

DERIVING 3D POINT CLOUDS FROM TERRESTRIAL PHOTOGRAPHS - COMPARISON OF DIFFERENT SENSORS AND SOFTWARE

Robert Niederheiser ^{a,*}, Martin Mokroš ^b, Julia Lange ^c, Helene Petschko ^d, Günther Prasicek ^e, Sander Oude Elberink ^f

^a Institute for Interdisciplinary Mountain Research - Austrian Academy of Sciences, Innsbruck, Austria,
robert.niederheiser@oeaw.ac.at

^b Department of Forest Management and Geodesy, Technical University in Zvolen, Slovak Republic

^c Department of Wind Energy, Technical University of Denmark, Copenhagen, Denmark

^d GIScience research group, Department of Geography, Friedrich Schiller University Jena, Germany

^e Department of Geography and Geology, University of Salzburg, Austria

^f Faculty of Geo-Information Science and Earth Observation, Universiteit Twente, Enschede, Netherlands

KEY WORDS: Close-range photogrammetry, structure from motion, dense matching, software

ABSTRACT:

Terrestrial photogrammetry nowadays offers a reasonably cheap, intuitive and effective approach to 3D-modelling. However, the important choice, which sensor and which software to use is not straight forward and needs consideration as the choice will have effects on the resulting 3D point cloud and its derivatives.

We compare five different sensors as well as four different state-of-the-art software packages for a single application, the modelling of a vegetated rock face. The five sensors represent different resolutions, sensor sizes and price segments of the cameras. The software packages used are: (1) Agisoft PhotoScan Pro (1.16), (2) Pix4D (2.0.89), (3) a combination of Visual SFM (V0.5.22) and SURE (1.2.0.286), and (4) MicMac (1.0). We took photos of a vegetated rock face from identical positions with all sensors. Then we compared the results of the different software packages regarding the ease of the workflow, visual appeal, similarity and quality of the point cloud.

While PhotoScan and Pix4D offer the user-friendliest workflows, they are also “black-box” programmes giving only little insight into their processing. Unsatisfying results may only be changed by modifying settings within a module. The combined workflow of Visual SFM, SURE and CloudCompare is just as simple but requires more user interaction. MicMac turned out to be the most challenging software as it is less user-friendly. However, MicMac offers the most possibilities to influence the processing workflow. The resulting point-clouds of PhotoScan and MicMac are the most appealing.

1. INTRODUCTION

There is high demand for 3D models of the Earth's surface to model Earth surface processes or depict surface structures such as vegetation accurately. A methodology widely used to derive dense 3D point clouds and digital surface models (DSMs) from airborne imagery taken with the help of unmanned aerial vehicles (UAVs) or from terrestrial imagery taken with consumer cameras is the Structure-from-Motion (SfM) approach combined with dense image matching (James and Robson, 2012; Remondino et al. 2014). SfM was developed in the field of computer vision for the automatic generation of 3D models from unordered datasets (Ullman, 1979; Brown and Lowe, 2005). Point clouds and DSMs derived by SfM and dense image matching are comparable to airborne and terrestrial laser scanning (LiDAR) and photogrammetry regarding data quality and achieved resolution (Westoby et al., 2012; Fonstad et al., 2013; Remondino et al. 2014). A big advantage of SfM is its simple data acquisition process and the possibility to derive a point cloud without knowing about the camera calibration or position in advance. This carries a huge potential in the use of old aerial photographs, analogue photos or photos taken with consumer cameras, e.g. provided through crowdsourcing from the internet (Snavely et al., 2008).

While terrestrial photogrammetry offers a reasonably cheap, intuitive and effective approach to 3D-modelling with many ready-to-use software packages available, only little is known about which sensor and which software will produce results that meet the user's requirements. The important choice of sensor and software is not straight forward and needs consideration as it will have effects on the resulting 3D point cloud and its derivatives.

Our objective was to compare five different sensors as well as four different state-of-the-art software packages regarding their ease of workflow, visual appeal, similarity and quality of the resulting point cloud. The five sensors represent different resolutions, sensor sizes and price segments of consumer cameras. The software packages used are: (1) Agisoft PhotoScan Pro (1.16), (2) Pix4D (2.0.89), (3) a combination of Visual SFM (V0.5.22) and SURE (1.2.0.286), and (4) MicMac (1.0). All software packages are available for the operating systems Microsoft Windows and Linux, except the software SURE, which used to be available for Linux systems until the release of version 2, and the software Pix4D, which is only developed for Windows and MAC OS. MAC OS is also supported by Agisoft PhotoScan and MicMac.

* Corresponding author

With this contribution we aim to support future users of terrestrial photogrammetry to make a well informed decision on which sensor (resolution of photos) and software to use for their requirements.

2. MATERIALS AND METHODS

We used cameras with different sensor sizes and lenses representing different price segments of consumer cameras and availability: Canon5D, Konica KD-310Z, GoPro Hero3+ Black Edition, Nikon D3000 and the smart-phone LG-D331. The Canon5D and Nikon D3000 are digital single-lens reflex (DSLR) cameras. These two cameras offer many possibilities in manual adjustment of camera settings. In contrast to the two DSLR cameras the GoPro and the smart-phone camera offer hardly any options to manually adjust the image quality but have a similarly high image resolution. They are made to work fully automatically. The Konica camera is the smallest and oldest camera used in this comparison. It offers the least resolution and also hardly any manual options. Table 1 shows the sensor details and lens settings for each camera.

Camera model	Sensor size (mm)	Resolution (pixel)	Focal length (mm)	Aperture
Canon5D	35.8 x 23.8	4368 x 2912	24	F/20
Nikon D3000	23.6 x 15.8	3872 x 2592	18	F/18
GoPro Hero3+ Black Edition	~ 6.17 x 4.55	4000 x 3000	3	F/2.8
LG-D331	unknown	3264 x 2448	4	F/2.4
Konica KD-310Z	~ 7.11 x 5.33	2048 x 1536	8	F/4.7

Table 1. Camera details

The object of interest for this paper was an easily accessible small and steep rock slope located next to a road and partially covered with vegetation (Figure 1).



Figure 1. Rocky slope with vegetation and wooden scale bar

Before describing the workflow for each individual software package in the following subsections, we summarize the main steps for dense point cloud generation independent from software package. Photos were taken from the same positions with all cameras. From 12 positions along the front of the vegetated rock slope a total of 18 images were captured with each camera, 15 in landscape and 3 in portrait format. The vertical images improve the results of the self-calibration process. The GoPro continuously captured 60 images while

walking a semi-circle along the front of the rock slope (Figure 2).



Figure 2. Example for a dense point cloud and the camera positions that were used for all cameras

We used the default parameters of each software package to compare their standard outputs. While these standard parameters are by no means optimized for the chosen setting and object of interest, this approach enables us to introduce a beginner to highlights and pitfalls in state of the art photogrammetry and to assess the ease of the standard processing workflow.

Within the software Cloud Compare we manually scaled the point clouds resulting from all software packages except MicMac to the same scale and centred them in the same arbitrary coordinate system around a wooden scale bar that we placed on the slope (Figure 1). Due to different absolute image sizes and fields of view the point clouds had different dimensions and fuzzy edges. Therefore, all point clouds were clipped to the same extent. All point clouds were registered using the wooden scale bar. We extracted horizontal profiles at the same location from all point clouds to compare their geometric similarity.

2.1 Agisoft PhotoScan workflow

Agisoft PhotoScan is a commercial software package for photogrammetric processing of digital images. Main outputs are dense point clouds, meshes, orthomosaics and digital elevation models. Only two major steps have to be taken to generate dense point clouds. The first step is the alignment of the images. One of the alignment settings is accuracy, which handles the image resolution for processing. High accuracy means full image resolution and further options are downscaling by a factor of 4 (medium) or 16 (low). The alignment can be done in three ways: Disabled, generic or referenced. The first option consumes the most memory because it searches tie points in every possible photo pair. The generic option searches tie points in two steps, of which the first is to make pairs or groups with subsampled photos and the second step is to merge these groups using photos with the desired accuracy. The third possible option uses the geolocation of the images as a first step and subsequently searches tie points based on the proximity of neighbouring images. We used manual pair preselection and the full resolution images. These settings work for all cameras. The second step after aligning images is to generate dense point clouds based on the afore mentioned alignment. These dense point clouds can be computed in five different qualities which mainly address the number of points in the results. Higher quality leads to a higher number of points. We used the high

quality setting for all cameras. The resulting point clouds are then scaled and merged within Cloud Compare.

2.2 Pix4D workflow

Pix4D is a proprietary software package for the generation of point clouds, models and orthomosaics from photos captured via a hand-held camera or UAV (Pix4D 2016). The workflow is similar to Agisoft PhotoScan. The first step is called initial processing and handles to image alignment. Different image resolutions may be used (2, 1, 1/2, 1/4, 1/8). In this study we used the full resolution of all images. The second step is the point cloud generation itself. Also, in this second step the original resolution of the images could be used but a resampling to half of the original resolution is recommended (Pix4D 2016). Other options are a quarter and an eighth of the original size of the images. Furthermore, the desired point density of the final point cloud can be set (i.e. low, optimal or high). For this study we chose the option of “optimal density” which includes every fourth pixel in 3D point cloud generation. For each input image a final point cloud is automatically saved to *.las format. The resulting point clouds are then scaled and merged within Cloud Compare.

2.3 Visual SFM and SURE workflow

The third workflow we tested was generating a dense point cloud using the freely available software packages Visual SFM (Wu et al., 2011, Wu, 2013), SURE (Rothermel et al., 2012) and Cloud Compare (version 2.6.0, 2015). The consecutive usage of all three software packages was necessary to obtain a dense point cloud. In Visual SFM an incremental SfM approach is used (Wu, 2013). The image matching identifying the tie points is executed by a feature detection and full pairwise image matching algorithm. Furthermore, a sparse reconstruction using a multicore bundle adjustment is performed to compute a sparse point cloud. The result is an NView Match. In SURE this match is loaded to prepare the dense cloud by a multi-view stereo (MVS) method based on the semi-global matching method merging the redundant depth estimations across single stereo models (Rothermel et al., 2012). The resulting point clouds are scaled, merged, scanned for duplicate points, and filtered within Cloud Compare. The identified duplicate points with a minimum distance between points of 0.001 m and statistical outliers were removed.

2.4 MicMac Workflow

The MicMac workflow requires the most knowledge about what is actually done during photogrammetric dense matching. What is commonly known as MicMac is a suite of software packages that need to be called in sequence. A good starting point may be the tutorial by Fried (2014) and the MicMac documentation (Institut national de l'information géographique et forestière 2016). First, tie-points are computed using Tapioca, adopting the SIFT++ algorithm. In our case tie-points are calculated between all available image pairs using a multi-scale approach. After the tie-points have been identified the inner and outer orientations of the camera positions are computed with Tapas. The result can be validated by visual inspection of the sparse point cloud and the camera positions with AperiCloud. In order to get quantitative results, the point clouds need to be scaled by selecting ground-control-points. Because accurate GPS data was not available we chose three points on the scale bar that was placed on the slope and is visible in several images, and gave them arbitrary coordinates in the metric system. With the

coordinate transformation completed the dense point clouds can be calculated using Malt and the results can be converted to point clouds running Nuage2Ply. In our case we computed dense point clouds for specific master images. Additionally, masks on these master images were used to define the area of interest and to mask sky and distant objects. This results in several point-clouds from different perspectives for each camera within a common reference system. In turn, these point-clouds can be loaded, merged, cleaned and filtered in CloudCompare, yielding one single dense point-cloud. Duplicate points with a minimum distance of 0.001 m and statistical outliers were removed.

3. RESULTS AND DISCUSSION

After clipping the point clouds to the same extent, the absolute number of resulting points representing a surface of approximately 60 m² ranged from 232.946 (Nikon D3000 and Agisoft PhotoScan) to 23.342.430 (Canon5D and SURE). The densities of the point clouds depend very much on the image resolution. The more pixels, the higher the point density and the resulting total point count. However, the distribution of the points in the resulting models is not homogeneous. Visible rock surfaces show higher point densities and are depicted relatively completely. Vegetation, on the other hand, is very difficult to reproduce by photogrammetric means and the software packages used. Vegetated areas show holes and lower point densities (Figure 3 and 4).



Figure 3. Dense point cloud for Canon5D, computed with MicMac. Vegetated areas show distinct holes.

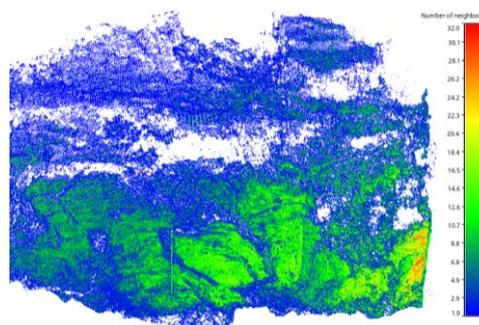


Figure 4. Number of neighbours within a sphere of 1 cm radius for each point as in Figure 3. Rock surfaces show a higher density (greens to red) than vegetated areas (blues).

The four workflows and 5 cameras resulted in 18 point clouds (exported with identical settings from Cloud Compare: Figures 5 – 22). The three workflows using Agisoft PhotoScan, Pix4D and SURE were successful for all five cameras used. MicMac in its standard configuration failed for the GoPro and the smartphone camera due to strong distortion and unknown lens parameters, respectively. However, there are other



Figure 5. Agisoft PhotoScan Canon5D dense point cloud



Figure 10. Pix4D Canon5D dense point cloud



Figure 6. Agisoft PhotoScan GoPro dense point cloud



Figure 11. Pix4D GoPro dense point cloud



Figure 7. Agisoft PhotoScan Konica dense point cloud



Figure 12. Pix4D Konica dense point cloud



Figure 8. Agisoft PhotoScan Smartphone dense point cloud



Figure 13. Pix4D Smartphone dense point cloud



Figure 9. Agisoft PhotoScan Nikon dense point cloud



Figure 14. Pix4D Nikon dense point cloud



Figure 15. SURE Canon5D dense point cloud



Figure 20. MicMac Canon5D dense point cloud



Figure 16. SURE GoPro dense point cloud



Figure 21. MicMac Konica dense point cloud



Figure 17. SURE Konica dense point cloud



Figure 22. MicMac Nikon dense point cloud



Figure 18. SURE Smartphone dense point cloud



Figure 19. SURE Nikon dense point cloud

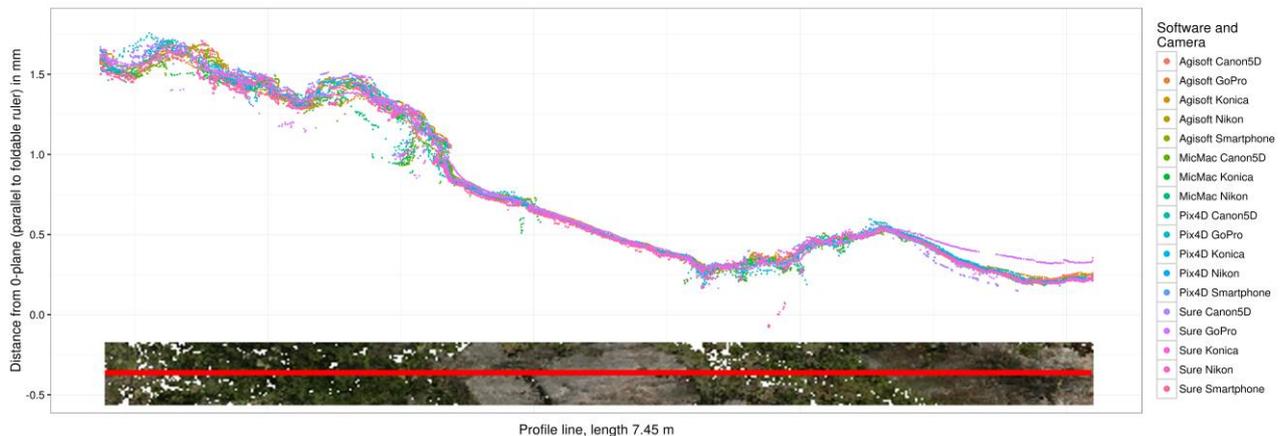


Figure 23. Scatter plot: Points along a horizontal profile for all dense point clouds resulting from all sensors and software packages. Below the scatter plot the profile (red) is presented on an exemplary point cloud covering vegetation and bare rock surfaces.

examples that use the GoPro camera with MicMac successfully (e.g. Zhuo et al. 2015, or the MicMac forum at <http://forum-micmac.forumprod.com/>). The possibility to manually add lens-specific calibration parameters is given in the software.

The profiles in Figure 23 depict the distance of each point from a plane parallel to the foldable ruler that was placed on the slope. The profiles from all point clouds in are smooth with a maximum offset between profiles of less than 0.1 mm along the rock surface. In vegetated areas the maximum offset reaches 0.45 mm. Profiles from the Canon5D show variations between results derived from different software packages (Figure 25). The rock surface in the centre of the profiles is very smooth with only little noise. The vegetation is noisier, especially for MicMac and SURE. The Agisoft and Pix4D results are very similar in general. If only profiles generated with the same software package are compared (e.g. Agisoft, Figure 26), the differences between sensors can be analysed. For Agisoft, no big difference between sensors is evident. Considering the vegetation, the results from the GoPro camera are not catching small changes and are smoother. This could be caused by the strong distortions of the fish-eye lens, by darker images or internal image processing from the GoPro camera.

All point clouds are visually appealing (Figures 5 – 22). Naturally, depending on the camera, RGB values vary. Comparing the point clouds visually, the Agisoft PhotoScan and MicMac results seem to be the most complete. Unlike the rocky textures, which were reliably depicted with all software products, the vegetation proved difficult to reconstruct. AgiSoft PhotoScan seemingly represents the vegetation best. However, since PhotoScan is a proprietary software, details on the dense point cloud construction remain confidential.

Dense point cloud could be generated with all sensors. Except MicMac, all software packages were able to reconstruct dense point clouds with all sensors. The professional cameras Canon5D and Nikon D3000 produced dense point clouds with highest point counts and highest densities. The GoPro results made a smooth impression, but the original images received a lot of camera-internal processing and compensation. Thus, the texture in the photos was very smooth and the software packages had difficulties to find distinct features for image matching (Figure 24).



Figure 24. Details of GoPro (left) and Canon 5D (right) images

The Windows packages of Agisoft PhotoScan, Pix4D, SURE and CloudCompare were the easiest to install. Visual SFM and MicMac need to be compiled on Windows and Linux, and also require some dependencies.

Of the four workflows, the Agisoft PhotoScan and Pix4D procedures were the easiest to carry out and required the least prior knowledge. However, both packages do not offer any detailed insights into their algorithms and options for adjustment. MicMac on the other hand required the most knowledge about the photogrammetric workflow and the handling of single processing steps in order to achieve decent results. Due to its open source policy MicMac its code is available and offers the most options for optimization. For this, a lot of knowledge about the single processing steps is needed. The Visual SFM/SURE workflow is equally easy to follow as the PhotoScan and Pix4D workflows as it also provides a graphical user interface (GUI). However, its usage is not as straight forward. Visual SFM offers very powerful tools (e.g. setting of ground control points), but details on most of the more advanced tools are missing in the documentation. MicMac also offers a GUI but this was not tested here.

4. CONCLUSIONS

The aim of this experiment was to evaluate the ease and performance of dense point cloud production with different cameras and software packages. For beginners in this field we intended to provide a guide in choosing a first hard- and software setup and workflow to be adapted for specialised applications.

For all four sensors with all their different specifications dense point clouds could be computed with nearly all software packages (exception MicMac with GoPro and LG Smartphone). However, the results showed a wide range of number of matched points and quality of represented vegetation. The extracted profiles demonstrated a strong geometric similarity of all point clouds. For shape-invariant objects with a detailed

texture all software packages with all camera sensors seem to produce reliable results. Vegetation that moves in the wind and has a lot of shadowy parts proves to be difficult to construct. For these areas the comparison of the extracted profiles showed a high variability between the different software packages and sensors.

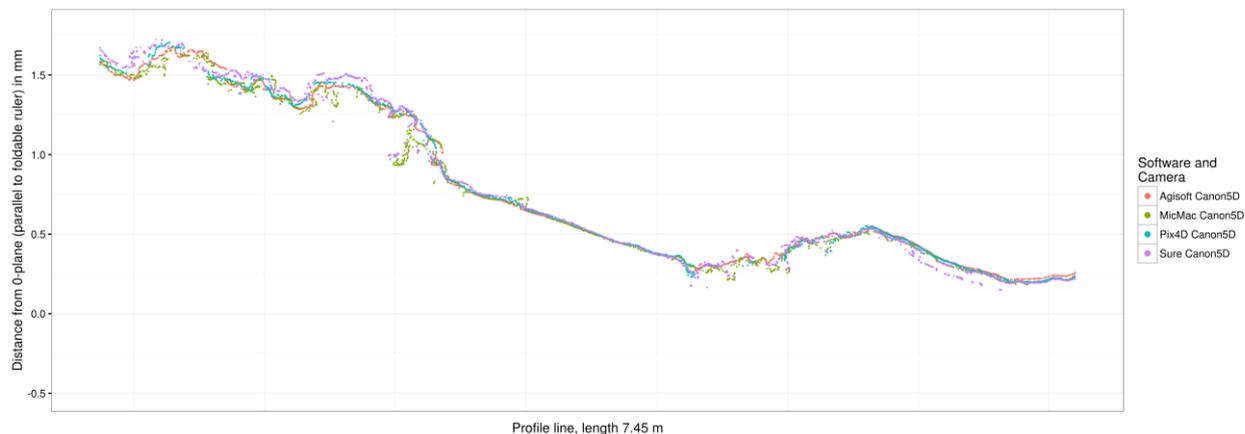


Figure 25. Profile lines of the point clouds generated by the different softwares using only the Canon5D images as input.

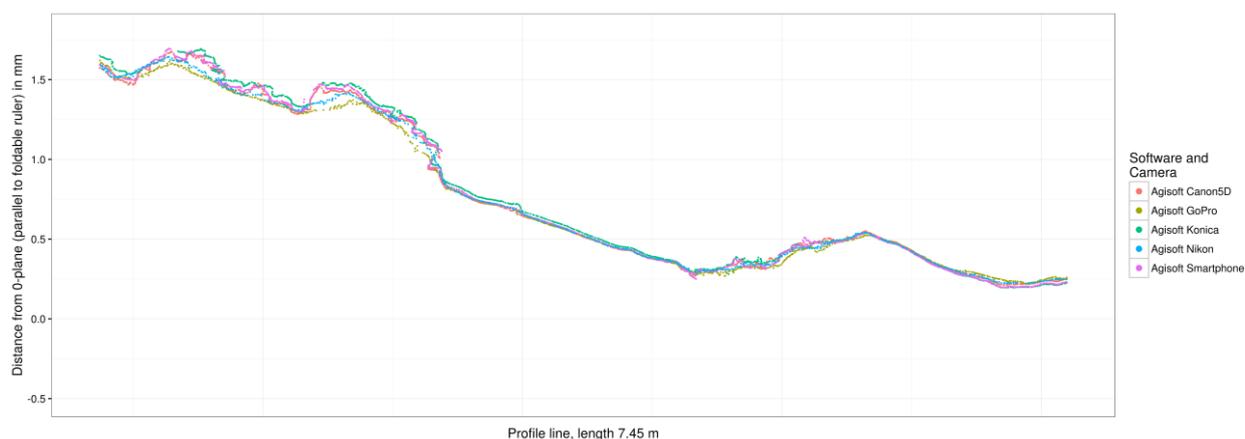


Figure 26. Profile lines produced by Agisoft PhotoScan for different sensors.

We conclude that, given the overall good visual appeal of all computed point clouds, but the higher completeness, point density and number of points of point clouds resulting from Agisoft PhotoScan and MicMac, these two software packages can be recommended for representing vegetation, in particular. Especially for beginners, some investment in the proprietary Agisoft package might be worth the simple and user-friendly handling. However, depending on the object of interest (rock or vegetation) the results of Visual SFM/SURE and Pix4D were comparable. In case of Pix4D we used recommended settings (half the image size and “optimal” point density). Nevertheless, we also tried to use the highest settings possible (original image size and high point density) to see whether the results improve. Figure 27 shows two resulting point clouds from Canon5D images. The number of points changes from 702.721 to 14.173.630. Also, the vegetated areas have fewer holes if highest quality settings are used. As a consequence, the

computing time becomes much longer but the results increase a lot.

The DSLR cameras, the Canon 5D and the Nikon D3000, produced the best point clouds in conjunction with the good input images. This is due to the possibility of disabling automatic options in the cameras themselves, such as the auto-focus, image stabilisation, and the good sensor qualities, e.g. sensor sizes and pixel counts, therefore reducing the degrees of freedom during self-calibration. However, DSLR cameras are expensive and the smartphone camera and even the cheap Konica camera produced accurate results. There are a number of bridge cameras on the market that fill the niche between the DSLRs and small hand-held digital cameras. They offer good quality lenses and large sensors. Therefore, they may offer a cheaper alternative to the expensive DSLRs.



Figure 27. Comparison of two point clouds from Pix4D and Canon5D with recommended settings (left) and highest settings (right).

ACKNOWLEDGEMENTS

We had the wonderful opportunity to start this research during the Innsbruck Summer School of Alpine Research on Close Range Sensing Techniques in Alpine Terrain in July 2015.

MicMac results were processed at the high performance cluster at the University Innsbruck.

REFERENCES

- AgiSoft PhotoScan Professional (version 1.16), 2015. Retrieved from <http://www.agisoft.com>
- Brown, M., Lowe, D.G., 2005. Unsupervised 3D object recognition and reconstruction in unordered datasets. In: *Fifth International Conference on 3D Digital Imaging and Modeling*, 2005. IEEE Comput. Soc., Ottawa, Canada, pp. 56–63.
- CloudCompare (version 2.6.3) [GPL software], 2015. Retrieved from <http://www.cloudcompare.org/>
- Fonstad, M.A., Dietrich, J.T., Courville, B.C., Jensen, J.L., Carbonneau, P.E., 2013. Topographic structure from motion: a new development in photogrammetric measurement: TOPOGRAPHIC STRUCTURE FROM MOTION. In: *Earth Surface Processes and Landforms* 38, pp. 421–430. doi:10.1002/esp.3366
- Fried, J.-M., 2015. Photogrammetric 3D structure reconstruction using Micmac. Available at <http://jmfriedt.free.fr/> (7.4.2016).
- Institut national de l'information géographique et forestière (IGN) (2016): MicMac, Aperro, Pastis and Other Beverages in a Nutshell. Available via Mercurial repository <https://culture3d:culture3d@geoportail.forge.ign.fr/hg/culture3d> Accessed 6.1.2016.
- James, M.R., Robson, S., 2012. Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. In: *Journal of Geophysical Research* 117. doi:10.1029/2011JF002289
- MicMac (1.0) [CeCILL-B software], Institut national de l'information géographique et forestière (IGN), 2016. Retrieved from Mercurial repository <https://culture3d:culture3d@geoportail.forge.ign.fr/hg/culture3d>
- Pix4D mapper (version 2.0.89), 2015. Retrieved from <https://www.pix4d.com/>
- Pix4Dmapper Software manual, 2016. Available at https://s3.amazonaws.com/mics.pix4d.com/manual_pdf/manual_2_1_march2016.pdf (12.04.2016).
- Remondino, F., Spera, M.G., Nocerino, E., Menna, F., Nex, F., 2014. State of the art in high density image matching. In: *The Photogrammetric Record* 29, pp. 144–166. doi:10.1111/phor.12063
- Rothermel, M., Wenzel, K., Fritsch, D., Haala, N., 2012. SURE: Photogrammetric surface reconstruction from imagery. In: *Proceedings LC3D Workshop*. Presented at the LC3D Workshop, Berlin.
- Snively, N., Seitz, S.M., Szeliski, R., 2008. Modeling the World from Internet Photo Collections. In: *International Journal of Computer Vision* 80, pp. 189–210. doi:10.1007/s11263-007-0107-3
- Ullman, S., 1979. The interpretation of structure from motion. In: *Proceedings of the Royal Society London, Ser. B*, pp. 405–426.
- Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., Reynolds, J.M., 2012. “Structure-from-Motion” photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* 179, 300–314. doi:10.1016/j.geomorph.2012.08.021
- Wu, C., Agarwal, S., Curless, B., Seitz, S.M., 2011. Multicore bundle adjustment. In: *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011 IEEE. IEEE, pp. 3057–3064.
- Wu, C., 2013. Towards linear-time incremental structure from motion. In: *International Conference on 3D Vision-3DV* 2013. IEEE, pp. 127–134.
- Zhuo, X., Kurz, F., Reinartz, P., 2015. Fusion of multi-view and multi-scale aerial imagery for real-time situation awareness applications. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Toronto, Canada, Vol. XL-1/W4, pp. 201–206.