Correspondence Rejection by Trilateration for 3D Point Cloud Registration

Kishan Lachhani, Jifang Duan, Hadi Baghsiahi, Eero Willman, David R. Selviah
Department of Electronic and Electrical Engineering, University College London (UCL)
Torrington Place, London, WC1E 7JE
kishan.lachhani.13@ucl.ac.uk, d.selviah@ucl.ac.uk

Abstract

Recent years have shown increases in virtual 3D perception and applications, many of these applications require 3D model reconstruction from high quality LIDAR scans. High quality 3D models may be acquired from a collection of overlapping LIDAR scans which need to be registered or aligned to a common coordinate system. This paper investigates the use of a novel implementation of trilateration for correspondence rejection in highly accurate 3D point cloud registration. It is shown that from a synthesized correspondence set of size 100 containing 85% outliers, all or most of the remaining 15% inliers can be retrieved. The trilateration problem is solved for all 4-combinations of correspondence elements from which the true correspondence subsets are easily identifiable. It is also shown that this method's performance may be greatly affected by noisy distance measurements, however the method works well for distance measurements typically acquired by LIDAR systems. Lastly, unnecessarily large sizes of correspondence sets can quickly make the method computationally expensive if all combination subsets require to be evaluated.

1 Introduction

A common need in computer vision and pattern recognition is to compute the 3D rigid body transformations that align two or more point clouds. The process of placing two or more 3D data sets or point clouds in the same coordinate system is a crucial step for the accurate reconstruction of 3D models for applications in architecture, gaming, medical imaging, heritage conservation, robotics and industrial automation.

Point clouds can be captured by a number of technologies, including depth cameras and terrestrial LIDAR scanners. LIDARs can capture scenes in high resolution and also tend to be very accurate (around ± 2.5 mm at 20 m) with some models able to also capture colour. Subsequently such scans can be registered with greater accuracy than what might be achieved with data captured from other less accurate devices. Also, since high-accuracy and high-resolution scanners must be stationary during each scan, their usability and utility rely on accurate registration.

There exist a number of methods to automatically register point clouds, one of the most popular and prominent in the literature is feature-based registration. This process requires the detection of features followed by a description (or feature vector) of each feature. The description allows the estimation of correspondences between multiple scans. Correspondence estimation is typically paired with correspondence rejection to ensure only true corresponding pairs of points are selected for reliable transformation estimation and ultimately, robust and accurate registration.

Robust descriptor and correspondence estimation algorithms are able to identify corresponding pairs of points captured at different positions and orientations; consequently, this also means that similar but non-corresponding points are likely to be incorrectly matched. This problem is particularly pertinent in indoor computer vision applications where many indistinguishable features such as tiles, switches, lighting and various other features are observed.

The final steps of the registration process can be greatly and negatively impacted by spurious correspondences, and subsequently rely heavily on correspondence rejection algorithms. Unlike feature extraction and description, there are only a handful of methods commonly implemented for correspondence rejection. Popular methods such as RANSAC [1] are discussed below and a novel implementation of trilateration for correspondence rejection is also proposed.

1.1. Correspondence Rejection

Correspondence rejection is a vital step in the registration process and there exist a number of commonly used methods [2], [3]. Distance-based rejection removes correspondences which are beyond a certain distance; trimmed correspondences retain a percentage of the best determined correspondences dependent on the metric of the descriptor and reciprocal rejection retains only those correspondences for which the query point is the best match for the matching point. Although these methods remove correspondences of poor quality, they do not actively remove incorrectly matched correspondences (outliers), and undesirably they may also remove inliers. Active outlier removal can be achieved using robust estimators.

Robust estimation can accurately compute model parameters from a data set containing a significant portion of outliers. The RANSAC algorithm is possibly the most widely used robust estimator in computer vision; it is an iterative and non-deterministic method in which the results are agreed by a certain probability which increases with iterations.

In general, RANSAC requires a certain model for which model parameters are estimated and a cost function is evaluated to determine the quality of the fitted model. The model used by RANSAC for correspondence rejection in 3D point cloud registration is a transformation estimation (translation and rotation) which can be efficiently evaluated by a number of closed form solutions [4]–[7] whereby the

cost is determined by the residual distance between corresponding points.

In this paper we explore the use of trilateration for correspondence rejection. Trilateration is typically used to solve localisation problems for positioning systems, navigation and tracking, but here we extend this for the application of correspondence rejection.

2 Trilateration

Trilateration is the process of determining the position of a point by distance measurements using the geometry of spheres in this case; there exist a number of methods to robustly estimate this position [8]–[11].

If it is known that a point lies of the surface of three spheres, the centres and the radii provide sufficient information to narrow possible locations down to no more than two (unless the centers lie on a straight line). A fourth sphere or other conditional information can be used to determine which of the two possible positions the correct location is.

A LIDAR scans a scene from two locations, M and D, which are respectively called scan M and scan D. Scan M and D also have feature points m_i and d_i where i is the index of the feature point. Features are extracted from the two scans and correspondences are identified. The distance from the scanner to each feature point is calculated/known from scan M.

For each distance measure evaluated, a sphere with a radius of that distance is considered for the corresponding points in scan D. The position of M, in relation to D, is then determined by the intersection of the four spheres.

The equation of a sphere is given by (1), where (x_0, y_0, z_0) define the centre of the sphere and r is the radius of the sphere.

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = r^2 \qquad (1)$$

In our case the center of a sphere is determined by a feature point d_i which has a corresponding feature point m_i, and the radius is given by the distance between m_i and the origin, m_{org}. Since, the origin of $\{m_i\}$ is arbitrary, for simplicity we choose (0,0,0). This simplifies the squared radius of the i^{th} point to $\|m_i\|^2$. As a result the new equation for a sphere is then given by (2) and illustrated by Figure 1 where the subscripts denote the correspondence/feature index and the x, y or z component.

$$\left(x - d_{i,x}\right)^2 + \left(y - d_{i,y}\right)^2 + \left(z - d_{i,z}\right)^2 = \|m_i\|^2 \qquad (2)$$

There are 3 unknowns (x, y, z), so a system of least 3 equations are required to determine a solution, and at least a system of 4 equations to determine a single solution. In practise, there are many more than 4 constraints. With more than 3 constraints, the system is over-determined; exact solutions can be found, though with the inclusion of noise in real systems, it may be that no exact solutions exist. The alternative is to seek approximate solutions.

The system of N non-linear equations (where N is the number of corresponding features) can be transformed into a system of $N - 1$ linear equations which can be expressed as a simple linear system (3) in which $\mathbf{x}$ is the unknown.

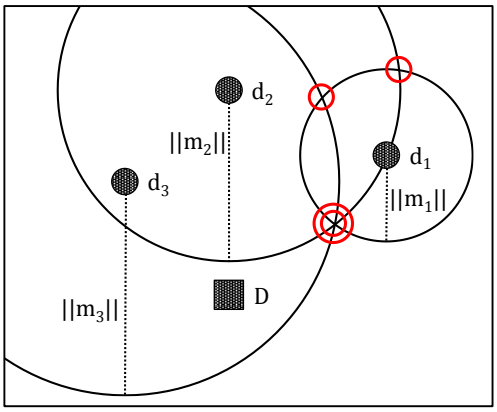

Figure 1. The two-dimensional position of scan M is identified in relation to scan D using a system of three equations for a circle which are similar to (2).

$$\mathbf{Ax} = \mathbf{b} \qquad (3)$$

The system of non-linear equations can be formed into a linear system (planar equations) by subtracting an arbitrarily chosen j^{th} constraint from the i^{th} constraints where $i \neq j$:

$$\begin{aligned} x\left(d_{j,x} - d_{i,x}\right) + y\left(d_{j,y} - d_{i,y}\right) + z\left(d_{j,z} - d_{i,z}\right) \\ = {}^1\!/_2 \left(\|m_i\|^2 - \|m_j\|^2 - d_{i,x}\right. \\ - d_{i,y} - d_{i,z} + d_{j,x} + d_{j,y} \\ \left. + d_{j,z}\right) = r_{i,j} \end{aligned} \qquad (4)$$

With j arbitrarily chosen to be N such that $i = 1 \ldots n$, where $n = N - 1$. The linear system (3) can be expressed by (5)-(7).

$$\mathbf{A} = \begin{pmatrix} d_{N,x} - d_{1,x} & d_{N,y} - d_{1,y} & d_{N,z} - d_{1,z} \\ d_{N,x} - d_{2,x} & d_{N,y} - d_{2,y} & d_{N,z} - d_{2,z} \\ \vdots & \vdots & \vdots \\ d_{N,x} - d_{n,x} & d_{N,y} - d_{n,y} & d_{N,z} - d_{n,z} \end{pmatrix} \qquad (5)$$

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \qquad (6)$$

$$\mathbf{b} = \begin{pmatrix} r_{1,N} \\ r_{2,N} \\ \vdots \\ r_{n,N} \end{pmatrix} \qquad (7)$$

Least squares minimization methods are notoriously ineffective with the inclusion of anomalies. In order to be able to robustly identify the correct position of the scanner, we adopt the methodology of Nishida et al. [10] with slight modifications. We solve for systems of 3 equations (which requires 4 corresponding points where the 4th element of the subset is the j^{th} constraint) as opposed to a single system of n equations.

The solution to each system produces a three-dimensional Cartesian point, this complete collection of points will from here on be referred to as the trilateration set. If a subset contains only true correspondences, the solution of

this system will correctly estimate the position of the scanner and will coincide with other truly corresponding subsets. As more true subsets are evaluated, the accumulation of their results will form a dense cluster. If a subset contains one or more spurious correspondences, its result will likely not coincide with the correct position and will not belong to the dense cluster. Spurious correspondences may easily be identified as those which repeatedly produce points which do not belong to the dense cluster.

Given that there are N correspondences and that subsets of only 4 elements are required, the total number of subsets is given by:

$$\frac{N!}{4!\,(N-4)!} = \frac{1}{24}(N^4 - 6N^3 + 11N^2 - 6N) \qquad (8)$$

A large number of elements in the correspondence set can produce a very large trilateration set and can make this method increasingly computationally expensive. However, with RANSAC being introduced over 30 years ago and computation power exponentially increasing since then, evaluating for all subsets and analyzing the trilateration set in its entirety is feasible and may also produce information not readily available otherwise.

3 An Example

Two randomly generated point sets which share a certain number of identical/cloned points are used as correspondence test data. The cloned points represent true correspondences (inliers) while the rest of the points are considered spurious (outliers).

The random point sets lie within a sphere of radius 5 metres and are generated from a uniform distribution. Additionally, the points from both sets are perturbed by a normally distributed variable $X \sim N(0, 0.005^2)$. A standard deviation of 0.005m used for the test case is an overestimation of terrestrial LIDAR range error which can be up to around 2 or more times smaller. Lastly, the second point cloud is translated by $(-5, -5, -5)$, therefore its position in relation to the first point cloud is $(5,5,5)$.

Trilateration sets are evaluated for varying number of inliers and is plotted as a histogram which intuitively shows the dense cluster formed by the subsets describing only true correspondences as a narrow and tall peak. With a large number of inliers, a very prominent narrow peak is expected.

Additionally, the trilateration set of three correspondence sets of size 20 containing only inliers with varying levels of noise are also evaluated. The three sets are also respectively translated by -10, 0 and 10 in the x-direction. Noisier points sets are expected to have shorter and wider peaks.

4 Results and Discussion

Trilateration sets are evaluated for 30, 15 and 10 inliers out of 100 correspondences. Their respective histograms for the x-component are shown by Figure 2, Figure 3 and Figure 4 (similar results can be observed for y- and z-components).

From the three cases, a distinct peak can be observed for the case of 30 and 15 inliers. The peak directly corresponds to the position of scanner M in the coordinate system of scan D. 30 and 15 are considered to be a relatively small number of inliers, however the density of the cluster formed by the true correspondences can still be overwhelming if the number of inliers in sufficient.

The peak in Figure 2 is significantly more distinct than the peak from Figure 3. Although there are only 15 more inliers, according to (8) the dense cluster in the trilateration set has around 20 times as many points.

As for the case of 10 inliers, no outstanding peak can be distinguished and determining the true set of correspondences would not be possible for this particular case; a distinct peak can be observed for systems with reduced noise.

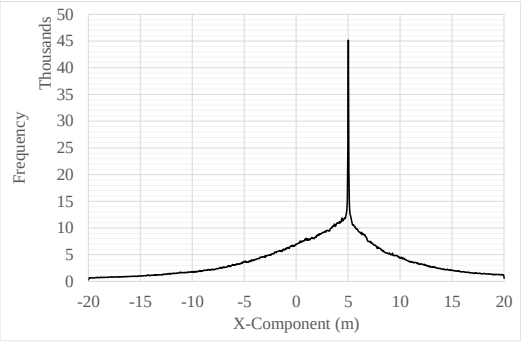

Figure 2. Histogram of x-component of trilateration set with bin width 0.05m. 100 correspondences, 30 inliers with normally distributed noise: $N(0, (0.005)^2)$.

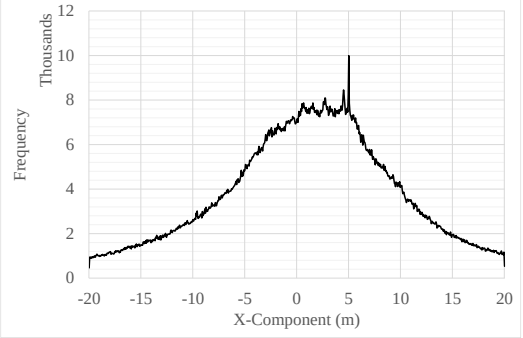

Figure 3. Histogram of x-component of trilateration set with bin width 0.05m. 100 correspondences, 15 inliers with normally distributed noise: $N(0, (0.005)^2)$.

Understandably if RANSAC was used, instead of exhaustively computing of all subsets, the likelihood of selecting a true subset is extremely low and RANSAC iterations may be curtailed before a good model can actually be found. The approach used here to analyze the entire trilateration set in order to determine the correct solution is shown here to be extremely resilient.

Figure 5 shows how the distribution of the trilateration set varies with noise. The consideration of noise in this method is crucial. If the noise is too large, the height of the peak is dramatically reduced which may be difficult to recover if the correspondence set is heavily contaminated.

Furthermore, peaks observed in complete trilateration sets are independent of the spurious correspondences. This

means that the correct position of the scanner may be determined confidently with a sufficiently high number of true correspondences regardless of the number of spurious correspondences.

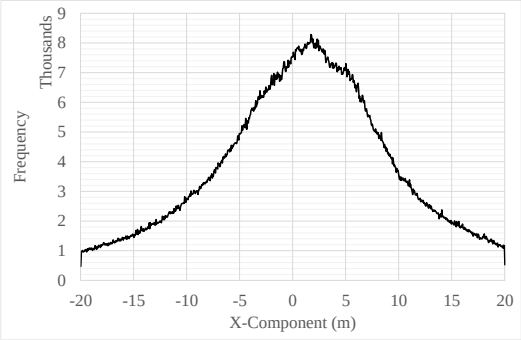

Figure 4. Histogram of x-component of trilateration set with bin width 0.05m. 100 correspondences, 10 inliers with normally distributed noise: $N(0,(0.005)^2)$.

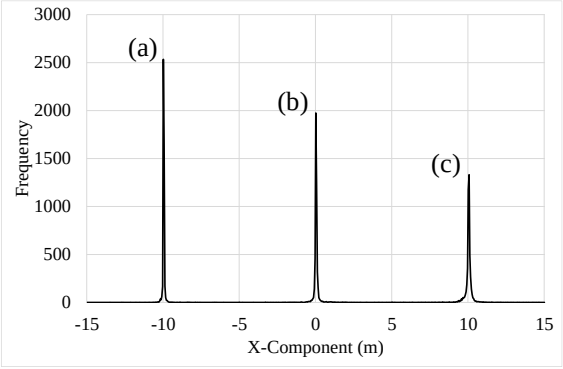

Figure 5. Histogram of x-component of trilateration set with bin width $0.05\,\mathrm{m}$. 3 sets of 20 correspondences with no outliers and with varying levels of noise and translation. From left to right: (a) $N(0,0.0025^2)$, (b) $N(0,0.005^2)$ and (c) $N(0,0.01^2)$.

Additionally, it should be noted that even a small increase in the number of true correspondences can make drastic differences in the trilateration set if the narrow peaks are indeed independent of spurious correspondences. For example, for the case of 20 true correspondences, there exist 4845 possible combinations (8). Increasing the number of true correspondences by only one results in 5985 combinations (8), over 1000 more which would result in a significantly taller peak in the trilateration set.

5 Conclusion

Trilateration for correspondence rejection is shown to be very robust by being able to identify the correct position of the scanner with great confidence when the correspondence set is highly contaminated. Moreover, the method shows promise in scenarios where the correspondence set is contaminated by even greater magnitudes if the number of inliers in sufficiently high. It succeeds in cases where RANSAC is likely to fail due to a very low probability of selecting a true subset.

Two drawbacks of the method are its susceptibility to noise and lengthy computation time if the correspondence set is large (8). For the case of LIDAR scanners, the noise in distance measurements does not pose a serious problem though for other noisier scanners it may. With regards to computation time, a correspondence set of around 180 points produces a trilateration set of around 44 million points which is very similar to the size of a typical scan and so the trilateration set is very manageable. Typically, this many correspondences are more than enough to reasonably estimate transformation.

To improve performance, additional approximate information such as scanner height, distance between scanners and size of the scanned scene may be used to refine the trilateration set to remove obvious outliers without the need for other difficult-to-obtain problem specific thresholds.

The method is still in its early stages and requires further testing and analysis is required to better understand effects of noise and percentage of inliers; identify strengths and weaknesses in comparison to other popular correspondence rejection methods such as RANSAC and to also test with real data.

References

[1] M. A. Fischler and R. C. Bolles, “Random Sample Consensus: A Paradigm for Model Fitting with Applicatlons to Image Analysis and Automated Cartography,” *Commun. ACM*, vol. 24, pp. 381 – 395, 1981.

[2] D. Holz, “Registering point clouds using the Point Cloud Library,” 2011. [Online]. Available: http://www.pointclouds.org/assets/files/presentations/holz_intern_presentation.pdf.

[3] R. B. Rusu, “Point Cloud Library (PCL).” 2011.

[4] K. S. Arun, T. S. Huang, and S. D. Blostein, “Least-Squares Fitting of Two 3-D Point Sets,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-9, no. 5, pp. 698–700, Sep. 1987.

[5] B. K. P. Horn, H. M. Hilden, and S. Negahdaripour, “Closed-form solution of absolute orientation using orthonormal matrices,” *Journal of the Optical Society of America A*, vol. 5. p. 1127, 1988.

[6] B. K. P. Horn, “Closed-form solution of absolute orientation using unit quaternions,” *Journal of the Optical Society of America A*, vol. 4. p. 629, 1987.

[7] M. Walker, L. Shao, and R. Volz, “Estimating 3-D location parameters using dual number quaternions,” *CVGIP image Underst.*, vol. 54, no. 3, pp. 358–367, 1991.

[8] W. Murphy and W. Hereman, “Determination of a position in three dimensions using trilateration and approximate distances,” *Department of Mathematical and Computer Sciences, Colorado School of Mines, Golden, Colorado*, 1995.

[9] W. Navidi, W. S. Murphy, and W. Hereman, “Statistical methods in surveying by trilateration,” *Computational Statistics & Data Analysis*, vol. 27. pp. 209–227, 1998.

[10] Y. Nishida, H. Aizawa, T. Hori, N. H. Hoffman, T. Kanade, and M. Kakikura, “Ultrasonic tagging system for observing human activity,” in Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453), 2003, vol. 1, pp. 785–791.

[11] F. Thomas and L. Ros, “Revisiting trilateration for robot localization,” IEEE Trans. Robot., vol. 21, no. 1, pp. 93–101, Feb. 2005.