



The accuracy of photo-based structure-from-motion DEMs

M. R. James (1) and S. Robson (2)

(1) Lancaster Environment Centre, Lancaster University, Lancaster, United Kingdom (m.james@lancs.ac.uk), (2) Dept. Civil, Env. and Geomatic Engineering, UCL, Gower Street, London, United Kingdom

Data for detailed digital elevation models (DEMs) are usually collected by expensive laser-based techniques, or by photogrammetric methods that require expertise and specialist software. However, recent advances in computer vision research now permit 3D models to be automatically derived from unordered collections of photographs, offering the potential for significantly cheaper and quicker DEM production. Here, we assess the accuracy of this approach for geomorphological applications using examples from a coastal cliff and a volcanic edifice.

The reconstruction process is based on a combination of structure-from-motion and multi-view stereo algorithms (SfM-MVS). Using multiple photographs of a scene taken from different positions with a consumer-grade camera, dense point clouds (millions of points) can be derived. Processing is carried out by automated 'reconstruction pipeline' software downloadable from the internet, e.g. <http://blog.neonascend.net/archives/bundler-photogrammetry-package/>. Unlike traditional photogrammetric approaches, the initial reconstruction process does not require the identification of any control points or initial camera calibration and is carried out with little or no operator intervention. However, such reconstructions are initially un-scaled and un-oriented so additional software (http://www.lancs.ac.uk/staff/jamesm/software/sfm_georef.htm) has been developed to permit georeferencing. Although this step requires the presence of some control points or features within the scene, it does not have the relatively strict image acquisition and control requirements of traditional photogrammetry. For accuracy, and to allow error analysis, georeferencing observations are made within the image set, rather than requiring feature matching within the point cloud.

In our coastal example, 133 photos taken with a Canon EOS 450D and 28 mm prime lens, from viewing distances of ~ 20 m, were used to reconstruct a ~ 60 m long section of eroding cliff. The resulting surface model was compared with data collected by a Riegl LMS-Z210ii terrestrial laser scanner. Differences between the surfaces were dominated by the varying effects of occlusions on the techniques, and systematic distortion of the SfM-MVS model along the length of the cliff could not be resolved over the ± 15 mm precision of the TLS data. For a larger-scale example, a ~ 1.6 km wide region over the summit of Piton de la Fournaise volcano was reconstructed using 133 photos taken with a Canon EOS D60 and 20 mm prime lens, from a microlight aircraft (with a representative viewing distance of 1.0 km). In this case, the resulting DEM showed an RMS error of 1.0 m when compared with the results from traditional photogrammetry and some areas of systematic error were evident. Such errors were minimised by reprocessing the SfM-MVS results with a more sophisticated camera model than is integrated into the reconstruction pipeline.

In combination, the results indicate that, with a good, convergent image set, SfM-MVS can be anticipated to deliver relative precisions of 1:1000 or better, for geomorphological applications. However, under certain conditions, the restricted camera model used can result in detectable error. We highlight the requirement for new network design tools that will help optimise image collection, facilitate error visualisation and allow a user to determine whether their image network is fit for purpose.