Int J Appl Earth Obs Geoinformation 70 (2018) 25–34

Contents lists available at ScienceDirect

Int J Appl Earth Obs Geoinformation

journal homepage: www.elsevier.com/locate/jag

A patch-based method for the evaluation of dense image matching quality

Zhenchao Zhang[a,*], Markus Gerke[b], George Vosselman[a], Michael Ying Yang[a]

[a] Department of Earth Observation Science, Faculty ITC, University of Twente, Enschede, The Netherlands
[b] Institute of Geodesy and Photogrammetry, Technical University of Brunswick, Germany

ARTICLE INFO

Keywords:
Quality evaluation
Photogrammetry
Dense image matching
Laser scanning
Point cloud
Digital Surface Model

ABSTRACT

Airborne laser scanning and photogrammetry are two main techniques to obtain 3D data representing the object surface. Due to the high cost of laser scanning, we want to explore the potential of using point clouds derived by dense image matching (DIM), as effective alternatives to laser scanning data. We present a framework to evaluate point clouds from dense image matching and derived Digital Surface Models (DSM) based on automatically extracted sample patches. Dense matching errors and noise level are evaluated quantitatively at both the local level and whole block level. In order to demonstrate its usability, the proposed framework has been used for several example studies identifying the impact of various factors onto the DIM quality. One example study proves that the overall quality on smooth ground areas improves when oblique images are used in addition. This framework is then used to compare the dense matching quality on three different terrain types. In another application of the framework, a bias between the point cloud and the DSM generated from a photogrammetric workflow is identified. The framework is also used to reveal inhomogeneity in the distribution of the dense matching errors caused by overfitting the bundle network to ground control points.

1. Introduction

Airborne laser scanning (ALS) and airborne photogrammetry are the two main techniques to obtain 3D data representing the earth surface (Höhle and Höhle, 2009). The properties of laser scanning and photogrammetry have been widely compared before (Baltsavias, 1999; Leberl et al., 2010; Haala et al., 2010; Remondino et al., 2014; Cavegn et al., 2014; Yang and Chen, 2015; Tian et al., 2017). Compared to airborne laser scanning, image acquisition in airborne photogrammetry is mostly cheaper and more efficient in data acquisition flights (Hobi and Ginzler, 2012; Nurminen et al., 2013; Maltezos et al., 2016). In many countries photogrammetric image blocks are captured anyway for administrative and planning purposes with decreasing time intervals, so the question is to what extent these data can be used to replace ALS data in various application domains such as Digital Elevation Model (DEM) acquisition (Ressl et al., 2016), forestry mapping (Mura et al., 2015), classification and object extraction (Tomljenovic et al., 2016; Dong et al., 2017), and 3D modeling (Xiong et al., 2015).

We want to explore the potential of using photogrammetric products as effective alternatives to laser scanning data. In order to judge this potential, it is necessary to evaluate the data quality of 3D products from dense image matching (DIM). Assessing the absolute accuracy of 3D data can be time-consuming and labor-intensive for two reasons. Firstly, the reference data must be verified as being more accurate than the compared data. Secondly, the sample size should be sufficiently large in order to arrive at sound conclusions. Previous work of evaluating the absolute accuracy of 3D data can be divided into two categories based on the reference data.

In some previous evaluation studies, the reference data was collected by Real Time Kinematic (RTK) GPS. However, the sample size was relatively small in this case. Jaud et al. (2016) evaluated point clouds generated from images obtained by Unmanned Aerial Vehicles (UAVs). Twenty-four ground targets were set in the study area which served as GCPs in the triangulation and as check points in the DIM evaluation. The coordinates of these targets were obtained by post-processed differential GPS. Hobi and Ginzler (2012) evaluated the quality of Digital Surface Models (DSMs) from stereo matching of WorldView-2 satellite images and ADS80 aerial images using 36 reference points obtained by sub-decimeter differential GPS. Nurminen et al. (2013) studied the accuracy of DSMs derived from ALS and DIM in the estimation of plot-level variables. The reference variables of the forest plots were obtained by field surveys.

In addition, the reference data may be obtained by laser scanning. The basic assumption is that the point clouds obtained by laser scanning are more accurate than point clouds from photogrammetry, at least concerning the height component. Mandlburger et al. (2017) calculated the deviation between DIM-DSM and Lidar-DSM at impervious surfaces and found a systematic deviation of 0.043 m and a dispersion of

* Corresponding author.
E-mail address: z.zhang-1@utwente.nl (Z. Zhang).

https://doi.org/10.1016/j.jag.2018.04.002
Received 29 January 2018; Received in revised form 3 April 2018; Accepted 4 April 2018
0303-2434/

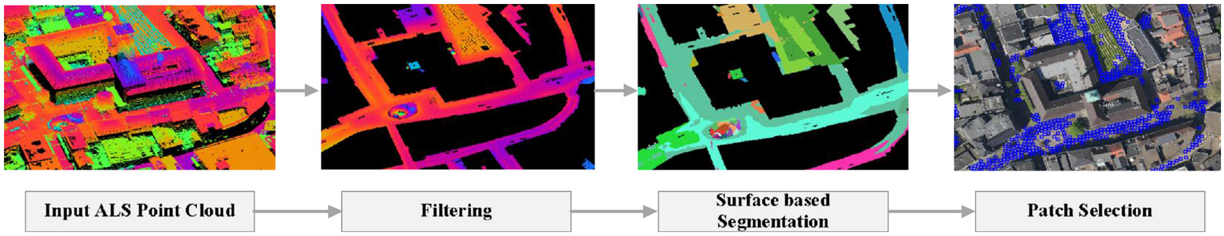

Fig. 1. Workflow for detecting candidate patches from ALS point cloud.

0.041 m. Tian et al. (2017) selected 184 inventory plots as the samples for DSM evaluation in a forest area. Two datasets from ALS were taken as reference data. Similar work taking laser scanning data as reference can also be found in (Poon et al., 2005; Gehrke et al., 2010; Moussa et al., 2013; Remondino et al., 2014; Nex et al., 2015; Jaud et al., 2016; Maltezos et al., 2016; Sofia et al., 2016; Ressl et al., 2016).

Some deficiencies of previous DIM evaluation work are summarized as follows: Firstly, some studies evaluated the point cloud derived from Semi-Global Matching (SGM) by making comparisons with ALS data or terrestrial laser scanning data on a planar sports field, complex castle or building façade (e.g. in Rothermel et al., 2012; Haala and Rothermel, 2012; Cavegn et al., 2014; Remondino et al., 2017). However, the small sample size or local area cannot properly represent the error distribution in the whole block. Secondly, when calculating quality measures, point-to-point distance (Kraus et al., 2006) and point-to-plane distance (Rothermel and Haala, 2011; Nex et al., 2015) were widely used as the measures to represent the accuracy. However, these measures are sensitive to blunders and random noise within the dense matching point clouds. Thirdly, the quality measures were less reliable or persuasive if calculated without consideration of the breaklines in natural scenes, such as bumpy terrain, edges of traffic islands or curbstones, and edges and ridges of roofs (e.g. in Ressl et al., 2016; Jaud et al., 2016).

In our previous work of evaluating point cloud from multi-view photogrammetry (Zhang et al., 2017), robust quality measures were calculated on roof segments. In this paper, a framework for evaluating point clouds and DSMs generated from a state-of-the-art dense matching algorithm is proposed. The contributions are as follows:

- The dense matching quality is evaluated robustly based on a very large number of planar patches of the same size extracted from planar ground surfaces in both the DIM point cloud and the ALS point cloud. Quantitative quality measures are proposed to represent the accuracy and precision at both the local patch level and the whole block level. After considering possible breaklines in natural scene and excluding patches with possible changes between the DIM data and reference data, the evaluation based on these planar patches reveals the distribution of DIM errors in the whole photogrammetric block for the first time. Compared to the previous point-to-point and point-to-plane comparisons, this framework computing the plane-to-plane distance is more robust to local blunders and artefacts.
- In order to test the usability of the proposed framework, several influencing factors related to the DIM quality are studied. One example is the additional use of oblique airborne imagery. This is not yet standard, but especially in urban applications it becomes more important (Toschi et al., 2017). Among other factors we evaluate how the additional use of oblique images influences the dense matching quality. We also compare the dense matching quality on different types of terrain. Meanwhile, suggestions are given on the photogrammetric quality control and dense matching parameter settings.

The paper is organized as follows: Section 2 presents the patch-based DIM evaluation framework. Section 3 gives details on the study area and experimental settings, while Section 4 focuses on experimental results. Section 5 discusses those results and Section 6 finally concludes the paper.

2. Methodology

In our evaluation framework, an ALS point cloud is taken as the reference data. The ALS data are assumed to be accurate with regards to the external reference and precise in consideration of random noise. The "patches" used as evaluation units are regular squares selected from the ALS data. Every patch is a sample for quality evaluation. Therefore, the densely selected patches on the ground can indicate the error distribution in the whole photogrammetric block. The proposed framework for DIM evaluation includes four steps: Firstly, square patches are detected from the ALS data and validated (Section 2.1); Secondly, corresponding DIM points are searched for each patch and the patches are further screened based on patch-based attributes (Section 2.2); Thirdly, quality measures are computed (Section 2.3); Finally, statistical analyses are performed on the valid patches.

2.1. Patch detection

The goal of patch detection is to localize candidate planar patches on the ALS point cloud. The patches taken as samples should be selected from the ALS data. The selection of patches should further avoid data gaps and breaklines. Planar patches of uniform size with acceptable noise level are considered valid and thus used for evaluation purpose. The examples used in this paper all make use of patches on the ground. The framework could, however, equally well be applied to planar non-ground patches.

In Fig. 1, a workflow is depicted for detecting ground patches. Firstly, ground points are identified from the ALS data using the method of (Axelsson, 2000). Then planar segments are extracted from ground points using a surface-based growing method (Vosselman, 2013). This approach employs the 3D Hough transform to detect seed surfaces. Then the nearby points are added to the surface if the distance from a certain point to the fitted plane is below a certain threshold. After new points are added to the segment, the plane parameters are recalculated before testing the next point. Slight over-segmentation is preferred over under-segmentation: over-segmentation can ensure better planarity and help avoiding breaklines in the segments.

After segmentation, the laser points with segment labels should be screened to discard small clusters or noisy segments. Features listed in

Table 1
Segment-based features for extracting smooth segments.

Feature	Description
Segment size	Number of points in the segment
linearity of segment	$(\lambda_1 - \lambda_2)/\lambda_1$, λ_1 is the maximum eigenvalue of the covariance matrix (Weinmann et al., 2015)
Plane slope	Normal direction of the fitted plane
Average angle	Mean of the angles between local point normals and the fitted plane normal
Residual of plane fitting (RPF)	Standard deviation of the distances between points and the plane fitted to the segment