ISPRS Journal of Photogrammetry and Remote Sensing 113 (2016) 43–58

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Automatic registration of large-scale urban scene point clouds based on semantic feature points

Bisheng Yang [a,b,*], Zhen Dong [a,b,*], Fuxun Liang [a], Yuan Liu [a]

[a] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China
[b] Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan 430079, China

ARTICLE INFO

Article history:
Received 9 July 2015
Received in revised form 15 December 2015
Accepted 16 December 2015
Available online 18 January 2016

Keywords:
Laser scanning point clouds
Point cloud segmentation
Semantic feature point extraction
Urban scene mapping
Multi-view registration

ABSTRACT

Point clouds collected by terrestrial laser scanning (TLS) from large-scale urban scenes contain a wide variety of objects (buildings, cars, pole-like objects, and others) with symmetric and incomplete structures, and relatively low-textured surfaces, all of which pose great challenges for automatic registration between scans. To address the challenges, this paper proposes a registration method to provide marker-free and multi-view registration based on the semantic feature points extracted. First, the method detects the semantic feature points within a detection scheme, which includes point cloud segmentation, vertical feature lines extraction and semantic information calculation and finally takes the intersections of these lines with the ground as the semantic feature points. Second, the proposed method matches the semantic feature points using geometrical constraints (3-point scheme) as well as semantic information (category and direction), resulting in exhaustive pairwise registration between scans. Finally, the proposed method implements multi-view registration by constructing a minimum spanning tree of the fully connected graph derived from exhaustive pairwise registration. Experiments have demonstrated that the proposed method performs well in various urban environments and indoor scenes with the accuracy at the centimeter level and improves the efficiency, robustness, and accuracy of registration in comparison with the feature plane-based methods.

1. Introduction

Terrestrial laser scanning (TLS) captures three-dimensional (3D) point clouds with high flexibility and precision and is widely used for various applications such as 3D model reconstruction (Vosselman et al., 2004; Pu and Vosselman, 2009), cultural heritage management (Guarnieri et al., 2006; Yang and Zang, 2014; Montuori et al., 2014), forest surveys (Kankare et al., 2013; García et al., 2015), landslide monitoring (Prokop and Panholzer, 2009), and urban planning (Pieraccini et al., 2006). TLS has line-of-sight instruments and limited measurement range, multiple scans from different viewpoints are needed for full coverage of a large urban environment. Because each scan refers to its own local coordinate reference, a registration step transforming all scans into a uniform coordinate reference system has to be carried out.

Point clouds collected by terrestrial laser scanning (TLS) from large-scale urban scenes contain a wide variety of objects (buildings, cars, pole-like objects, and others) with symmetric and incomplete structures, and relatively low-textured surfaces, all of which pose great challenges for automatic registration between scans. To address these challenges, many scientific studies have been carried out to register point clouds in urban scenes, mostly with a coarse to fine strategy (Salvi et al., 2007; Guo et al., 2014; Restrepo et al., 2014). For fine registration, the iterative closest point (ICP) algorithm and its variants, described in Besl and McKay (1992) and Rusinkiewicz and Levoy (2001), are the commonly used standard approaches (Weinmann et al., 2011). However, the ICP algorithm may converge to a local minimum or even be non-convergent without a priori alignment of the point clouds. Therefore, coarse registration methods are needed to search for good initial alignments between scans. Automatic coarse registration methods can be classified into three categories: feature point-based methods, feature line-based methods, and feature plane-based methods, according to the kind of features used for registration.

* Corresponding authors at: State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China.
E-mail addresses: bshyang@whu.edu.cn (B. Yang), dongzhenwhu@whu.edu.cn (Z. Dong).

http://dx.doi.org/10.1016/j.isprsjprs.2015.12.005
0924-2716/

In general, point features are most commonly used for point cloud registration. Böhm and Becker (2007) explored the application of the SIFT method on the automatic marker-free registration of TLS data, by using SIFT key points extracted from the reflectance data of the scans. Barnea and Filin (2008) proposed a key point-based autonomous registration method, which extracts the key points from a panoramic range image and establishes the correspondences between them using 3D Euclidean distance. Rusu et al. (2008) estimated a set of robust 16D features to describe the geometry of each point feature locally to find an initial alignment. Experiments have shown that the algorithm is invariant to pose and sampling density and can cope well with noisy laser scans. Weinmann et al. (2011) extracted characteristic two-dimensional (2D) points based on SIFT features and used radiometric and geometric information on 3D conjugate points to estimate the transformation parameters between two adjacent scans. The proposed approach was successfully applied to a benchmarked data set, resulting in fast and accurate estimation of the transformation parameters. Theiler et al. (2014) proposed a marker-free registration scheme that matches the extracted DoG (difference-of-Gaussians) or Harris key points using 4PCS (4-point congruent sets). Weber et al. (2015) proposed a pipeline to register unordered point clouds acquired by Kinect sensors automatically using fast point-feature histograms (FPFH) as described in Rusu et al. (2008) for initial alignment between two adjacent scans.

Many studies using line or plane features to register point clouds have also been widely reported. Stamos and Leordeanu (2003) proposed an autonomous registration method based on line features, which extracts the intersection lines of neighboring planes and calculates the transformation parameters between adjacent scans using at least two corresponding line pairs. Habib et al. (2005) also reported a methodology for registering photogrammetric and LiDAR data using 3D straight-line features instead of points. Yang and Zang (2014) used spatial curves as matching primitives to calculate the initial transformation parameters between the scanned point clouds of freeform surfaces (e.g., statues, cultural heritage artifacts).

Dold and Brenner (2006) presented a registration method based on plane patches that calculates rotation and translation parameters separately between two point clouds using at least three corresponding pairs of planar patches. Von Hansen (2006) presented a method for automatic and marker-free coarse registration of TLS data, which calculates the coarse parameters between two point clouds based on single plane correspondences. Theiler and Schindler (2012) developed a registration approach based on virtual tie points generated by intersecting triples of planes. Virtual tie points are matched using their descriptors, such as one divided by the condition number, the intersection angles between planes, the extent of planar segments, and the smoothness of planes.

Although the reported methods are generally able to align TLS point clouds based on the corresponding points, lines, or planes, they still have more or less difficulties in dealing with data sets from large-scale urban scenes. Most of the feature point-based methods are more sensitive to point densities and noises in comparison with the feature line-based and feature plane-based methods. The feature line-based methods take advantage only of lines derived from buildings and rarely use line features derived from pole-like objects, these methods may have difficulty dealing with data sets in the suburbs with few buildings. The feature plane-based methods requires more stringent overlapping conditions and works well in cases there are at least three corresponding pairs of planar patches between neighboring scans. Moreover, the semantic attributes (e.g., shape, height, category, direction and so on) of the extracted points, lines and planes are seldom used in the existing methods. To overcome these limitations, this paper proposes a marker-free and multi-view registration method for large-scale urban scene point clouds based on the semantic feature points. The main contribution of the proposed method is to detect a small set of semantic feature points and match them using both geometrical constraints (3-point scheme) and their semantic information (category and direction).

Following this introduction, the key steps of the proposed method will be elaborated. Then the proposed method will be validated in experimental studies, after which conclusions will be drawn.

2. Automatic registration of large-scale urban scene point clouds

The proposed method provides a marker-free and multi-view registration method for large-scale urban scene point clouds based on semantic feature points extracted. The proposed method consists of two key steps: semantic feature point extraction and point cloud registration. First, the method detects the semantic feature points within a detection scheme, which consists of three sequent steps: point cloud segmentation, vertical feature lines extraction (lines derived from pole-like objects and vertical planes) and semantic information calculation (category and direction) and finally takes the intersections of the vertical feature lines with the ground as the semantic feature points. Second, the proposed method matches the semantic feature points using geometrical constraints (3-point scheme) as well as semantic information (category and direction), and eliminates the mismatches by geometric consistency tests to calculate pairwise transformations; next, it implements multi-view registration by constructing a minimum spanning tree of the fully connected graph derived from exhaustive pairwise transformations. The framework of the proposed method is illustrated in Fig. 1.

2.1. Semantic feature points extraction

The pipeline of the proposed semantic feature point extraction method consists of three sequent steps: point cloud segmentation (Section 2.1.1), vertical feature line extraction (Section 2.1.2) and semantic information calculation (Section 2.1.3).

2.1.1. Point cloud segmentation

Segmenting point clouds into meaningful segments is a precondition of vertical feature line extraction. The proposed method segments point cloud using the following steps. First, ground points

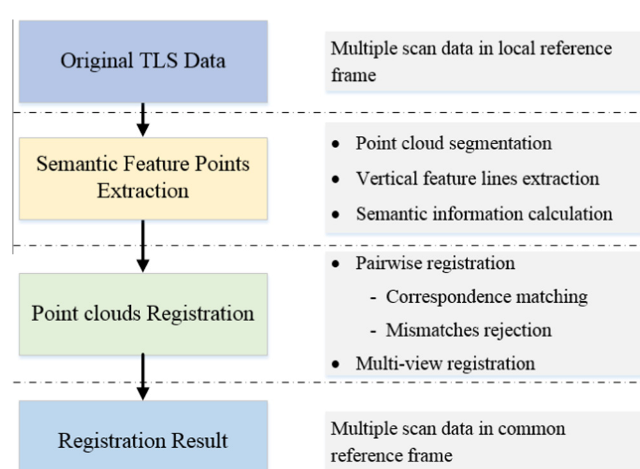

Fig. 1. Point cloud registration framework.