

# A Semantic-Based Method for Point Cloud Angular Consistency in Urban Environments

Anton Lohr, Jose Guivant

University of New South Wales, Australia  
anton@unsw.edu.au, j.guivant@unsw.edu.au

## Abstract

Building locally and globally consistent maps from sensor data is an important step in any robotic navigation task. In large 3D maps, incremental error in scan registration can accumulate causing large absolute errors which can hinder global pose optimisation. Urban scenes often contain a large number of vertical surfaces which can be used to improve angular consistency in arbitrarily large maps. We propose a method to align surface normals in 3D scans to an absolute frame of reference, and to distribute this error across the pose history of the robot. We offer a closed form solution to the single-frame angular optimisation, as well as a heuristic method for correcting the pose trajectory. Our results show a reduction in accumulated error in loops before loop closure or global optimisation.

## 1 Introduction

In recent years, unmanned and autonomous vehicles have received a lot of attention in both media and academia. Both teleoperated and autonomous vehicles benefit enormously from generating accurate and consistent maps of their environments for decision making and navigation. Traditionally, 2D range sensors and cameras have been used due to their simplicity and economy, however the increasing computer processing power and decreasing cost of 3D range sensors have made 3D sensing and point cloud mapping an increasingly attractive option.

One popular approach involves using point cloud registration to estimate changes in robot pose and using these pose relations to generate a large map, however, small incremental errors in registration can accumulate and cause large global errors. Loop closure can reduce global error by redistributing registration errors when returning to previously visited locations, however it can

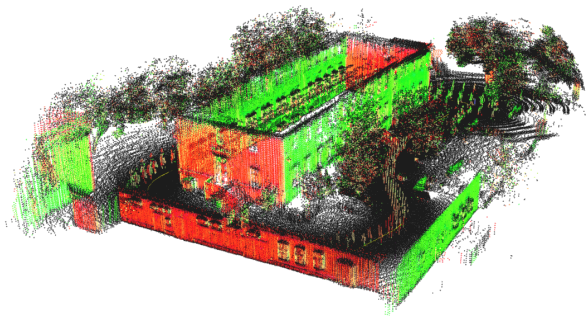


Figure 1: Point Cloud excerpt from the HANNOVER2 dataset with near-vertical surfaces labelled in green and red.

be hard to detect large loops due to large incremental error. By extracting additional semantic information, global error can be reduced to aid proper detection of large loops.

In this paper we propose a method for using the structured nature of urban environments to improve angular consistency in point cloud maps. We make the assumption that urban environments contain a high population of truly vertical surfaces, and we can use the alignment of these surfaces to maintain absolute angular consistency in the map. Since not every frame will contain sufficient semantic information for absolute alignment, we also propose a method for propagating the alignment error along the path trajectory of the robot. An example of an urban scene with prominent near-vertical surfaces is shown in Figure 1, taken from the HANNOVER2 dataset courtesy of Nüchter and Lingemann [2011].

## 2 Related Work

Building locally and globally consistent maps during robotic navigation has seen a lot of interest in recent years, with many new approaches being made feasible by increases in processing power and availability of cheap

sensors. The problem is often posed as *Simultaneous Localisation And Mapping* (SLAM), as sensor data is used to improve the robot’s knowledge of its location at the same time as generating a consistent map of its environment. Traditional approaches to SLAM such as EKF-SLAM [Smith *et al.*, 1987; Bailey *et al.*, 2006], FAST-SLAM [Montemerlo *et al.*, 2002], and DP-SLAM [Eliazar and Parr, 2003] extract sparse features (such as corners or other salient regions) from 2D planar range data and correlate these with features already present in the map to improve localisation. These approaches rely on the assumption that the robot’s trajectory lies in only three planar degrees of freedom (3DoF), which does not hold when considering non-planar ground surfaces such as those found in normal urban driving conditions. To extend these techniques to motion in six degrees of freedom (6DoF), sensors capable of detecting features within 3D space are necessary, leading to the dominance of camera based SLAM techniques [Karlsson *et al.*, 2005; Davison *et al.*, 2007]. Some limitations of camera based SLAM are that it is heavily dependent on lighting conditions, and can perform poorly in visually homogeneous environments such as natural caves.

More recently, 3D sensors such as 3D LiDAR and depth cameras have become popular for mapping as they allow for generating dense environment representations known as point clouds. better information about traversability [Thrun *et al.*, 2006], and performance independent of lighting conditions. Existing feature-based SLAM techniques such as FAST-SLAM are difficult to implement using point cloud data, mainly due to the difficulty of extracting features that are invariant to view-point. Some proposed features used for feature-based SLAM include point feature histograms [Rusu *et al.*, 2009] and planar patches [Rusu *et al.*, 2009; Trevor, 2012]. Becoming more popular are SLAM techniques that use the dense point cloud data directly to compute pose relationships along the trajectory of the robot [Nüchter *et al.*, 2007; Engelhard *et al.*, 2011]. Pose relationships are determined by matching scan geometry using pairwise registration such as ICP [Chen and Medioni, 1991; Besl and McKay, 1992] or NDT [Magnusson *et al.*, 2007], and performing loop closure and global optimisation to maintain global consistency [Lu and Milios, 1997; Borrmann *et al.*, 2008].

Existing 3D point cloud SLAM techniques make use of geometric constraints such as surface continuity, but most do not utilise semantic constraints which can be extracted from expert knowledge of ones environment, such as the presence of vertical walls in urban environments, or flat ground surfaces in indoor environments. Nüchter and Hertzberg [2008] used a semantic labelling scheme to classify regions of indoor environments, and used a flat ground plane constraint to improve scan reg-

istration. Harrison and Newman [2008] used alignment of near-vertical wall planes to estimate vehicle pitch and roll from point clouds, but assume these near-vertical planar regions are present in all scans. If good quality planar surfaces are only present in a small subset of captured laser scans, the change in vehicle pose due to this alignment may be large and cause local inconsistency.

In the image processing community, the prevalence of vertical structures in urban environments has been used to constrain stereo correspondences [Zeisl *et al.*, 2011] by constraining depth maps to lie on vertical and horizontal planes. Vertical planar constraints, specifically a Manhattan world model, has been used to constrain multi-view stereo 3D reconstruction with good results [Furukawa *et al.*, 2009].

### 3 Pose Alignment using verticals

#### 3.1 Frame alignment

We propose an algorithm that aims to improve the global consistency of a set of point clouds captured by a vehicle moving through an urban environment. Point clouds are first matched pairwise using ICP or NDT to develop initial correspondences, then optimise rotations in specific point clouds where salient near-vertical surfaces are present. For future sections we refer to a co-ordinate frame where  $\hat{z}$  represents the vertical ‘up’ direction vector.

Harrison and Newman [2008] propose improving laser scan quality on a moving vehicle by adjusting pitch and roll by assuming near-vertical surfaces in urban environments should be vertical in most cases. For each scan they propose to find a new pitch and roll  $(\phi, \psi)$  which minimises the cost function

$$F(\Pi_i) = \frac{\sum_{j=1}^n A_j [R(\phi, \psi) \hat{\pi}_{i,j} \cdot \hat{z}]^2}{\sum_{j=1}^n A_j} + \omega(\phi, \psi) \quad (1)$$

$$\phi, \psi = \underset{\phi, \psi}{\operatorname{argmin}} [F(\Pi_i)] \quad (2)$$

where  $\Pi_i$  is the set of near-vertical planes, each with normal  $\pi_{i,j}$ ;  $A_j$  is the area of plane  $j$ , and  $R(\phi, \psi)$  is the rotation matrix corresponding to the adjustment in pitch and roll, and  $\hat{z}$  is the unit vector in the vertical direction.  $\omega(\phi, \psi)$  is a regulariser which penalises changes due to underconstrained geometry (such as when there is only a single dominant plane). The  $\omega$  regulariser is not necessary if we apply the optimisation if and only if we know the geometry is not underconstrained; in this case, we also need a method for distributing error from vertical alignment across the unaligned point clouds of the vehicle trajectory. This error distribution is covered in section 3.2.

Rather than aligning verticals before registration, we assume the scans have already been registered pairwise using an algorithm such as ICP or NDT and we apply a vertical alignment that tries to minimize the impact on registration quality.

If we remove the regularization term from Equation 1, and rather than extracting planar regions directly we attempt to optimise over all the near-vertical normal vectors associated with points in the point cloud, we can also remove the area weighting term  $A_j$  as each point will provide equal weighting, and large surfaces will be comprised of more points. We arrive at

$$G(\chi_i) = \sum_{j=1}^n [\mathbf{R}\hat{\boldsymbol{\pi}}_{i,j} \cdot \hat{\mathbf{z}}]^2 \quad (3)$$

where we desire to find a rotation matrix  $\mathbf{R}$  to minimise the function  $G(\chi_i)$ . If we express  $R_z = \mathbf{R} \cdot \hat{\mathbf{z}}$ , this can be shown to reduce Equation 3 to finding  $\mathbf{R}$  such that  $R_z$  minimizes the following

$$G(\chi_i) = \left\| \bar{\boldsymbol{\Pi}} \hat{R}_z \right\|_2^2 \quad (4)$$

where  $\bar{\boldsymbol{\Pi}}$  is the concatenated matrix of normals

$$\bar{\boldsymbol{\Pi}} = [\hat{\pi}_{i,1} \quad \hat{\pi}_{i,2} \quad \dots \quad \hat{\pi}_{i,n}] \quad (5)$$

Unlike Harrison and Newman's Equation 1, Equation 4 has a closed form solution. It can be shown that the minimum value of

$$\left\| \bar{\boldsymbol{\Pi}} \hat{R}_z \right\|_2^2 \quad (6)$$

subject to the constraint

$$\left\| \hat{R}_z \right\|_2^2 = 1 \quad (7)$$

is given by

$$\hat{R}_z = \hat{v}_n R_z = \hat{v}_n \quad (8)$$

where  $\hat{v}_n$  is the last column of  $\mathbf{V}$  in the SVD decomposition of  $\bar{\boldsymbol{\Pi}}$ . We now have the row  $R_z$  of  $\mathbf{R}$  which minimises the z component of near-vertical normals, and must find the other two rows  $\hat{R}_x$ ,  $\hat{R}_y$  which form an orthogonal matrix  $\mathbf{R}$ .

Since point clouds are already registered pairwise, and we cannot extract any information about yaw rotation from the vertical wall semantic information, we want to choose  $\hat{R}_x$ ,  $\hat{R}_y$  to minimise the change in yaw rotation, e.g. minimise the change in  $x$  and  $y$  coordinates of the normal vectors. Since the change in  $x$  coordinate is proportional to  $\hat{R}_{1,1}$  and change in  $y$  coordinate is proportional to  $\hat{R}_{2,2}$ , we are trying to find  $\mathbf{R}$  such that  $\hat{R}_{1,1}$  and  $\hat{R}_{2,2}$  are as close as possible to 1.

We choose to maximise

$$f(\mathbf{R}) = R_{1,1} + R_{2,2} \quad (9)$$

given  $\hat{R}_z$ , and subject to the constraints

$$\|\hat{R}_x\| = \|\hat{R}_y\| = 1 \quad (10)$$

$$\hat{R}_x \cdot \hat{R}_z = 0 \quad (11)$$

$$\hat{R}_y = \hat{R}_x \times \hat{R}_z \quad (12)$$

Equation 12 implies that

$$R_{2,2} = R_{1,1}R_{3,3} - R_{1,3}R_{3,1} \quad (13)$$

so Equation 9 becomes

$$f(\mathbf{R}) = R_{1,1} + R_{1,1}R_{3,3} - R_{1,3}R_{3,1} \quad (14)$$

which can be shown, using the method of lagrange multipliers, to take a maximum value for

$$\hat{R}_x = \left[ \frac{\mathbf{R}_{3,3} + 1 - \mathbf{R}_{3,1}^2}{\mathbf{R}_{3,3} + 1} \quad \frac{\mathbf{R}_{3,1}\mathbf{R}_{3,2}}{\mathbf{R}_{3,3} + 1} \quad \frac{-\mathbf{R}_{3,1}\mathbf{R}_{3,3} - \mathbf{R}_{3,1}}{\mathbf{R}_{3,3} + 1} \right] \quad (15)$$

$$\hat{R}_y = \hat{R}_x \times \hat{R}_z \quad (16)$$

We now have a closed form expression which gives us the optimal rotation to minimize the vertical component of near-vertical surface normal vectors. What we cannot extract from these normal vectors is the required translation which keeps the pairwise scan registration as consistent as possible after applying the vertical alignment. To find the translation we also need to consider the pose trajectory of the robot, and the distribution of the rotational alignment along this trajectory.

### 3.2 Error distribution

Unlike the method of Harrison and Newman, we do not assume every single laser frame contains sufficient saliency to perform vertical alignment. If vertical alignment is applied to single scans infrequently in the map, local consistency will be lost due to gradually accumulated error being corrected suddenly. If it is the first time any scan has been aligned using semantic knowledge, we can safely apply this rotation to all scans in the map without compromising consistency; however if any previous scans have already been aligned, there are conflicting constraints to take into account.

To improve local consistency among consecutive scans, we apply an algorithm similar to Explicit Loop Closing Heuristic (ELCH) of Sprickerhof and Nüchter [2009]. First, the robot's pose trajectory is expressed as an undirected graph, with covariances of registration [Censi, 2007] labelling the edges of the graph of pose relations.

The edge covariances angular components are used to compute weights  $w_i \in [0, 1]$  which represent the proportion of trajectory error associated with individual pose relations, as described in Sprickerhof and Nüchter [2009]. Since we assume the vertical alignment absolutely aligns the frames on which it is applied, we choose the most recent previously aligned scan to be considered having 0 weighted error, and use the Loop Optimisation Algorithm [Sprickerhof and Nüchter, 2009] to find weights for all subsequent scans.

The rotational adjustment  $\mathbf{R}$  can then be expressed as a quaternion and SLERP [Shoemake, 1985] used to distribute this adjustment along the pose trajectory proportional to the error weight  $w_i$ . If only the rotational components of the pose relations are adjusted, the map will be inconsistent due to the lack of translational information derived from the vertical alignment procedure. We choose to move along the pose trajectory, applying the original pose transformation from registration to the newly rotated scans to find the new translation of the next scan, e.g.

$$t_k = R'_{k-1} t_{reg} \quad (17)$$

where  $t_k$  is the new translation,  $R'_{k-1}$  is the rotation of the previous scan after vertical alignment, and  $t_{reg}$  is the translational relation between scan  $k$  and  $k-1$  from pairwise registration (e.g. ICP).

An alternative but more computationally expensive method to find the translation would be to recompute point relationships and apply a translation-only ICP-like registration.

## 4 Results

### 4.1 Monte Carlo Simulations

To validate that the vertical alignment proposed in Section 3.1 correctly orientates a set of near-vertical normals, we apply a Monte Carlo simulation to investigate the effect of noise in normal estimation on the closed form alignment.

We generate a set of purely vertical normals (zero vertical component), apply zero-mean gaussian noise to each of the components and re-normalize to generate a new set of simulated normals:

$$\hat{\pi}' = \hat{\pi} + \boldsymbol{\eta}, \quad \eta_i \sim N(0, \sigma_\eta) \quad (18)$$

We then select pitch and yaw rotations from a uniform distribution to apply to these normals, and then use the new vertical alignment method to recover the original orientation. Comparing the pitch and roll from the recovered orientation with the originally applied rotations, we can investigate the effect of noisy normal estimations on the quality of alignment.

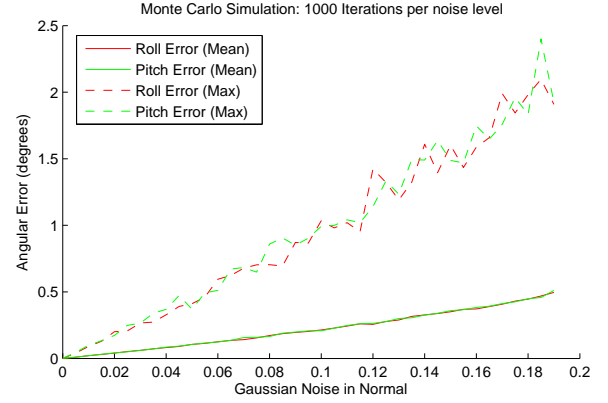


Figure 2: Results from 1000 iteration Monte Carlo simulation at 40 noise increments, with a maximum simulated angle of pitch and roll was  $30^\circ$

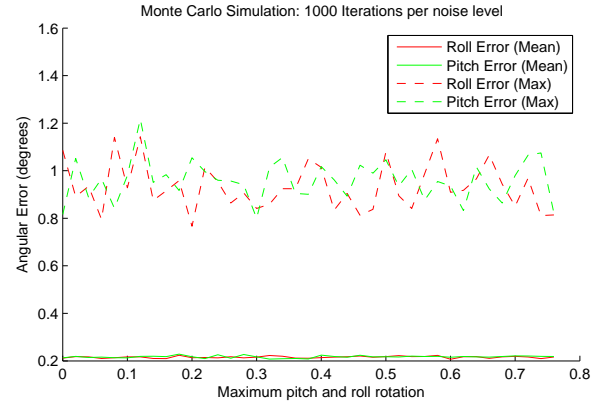


Figure 3: Results from 1000 iteration Monte Carlo simulation at 40 levels of rotational variance, with a noise level of 0.1

Figure 2 shows the results of 1000 iteration simulations at 40 different levels of noise ( $\sigma_\eta$ ), each with a random rotation between  $\pm 30^\circ$ . Both maximum and mean error in angular estimation are strongly linear with respect to the error in normals. The maximum error falls below  $2.5^\circ$  for reasonable noise levels, and mean error falls well below  $0.5^\circ$ .

Figure 3 shows the results of 1000 iteration simulations at a fixed noise level  $\sigma_\eta = 0.1$  with 40 different maximum angular rotations. Unlike noise level, the method seems invariant to angular rotation; however, in real world datasets not all near-vertical normals will actually correspond to vertical surfaces, and this data association problem will degrade performance when the error in pitch and roll is large.

	$\Delta\phi$ (degrees)		$\Delta t $ (metres)	
	ICP	ICP+VA	ICP	ICP+VA
MANLY	2.7476	1.5954	1.0699	0.4911

Table 1: Comparison between rotational and translational error before loop closure using ICP alone and ICP with Vertical Alignment (VA).

## 4.2 Experiments

In Section 4.1 we verified that the single-frame vertical adjustment does tend to correct orientation even in the presence of noise. To verify that the overall method with pose history adjustment tends to create a consistent map, we applied the method on real-world point cloud datasets. The dataset, MANLY, was taken by UNSW Robotics driving a consumer vehicle equipped with a rotating SICK LMS151 LiDAR scanner, with IMU and wheel odometry providing apriori estimates of vehicle pose for frame construction. A loop from this dataset was isolated and used for the following experiments; This loop is presented in Figure 4, consisting of 51 laser scans each of which comprises approximately 10,000 3D points, covering a path distance of approximately 300 metres.

Global consistency was tested by inspecting revisited regions, and seeing how closely the point clouds align before loop closure, and was measured by comparing the registration error before and after loop closure. Our new alignment method was compared with pure pairwise ICP matching (using `libpointmatcher` [Pomerleau *et al.*, 2013]), and a comparison between the two is supplied in 5, where our method is more visually consistent when compared to ICP registration alone.

Table 1 shows the rotational and translational error between scans at B and A before loop closure. The vertical alignment technique is shown to reduce error in the loop, increasing the chance of convergence during loop closure in long loops.

Preliminary results on publicly available datasets are promising, but due to the need for accurate surface normal estimates and salient vertical surfaces, more work is being done to select appropriate datasets from the public domain.

## 5 Conclusions and Future Work

We have proposed a method for aligning point clouds using semantic knowledge of the orientation of vertical planes in urban or constructed environments.

The vertical alignment method will be improved by applying additional filtering to suspected vertical surface normals to remove erroneous surfaces.

Semantic alignment of vertical surfaces provides absolute information about the environment which should aid global consistency in long unclosed loops. The infor-

mation provided by the vertical alignment will provide more constraints to a global optimisation process such as Lu and Milios style GraphSLAM [Lu and Milios, 1997], further enhancing map quality for large maps. We propose to attempt integration of this and other semantic information derived from point clouds for global optimisation and relaxation. To facilitate this, a good estimate of the covariance of alignment will need to be found.

The MANLY dataset contains only small loops, which allows too little time for incremental error to accumulate enough to prohibit loop detection. We hope to apply our method to datasets containing much larger loops, such that the error in registration makes loop detection impossible without some semantic knowledge.

## References

- [Bailey *et al.*, 2006] Tim Bailey, Juan Nieto, Jose Guivant, Michael Stevens, and Eduardo Nebot. Consistency of the EKF-SLAM Algorithm. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, number 1, pages 3562–3568. Ieee, October 2006.
- [Besl and McKay, 1992] PJ Besl and ND McKay. A Method for Registration of 3-D Shapes. In *Robotics*, pages 586–606. 1992.
- [Borrmann *et al.*, 2008] Dorit Borrmann, Jan Elseberg, Kai Lingemann, N Andreas, and Joachim Hertzberg. The Efficient Extension of Lu and Milios SLAM to 6 DoF. In *Proceedings of the 4th International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT '08)*, 2008.
- [Censi, 2007] A Censi. An accurate closed-form estimate of ICP’s covariance. In *IEEE International Conference on Robotics and Automation (ICRA '07)*, 2007.
- [Chen and Medioni, 1991] Y Chen and G Medioni. Object modelling by registration of multiple range images. In *IEEE International Conference on Robotics and Automation (ICRA 1991)*, Sacramento, California, 1991. IEEE.
- [Davison *et al.*, 2007] Andrew J Davison, Ian D Reid, Nicholas D Molton, and Olivier Stasse. MonoSLAM: real-time single camera SLAM. *IEEE transactions on pattern analysis and machine intelligence*, 29(6):1052–67, June 2007.
- [Eliazar and Parr, 2003] Austin Eliazar and Ronald Parr. DP-SLAM: Fast, robust simultaneous localization and mapping without predetermined landmarks. In *International Joint Conferences on Artificial Intelligence (IJCAI '03)*, volume 3, pages 1135–1142, 2003.
- [Engelhard *et al.*, 2011] Nikolas Engelhard, F Endres, Jurgen Hess, and Jurgen Sturm. Real-time 3D visual SLAM with a hand-held RGB-D camera. In *Proc. of*



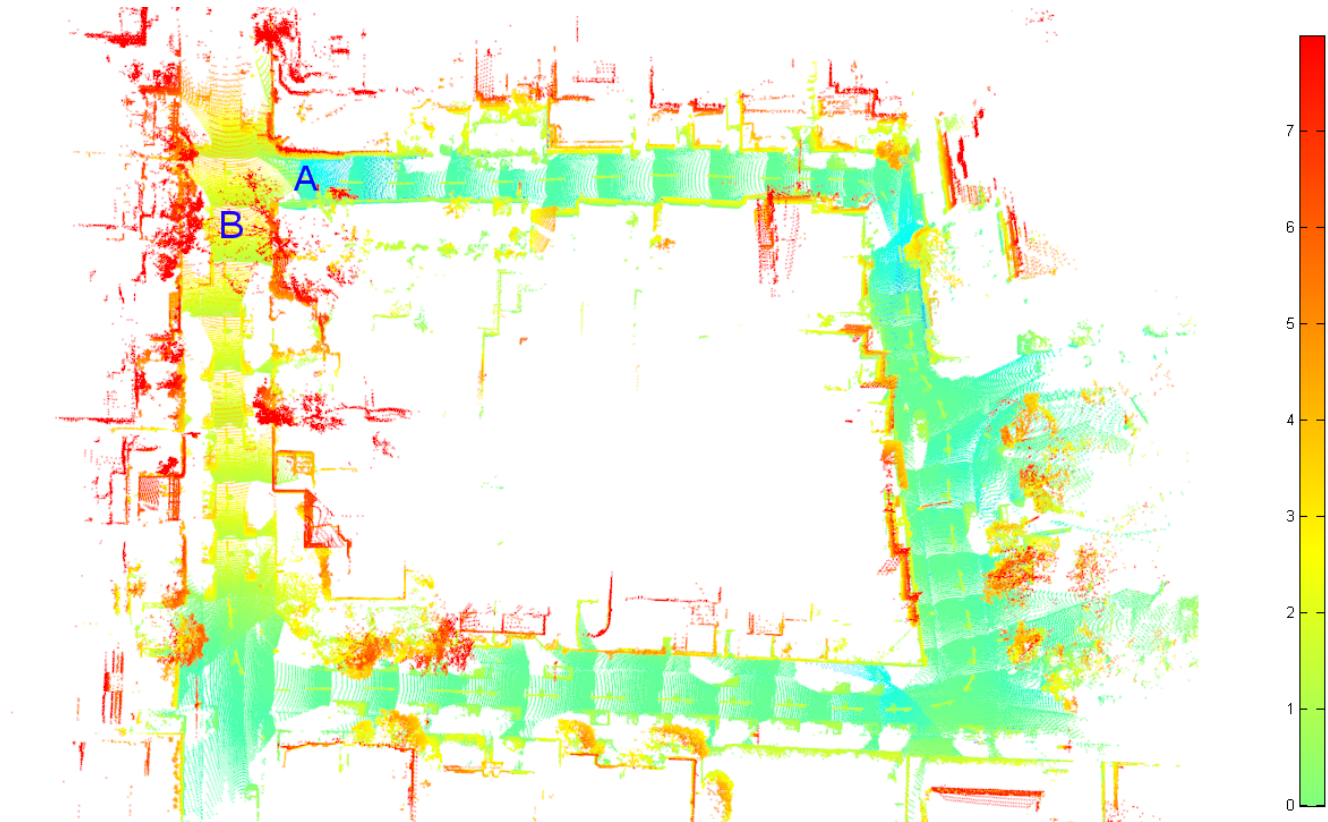


Figure 4: Top down view of loop extracted from the MANLY dataset. The vehicle began at point A and travelled clockwise around the loop until point B, where the loop has been closed after revisiting a prior location. Points are coloured by height.

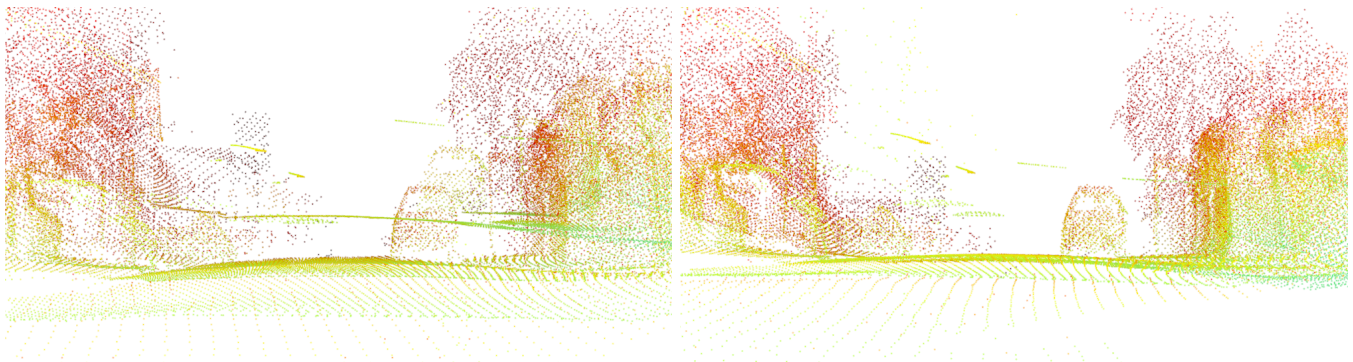


Figure 5: View from point B in the MANLY dataset, coloured by height LEFT: Pairwise ICP alone gives a misaligned ground plane when returning to point A. RIGHT: Vertical alignment method ground plane is more consistent.

*the RGB-D Workshop on 3D Perception in Robotics at the European Robotics Forum*, volume 180, 2011.

- [Furukawa *et al.*, 2009] Yasutaka Furukawa, Brian Curless, Steven M. Seitz, and Richard Szeliski. Reconstructing building interiors from images. In *2009 IEEE 12th International Conference on Computer Vision*, pages 80–87. Ieee, September 2009.
- [Harrison and Newman, 2008] Alastair Harrison and Paul Newman. High quality 3D laser ranging under general vehicle motion. *2008 IEEE International Conference on Robotics and Automation*, pages 7–12, May 2008.
- [Karlsson *et al.*, 2005] N. Karlsson, E. di Bernardo, J. Ostrowski, L. Goncalves, P. Pirjanian, and M.E. Munich. The vSLAM Algorithm for Robust Localization and Mapping. In *IEEE International Conference on Robotics and Automation (ICRA '05)*, pages 24–29. Ieee, 2005.
- [Lu and Milios, 1997] F Lu and E Milios. Globally consistent range scan alignment for environment mapping. *Autonomous robots*, 4(4):333–349, 1997.
- [Magnusson *et al.*, 2007] Martin Magnusson, Achim Lilienthal, and Tom Duckett. Scan registration for autonomous mining vehicles using 3D-NDT. *Journal of Field Robotics*, 24(10):803–827, 2007.
- [Montemerlo *et al.*, 2002] Michael Montemerlo, Sebastian Thrun, Daphne Koller, and Ben Wegbreit. Fast-SLAM: A factored solution to the simultaneous localization and mapping problem. In *AIII National conference on Artificial Intelligence*, number July, pages 593–598, Edmonton, Canada, 2002. AIII.
- [Nüchter and Hertzberg, 2008] Andreas Nüchter and Joachim Hertzberg. Towards semantic maps for mobile robots. *Robotics and Autonomous Systems*, 56(11):915–926, 2008.
- [Nüchter and Lingemann, 2011] Andreas Nüchter and Kai Lingemann. Robotic 3D Scan Repository, 2011.
- [Nüchter *et al.*, 2007] Andreas Nüchter, Kai Lingemann, Joachim Hertzberg, Schloss Birlinghoven, and D-Sankt Augustin. 6D SLAM 3D Mapping Outdoor Environments. *Journal of Field Robotics*, 24(8-9):699–722, 2007.
- [Pomerleau *et al.*, 2013] François Pomerleau, Francis Colas, Roland Siegwart, and Stéphane Magnenat. Comparing ICP variants on real-world data sets. *Autonomous Robots*, 34(3):133–148, February 2013.
- [Rusu *et al.*, 2009] Radu Bogdan Rusu, Nico Blodow, and Michael Beetz. Fast Point Feature Histograms (FPFH) for 3D registration. In *IEEE International Conference on Robotics and Automation (ICRA 2009)*, pages 3212–3217, Kobe, Japan, May 2009. IEEE.
- [Shoemake, 1985] K Shoemake. Animating rotation with quaternion curves. In *ACM SIGGRAPH computer graphics*, 1985.
- [Smith *et al.*, 1987] Randall Smith, Matthew Self, and P Cheeseman. Estimating uncertain spatial relationships in robotics. In *International Symposium of Robotics Research*, pages 467–474, 1987.
- [Sprickerhof and Nüchter, 2009] Jochen Sprickerhof and Andreas Nüchter. An Explicit Loop Closing Technique for 6D SLAM. In *Proceedings of the 4th European Conference on Mobile Robots*, pages 1–6, 2009.
- [Thrun *et al.*, 2006] Sebastian Thrun, Mike Montemerlo, and Andrei Aron. Probabilistic Terrain Analysis For High-Speed Desert Driving. In *Robotics: Science and Systems*, pages 16–19. 2006.
- [Trevor, 2012] AJB Trevor. Planar surface SLAM with 3D and 2D sensors. In *IEEE International Conference on Robotics and Automation (ICRA '12)*, pages 2–9, 2012.
- [Zeisl *et al.*, 2011] Bernhard Zeisl, Christopher Zach, and Marc Pollefeys. Stereo Reconstruction of Building Interiors with a Vertical Structure Prior. In *2011 International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission*, pages 366–373. Ieee, May 2011.