TEASER: Fast and Certifiable Point Cloud Registration

Heng Yang, Jingnan Shi, Luca Carlone

***Abstract*—We propose TEASER, the first fast and certifiable algorithm, for 3D point cloud registration with large amounts of outlier correspondences. We decouple scale, rotation, and translation estimation, and adopt a *Truncated Least Squares* (TLS) formulation for each subproblem. Despite being non-convex and combinatorial, we show that (i) TLS scale and translation estimation can be solved exactly in polynomial time via *adaptive voting*, (ii) TLS rotation estimation can be solved by a *tight semidefinite programming (SDP) relaxation*, with a *certifiable* global optimality guarantee. We also develop a second certifiable algorithm, named TEASER++, that circumvents solving an SDP and runs in *milliseconds*. We provide theoretical bounds on the estimation errors for both algorithms. Experiments show that both algorithms dominate the state of the art and are robust against 99% outliers. We release a fast open-source C++ implementation of TEASER++ at `https://github.com/MIT-SPARK/TEASER-plusplus`.**

I. Introduction

Point cloud registration is a fundamental problem in robotics and computer vision [1, 2, 3, 4]. It consists in finding the best transformation (rotation, translation and potentially scale) that aligns two point clouds.

A typical approach consists of extracting and matching features in the two point clouds, and computing the transformation that aligns corresponding features. When the correspondences found are correct, the registration problem can be solved with closed-form solutions [5, 6]. In practice, the correspondences may contain many outliers, leading these solvers to produce poor estimates [7, 8].

This paper presents an overview of two recently proposed certifiable algorithms for 3D registration with outliers. A full technical description is given in [9]. By *certifiable algorithm*, we mean an algorithm that attempts to solve an intractable problem (e.g., robust estimation with outliers) and provides checkable conditions on whether it succeeded. The first algorithm, TEASER, is accurate and robust but requires solving a large Semidefinite Program (SDP). The second algorithm, TEASER++, has similar performance in practice but circumvents the need to solve an SDP and can run in milliseconds. Both algorithms formulate the problem using a *Truncated Least Squares* (TLS) cost, and use a graph-theoretic framework to decouple scale, rotation, and translation estimation that also allows pruning outliers by finding a *maximum clique*. We develop theoretical results certifying the quality of the solutions returned by our algorithms. Moreover, we show that the proposed algorithms dominate the state of the art in terms of robustness, accuracy and speed.

H. Yang, J. Shi, and L. Carlone are with the Laboratory for Information & Decision Systems (LIDS), Massachusetts Institute of Technology, Cambridge, MA 02139, USA, Email: {hankyang,jnshi,lcarlone}@mit.edu

II. Truncated least squares Estimation And Semidefinite Relaxation (TEASER) Overview

A. Robust Registration with Truncated Least Squares Cost

In the robust registration problem, we are given two 3D point clouds $\mathcal{A} = \{\boldsymbol{a}_i\}_{i=1}^N$ and $\mathcal{B} = \{\boldsymbol{b}_i\}_{i=1}^N$, with $\boldsymbol{a}_i, \boldsymbol{b}_i \in \mathbb{R}^3$. We consider a *correspondence-based* setup, where we are given putative correspondences $(\boldsymbol{a}_i, \boldsymbol{b}_i), i = 1, \ldots, N$, that obey the following generative model:

$$\boldsymbol{b}_i = s^\circ \boldsymbol{R}^\circ \boldsymbol{a}_i + \boldsymbol{t}^\circ + \boldsymbol{o}_i + \boldsymbol{\epsilon}_i, \tag{1}$$

where $s^\circ > 0$, $\boldsymbol{R}^\circ \in \mathrm{SO}(3)$, and $\boldsymbol{t}^\circ \in \mathbb{R}^3$ are the unknown scale, rotation, and translation, $\boldsymbol{\epsilon}_i$ models the measurement noise, and $\boldsymbol{o}_i$ is a vector of zeros if the pair $(\boldsymbol{a}_i, \boldsymbol{b}_i)$ is an *inlier*, or a vector of arbitrary numbers for *outlier correspondences*.

In our formulation, we estimate the unknown scale, rotation and translation using a *Truncated Least Squares (*TLS*)* cost:

$$\min_{s>0, \boldsymbol{R}\in\mathrm{SO}(3), \boldsymbol{t}\in\mathbb{R}^3} \sum_{i=1}^{N} \min\left(\frac{1}{\beta_i^2}\left\|\boldsymbol{b}_i - s\boldsymbol{R}\boldsymbol{a}_i - \boldsymbol{t}\right\|^2,\ \bar{c}^2\right), \tag{2}$$

where β_i is the bound on the i-th measurement noise, and $\bar{c}^2$ is a constant typically chosen to be 1. Essentially, (2) discards outlier measurements with residuals greater than $\bar{c}^2$.

B. Decoupled Scale, Rotation, and Translation Estimation

We propose a general approach to decouple scale, translation, and rotation estimation in problem (2) using quantities that are invariant to a subset of the transformations (scaling, rotation, translation). First, we obtain a *Translation Invariant Measurement* (TIM) from (1) by computing $\bar{\boldsymbol{a}}_{ij} \doteq \boldsymbol{a}_j - \boldsymbol{a}_i$ and $\bar{\boldsymbol{b}}_{ij} \doteq \boldsymbol{b}_j - \boldsymbol{b}_i$, and the TIM satisfies the following generative model:

$$\bar{\boldsymbol{b}}_{ij} = s\boldsymbol{R}\bar{\boldsymbol{a}}_{ij} + \boldsymbol{o}_{ij} + \boldsymbol{\epsilon}_{ij}, \tag{TIM}$$

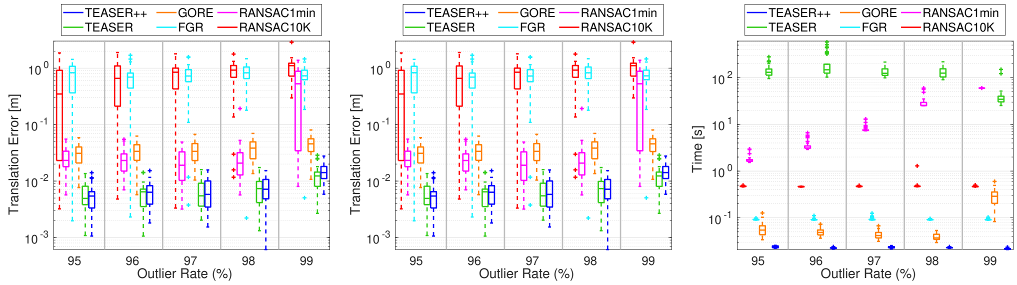
where $\boldsymbol{o}_{ij} \doteq \boldsymbol{o}_j - \boldsymbol{o}_i$ is zero if *both* the i-th *and* the j-th measurements are inliers (or arbitrary otherwise), while $\boldsymbol{\epsilon}_{ij} \doteq \boldsymbol{\epsilon}_j - \boldsymbol{\epsilon}_i$ is the measurement noise. If $\|\boldsymbol{\epsilon}_i\| \leq \beta_i$ and $\|\boldsymbol{\epsilon}_j\| \leq \beta_j$, then $\|\boldsymbol{\epsilon}_{ij}\| \leq \beta_i + \beta_j \doteq \delta_{ij}$.

Secondly, we define *Translation and Rotation Invariant Measurement* (TRIM) as $s_{ij} = \frac{\|\bar{\boldsymbol{b}}_{ij}\|}{\|\bar{\boldsymbol{a}}_{ij}\|}$ which are invariant to both $\boldsymbol{R}$ and $\boldsymbol{t}$:

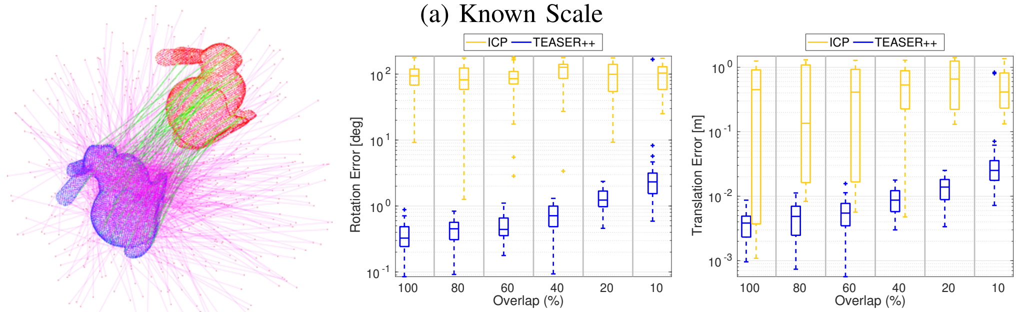
$$s_{ij} = s + o_{ij}^s + \epsilon_{ij}^s, \tag{TRIM}$$

where $\epsilon_{ij}^s \doteq \frac{\tilde{\epsilon}_{ij}}{\|\bar{\boldsymbol{a}}_{ij}\|} \leq \frac{\delta_{ij}}{\|\bar{\boldsymbol{a}}_{ij}\|}$, and $o_{ij}^s = 0$ if both i and j are inliers and arbitrary scalar otherwise. Notice that $|\epsilon_{ij}^s| \leq \delta_{ij}/\|\bar{\boldsymbol{a}}_{ij}\|$ since $|\tilde{\epsilon}_{ij}| \leq \delta_{ij}$.

Using TRIMs and TIMs, we are able to estimate scale, rotation and translation in a cascade, using TLS estimation. To estimate scale, we propose an adaptive voting algorithm that enumerates all possible consensus sets in polynomial time, and selects the one with the lowest TLS cost.



(a) Known Scale



(b) Correspondence-free

Fig. 1. Benchmark results. *(a) from left to right*: Boxplots of rotation errors, translation errors, and timing for the six compared methods on the Bunny dataset with known scale. *(b) from left to right*: An example with 95% outlier correspondences,boxplots of rotation and translation errors for TEASER++ and ICP on the Bunny dataset with the correspondence-free problem setup.

Scenes	Kitchen (%)	Home 1 (%)	Home 2 (%)	Hotel 1 (%)	Hotel 2 (%)	Hotel 3 (%)	Study Room (%)	MIT Lab (%)	Avg. Runtime [ms]
RANSAC-1K	90.9	91.0	73.1	88.1	80.8	87.0	79.1	81.8	9.7
RANSAC-10K	96.4	92.3	73.1	92.0	84.6	90.7	82.2	81.8	96.5
TEASER++	97.7	92.3	82.7	96.9	88.5	94.4	88.7	84.4	85.1
TEASER++ (CERT)	99.2	97.5	90.0	98.8	94.9	97.7	94.8	93.9	-

TABLE I

PERCENTAGE OF CORRECT REGISTRATION RESULTS USING TEASER++, TEASER++ CERTIFIED, AND RANSAC ON THE *3DMatch* DATASET.

To estimate rotation, we formulate the TLS rotation estimation problem as a a Quadratically Constrained Quadratic Program (QCQP). We then develop a tight convex semidefinite programming (SDP) relaxation. We further develop a fast global optimality certification algorithm that computes a sub-optimality gap for any rotation estimate.

After obtaining the scale and rotation estimates, we can estimate the translation $\boldsymbol{t}$ component-wise by substituting them back into problem (2).

For a discussion of the theoretical bounds on the estimation errors, please refer to our full technical description in [9].

III. TEASER++: A FAST C++ IMPLEMENTATION

We have also developed a fast C++ implementation of TEASER, named TEASER++. TEASER++ has been released as an open-source library and can be found at `https://github.com/MIT-SPARK/TEASER-plusplus`. TEASER++ follows the same decoupled approach described in Section II, except that it circumvents solving a large-scale SDP by using the GNC approach described in [10] for TLS rotation estimation.

IV. EXPERIMENTAL RESULTS

We benchmark TEASER and TEASER++ against two state-of-the-art robust registration techniques: *Fast Global Registration* (FGR) [11] and *Guaranteed Outlier REmoval* (GORE) [12]. In addition, we test two RANSAC variants: a fast version where we terminate RANSAC after a maximum of 10,000 iterations (RANSAC10K) and a slow version where we terminate RANSAC after 60 seconds (RANSAC1min). Fig. 1 shows the benchmarking results for the Stanford Bunny dataset [13]. TEASER, TEASER++, and GORE are robust against up to 99% outliers, while RANSAC1min with 60s timeout can resist 98% outliers with about 10^6 iterations. RANSAC10K and FGR perform poorly under extreme outlier rates. While GORE, TEASER and TEASER++ are both robust against 99% outliers, TEASER, and TEASER++ produce lower estimation errors, with TEASER++ being one order of magnitude faster than GORE.

In addition, we tested the performance of TEASER++ on the 3DMatch dataset [14]. We use 3DSmoothNet [15] to compute descriptors, and generate correspondences using nearest-neighbor matching. We then feed the correspondences to TEASER++ and RANSAC with 1K and 10K iterations (using Open3D [16]) and compare their performances in terms of percentages of successfully matched scans and runtime (Table I). Two scans are successful matched when the transformation computed by a technique has (i) rotation error smaller than $10°$, and (ii) translation error less than $30\,\mathrm{cm}$. TEASER++ dominates all RANSAC variants in success rates. The last row in Table I shows the success rate for the poses certified as optimal by TEASER++. The success rate strictly dominates both RANSAC variants and TEASER++, since TEASER++ (CERT) is able to identify and reject unreliable registration results.

References

[1] P. Henry, M. Krainin, E. Herbst, X. Ren, and D. Fox, "Rgb-d mapping: Using kinect-style depth cameras for dense 3d modeling of indoor environments," *Intl. J. of Robotics Research*, vol. 31, no. 5, pp. 647–663, 2012.

[2] G. Blais and M. D. Levine, "Registering multiview range data to create 3d computer objects," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, no. 8, pp. 820–824, 1995.

[3] A. Zeng, K. T. Yu, S. Song, D. Suo, E. Walker, A. Rodriguez, and J. Xiao, "Multi-view self-supervised deep learning for 6d pose estimation in the amazon picking challenge," in *IEEE Intl. Conf. on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 1386–1383.

[4] M. A. Audette, F. P. Ferrie, and T. M. Peters, "An algorithmic overview of surface registration techniques for medical imaging," *Med. Image Anal.*, vol. 4, no. 3, pp. 201–217, 2000.

[5] B. K. P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *J. Opt. Soc. Amer.*, vol. 4, no. 4, pp. 629–642, Apr 1987.

[6] K. Arun, T. Huang, and S. Blostein, "Least-squares fitting of two 3-D point sets," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 9, no. 5, pp. 698–700, sept. 1987.

[7] F. Tombari, S. Salti, and L. D. Stefano, "Performance evaluation of 3d keypoint detectors," *Intl. J. of Computer Vision*, vol. 102, no. 1-3, pp. 198–220, 2013.

[8] R. B. Rusu, N. Blodow, Z. C. Marton, and M. Beetz, "Aligning point cloud views using persistent feature histograms," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*. IEEE, 2008, pp. 3384–3391.

[9] H. Yang, J. Shi, and L. Carlone, "TEASER: Fast and Certifiable Point Cloud Registration," *arXiv preprint arXiv: 2001.07715*, 2020, (pdf).

[10] H. Yang, P. Antonante, V. Tzoumas, and L. Carlone, "Graduated non-convexity for robust spatial perception: From non-minimal solvers to global outlier rejection," *IEEE Robotics and Automation Letters (RA-L)*, vol. 5, no. 2, pp. 1127–1134, 2020, arXiv preprint arXiv:1909.08605 (with supplemental material), (pdf), .

[11] Q. Zhou, J. Park, and V. Koltun, "Fast global registration," in *European Conf. on Computer Vision (ECCV)*. Springer, 2016, pp. 766–782.

[12] Á. Parra Bustos and T. J. Chin, "Guaranteed outlier removal for point cloud registration with correspondences," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 40, no. 12, pp. 2868–2882, 2018.

[13] B. Curless and M. Levoy, "A volumetric method for building complex models from range images," in *SIGGRAPH*, 1996, pp. 303–312.

[14] A. Zeng, S. Song, M. Nießner, M. Fisher, J. Xiao, and T. Funkhouser, "3dmatch: Learning the matching of local 3d geometry in range scans," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, no. 2, 2017, p. 4.

[15] Z. Gojcic, C. Zhou, J. D. Wegner, and A. Wieser, "The perfect match: 3d point cloud matching with smoothed densities," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 5545–5554.

[16] Q.-Y. Zhou, J. Park, and V. Koltun, "Open3D: A modern library for 3D data processing," *arXiv:1801.09847*, 2018.