Structure from motion (SFM) photogrammetry vs terrestrial laser scanning

This item was submitted to Loughborough University’s Institutional Repository by the author.


Additional Information:

- This is a book chapter.

Metadata Record: https://dspace.lboro.ac.uk/2134/21673

Version: Accepted for publication

Publisher: ©2016 American Geosciences Institute and used with their permission.

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
Structure from motion (SfM) photogrammetry vs terrestrial laser scanning- Jim Chandler¹ & Simon Buckley²

1. School of Civil & Building Eng., Loughborough University, LE11 3TU, UK: j.h.chandler@lboro.ac.uk
2. Uni Research CIPR, Bergen, Norway: Simon.Buckley@uni.no

Structure from Motion (SfM) has its roots in the well-established spatial measurement method of photogrammetry, but is becoming increasingly recognised as a means to capture dense 3D data to represent real-world objects, both natural and man-made. This capability has conventionally been the domain of the terrestrial laser scanner (TLS), a mature and easy to understand method used to generate millions of 3D point coordinates in a form known as a “point cloud”. Each technique is described and noted for its strengths and weaknesses.

Terrestrial Laser Scanning

TLS, commonly known by the technique’s measurement principle lidar (light detection and ranging), has been used for topographic mapping since the mid-late 1990s (e.g. Kraus, 2007). Lidar itself uses a number of laser-based measurement techniques to determine 3D point coordinates on a surface object relative to the instrument. For earth science and topographic mapping, the “time-of-flight” principle is most commonly used, as it allows for longer ranges than phase-based (very fast capture and dense point clouds, but range limited) and triangulation (high accuracy and density, but very short range, <2 m) methods. Time-of-flight implementations on TLS instruments use pulses of laser light and reflectorless electromagnetic distance measurement (EDM) to determine a range to the object, while a scanning mechanism provides deflection angles using a mirror system and/or rotating head. These known vectors allow individual 3D coordinates to be determined and when combined, enable a dense point cloud to be captured quickly in an arbitrary but scaled 3D coordinate system (Buckley et al., 2008). Tens to hundreds of thousands of points are collected per second in modern pulsed time-of-flight instruments. Multiple scans are collected from different positions to obtain full coverage of an object or landform.

Early development to TLS systems, pioneered by companies such as Cyra (USA, now Leica Geosystems) and Riegl (Austria), was characterised by rapid and significant advances, such as digital camera integration to provide colour-coded geometry, massive increases in data acquisition rates, and increased portability, ruggedness, and battery life. Recently, developments have been more incremental, with longer measurement ranges (>5 km for scanners optimised for snow and ice measurement) and full waveform technology to allow multiple returns within a single laser footprint to be analysed. This is especially useful for vegetation studies (Mallet and Bretar, 2009). Integration with other sensors, such as higher spectral resolution cameras, is also a developing area (Kurz et al., 2012; Kurz et al., 2013). Because of the different measurement principles and laser classes used, no single TLS instrument will fit all applications over the range of scales and accuracy requirements, making it less flexible as an overall approach. A compromise of range versus point precision is required.

TLS technology and market penetration has matured to the extent that many see the technology as the obvious choice for capturing our world in 3D at close range, with applications including city modelling and building information modelling (BIM), architecture and crime scene recording. Within the geosciences, TLS has been heavily used in landform measurement and monitoring (Montreuil et al., 2013; Dewez, et al., 2013), geology (Bellian et al., 2005; Remondino et al., 2010), and change detection (Rosser et al., 2005; Nield et al., 2011). Despite this increasing number of applications, significant disadvantages of the method remain. These are largely practical and relate to the cost of equipment (typically US$30,000-80,000) and the size and weight of equipment can mean that some field sites remain inaccessible.
Structure from Motion

The phrase “Structure-from-Motion” evolved from the machine vision community, specifically for tracking points across sequences of images captured at different positions (e.g. Szeliski and Kang, 1994). However, SfM owes its existence to mathematical models developed many years ago in photogrammetry, including: coplanarity and collinearity, specifically the self-calibrating bundle adjustment (Kenefick et al., 1972; Faig, 1975). Photogrammetry is a long established method that has been used to compile the world’s maps since the 1920s and has evolved to take advantage of digital image processing to automatically generate digital elevation models and orthophotos, such as those now used in Google Earth. The SfM approach has been explained by a range of authors (e.g. Westoby et al., 2012; James and Robson, 2012; Fonstad et al., 2013; Micheletti et al., 2015), but in essence it involves acquiring images from a number of positions relative to the object of interest. An interest operator, such as the scale invariant feature transform (SIFT) identifies distinctive features appearing upon multiple images and establishes the spatial relationships between the original camera positions in an arbitrary and unscaled coordinate system. A self-calibrating bundle adjustment is then used to calibrate the camera(s) and derive a sparse set of coordinates to represent the object. Point density is then intensified using “multi-view stereo” (MVS) techniques, to generate a very high resolution point cloud, which is colour-coded using the original image data.

Figures 1-3 demonstrate the type of high resolution point clouds achievable to represent a natural feature, here an eroded dune system (Figure 1), using both TLS (Figure 2) and SfM photogrammetry (Figure 3). SfM photogrammetry can eradicate two of the challenges associated with TLS: cost and occlusions. As the method simply requires images acquired from a digital camera or smartphone and access to cheap or even free software solutions, it provides a very cost-effective solution to create a 3D point cloud. The need to acquire images from multiple positions also addresses some of the issues related to occlusions. To this, should be added the flexibility of scale. The precision of any photogrammetric solution is directly related to scale of the imagery and the geometry of the images captured. Precision is therefore commensurate with scale and if image geometry is appropriate and camera calibration effective, then an SfM photogrammetric solution can generate data that is of higher accuracy than achievable using
Despite these apparent advantages there are some dangers. To be effective, imagery has to be obtained which is of high image quality (i.e. sharply focused, no motion blur) and obtained from locations that provide appropriate coverage (each desired point on the object must appear on at least three frames) and with an appropriate geometry (images must be acquired from spatially different positions) (Micheletti et al., 2015). Failure to achieve the latter is particularly problematic, with many users using UAVs or drones to acquire vertical aerial imagery of an object. This configuration generates imagery which has inherently weak geometry, and inaccuracies in the calibration of the camera (specifically the lens model determined) will generate inaccuracies within the object (Wackrow and Chandler, 2008). This may not always be detected but will manifest itself as a highly systematic domed surface, which has been reported recently by a number of authors (James and Robson, 2014; Woodgett et al., 2014). Finally, as SFM photogrammetry relies on identical features being found between multiple images, object surfaces with uniform colour or texture may not be suitable for generating 3D data, unlike the active TLS method (compare bottom parts of Figures 2 and 3).

Conclusion

The purpose of this short review has been to outline the basic approaches used by these two technologies and identify strengths and weaknesses between the two. Speed of acquisition is significantly slower for TLS than image-based techniques (often hours vs minutes) for comparable data resolutions. However, as an active range measurement technique TLS offers advantages in terms of accuracy, repeatability and reliability, and can still be viewed as the "gold standard" for 3D measurement. SFM is both cost-effective, automated and allows consistent image-to-geometry registration, suggesting that structure from motion photogrammetry can rival terrestrial laser scanning for many applications.

Both techniques can generate very high resolution point clouds consisting of millions of 3-D points. This creates challenges in terms of data storage and processing hardware and users must consider what information is to be extracted beyond simply visualisation.

References


