

PAPER • OPEN ACCESS

A Structure-from-Motion Pipeline for Generating Digital Elevation Models for Surface-Runoff Analysis

To cite this article: Jhacson Meza *et al* 2019 *J. Phys.: Conf. Ser.* **1247** 012039

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

A Structure-from-Motion Pipeline for Generating Digital Elevation Models for Surface-Runoff Analysis

Jhacson Meza¹, Andres G. Marrugo¹, Gabriel Ospina¹,
Milton Guerrero¹, and Lenny A. Romero²

¹Facultad de Ingenieria, Universidad Tecnologica de Bolivar, Cartagena, Colombia.

²Facultad de Ciencias Basicas, Universidad Tecnologica de Bolivar, Cartagena, Colombia.

E-mail: jhacsonmeza@outlook.com

Abstract. Digital Elevation Models (DEMs) are used to derive information from the morphology of a land. The topographic attributes obtained from the DEM data allow the construction of watershed delineation useful for predicting the behavior of systems and for studying hydrological processes. Imagery acquired from Unmanned Aerial Vehicles (UAVs) and 3D photogrammetry techniques offer cost-effective advantages over other remote sensing methods such as LIDAR or RADAR. In particular, a high spatial resolution for measuring the terrain microtopography. In this work, we propose a Structure from Motion (SfM) pipeline using UAVs for generating high-resolution, high-quality DEMs for developing a rainfall-runoff model to study flood areas. SfM is a computer vision technique that simultaneously estimates the 3D coordinates of a scene and the pose of a camera that moves around it. The result is a 3D point cloud which we process to obtain a georeference model from the GPS information of the camera and ground control points. The pipeline is based on open source software OpenSfM and OpenDroneMap. Encouraging experimental results on a test land show that the produced DEMs meet the metrological requirements for developing a surface-runoff model.

1. Introduction

Advances in survey technology have improved acquisition of Digital Elevation Models (DEMs) datasets, and in turn, these have improved its quality and its spatial resolution. Aerial LiDAR (Laser Imaging Detection and Ranging) and Terrestrial Laser Scanning (TLS) has increased the spatial coverage and density of traditional techniques such as Differential Global Positioning Systems (dGPS) and Total Stations (TS) [1]. A disadvantage of LiDAR and TLS is the investment in personal time and their high-cost [2], and while their spatial resolution is better than traditional techniques, it is still not enough to measure the characteristics of the microtopography of a terrain [3, 4].

Structure from Motion (SfM) is a low-cost photogrammetry technique which produces high-resolution DEM datasets without the need for expensive equipment, only by using a camera [5]. The easy accessibility to Unmanned Aerial Vehicles (UAVs) for image acquisition and the development of SfM has democratized the 3D topography survey [1]. SfM estimates the camera pose, given by a translation vector and a rotation matrix, and the 3D coordinates of some points in the scene using a dataset of overlapping images. From this process we obtain a sparse point cloud, that later using Multi-View Stereo (MVS) [6] we can increase the number of points. In other words, MVS allows us to densify the point cloud obtained with SfM.



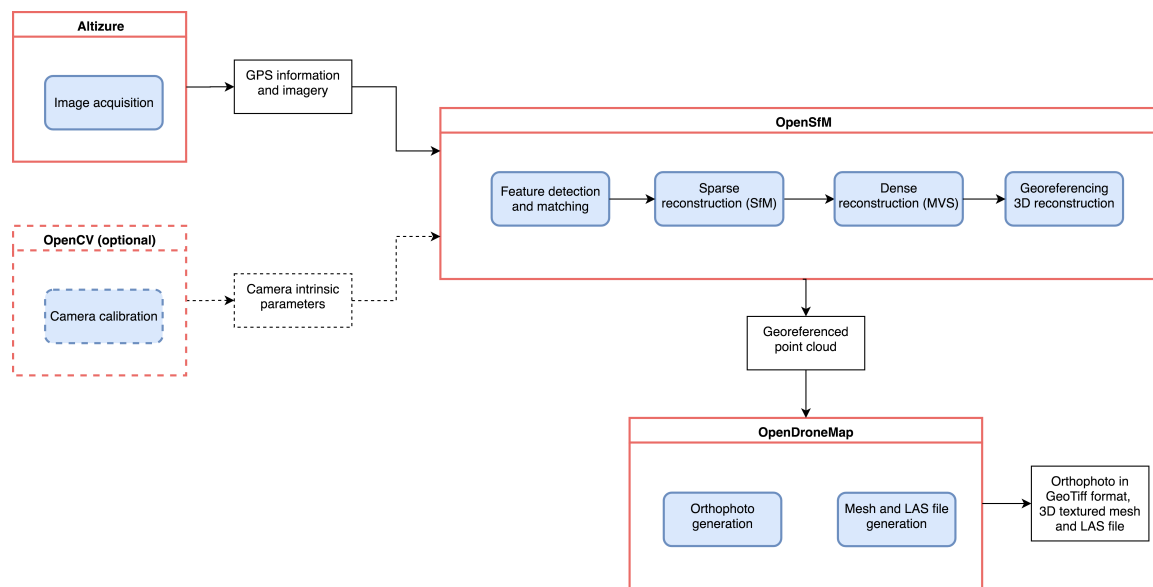


Figure 1: Reconstruction process for DEM generation: Camera calibration with OpenCV (optional step). First stage consists of image acquisition with a drone and Altizure app. The second stage is based on the OpenSfM library for the 3D reconstruction. The final stage is based on the OpenDroneMap library for post-processing the point cloud.

This dataset of points obtained with the SfM-MVS approach has allowed the studies of different problems such as monitoring glacier movement, quantifying soil loss and gully erosion, surveying fluvial morphology, among others [1]. There are several commercial software like Agisoft [7] or Pix4D [8] for obtaining 3D dense points clouds but, being closed code applications they do not favor research reproducibility. In this work, we proposed a Structure from Motion pipeline using UAVs for generating DEMs for developing a rainfall-runoff model to study a flood area. A rainfall-runoff model aims to simulate the flow of water induced by an observed or a hypothetical rainfall of a rainfall catchment area, that is, this model calculates the conversion of rainfall into runoff.

2. Structure from Motion pipeline

The proposed pipeline [9] (Fig. 1) is based on open source software, mainly on OpenSfM [10] and OpenDroneMap [11], and this one has four stages (one of them is optional). The camera calibration stage is an optional step in the method. For this reason, we show this block in a dotted line in Fig. 1. We carried out the camera calibration with the OpenCV library [12] for estimating the intrinsic camera parameters for the drone camera.

2.1. Altizure Stage: Image Acquisition

Three-dimensional reconstruction algorithms require a set of overlapping, offset images of the object or scene of interest acquired from different positions. The first stage consists in setting the flight path for the drone to carry out the image acquisition. We used the Altizure application [13] to set a flight strategy or flight mission.

2.2. OpenSfM Stage: Pipeline Reconstruction

The second stage consists in performing the 3D reconstruction. The workflow of this stage is based mainly on the SfM technique implemented with the OpenSfM library, where the input

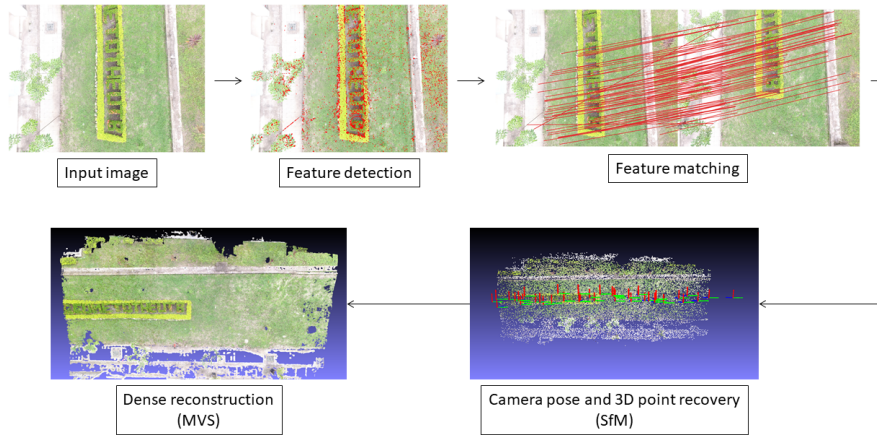


Figure 2: Pipeline reconstruction: each image goes through the stages of feature detection, point matching, camera pose estimation (motion), sparse reconstruction (structure) and finally, dense reconstruction using multi-view stereo (MVS).

is a set of overlapping images, and the output is a scene point cloud. This point cloud is geo-referenced, i.e., each point has a GPS coordinate assigned. During the SfM process, a point cloud is produced in a projective coordinate system, which needs to be aligned to a real-world coordinate system. For this purpose, we use Ground Control Points (GCPs), which are points on the terrain with known coordinates that allow us to align the reconstruction and to geo-reference the point cloud. In the workflow shown in Fig. 2, the algorithm initially detects image features or key-points. A key point is a distinctive feature which must be identifiable in another image of the same scene. The feature detection process must be repeatable to identify common points between images. Then with the detected points, we find correspondences between image pairs so that we identify the 3D points of the same physical object which appear in more than one image. Based on the correspondences and the SfM technique, we recover the camera poses, and the 3D point coordinates producing a sparse point cloud. These outputs are used as a first estimation and are refined in an iterative non-linear optimization process known as Bundle Adjustment [14] to minimize the reprojection error given by

$$E(\mathbf{P}, \mathbf{X}) = \sum_{i=1}^n \sum_{j=1}^m \|\mathbf{u}_{ij} - \mathbf{P}_i \mathbf{X}_j\|^2, \quad (1)$$

where \mathbf{u}_{ij} are the point correspondences, \mathbf{P}_i is the projection matrix of the camera in the image i , and \mathbf{X}_j is the 3D point of the correspondence j . This reprojection error is the distance d_{ij} between the measured point \mathbf{u}_{ij} in the keypoint detection process and the reprojected 3D point \mathbf{X}_j onto the image plane, as seen in Fig. 3. The 2D coordinates of a 3D point projected in the image plane is given by $\mathbf{P}_i \mathbf{X}_j$, where \mathbf{P}_i is the projection matrix

$$\mathbf{P}_i = \mathbf{K}_i [\mathbf{R}_i \mid \mathbf{t}_i], \quad (2)$$

\mathbf{K}_i being the intrinsic matrix that contains the camera internal parameters, and $[\mathbf{R}_i \mid \mathbf{t}_i]$ the camera pose that consists of a rotation matrix and a translation vector.

Finally, with the sparse point cloud and MVS technique, we densify the sparse point cloud using multiple depth maps [15], one for each image pair, and then they are merged into a single representation [16] obtaining a dense point cloud.

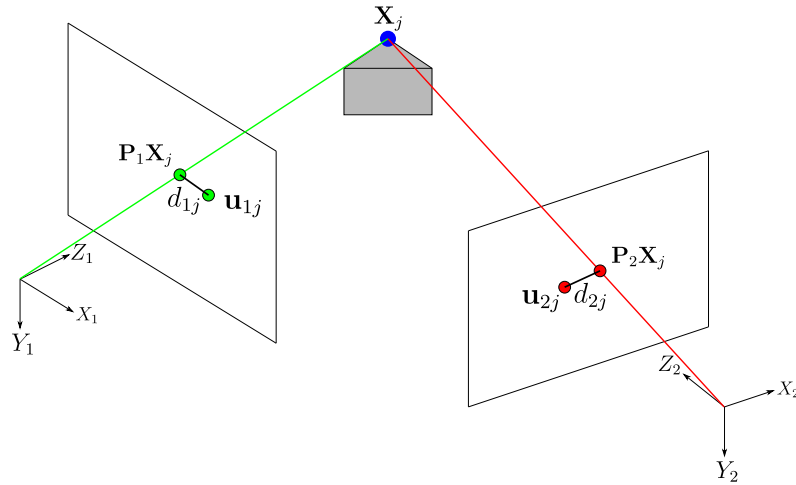


Figure 3: Bundle Adjustment procedure: reprojection is minimized using an iterative non-linear optimization process.

2.3. OpenDroneMap (ODM) stage: Post-processing

With the sparse and dense reconstruction generated with OpenSfM in PLY format, we use OpenDroneMap (ODM) to post-process the point cloud. The post-processing consists in generating a geo-referenced point cloud in LAS format, a geo-referenced 3D textured mesh and an orthophoto mosaic in GeoTIFF format. The LAS format is a standard binary format for the storage of LIDAR data and point clouds. The 3D textured mesh is a surface representation of the terrain that consists of vertices, edges, faces and the texture from the input images that is projected onto it. An orthophoto is an orthorectified aerial image, i.e., there are no geometrical distortions, and the scale is uniform throughout the image. The GeoTIFF format allows embedding the georeferencing information within an orthophoto in TIFF format generated with all images used in the reconstruction process.

3. Digital Terrain Model and Digital Surface Model generation

The Digital Surface Model (DSM) represents the surface of the earth and all objects on it. The DSM includes the top of buildings, tree canopy, powerlines, and other features. Unlike the DSM, a Digital Terrain Model (DTM) represents the bare ground surface without any objects like plants or buildings, as seen in Figure 4(a).

With the point cloud in LAS format obtained with ODM, we have ground and non-ground points of the scene so that our output point cloud from SfM pipeline represents a DSM of the scene. Our goal is to generate a DTM using the LAS point cloud, which can be used to generate a rainfall-runoff model. To do this, first, we filter the point cloud to remove the non-ground points that represent objects like plants or buildings. For example, as seen in Fig. 4(b), the red dots represent the ground points, and the green dots the non-ground points. With the filtering process, we remove the green points and therefore we have some holes left (red line). The filtering was carried out with the PDAL open source library [17], using the Extended Local Minimum method [18] to identify low noise points, and the Simple Morphological Filter (SMRF) approach [19] implementing the nearest neighbor void filling to segment ground and non-ground points. Then with the segmented point cloud, using PDAL we generate the DTM in GeoTIFF format, and from this process, we obtain a TIF image with holes due to segmentation, for this reason, we use ODM for filling the gaps in the image based on nearest-neighbor interpolation, producing a TIF image without gaps.

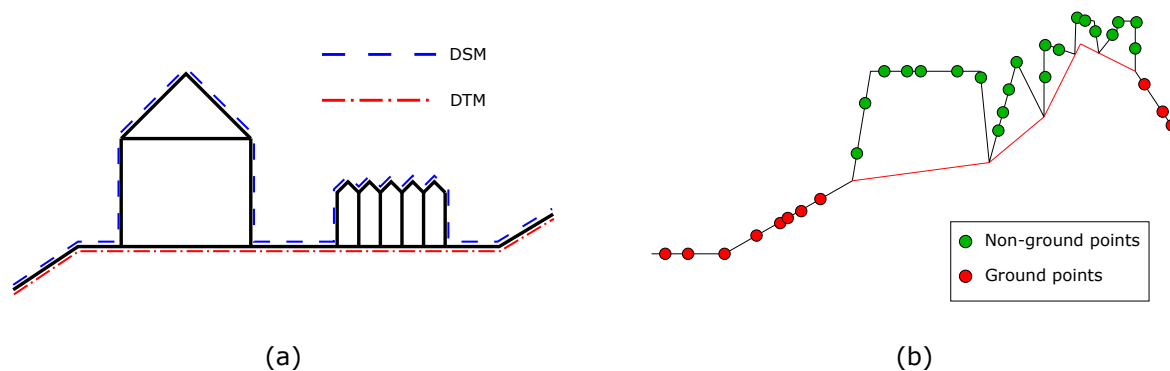


Figure 4: (a) Difference between DTM and DSM: DSM includes all objects on the ground and DTM does not. (b) The filtering consists of removing the non-ground points (green) and maintaining the ground points (red).



Figure 5: The testing site.

4. Results and discussion

We have chosen a locality in the municipality of Turbaco, Colombia as the testing site (Fig. 5) because this is a flood area due to rainfall. From this zone of approximately $144\,566\text{ m}^2$, we acquired 287 images with a DJI Phantom 3 Professional drone and 9 GCPs distributed on the ground for the geo-referencing process, measured with a differential GPS (dGPS). From the SfM pipeline discussed in section 2 we obtained a geo-referenced point cloud in LAS format as shown in Fig. 6(a). Based on this point cloud and using the previous segmentation strategy we removed the non-ground points of the reconstruction and keep the ground points as seen in Fig. 6(b), where we only keep the earth points.

In Fig. 7(a) we show the digital surface model generated with the unsegmented point cloud of Fig. 6(a) with ODM. Using the segmented point cloud of Fig. 6(b), we generated a GeoTIFF image (Fig. 7(b)) which has many gaps. By filtering, we obtain the DTM with the gaps filled using nearest-neighbor interpolation, implemented with a function provided by ODM. Fig. 7(c). In Fig. 7(d) we show a 1D profile of the DSM and the DTM, which have been extracted from

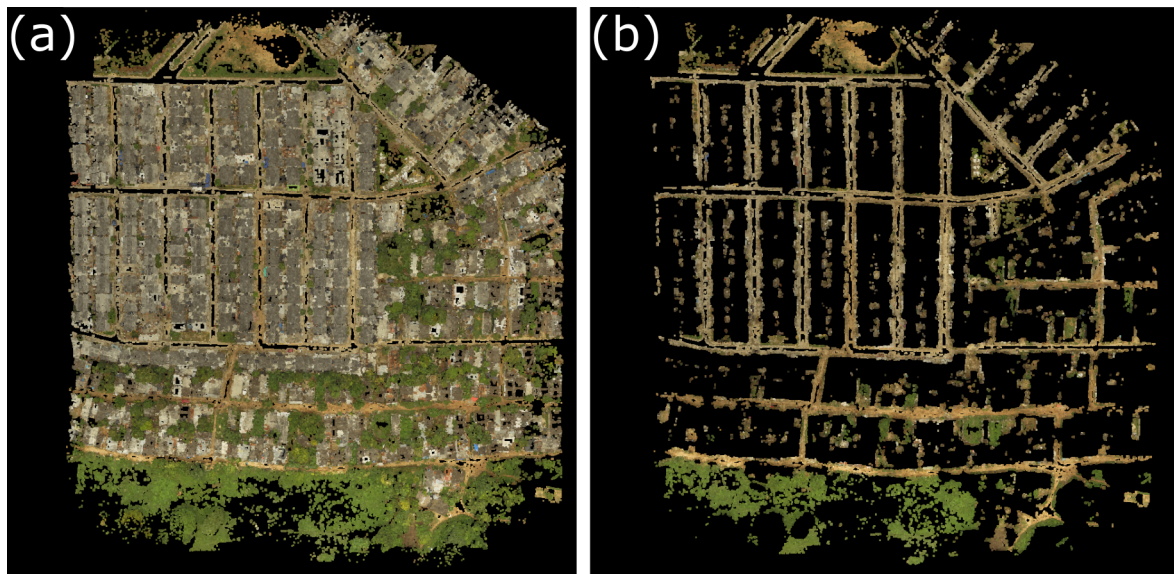


Figure 6: (a) output LAS point cloud from the SfM pipeline. (b) Ground segmentation results in LAS format.

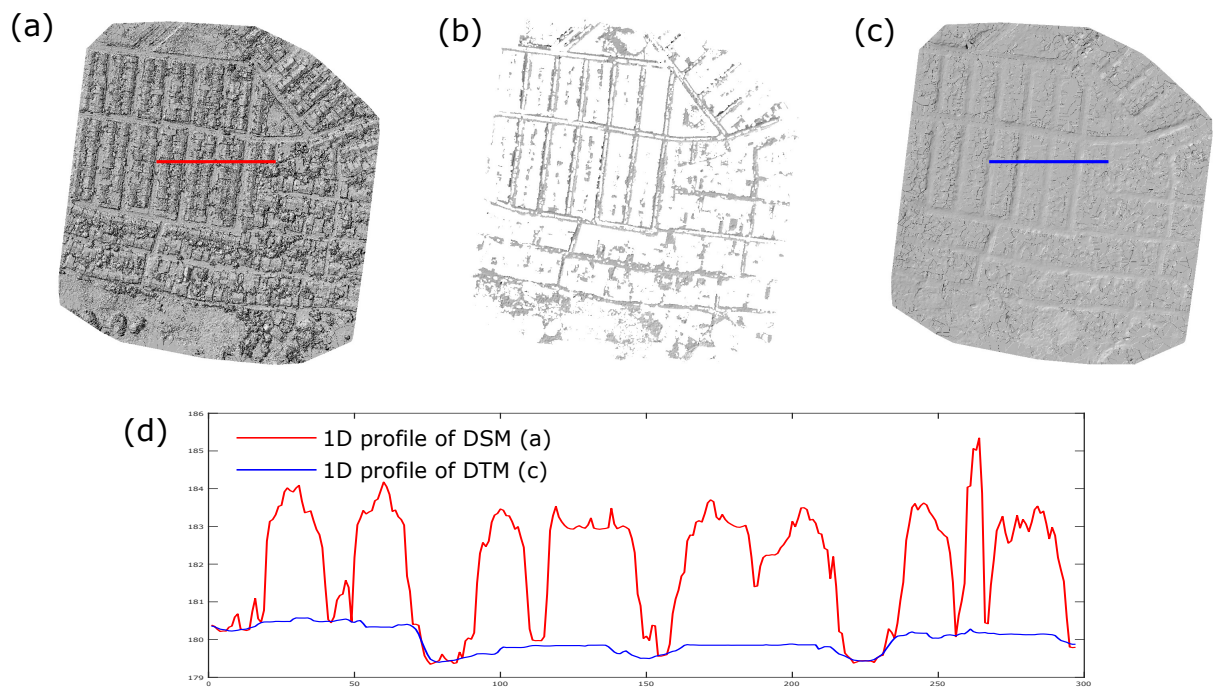


Figure 7: (a) DSM from unsegmented point cloud. (b) GeoTIFF image with gaps due to filtering of the segmented point cloud. (c) DTM: gaps filled based on nearest-neighbor interpolation. (d) 1D profile of DSM and DTM.

the red an blue line over the Fig. 7(a) and Fig. 7(c), respectively. In this plot, it is clear that the DSM includes the buildings the terrain, and after filtering and gap-filling, the DTM maintains the bare terrain surface with little influence from outlier points.



Figure 8: Watershed delineation of the test site.

4.1. Watershed delineation

As previously stated the test site is a flood area, this is because it is a relatively flat zone with a little slope and this makes the terrain has little capacity to drain water from precipitations. With the DTM of land, we make the hydrographic watershed delineation (Fig. 8) taking into account a hypothetical rainfall in the area. A watershed is an area of terrain where rainfall collects and drains off into a common outlet, as a result of water from rain runoff that runs downslope towards the shared outlet. In Fig. 8 in red color is the largest catchment area with approximately $123\,684\text{ m}^2$ and the other ones represent an area of approximately $20\,881\text{ m}^2$.

5. Conclusions

In this work, we have proposed a software pipeline using UAVs for generating DEMs to derive information from the morphology of a land. The topographic attributes obtained from the DEM data allowed us to determine the watershed delineation useful for predicting the behavior of systems and for studying hydrological processes. In particular, for developing a rainfall-runoff model to study flood areas. Imagery acquired from UAVs and 3D photogrammetry techniques offer cost-effective advantages over other remote sensing methods such as LIDAR or RADAR. In particular, a high spatial resolution for measuring the terrain microtopography. The resulting the 3D point cloud enabled us to obtain a georeference model from the GPS information of the camera and ground control points. Encouraging experimental results on a test land show that the produced DEM meets the metrological requirements for developing a surface-runoff model. However, we need to reconstruct a larger area of the zone to make a better analysis of watershed due to the boundaries of the red catchment area in Fig. 8 is truncated, i.e., we do not have the whole red catchment area with acquired image dataset.

Acknowledgement

This work has been partly funded by Universidad Tecnológica de Bolívar project (FI2006T2001). The authors thank Direccion de Investigaciones Universidad Tecnologica de Bolivar for their support.

References

- [1] Smith M, Carrivick J and Quincey D 2016 *Progress in Physical Geography* **40** 247–275
- [2] Marcus W A and Fonstad M A 2008 *Earth Surface Processes and Landforms* **33** 4–24
- [3] Westoby M, Brasington J, Glasser N, Hambrey M and Reynolds J 2012 *Geomorphology* **179** 300–314
- [4] Enciso J, Jung J and Chang A 2018 *Journal of Applied Remote Sensing* **12** 1–9
- [5] Nobajas A, Waller R I, Robinson Z P and Sangonzalo R 2017 *International journal of remote sensing* **38** 2844–2860
- [6] Goesele M, Curless B and Seitz S M 2006 *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on (IEEE)* pp 2402–2409
- [7] Agisoft photostan professional <http://www.agisoft.com/downloads/installer/>
- [8] Pix4d <https://pix4d.com/>
- [9] Meza J, Marrugo A G, Sierra E, Guerrero M, Meneses J and Romero L A 2018 *Communications in Computer and Information Science* **885** 213–225
- [10] Mapillary: Opensfm <https://github.com/mapillary/OpenSfM>
- [11] Opendrone map <https://github.com/OpenDroneMap/OpenDroneMap>
- [12] Bradski G and Kaehler A 2000 *Dr. Dobb's journal of software tools* **3**
- [13] Altizure <https://www.altizure.com>
- [14] Triggs B, McLauchlan P F, Hartley R I and Fitzgibbon A W 1999 *International workshop on vision algorithms (Springer)* pp 298–372
- [15] Furukawa Y and Ponce J 2010 *IEEE transactions on pattern analysis and machine intelligence* **32** 1362–1376
- [16] Adorjan M 2016 *Ein kollaboratives Structure-from-Motion System* Master's thesis Technischen Universität Wien
- [17] PDAL contributors 2018 *PDAL: The Point Data Abstraction Library* URL <https://pdal.io/>
- [18] Chen Z, Devereux B, Gao B and Amable G 2012 *ISPRS journal of photogrammetry and remote sensing* **72** 121–130
- [19] Pingel T J, Clarke K C and McBride W A 2013 *ISPRS Journal of Photogrammetry and Remote Sensing* **77** 21–30