[DC-04-038] Structure from Motion Photogrammetry

Abstract

In recent years, structure-from-motion (SfM) photogrammetry has been demonstrated to have the potential to democratize image-based three-dimensional (3D) reconstruction for a wide range of applications. SfM is the process of estimating the 3D structure of a scene from a set of overlapping images acquired from different viewpoints. SfM is not a single technique but rather a workflow exploiting multiple algorithms originally developed in computer vision and stereo photogrammetry. The SfM techniques along with multi-view stereo (MVS) algorithms can lead to a hyper-spatial resolution sampling of the real-world object or environment, which is represented via a dense set of 3D points with color or spectral information derived from the input images. SfM-MVS techniques have been implemented in different commercial and open-source suites, making the digital documentation of 3D objects or environments fully automated, highly accurate, and unprecedently inexpensive.

Author & citation

Pashaei, M. and Starek, M. J. (2024). Structure-from-Motion Photogrammetry. The Geographic Information Science & Technology Body of Knowledge (1st Quarter 2024 Edition), John P. Wilson (Ed.). DOI: 10.22224/gistbok/2024.1.2.

Explanation

1. Introduction
2. Approach
3. Point Cloud Georeferencing
4. Error Sources
5. Applications

1. Introduction

The production of high-resolution and accurate digital 3D models of real-world environments or objects is fundamental for a wide range of applications. In geoscience studies, for example, the availability of a high-resolution topographic model of the surveyed area, commonly called a digital surface model (DSM), is crucial for accurate representation of the landform and its exploitation for subsequent spatial analyses.

Traditionally, dense digital sampling of real-world scenes for surveying and mapping applications, e.g., the surface of the Earth, has been performed with airborne, mobile, or terrestrial light detection and ranging (LiDAR) systems. Such systems efficiently collect a huge number of range measurements, and this information is converted into a dense set of points, commonly called a point cloud, representing the 3D structure of the object or surveyed environment. Although LiDAR systems are highly effective in providing 3D models of the scanned object or environment, employing such systems for small-scale 3D

reconstruction scenarios is limited due to the high capital expenditure cost and portability issues (Carrivick et al., 2016). Likewise, traditional airborne photogrammetry also has limitations as an efficient and cost-effective method for localized mapping due to reliance on piloted aircraft, expensive metric-grade cameras, and highly specialized photogrammetric software and processing workflows.

On the other hand, advances in structure-from-motion (SfM) photogrammetry and multi-view stereo (MVS) algorithms have enabled highly efficient image-based 3D reconstruction where the quality of digital sampling of the real world has been reported to be comparable with LiDAR-based surveying techniques regarding sampling density and accuracy (Westoby et al., 2012). Additionally, in comparison to other techniques and approaches, SfM-MVS is recognized as a cost-effective technique for small-scale 3D reconstruction scenarios (Jones and Church, 2020).

Developed in the 1990s, the SfM-MVS techniques have its origins in the computer vision and photogrammetry communities (Westoby et al., 2012). SfM is the process of reconstructing the 3D scene via its projections into a series of 2D overlapping images collected from different viewpoints. This process reconstructs real world scenes based on estimating the pose (position and orientation) parameters of the camera in each imaging station through corresponding image features, commonly called keypoints, identified in overlapping images (Westoby et al., 2012). The development of automatic image feature detection and matching in computer vision has had an essential role in the efficiency of the SfM technique in image-based 3D reconstruction (James et al., 2017). The SfM approach has been popularized through a range of commercial software for processing collected imagery and cloud-processing engines, such as Microsoft® Photosynth™, which used SfM techniques to process user-uploaded and crowd-sourced photography to generate coverages of a certain scene and automatic generation of sparse 3D point clouds (Westoby et al., 2012).

On the other hand, photogrammetry is a relatively old technique in which 3D reconstruction efforts of pioneers in 1840s initially attempted to use a pair of overlapping photos captured from ground-based imaging stations with a fixed baseline and was followed by airborne 3D mapping applications using expensive metric cameras onboard an airplane for collecting a series of overlapping images (Wolf et al., 2014). The development of digital photogrammetry and improvement in the cost and quality of single-lens reflex (SLR) cameras and techniques for accurate calibration of such non-metric cameras have democratized access to photogrammetric modeling in a wide range of applications (Carbonneau et al., 2003; Haneberg 2008; Woodget et al., 2017). By integration of advanced image analysis algorithms developed in computer vision, such as automatic feature extraction and keypoint matching, as well as rigorous algorithms and techniques developed in photogrammetry, such as aerial triangulation and self-calibration bundle adjustment, the geometry of the stationary scene, i.e., the camera pose parameters, usually referred to as exterior orientation (EO) parameters, the camera interior orientation (IO) parameters, such as the focal length, lens distortion, etc., and the 3D location of each keypoint within the reconstructed scene are automatically estimated (Iglhaut et al., 2019). It should be mentioned, however, that traditional airborne photogrammetry has limitations as an efficient and cost-effective method for 3D mapping, particularly at localized geographic scales, due to reliance on piloted aircraft, expensive metric-grade digital cameras, and highly specialized photogrammetric software and processing workflows.

Over the past few years, advancements in small uncrewed aircraft systems (UAS) technology and digital cameras have opened a new paradigm for performing aerial surveying with UAS and SfM (called UAS-SfM). Image acquisition is generally done autonomously by using pre-programmed flight designs that target specific image sidelap and endlap for enabling 3D reconstruction (Westoby et al. 2012; Starek and Wilkinson, 2022). Several commercial and open-source software suites are available for processing overlapping UAS image sequences to derive a wide range of mapping products including a dense 3D point cloud, DSM, and high-resolution orthomosaic. Two of the most widely used commercial software suites at present are Agisoft Metashape and Pix4Dmapper™. Alternatively, OpenDroneMap is an open-source SfM and MVS packages for processing UAS images. Each particular software implementation of SfM-MVS may be slightly different. With this variability in mind and considering the black-box nature of the implemented details of the workflow, the following section describes major steps required in any implemented SfM-MVS technique to successfully reconstruct the scene using a set of overlapping images.

2. Approach

SfM photogrammetry is commonly used to refer to the entire image-based 3D reconstruction workflow from the acquisition of an image set to the generation of a 3D point cloud data; however, strictly speaking, SfM only refers to a certain step in this processing workflow, which leads to the derivation of relative camera pose (position and orientation) parameters and a sparse point cloud in an arbitrary 3D coordinate system. Dense image matching algorithms, such as MVS techniques, are exploited in a subsequent step to densify the final point cloud. Finally, derivate mapping products are produced, such as a DSM and orthomosaic. Thus, the whole process is strictly referred to as SfM-MVS. The typical SfM-MVS image processing workflow implemented with UAS imagery is summarized as follows and shown in Figure 1.

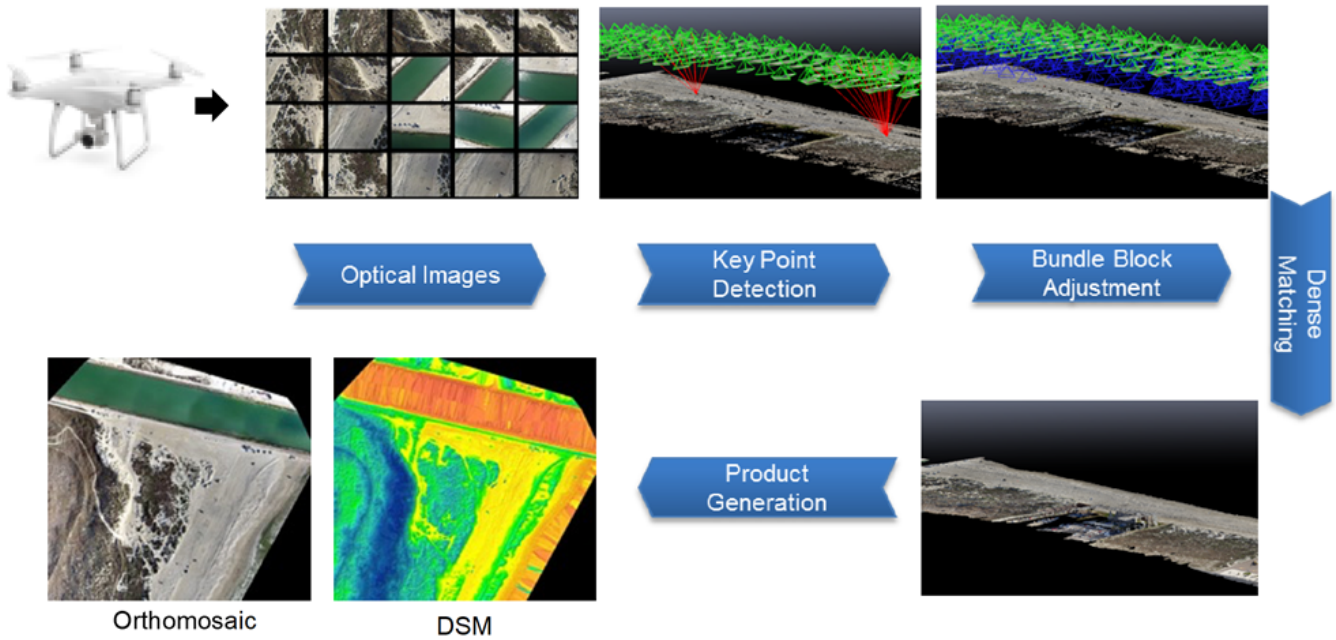


Figure 1. Typical SfM-MVS workflow to process overlapping UAS image sequences into a densified 3D point cloud, DSM, and orthomosaic. The example shown here was processed using Pix4Dmapper™ (Starek et al., 2019, reproduced with permission from Taylor &

Francis Group).

Feature detection

The SfM part of the workflow is first initiated by executing the feature detection algorithm to identify image features (or ‘keypoints) in each image. For each detected 2D feature a unique identifier or descriptor is assigned. While there are several alternative feature detection algorithms, the well-known scale-invariant feature transform (SIFT) algorithm (Lowe, 2004), developed in computer vision, is used most widely in SfM. It allows for feature detection regardless of scale, camera rotations, camera perspectives, and changes in illumination.

Keypoint correspondence

The next step requires the identification of keypoint correspondences (matching keypoints) in multiple images. Feature descriptor, which is typically a vector including a large number of entries (usually >100) is the main source to find keypoint correspondences.

Bundle adjustment and sparse point cloud generation

Keypoint correspondences extracted from overlapping images are the fundamental resource to estimate the camera EO parameters, including three shifts (positions) and three rotations for all camera stations at which images have been captured. Additionally, through the integrated self-calibration bundle adjustment technique within the SfM techniques, the camera IO parameters, including focal length, principal point offsets, and additional parameters such as lens distortion coefficients are estimated. Typically, the iterative bundle adjustment technique based on the nonlinear least-squares adjustment is performed to minimize the error at keypoint locations in image space, commonly referred to as the reprojection error. In addition to the estimation of camera parameters, exterior and interior, 3D object point coordinates of 2D keypoints are calculated in an arbitrary coordinate system creating a sparse point cloud.

Point cloud densification

With the knowledge about the image network (scene) geometry, i.e., camera parameters, a dense point cloud can be calculated by performing an MVS algorithm in this step which comprises the calculation of the 3D location of each object point which appears in at least three overlapping images.

Generation of derivative mapping products

The raw data product output from a UAS-SfM survey is a densified point cloud of X-Y-Z coordinates of the imaged scene. This point cloud is typically colorized by the RGB pixel values of the digital camera. UAS-SfM point clouds can have high point density (easily exceeding 1000 points/m2) due to the high camera resolutions and typical low altitudes at which data are collected. The 3D point cloud can then be used to generate a DSM of the terrain, orthomosaic image, or a 3D textured mesh.

3. Point Cloud Georeferencing

The resulting point cloud and derivative mapping products generated from SfM photogrammetry are often required to be represented in a certain local or global 3D coordinate system. For example, in geoscience or surveying applications, UAS-SfM products may need to be given in a global or real-world coordinate system such as the state plane coordinate system.

The georeferencing procedure can be integrated into the SfM-MVS approach to constrain the problem if camera geolocations, i.e., camera offsets with respect to a real-world coordinate system are available, for example, through an onboard global navigation satellite system (GNSS) receiver. Orientation information provided by an onboard inertial measurement unit (IMU), if available, can also be used by the SfM software to weigh the solution and potentially aid it, but in general, this information is not necessary since orientation is usually solved efficiently via bundle adjustment (Starek and Wilkinson, 2022). An alternative approach to constrain the SfM-MVS solution can be through introducing an appropriate number of ground control points (GCPs) to the SfM-MVS procedure. GCP coordinates in the real-world coordinate system are usually surveyed via GNSS receivers and along with their corresponding locations on overlapping images are input to the SfM-MVS computations. In both cases, the resulting 3D point cloud is georeferenced during the adjustment and the provided additional information, i.e., camera pose and/or GCP information, can be considered to optimize the SfM-MVS solution.

It is worth noting that point cloud georeferencing can be implemented independently through applying a 3D similarity (conformal) transformation on the resulting dense point cloud. The transformation includes seven parameters (three translations, three rotations, and one scale) which translates, rotates, and scales the dense point cloud. The transformation parameters are usually estimated through solving a system of equations which relates the coordinates of identified GCPs in the resulting dense point cloud to their coordinates in the real-world coordinate system where minimization of errors is achieved via the least squares.

4. Error Sources

The accuracy of the final 3D model derived from SfM photogrammetry is affected by a large number of variables that propagate into the quality of SfM-MVS solutions. For example, understanding and modeling the uncertainty involved in UAS-SfM applications in real-world reconstruction scenarios is an active area of research. In general, the main variables in SfM-MVS include: 1) scene variables, such as surface texture, repetitive pattens, moving objects, and occlusions; 2) lighting conditions, such as sun angle, changing illumination, and shadows; 3) camera parameters, such as focal length, lens quality, aperture, shutter speed, ISO, and pixel pitch; 4) survey characteristics, such as overlap, viewing angle, and distance to the object; and 5) processing variables, such as SfM feature extraction and keypoint identification, and MVS densification technique; and 6) georeferencing (if applicable, e.g., in geoscience applications using UAS-SfM photogrammetry).

5. Applications

In the last two decades, SfM-MVS techniques have been recognized as one of the most popular and efficient image-based 3D reconstruction approach exploited in a wide range of applications.

In computer vision and artificial intelligence for 3D scanning, visual perception, augmented reality, visual simultaneous localization and mapping (vSLAM), etc. UAS-SfM photogrammetry is used in a wide variety of geoscience applications for topographic mapping, landform change detection, forestry, etc. In addition, SfM-MVS is being exploited for digital documentation of cultural heritage objects and monuments.

Other applications of SfM-MVS photogrammetry include but are not limited to robotics, GIS, surveying, environmental science, geomorphology, vision meteorology, archeology, glaciology, engineering design and analysis, inspection surveying, construction monitoring, and photo-tourism.

References

Carbonneau, P. E., Lane, S. N., and Bergeron, N. (2003). Cost-effective non-metric close-range digital photogrammetry and its application to a study of coarse gravel river beds. International Journal of Remote Sensing 24(14): 2837-2854.

Carrivick, J. L., Smith, M. W., and Quincey, D. J. (2016). Structure from Motion in the Geosciences. John Wiley & Sons.

Haneberg, W.C. (2008). Using close range terrestrial digital photogrammetry for 3-D rock slope modeling and discontinuity mapping in the United States. Bulletin of Engineering Geology and the Environment, 67(4), 457-469.

Iglhaut, J., Cabo, C., Puliti, S., Piermattei, L., O'Connor, J., and Rosette, J. (2019). Structure from motion photogrammetry in forestry: A review. Current Forestry Reports, 5(3), 155-168.

James, M.R., Robson, S., d'Oleire-Oltmanns, S., and Niethammer, U. (2017). Optimising UAV topographic surveys processed with structure-from-motion: Ground control quality, quantity and bundle adjustment. Geomorphology, 280, 51-66

Jones, C.A. and Church, E. (2020). Photogrammetry is for everyone: Structure-from-motion software user experiences in archaeology. Journal of Archaeological Science: Reports, 30, 102261.

Lowe, D. (2004). Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2), 91-110.

Starek, M. J. and Wilkinson, B. E. (2022). Aerial surveying technology. In D.T. Gillins, M.L. Dennis, and A.Y. Ng (Eds.), Surveying and Geomatics Engineering: Principles, Technologies, and Applications (pp. 341-392). Reston, VA: American Society of Civil Engineers.

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., and Reynolds, J. M. (2012). 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology, 179, 300-314.

Wolf, P. R., Dewitt, B. A., and Wilkinson, B. E. (2014). Elements of Photogrammetry with

Applications in GIS, 4th Edition. McGraw-Hill Education.

Woodget, A. S., Austrums, R., Maddock, I. P., and Habit, E. (2017). Drones and digital photogrammetry: From classifications to continuums for monitoring river habitat and hydromorphology. Wiley Interdisciplinary Reviews: Water, 4(4): e1222