ISPRS Journal of Photogrammetry and Remote Sensing 109 (2015) 126–138

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Globally consistent registration of terrestrial laser scans via graph optimization

Pascal Willy Theiler *, Jan Dirk Wegner, Konrad Schindler

Photogrammetry and Remote Sensing, ETH Zurich, Switzerland

ARTICLE INFO

Article history:
Received 10 June 2015
Received in revised form 18 August 2015
Accepted 25 August 2015

Keywords:
Point cloud registration
Terrestrial laser scanning
Geometric matching
Global consistency
Loop closure
Energy minimization

ABSTRACT

In this paper we present a framework for the automatic registration of multiple terrestrial laser scans. The proposed method can handle arbitrary point clouds with reasonable pairwise overlap, without knowledge about their initial orientation and without the need for artificial markers or other specific objects. The framework is divided into a coarse and a fine registration part, which each start with pairwise registration and then enforce consistent global alignment across all scans. While we put forward a complete, functional registration system, the novel contribution of the paper lies in the coarse global alignment step. Merging multiple scans into a consistent network creates loops along which the relative transformations must add up. We pose the task of finding a global alignment as picking the best candidates from a set of putative pairwise registrations, such that they satisfy the loop constraints. This yields a discrete optimization problem that can be solved efficiently with modern combinatorial methods. Having found a coarse global alignment in this way, the framework proceeds by pairwise refinement with standard ICP, followed by global refinement to evenly spread the residual errors.

The framework was tested on six challenging, real-world datasets. The discrete global alignment step effectively detects, removes and corrects failures of the pairwise registration procedure, finally producing a globally consistent coarse scan network which can be used as initial guess for the highly non-convex refinement. Our overall system reaches success rates close to 100% at acceptable runtimes < 1 h, even in challenging conditions such as scanning in the forest.

1. Introduction

Static terrestrial laser scanners (TLS) are widely used to acquire dense 3D point clouds for various applications in geo-sciences, robotics, entertainment, archaeology, and many more. As laser scanning is a line-of-sight technology, multiple scans from different viewpoints are usually necessary to cover the geometry of a 3D scene. A prerequisite for further processing is thus to align all individual scans in a common coordinate system, to obtain one large point cloud of the complete scene. In practice, scan registration is at present mostly based on artificial markers (e.g., retro-reflective cylinders, spheres, checker-board targets) that are placed in the scene during scan acquisition. These markers are extracted from different scans either manually or automatically (e.g. Akca, 2003; Franaszek et al., 2009), and used to determine the desired 6DoF rigid body transformations.[1] Marker-based scan registration is very reliable, but has several disadvantages. Positioning the markers is time consuming and requires careful planning, to ensure visibility of the markers in different scans, and to avoid degenerate constellations. Additionally, one must ensure that all markers remain stable for the length of the measurement campaign. Finally, markers occlude (small) parts of the scene and often have to be removed from the data for further analysis or visualization.

In order to circumvent markers altogether, this paper presents a method for marker-less TLS point cloud registration based on natural 3D keypoints. The method can handle multiple scans acquired from arbitrary positions, without any prior knowledge about their relative orientations. Compared to the generic problem of aligning point clouds, specific challenges of TLS are (i) Near-field bias: the polar measurement principle of static TLS causes a quadratic decrease of the point density with increasing distance from the

* Corresponding author.
E-mail addresses: pascal.theiler@geod.baug.ethz.ch (P.W. Theiler), jan.wegner@geod.baug.ethz.ch (J.D. Wegner), schindler@geod.baug.ethz.ch (K. Schindler).

[1] TLS directly measures absolute distances, thus only 3D translations and rotations must be estimated.

http://dx.doi.org/10.1016/j.isprsjprs.2015.08.007

sensor. This results in a very uneven point density (respectively resolution) within a scan. (ii) Outliers: real point clouds are contaminated by gross errors, e.g. due to moving objects, reflections or complicated surfaces such as vegetation.[2] (iii) Sheer point cloud size: current TLS operate at frequencies up to 1 million points per second. Scan projects featuring several 100 million points are common and have to be handled in reasonable time.

In most cases TLS registration is divided into an initial *coarse alignment* and a subsequent *fine registration*. Coarse alignment uses only a sparse set of corresponding points whereas fine registration, initialized with the coarse solution, typically uses a much larger subset or even all points. Fine registration amounts to minimizing the point-to-point or point-to-surface distances between the point clouds, usually with the *Iterative Closest Point* algorithm (ICP; Besl and McKay, 1992; Chen and Medioni, 1992) or some variant of it (e.g. Bergevin et al., 1996; Bae and Lichti, 2004; Minguez et al., 2006; Bouaziz et al., 2013). These methods locally minimize the sum of per-point residuals, which is conceptually straight-forward, but highly non-convex, and notoriously prone to converge to weak local minima. Arguably the more critical part of the pipeline is *coarse registration*, which aims to provide a rough initial transformation that is good enough as starting value for ICP, so that the latter reaches a useful minimum. An interesting one-step approach without separate coarse registration has been introduced by Yang et al. (2013). They propose to globally solve the non-convex objective function with a branch-and-bound scheme. The resulting *Globally Optimal ICP* (Go-ICP) works very well with small point clouds, but is intractable for large point clouds like those produced by TLS.

In this paper we describe a *complete framework for fully automated, marker-less registration* of multiple TLS point clouds acquired from arbitrary positions. The framework follows the common procedure to split the registration into coarse initialization and fine registration, and uses existing methods for the latter – in particular a combination of standard pairwise ICP and the global refinement algorithm of Lu and Milios (1997), from now on called LUM.

The main methodological novelty of the proposed framework is an efficient way to find a coarse alignment that is *consistent across all overlapping scans* in a project. Note that this step, while often overlooked, is crucial for the entire registration pipeline: for a successful refinement with local methods the initial alignment must be roughly correct for *all* scans. In other words a possibly quite inaccurate, but reasonably consistent network must be established *before* fine registration. Statistically speaking, coarse registration does not need to have high accuracy, but it must have high reliability. Consequently, we design a scheme that exploits the available redundancy already at this stage, but without the computational cost associated with high accuracy.

Our global alignment method assumes that some *hypothesis generator* is available, which returns putative pairwise alignments between two scans. I.e., for any two scans with reasonable overlap, the hypothesis generator can find a set of relative transformations that are plausible in the light of the observed points. We start from our previous work on pairwise alignment with *Keypoint-based 4-Points Congruent Sets* (K-4PCS, Theiler et al., 2014a), but the framework is generic and can use any other hypothesis generator instead. Given a discrete set of putative alignments between scan pairs, we develop a graph-based energy minimization scheme which selects one of the putative transformations for each pair in such a way that the total loop closure error across all scan pairs in the project is as low as possible. The formulation allows for an efficient solution with combinatorial optimization algorithms. The global coarse registration, in conjunction with standard fine registration algorithms, yields correct scan alignment ($< 3\%$ failures) without any manual intervention across a range of applications scenarios (urban, forest, archaeological sites, indoors). An open-source implementation of the complete registration pipeline will be made available in conjunction with the paper.

[2] Note, the fact that TLS is a rather high-end sensor does not mean that it is less prone to outliers. Low-cost alternatives like robotic scanners or even the KINECT will usually have higher random noise and systematic errors, but not necessarily more gross errors.

2. Related work

2.1. Pairwise coarse registration

Coarse, pairwise point cloud registration typically has two consecutive parts. First, raw point clouds are reduced to sets of sparse features (a.k.a. keypoints) and, second, correspondences are sought. Transformation parameters based on groups of 3 or more corresponding features are estimated, which establish the relative orientation[3] of both point clouds in a common reference frame.

Typically, salient geometric entities are used to establish correspondences (although there are methods which randomly select a small set of features, e.g. Masuda and Yokoya, 1995; Barnea and Filin, 2008; Leng et al., 2014). The most popular features are either 3D keypoints, or planar surfaces, respectively surface normals.

A straight-forward approach is to leverage the power of interest point extraction in conventional images. To that end, 2D interest points are extracted from intensity or range images and lifted to 3D with the known range measurement. E.g., Böhm and Becker (2007) detect SIFT features (Lowe, 2004) in intensity images to derive pixel correspondences and estimate the rigid-body transformation from the associated 3D points. Kang et al. (2009) propose a similar approach, but add an outlier detection step based on 3D distances between putative feature point pairs.

3D keypoint detectors are often direct extensions of standard 2D methods like *Differences-of-Gaussians* (DoG) in 3D space (e.g. Allaire et al., 2008; Flitton et al., 2010). Another 3D keypoint extractor is adapted from the *Harris* corner detector (Harris and Stephens, 1988) by replacing image gradients with point normals (e.g. Sipiran and Bustos, 2011). Feature points are often encoded with descriptors which serve to measure similarity when searching for correspondences. For example, Flint et al. (2007) introduce a 3D version of the SURF descriptor (Bay et al., 2006) to compare keypoints in range images.

Descriptors specifically for 3D point cloud data typically describe a point's neighbourhood by histograms of the point distribution and/or the variability of the normals. Perhaps the first instance of this idea are *Spin Images* (Johnson and Hebert, 1999), more recent versions include *Point Feature Histograms* (PFH; Rusu et al., 2008) and their accelerated version (FPFH; Rusu et al., 2009). Overviews on the state-of-the-art keypoint detection and matching in 3D space, as well as performance evaluations, can be found in (Tombari et al., 2013; Hänsch et al., 2014).

An alternative strategy, which is however limited to man-made environments, relies on surfaces rather than keypoints, such as planes (Dold and Brenner, 2006; Brenner et al., 2008; Theiler and Schindler, 2012) or other geometrical primitives (e.g., spheres, cones) as in Rabbani et al. (2007). A more robust variant does not detect explicit planes, but uses salient directions of the (point-based) surface normals for alignment (Makadia et al., 2006; Zeisl et al., 2013; Novák and Schindler, 2013).

[3] In this paper we follow the terminology of photogrammetry: "relative orientation" refers to the full transformation between two instrument coordinate systems, including both the rotation *and* the translation.