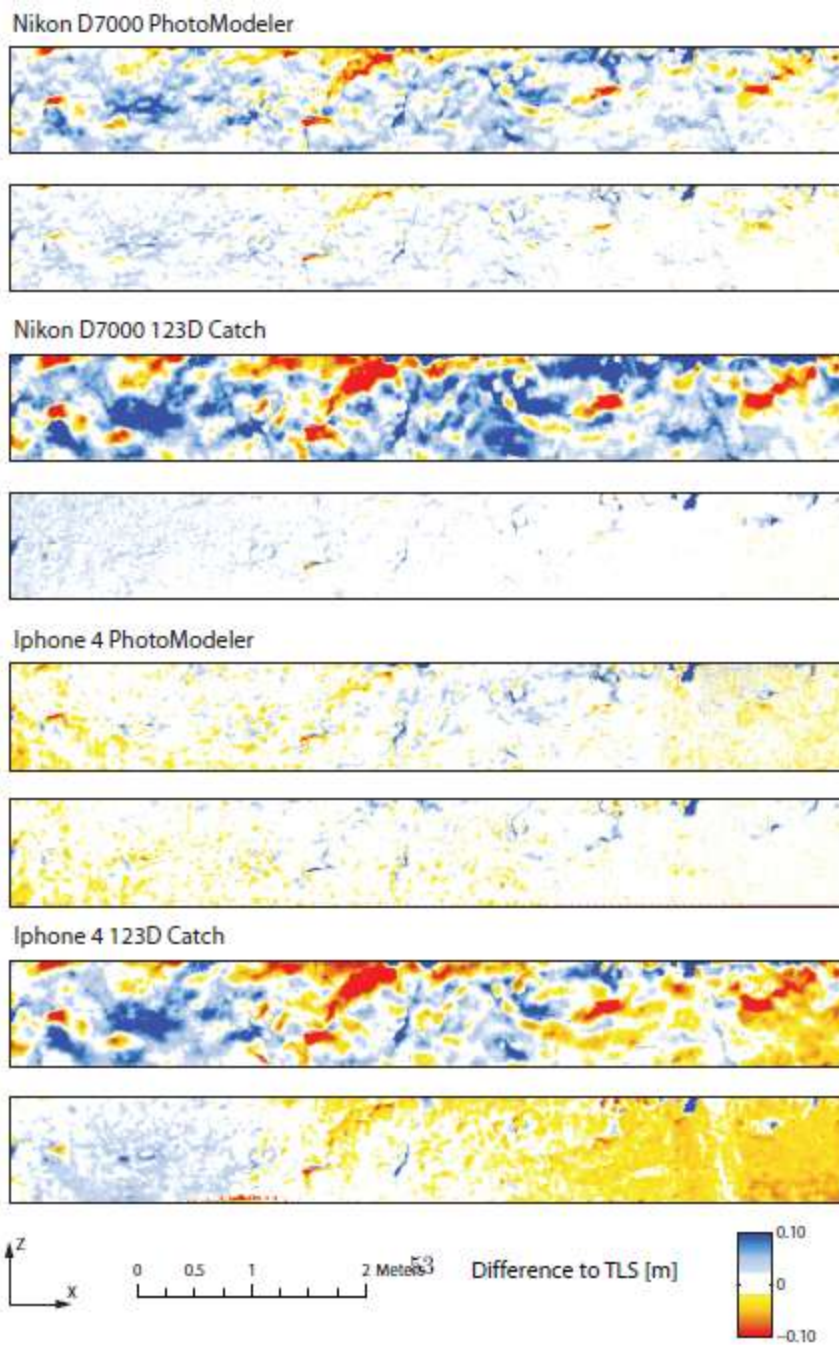


Figure 10: DEMs of Differences between SfM and PhotoModeler models and TLS, prior (above) and post (below) ICP application (0.005 m grid resolution).

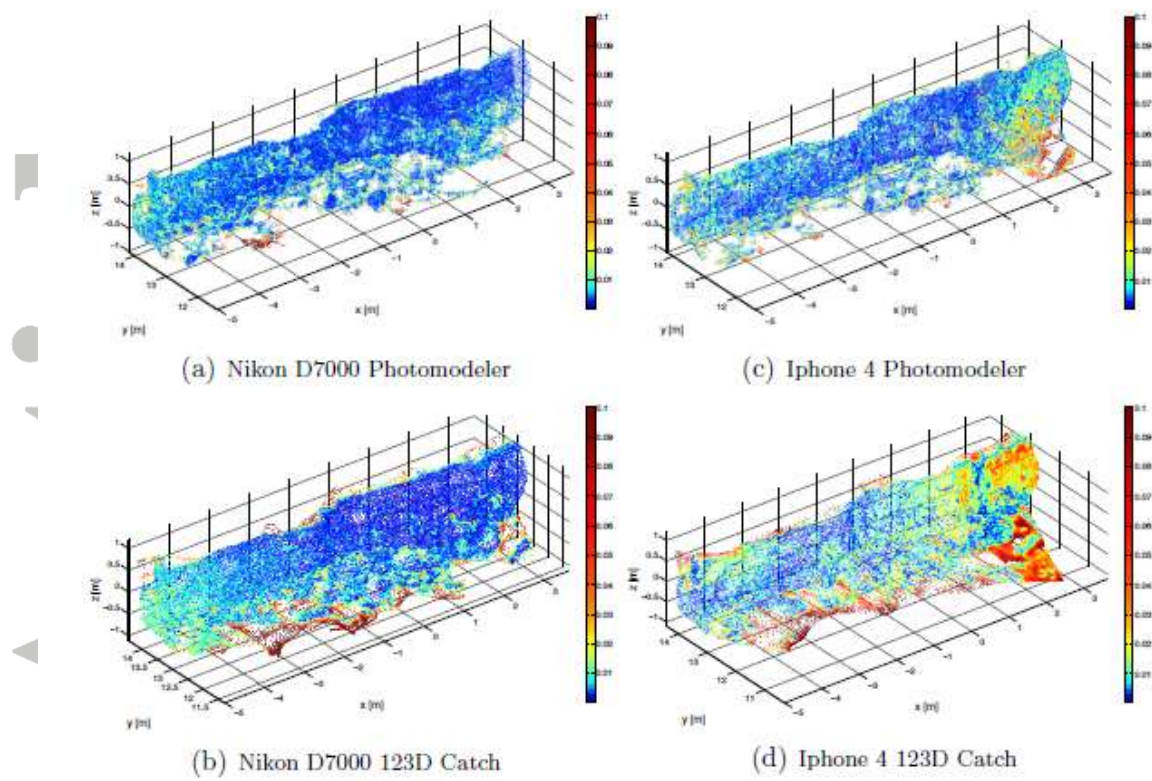


Figure 11: Point cloud errors for the Structure-from-Motion and PhotoModeler models and TLS comparison for the Borgne d'Arolla riverbank [m]

Accepted

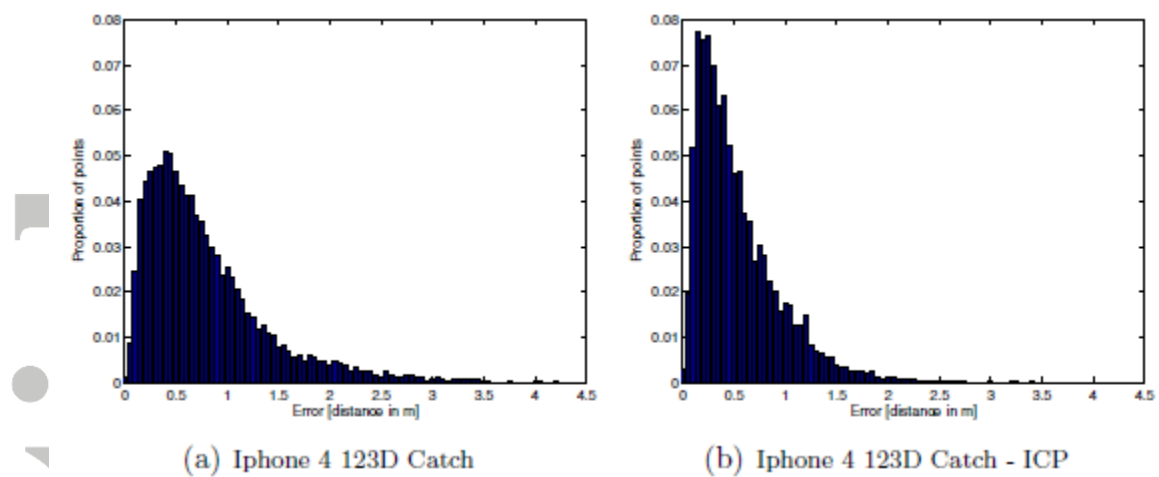


Figure 12: Histograms of distances distributions for Satarma alluvial fan SfM models compared to TLS data

Accepted Article

minis

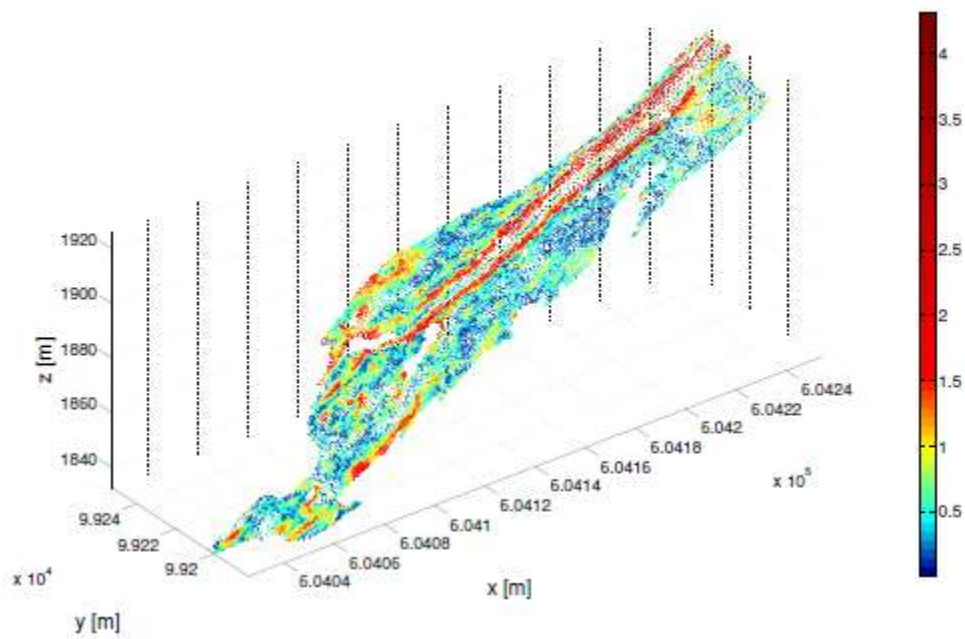


Figure 13: Point clouds errors for the Iphone 4 123D Catch model and TLS comparison for the main channel area of the Satarma alluvial fan [m].

Accepted

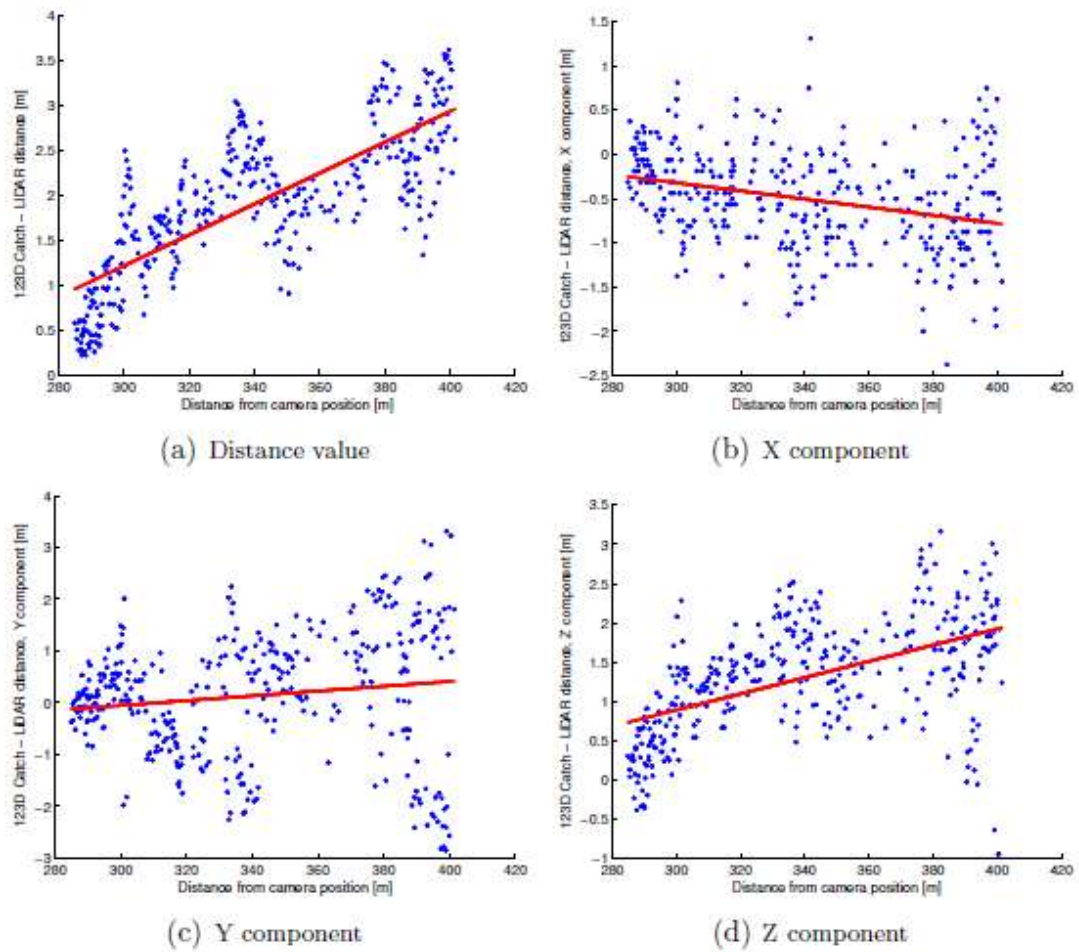


Figure 14: Distance errors between the 123D Catch model and the TLS data as a function of distance from the camera on the main channel flanks area of the Satarma alluvial fan.

Accept

Table 1: Examples of sensor and image processing alternatives for image-based high-resolution topographic data acquisition at close and intermediate scale with a hand-held camera.

Sensor	Smartphone ¹ High-quality sensor ²
Image processing approach	Internet-based SfM system ^a Locally-based SfM-MVS software ^b Close-range (oblique) photogrammetry ^c SfM-MVS photogrammetry ^d

Table 2: Root mean squared error of targets position transformation for Autodesk 123D Catch models [m].

Device	X	Y	Z
Nikon D7000	0.0043	0.0122	0.0035
Iphone 4	0.0146	0.0305	0.0018

Table 3: Distance errors between the riverbank SfM models and the TLS data for both direct registration and ICP algorithm [m].

DTM	Median ($Q_{.50}$)	$Q_{.75}$	$Q_{.90}$	$Q_{.95}$	$Q_{.99}$	Mean	Mode
N. D7000 - PM	0.0038	0.0070	0.0123	0.0170	0.0323	0.0061	0.0027
N. D7000 - PM - ICP	0.0031	0.0047	0.0089	0.0162	0.0717	0.0056	0.0027
N. D7000 - 123D C.	0.0044	0.0086	0.0159	0.0297	> 0.1	0.0090	0.0023
N. D7000 - 123D C. - ICP	0.0034	0.0059	0.0104	0.0141	0.0249	0.0053	0.0023
Iphone4 - PM	0.0053	0.0114	0.0203	0.0274	0.0455	0.0089	0.0028
Iphone4 - PM - ICP	0.0032	0.0051	0.0102	0.0152	0.0539	0.0054	0.0023
Iphone4 - 123D C.	0.0148	0.0274	0.0449	0.0615	> 0.1	0.0207	0.0044
Iphone4 - 123D C. - ICP	0.0079	0.0149	0.0225	0.0271	0.0382	0.0105	0.0030

Table 4: Root-Mean-Squared-Error between the riverbank and the TLS interpolated DEMs for both direct registration and ICP algorithm [m].

DEM	RMSE
N. D7000 - PM	0.0381
N. D7000 - PM - ICP	0.0197
N. D7000 - 123D C.	0.0647
N. D7000 - 123D C. - ICP	0.0168
Iphone4 - PM	0.0213
Iphone4 - PM - ICP	0.0174
Iphone4 - 123D C.	0.0502
Iphone4 - 123D C. - ICP	0.0306

Table 5: Distance errors between the alluvial fan 123D Catch model and the TLS for both direct registration and after application of the ICP algorithm [m].

DTM	Median ($Q_{.50}$)	$Q_{.75}$	$Q_{.90}$	$Q_{.95}$	$Q_{.99}$	Mean	Mode
Iphone4 - 123D C.	0.5998	0.9493	1.1333	1.5637	1.8882	0.6892	0.44
Iphone4 - 123D C. - ICP	0.4226	0.7282	1.0740	1.2912	1.7320	0.5263	0.15

Table 6: Distance errors between the flat surface models and best fitting plane [mm].

Number of images	Median ($Q_{.50}$)	$Q_{.75}$	$Q_{.95}$	Mean
7	0.171	0.279	0.529	0.207
8	0.114	0.193	0.528	0.168
10	0.115	0.211	0.498	0.166
13	0.128	0.215	0.423	0.161
20	0.117	0.203	0.394	0.149
30	0.090	0.176	0.415	0.136
40	0.095	0.181	0.397	0.134
53	0.091	0.167	0.392	0.137

Table 7: Technical aspects and some possible alternatives for the acquisition of high-resolution topographic and terrain data.

Technical aspect	Options	Main characteristics
Survey	Smartphone High-quality sensor TLS	Low cost, portable, potential wireless internet access, low quality High quality, portable, moderate cost, no internet access High precision, expensive, less portable
Image Processing	Internet-based SfM Local software SfM Traditional “stereo” photogrammetry SfM-MVS photogrammetry	Free, near real time, fully automatic, lower quality Mostly free and automatic, better quality expected High quality, subscription cost, expert knowledge as above but also greater automation and reliability
Co-registration	Scaling Tie points Targets + GCP ICP	Fast and easy, comparison with other datasets not possible Comparison with other datasets only, average precision Any coordinate system, high precision, not always possible Refinement of alignment, needs two co-registered datasets