ISPRS Journal of Photogrammetry and Remote Sensing 136 (2018) 41–57

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Automatic registration of panoramic image sequence and mobile laser scanning data using semantic features

Jianping Li [a], Bisheng Yang [a,*], Chi Chen [a], Ronggang Huang [b], Zhen Dong [a], Wen Xiao [c]

[a] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China
[b] Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Wuhan, China
[c] School of Engineering, Newcastle University, Newcastle upon Tyne NE1 7RU, UK

ARTICLE INFO

Article history:
Received 3 July 2017
Received in revised form 28 November 2017
Accepted 9 December 2017

Keywords:
Panoramic image sequence
Mobile laser scanning data
Semantic features
Registration

ABSTRACT

Inaccurate exterior orientation parameters (EoPs) between sensors obtained by pre-calibration leads to failure of registration between panoramic image sequence and mobile laser scanning data. To address this challenge, this paper proposes an automatic registration method based on semantic features extracted from panoramic images and point clouds. Firstly, accurate rotation parameters between the panoramic camera and the laser scanner are estimated using GPS and IMU aided structure from motion (SfM). The initial EoPs of panoramic images are obtained at the same time. Secondly, vehicles in panoramic images are extracted by the Faster-RCNN as candidate primitives to be matched with potential corresponding primitives in point clouds according to the initial EoPs. Finally, translation between the panoramic camera and the laser scanner is refined by maximizing the overlapping area of corresponding primitive pairs based on the Particle Swarm Optimization (PSO), resulting in a finer registration between panoramic image sequences and point clouds. Two challenging urban scenes were experimented to assess the proposed method, and the final registration errors of these two scenes were both less than three pixels, which demonstrates a high level of automation, robustness and accuracy.

1. Introduction

A Mobile Mapping System (MMS), usually equipped with laser scanners, panoramic cameras, Inertial Measurement Unit (IMU) and Global Positioning System (GPS) for acquisitions of images, point clouds and orientations, has been widely used for street inventory. The integration of complementary ranging and imaging data such as laser-scanning and electro-optical sensors provides new solutions to map the Earth's surface. The acquired datasets include rich spectrum and geometry information. Image-to-range registration is a critical step for many applications, such as texturing 3D models of large-scale scenes (Abayowa et al., 2015), building extraction (Yang et al., 2013), forest biomass inventory (Pflugmacher et al., 2014), and point cloud classification using the color information (Barnea and Filin, 2013; Yang and Dong, 2013).

Extrinsic calibration of the panoramic camera and the laser scanner is usually performed before data collection to achieve image-to-range alignment. However, there is often a considerable misalignment between the panoramic images and laser scanning data according to the calibrated extrinsic parameters (Miled et al., 2016; Cui et al., 2017). The main reason for the misalignment is the unforeseen movement of sensors (Levinson and Thrun, 2013). Since the mounting of the sensors will become less stable over a certain period of time, the relative extrinsic parameters should be calibrated manually frequently, which is laborious (Brenner, 2014).

An automated method is proposed in this paper to accurately register panoramic image sequence and mobile laser-scanning point clouds in an urban environment by estimating the transformation parameters between the panoramic camera and laser scanner using parked vehicles as registration primitives. Firstly, the original EoPs of panoramic images are adjusted in a GPS and IMU aided panoramic structure from motion (SfM). The accurate rotation between the panoramic camera and the laser scanner is obtained. Secondly, parked vehicles are extracted from both panoramic images and point clouds. To extract vehicles in panoramic images, the Faster-RCNN (Ren et al., 2015) is used to detect vehicle candidate areas, which are then segmented based on CRFASRNN (Zheng et al., 2015) and refined by image matting (Levin et al., 2008). Next, corresponding areas of the extracted

* Corresponding author.
E-mail address: bshyang@whu.edu.cn (B. Yang).

https://doi.org/10.1016/j.isprsjprs.2017.12.005
0924-2716/

vehicles are determined in the point clouds according to the initial EoPs of the panoramic images obtained in the initialization step. Then, vehicles in those corresponding areas of the point clouds are extracted. Finally, the relative translation between the panoramic camera and the laser scanner is refined by maximizing the overlapping area of the vehicle primitive pairs based on the Particle Swarm Optimization (PSO). The main contribution of the proposed method is that it refines the initial bundle adjustment based alignment by the Particle Swarm Optimization utilizing semantic features, resulting in an automatic and accurate registration between panoramic image sequence and mobile laser-scanning point clouds.

The remainder of this paper is organized as follows. Related literature is reviewed in Section 2. The proposed method is elaborated in Section 3. In Section 4, the experimental studies are undertaken to evaluate the proposed method, after which discussions are presented and conclusions are drawn at the end.

2. Literature review

Extensive studies on the registration of airborne laser-scanning data and aerial frame-images have been reported, including feature-based methods and area-based methods (Parmehr et al., 2014). Typical features used in feature-based methods are points (Palenichka and Zaremba, 2010), lines (Habib et al., 2005; Lee and Yu, 2006; Wang and Neumann, 2009), and planes (Kwak et al., 2006; Yang and Chen, 2015). Area-based approaches are usually relying on maximizing the statistical correlation (e.g., mutual information) or grayscale similarity (Parmehr et al., 2014). However, methods designed for airborne platform registration may not be applicable to the registration between terrestrial platforms, due to the substantial difference in scales, complexities, scanning perspectives, etc. For example, airborne and vehicle-borne laser scanning platforms clearly differ in data capture mode, typical project size, scanning orientation, and spatial resolution. Moreover, compared with images captured by aerial photogrammetry systems, images captured by panoramic cameras mounted on a vehicle are more complicated and challenging for registration due to occlusions caused by moving objects or nearby objects, smaller overlapping rate, and drastic changes in depth of view.

In fact, many methods have been explored to register mobile laser-scanning data and panoramic images, including features-based methods, statistical-based methods, and multi-view-based methods, which are reviewed in detail as follows.

Features-based methods: Böhm and Becker (2007) adopted the Scale-Invariant Feature Transform (SIFT) (Lowe, 2004) feature detector to extract the corresponding point features to register optical imagery and range images. W. Moussa et al. (2012) achieved registration of close-range images with TLS data by utilizing the affine SIFT (ASIFT) (Morel and Yu, 2009) feature matching strategy based on point-based environment model. Lines, bounding boxes, and skylines were also adopted by many studies for registration (Christy and Horaud, 1999; Liu and Stamos, 2005; Ramalingam et al., 2009). The key factor for feature-based registration is the strategy to find the correct matching feature pairs from point clouds and images.

Statistical-based methods: This kind of method aims to optimize the initial EoPs of images by maximizing the statistical correlation between images and point clouds. Miled et al. (2016) optimized the inaccurate transformation between camera and laser scanner by maximizing the mutual information (MI) between these two kinds of data resulting in an accurate online calibration. An extension of MI, normalized mutual information (NMI) was introduced by Wang et al. (2012) to register panoramic image and point clouds. Corsini et al. (2013) used a global refinement algorithm based on MI to optimize the color projection of aligned photos on the 3D object. The definition of mutual information can significantly affect registration result. When computing the MI, Yipu et al. (2016) set weights to different data sources according to different scenes, which improved the robustness of registration. However, the correlation between depth information and color information on which MI-based methods are built is not significant in many scenes (Wang et al., 2012).

Multi-view-based methods: This kind of method generates image point clouds from adjacent images using SfM and multi-view stereo (MVS), transforming the 2D-3D registration to a 3D-3D registration. The point clouds generated by photogrammetry are influenced by drift, leading to non-rigid transformation between photogrammetry point clouds and laser scanning data.

Generally, the complexity of urban environment scenes poses challenges for robust registration, especially for the registration methods based on points and linear features, which have difficulties in finding correspondences between panoramic images and point clouds. To overcome the difficulties, parked vehicles are proposed as registration features and corresponding vehicle candidates are paired automatically to fulfill the registration. The advantages of using parked vehicles are that they are often found in an urban environment and are relatively easy to be recognized in images and point clouds. Moreover, street-side vehicles in both panoramic images and point clouds are foreground objects, which are rarely occluded by other objects, meaning reliable detection and segmentation can be expected.

3. Methodology

This aim is to accurately estimate the transformation parameters, i.e. rotation and translation parameters, between the panoramic camera and the laser scanner. The proposed method consists of three main steps: (1) Accurate rotation parameters estimation by GPS/IMU aided SfM bundle adjustment. (2) Parked vehicle extraction from point clouds and panoramic images. (3) Accurate translation parameters estimation by maximizing the overlapping area of corresponding pairs. The key steps of the proposed method are illustrated in Fig. 1 and elaborated as follows.

3.1. Notations and definitions for the registration model of panoramic images and point clouds

We employ the following notations proposed by Barfoot (2016): $\mathcal{F}_A$ denotes a reference frame A; a point P in frame $\mathcal{F}_A$ is written as a vector $\boldsymbol{r}_A^{PA}$. $\mathbf{T}_{AB}$ denotes the transformation between $\mathcal{F}_A$ and $\mathcal{F}_B$, and the rotation matrix is represented by $\mathbf{C}_{AB}$; the corresponding quaternion is represented by $\mathbf{q}_{AB}$.

The coordinate system of the proposed registration model is shown in Fig. 2. $\mathcal{F}_W$ denotes the reference frame of the point cloud, $\mathcal{F}_{C_k}$ denotes the reference frame of the k^{th} panoramic image, and $\mathcal{F}_{\tilde{C}_k}$ denotes the reference frame with the original EoPs of the k^{th} panoramic image. The original EoPs of panoramic images are calculated by the pre-calibrated transformation with respect to the laser scanner, and are not accurate due to unforeseen sensor movements.

The EoPs of the k^{th} panoramic image is written as $\mathbf{x}_{C_k}$:

$$\mathbf{x}_{C_k} := [\mathbf{r}_W^{C_k W^T}, \mathbf{q}_{WC_k}^T]^T \in \mathbb{R}^3 \times SO^3 \quad (1)$$

where SO^3 is the Special Orthogonal Group. The inaccurate EoPs of the k^{th} panoramic image is written as $\mathbf{x}_{\tilde{C}_k}$: