Comments on :

“‘Structure-from-Motion’ photogrammetry”

Virgile Daugé*

University of Lorraine

virgile.dauge@inria.fr

January 29, 2018

Abstract

This paper outlines a low-cost, user-friendly photogrammetric technique for obtaining high-resolution datasets at a range of scales, termed ‘ Structure-from-Motion ’ (SfM). The goal of my little review was to gain a basic understanding of known photogrammetry techniques, to be able indentify the impact of our positioning method in concrete applications. I was mainly searching for an estimation of motion needed to cover and reconstruct 3D environement, in order to compare our work with well known odometry/SLAM solutions.

I. Introduction

In this paper, we report on an emerging, low-cost photogrammet- ric method for high resolution topographic reconstruction, ideally suited for low-budget research and application in remote areas. ‘ Structure-from-Motion ’ (SfM) operates under the same basic tenets as stereoscopic photogrammetry, namely that 3-D structure can be resolved from a series of overlapping, offset images.However, it differs fundamentally from conventional photogrammetry, in that the geometry of the scene, camera positions and orientation is solved automatically without the need to specify a priori, a network of tar- gets which have known 3-D positions. Instead, these are solved si- multaneously using a highly redundant, iterative bundle adjustment procedure, based on a database of features automatically extracted from a set of multiple overlapping image.
the approach is most suited to sets of images with a high degree of overlap

there exist few quantitative assessments of the quality of terrain products derived from this approach. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

*A thank you or further information

II. Methods

Camera pose and scene geometry are reconstructed simultaneously through the automatic identification of matching features in multiple images. Scale Invariant Feature Transform (SIFT) is used for features extraction.
Following keypoint identification and descriptor assignment, the sparse bundle adjustment

system Bundler (Snavely et al., 2008)is used to estimate camera pose and extract a low-density or ' sparse ' point cloud.
Keypoints in multiple images are matched using approximate nearest neighbour (Arya et al., 1998) and Random Sample Consensus (RANSAC;Fischler and Bolles, 1987) algorithms, and ' tracks ' linking specific keypoints in a set of pictures are established.
Tracks comprising a minimum of two keypoints and three images are used for point-cloud reconstruction, with those which fail to meet these criteria being automatically discarded
transient features such as people moving across the area of interest are automatically removed from the dataset before 3-D reconstruction begins
Keypoint correspondences place constraints on camera pose orientation, which is reconstructed using a similarity transformation, while minimisation of errors is achieved using a non-linear least- squares solution
Finally, triangulation is used to estimate the 3-D point positions and incrementally reconstruct scene geometry, fixed into a relative coordinate system
the transformation of SfM image-space coordinates to an absolute coordinate system can be achieved using a 3-D similarity transform based on a small number of known ground-control points (GCPs)
it is often easier to deploy physical targets with a high contrast and clearly defined centroid in the field before acquiring images

III. Results

After manual editing, the sparse dataset comprised 1.710e5 points, while the dense reconstruction produced 11.310e6 points; a 64-fold increase. This is comparable to the TLS survey density with 11.710e6 survey points over the same area.
No real error comparison ?

IV. Discussion

As the above examples demonstrate, the apparent logistical advan- tages of SfM (limited hardware needs and portability) are, at least in part, offset by the lengthy processing times compared to ' data-ready ' methods such as TLS or GPS. Keypoint descriptor extraction, matching, and sparse and dense reconstruction algorithms are computationally demanding.

References

[1] M. Westoby, J. Brasington, N. Glasser, M. Hambrey, and J. Reynolds, "'structure-from-motion' photogrammetry: A low-cost, effective tool for geoscience applications", *Geomorphology*, vol. 179, pp. 300–314, Dec. 2012, ISSN: 0169555X. DOI: `10.1016/j.geomorph.2012.08.021`. [Online]. Available: `http://linkinghub.elsevier.com/retrieve/pii/S0169555X12004217` (visited on 01/29/2018).