Automatic Segmentation of 3D Laser Point Clouds by Ellipsoidal Region Growing

Frederick Pauling
University of Queensland
Brisbane, Australia
frederick.pauling@uqconnect.edu.au

Michael Bosse
CSIRO ICT Centre
Brisbane, Australia
mike.bosse@csiro.au

Robert Zlot
CSIRO ICT Centre
Brisbane, Australia
robert.zlot@csiro.au

Abstract

We present and evaluate two variants of an algorithm for simultaneously segmenting and modeling a mixed-density unstructured 3D point cloud by ellipsoidal (Gaussian) region growing. The base algorithm merges initial ellipsoids into larger ellipsoidal segments with a minimum spanning tree algorithm. The variants differ only in the merge criterion used—a threshold on a generalised distance measure defined on the merge candidates. The first variant (*shape-distance*) considers the relative shape, orientation and position of the ellipsoids, and can grow regions across missing or sparse data, whilst the second (*density-distance*) attempts to maintain a good fit to the data by setting a minimum sample density threshold on the merged ellipsoid. Adjusting the threshold in each case changes the quality and degree of segmentation achieved. The threshold parameter is tuned by minimising Akaike's Information Criterion (AIC) with respect to the threshold value. Experiments show that thresholds selected in this way lead to low complexity models and are stable across different environments. The *shape-distance* measure segments large-scale structures more readily than the *density-distance* measure, but leads to higher AIC scores, and higher model complexity.

1 Introduction

Three-dimensional lidar point clouds are an important type of perceptual model for many robotics applications such as autonomous mapping and navigation. However, they often have high redundancy, which incurs processing overhead, and they lack explicit structure, which limits its semantic processes such as object classification. Segmenting a point cloud into connected regions, and modelling those segments with parameterised shape primitives provides both data compression and explicit structure. An appropriate model can also assist in the task of object classification and identification, and can help to solve a key problem in mobile robotics—Simultaneous Localisation and Mapping (SLAM)—by simplifying data association and reducing feature complexity.

The segmentation and modelling algorithm described in this paper has been designed for use on mixed-density point-clouds obtained in outdoor environments; for example those containing planar building and ground-plane structures, cylindrical tree-trunks and poles, and diffuse, amorphous vegetation. It has also been developed specifically to deal with the uneven sampling density and noise present in our dataset due to the acquisition method.

The algorithm simultaneously segments and models a point cloud by growing ellipsoidal regions via a minimum spanning-tree. The region-growing criteria is formulated as a threshold on a distance measure, defined on the merge candidate ellipsoids. Two distance measures are proposed in this paper, each based on different principles, and each producing a different style of segmentation. Varying the threshold parameter gives different degrees of segmentation, and we select the threshold value by minimising Akaike's Information Criterion (AIC) [Akaike, 1974] with respect to the threshold parameter. AIC ranks models according to how they balance variance with bias, and leads to the selection of the simplest model which best explains the data—the *principle of parsimony* or *Occam's Razor*. Using AIC also allows meaningful comparison of the variants at their optimal threshold values, and hence forms a relatively objective method of selecting the best algorithm or measure for the segmentation.

By using ellipsoids as the primary data abstraction, we are able to natively represent planar, cylindrical and spherical regions in the data. Each ellipsoid is effectively the multivariate normal PDF of the modelled segment, and this allows the algorithm to deal with uneven sampling density and noise by defining the distance mea-

sures to operate on the principal distribution components, rather than purely local properties. Ellipsoids are generated from the moment statistics of the modelled segments, and region growing is achieved by incrementally merging these statistics (rather than re-computing them each time). We have selected efficient methods wherever possible with a view towards online implementation in the future, however the implementation is not optimised and the computational efficiency is not evaluated in this paper.

Many techniques have been developed for segmenting high-resolution point clouds derived from terrestrial 3D lidar scanners, such as those used for surveying and reverse-engineering civil works [Haala and Brenner, 1999; Slob and Hack, 2004]. However, these techniques typically rely on the terrestrial scanner's high sampling density, low-error and regular acquisition scan pattern for accurate results [Klasing *et al.*, 2008; Rabbania *et al.*, 2006]. Because fixed 3D terrestrial scanners are expensive, relatively slow, and produce very large amounts of data, many roboticists use cheaper 2D scanners to acquire 3D point clouds. This is usually achieved by mounting the scanner on a vehicle and moving through the environment. Successive 2D scans are registered in the world-frame to produce a 3D point cloud. Typical approaches include mounting the scanner in a fixed vertical plane [Fruh and Zakhor, 2001; Thrun *et al.*, 2000], or actuating the scanner about some axis: either nodding [Cole *et al.*, 2005] or rotating [Bosse and Zlot, 2009a].

The point clouds we use are derived from an in-house spinning 2D SICK laser scanner with the spin axis along the centre beam of the scan plane [Bosse and Zlot, 2009a]. The rotating platform is mounted on a vehicle (Figure 2), and driven around semi-industrial and off-road woodland environments (Figure 1). The scanner generates a forward-facing hemispherical scan sweep from each one-second half-revolution. This unusual scan pattern generates very uneven sampling densities when the scans are globally registered [Bosse and Zlot, 2009b], and some slight registration errors persist from the scan matching process thereby introducing additional noise. The approach introduced in this paper, however, is not dependent on the scanning modality, and has been successfully applied to 3D point cloud data acquired using other laser configurations as well.

In general, this method of acquiring 3D point clouds can introduce significant noise into the mapping process due to vibration and uncertainty in the pose of the robot. The sampling density is also typically uneven due to varying linear and angular rates of the vehicle, as well as density drop-off with distance for actuated scanners. Pose error can also lead to ghosting and surface diffusion when a feature is composed of multiple overlapping scans. These problems present a difficult challenge when designing a segmentation algorithm. Indeed, many existing segmentation methods are rendered unsuitable for use on such data as they rely too heavily on even and dense sampling in the point cloud; for example, segmenting based on local surface normal and curvature estimates [Huang and Menq, 2001; Rabbania *et al.*, 2006]. In addition, many such methods employ intensive pre-processing steps to clean the data which can limit their efficiency. Therefore, the challenge we faced in this work was to develop a segmentation method which naturally dealt with the mixed sampling density and noise present in our dataset, yet was efficient enough to bring it within reach of on-line application.

(a) Woodland Scene (b) Light-industrial Scene

Figure 1: Two representative environments for which segmentations are produced and evaluated in the paper

We now review some recent related work in the area. Rusu *et al.* [2008] use outlier removal, robust plane estimation, persistent feature histograms and re-sampling to limit the effects of noise and uneven sampling density. However their work is mainly concerned with modelling planar surfaces in high resolution scans of indoor environments, and the processing is conceived of as initially taking place offline. Our work foregoes such preprocessing steps in favour of efficiency.

Unnikrishnan and Hebert [2003] use projection pursuit and robust estimators to extract planar structures from point clouds similar to those we examine. Clustering is performed at the voxel level via a mean-shift in parameter space, followed by further mode extraction in projection space to separate aligned but offset planes. By contrast, our *shape-distance* measure natively separates dominant modes in normal and Euclidean space as regions grow, without sequential processing steps (these measures will be introduced in Section 2.4). Their approach is also less general than ours since it considers non-planar structure to be clutter, where we extract cylindrical and spheroidal structure explicitly.

Vandapel *et al.* [2004] use neighbourhood covariance and filtering techniques to first classify points and then generate segments. Whilst our approach uses a similar local-shape analysis for the initial step, we do not classify shape, but instead use the descriptors as parameters in one of our distance measures to estimate neighbourhood

Figure 2: Data is collected from the spinning laser mounted on the roof of a Toyota Prado 4WD.

similarity thereby guiding the region growing process.

Stamos and Allen [2002] segment relatively dense point clouds using a local planar similarity constraint with the goal of discovering lines of intersection between planes for image registration purposes. The planarity constraint is similar to one component of our *shape-distance* measure, and is used to segment smoothly varying surfaces of buildings from the point cloud. The authors fit a plane to each of the produced segments under an assumption of global planarity, however, the planarity constraints only act locally (where ours are global), and the authors do not mention whether global planarity tests are performed. Their paper also does not consider cylindrical or diffuse distributions, which are prominent in natural and outdoor environments.

Melkumyan [2009] uses an efficient surface segmentation method based on region continuity and edge-point detection. The approach requires the scan order of the points to be known, unlike the present method, and it is unclear how effective the technique would be for the spinning laser configuration used in this paper.

The remainder of the paper is organised as follows. In Section 2, we cover our approach to the problem. We first provide background on the ellipsoid modelling, then define the distance measures used, and finally describe an overview of the algorithm. In Section 3 we cover the AIC ranking methodology and in Section 4 discuss some results in the context of their AIC ranking and their subjective qualities. Examples illustrate the results where appropriate. Finally in Section 5 we summarise the work and look at future research directions.

2 Approach

Our method simultaneously segments and models a point cloud through a graph-based ellipsoidal region-growing process. At the heart of the method is a minimum spanning-tree (MST) implementation which grows ellipsoidal segments from initial ellipsoids as the tree expands. A maximum edge-weight condition is imposed to

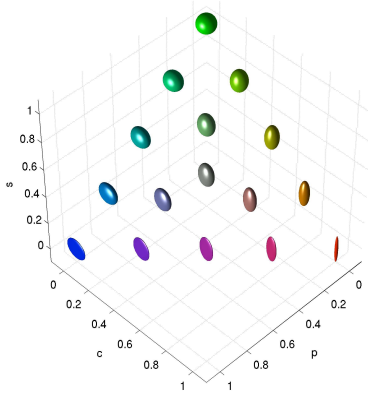


Figure 3: Ellipsoid shape-spectrum, coloured by $RGB = \sqrt{[c, s, p]}$. Spherical shapes are at top, planar shapes lower-left, cylindrical shapes lower-right, and scalene (amorphous) ellipsoids are near the centre.

prevent bad merges from occurring and this truncates the tree leaving the graph (and hence the point cloud) segmented into discrete regions. These regions typically represent objects or sub-components of objects in the environment. The resulting set of ellipsoids is a parameterised model of the point cloud data, and the labelling of points belonging to each ellipsoid is a segmentation of the point cloud. Two methods of edge-weight calculation are presented, and defined as generalised distance measures on the ellipsoids in Section 2.4. These measures lead to different segmentation styles, and varying the maximum edge-weight threshold for each leads to different granularities of segmentation.

2.1 Ellipsoids

Ellipsoids are used as the only shape primitive because they naturally model planar, cylindrical, and spheroidal distributions such as walls, tree-trunks and foliage, respectively. Another compelling reason for modelling with ellipsoids is that they can be efficiently generated directly from the moment statistics of the underlying point distribution.

An ellipsoid is defined as the set of points x which fall within the region:

$$(x-\mu)^T \Sigma^{-1} (x-\mu) \quad < \quad 1 \tag{1}$$

where μ is the centroid and Σ is the covariance of the ellipsoid, which is given by

$$\Sigma = \frac{1}{k} \sum_{i=1}^{k} (p_i - \mu)^T (p_i - \mu) \tag{2}$$

where $p_i = (x_i, y_i, z_i)$ is the i^{th} point, and k is the total number points in the point set. The eigenvectors of Σ

give the principal axes of the ellipsoid, and the square-root of the eigenvalues gives the one-sigma axis-lengths. Furthermore, we can classify how planar (p), cylindrical (c), or spherical (s) an ellipsoid is by looking at the the relative sizes of the eigenvalues as:

$$p = 2\frac{e_2 - e_3}{e_1 + e_2 + e_3} \quad (3)$$

$$c = \frac{e_1 - e_2}{e_1 + e_2 + e_3} \quad (4)$$

$$s = 1 - p - c = \frac{3e_3}{e_1 + e_2 + e_3} \quad (5)$$

where e_1, e_2, e_3 are the descending sorted eigenvalues of the ellipsoid's covariance matrix (equivalent to the squared axis lengths). These shape descriptors were originally developed for the purpose of 3D scan matching [Bosse and Zlot, 2009a], and in this paper we apply them to define a distance measure between two ellipsoids in Section 2.4.

The ellipsoid shape spectrum (Figure 3) demonstrates the versatility of using ellipsoids as a model. More complex shapes can be approximated reasonably well by groupings of these shapes, and thus we consider the ellipsoid shape spectrum to be a sufficient basis on which to model the world. A limitation of our approach is that smoothly varying non-planar regions, or curved cylinders are typically either broken into multiple segments, or approximated with a high variance by a single ellipsoid. Other researchers have used local smoothness criteria to perform the segmentation [Stamos and Allen, 2002; Rabbania *et al.*, 2006] and achieve a low-variance model but at the expense of a high number of parameters.

Two or more ellipsoids can be easily and efficiently merged into a new ellipsoid by combining their moment statistics and re-centralising the covariance about the new centroid. This ability to incrementally combine the model parameters, instead of recomputing them at each merge, is fundamental to the efficiency of our region-growing approach. Also, by simultaneously modeling and segmenting with the same abstraction, we preclude the need for separate segmentation and modeling stages, and thus obtain a further increase in efficiency.

This method of (directly) generating the model parameters is not robust to outliers and these can arbitrarily attract the mean and inflate/skew the principal axes of the covariance, affecting the segmentation process.

We intend to look into using a more robust parameter estimator, such as Minimum Volume Ellipsoid [Rousseeuw and Zomeren, 1990], but this approach would be less efficient to compute, and would require a significant improvement in the results to be justified.

2.2 Initial Processing

The ten nearest neighbours of each point are found using a *kd*-tree, and we remove isolated points whose closest neighbour distance is above two metres. This reduces the effect of outliers on the segmentation process and also reduces the computation time. Initial ellipsoids are generated about each point from the eigendecomposition of the sample covariance Σ of the local neighbourhood. The union of intersecting local neighbourhoods forms a graph on the point cloud, where each point is a vertex and neighbouring vertices defining an edge. Typical scenes contain disjoint graphs where regions are spatially separated. The MST algorithm operates on these graphs.

Initial edge weights for the MST algorithm are computed based on the Euclidean distance between the centroids of the neighbouring ellipsoids, with additional weight applied to planar ellipsoids depending on their planarity, degree of normal misalignment and planar offset. Thus, lower edge weights correspond to ellipsoids likely to be part of the same region (and vice versa).

2.3 Segmentation and Modelling

The point cloud is segmented and modeled by progressively and selectively merging connected ellipsoids into larger surfaces or regions. We use a thresholded version of Kruskal's MST algorithm—a greedy region growing algorithm—to perform the segmentation.

Edges are processed in ascending order of initial weight, and ellipsoids are merged only if the generalised distance (edge-weight) between the merge candidates does not exceed a user-defined threshold. Emerging subtrees of the MST are modelled as ellipsoids resulting from the merging of all the initial ellipsoids at the sub-tree vertices. Hence the algorithm attempts to construct maximally similar regions by only merging sufficiently similar ellipsoids. Enforcing a maximum edge-weight (distance) condition truncates the growth of the minimum spanning tree, and causes the point cloud to remain segmented into separate connected regions.

There are four major design decisions for segmenting a point cloud with a MST: the representation, the distance measure for calculating the initial edge weights, the distance measure for re-calculating the edge weights based on bulk properties (if any), and finally, the distance threshold for adding an edge to the tree (*i.e.*, for merging regions). The first three are structural choices which determine the 'style' of segmentation, and the fourth is a user tunable parameter (or set of parameters) which determines the granularity of segmentation. Finding the best set of choices is essentially a multi-dimensional optimisation problem. We use an information theoretic measure to help navigate the solution space as it provides a consistent method of scoring the resulting models.

2.4 Distance Measures

We present and evaluate two different distance measures for comparing two ellipsoids. Each measure leads to

qualitatively and quantitatively different segmentations when used in the present algorithm. The first measure, *density-distance*, sets a threshold on the minimum 'density' of the merged ellipsoid and attempts to maintain a good fit to the data. The second measure, *shape-distance*, utilises relative shape, orientation, position and mass information, and can grow regions across missing or sparse data.

Density Distance Since we are segmenting a point cloud via an evolving model of the data, we consider that the model should remain faithful to the underlying point distribution by fitting it closely with minimal empty space. Analogous to maintaining a high sampling density in the ellipsoids, *density-distance* is defined so that ellipsoids are allowed to merge only if the post-merge sample density remains above a threshold.

The density ρ of an ellipsoid is defined as $\rho = m/V$, where m, V are the mass (a function of the sample count) and volume of the ellipsoid respectively. Uneven sampling density in our data means that setting the mass equal to the sample count does not lead to sensible segmentations. To mitigate this problem we assume that most points lie on a surface, and normalise the surface sample density by setting the mass (m') of each point to the square of the mean neighbour distance

$$m' = \left(\frac{1}{k} \sum_{i=1}^{k} |p - p_i| \right)^2 \tag{6}$$

where k is the number of neighbours (excluding self), p is the point under consideration, and p_i is the i^{th} neighbouring point

Then, given two ellipsoids a and b to be merged, the density of the merged ellipsoid ρ_{ab} is given by $\rho_{ab} = (m'_a + m'_b)/V_{ab}$, where m'_a and m'_b are the normalised masses of ellipsoids a and b and V_{ab} is the volume of the merged ellipsoid.

Shape-Distance This measure adapts the s, p and c shape parameters (defined in Section 2.1) to a three-component distance measure that generates a dissimilarity (distance) between two ellipsoids. This measure was designed to overcome the problems of uneven sampling density and noise by comparing merge candidates based on their dominant distribution characteristics: mass, location, shape, and orientation. It is more general than typical measures for plane extraction since it natively handles comparisons between any ellipsoids and hence between any combination of planar, cylindrical, spherical, and intermediate distributions. By definition, this measure assumes that the segments should be topologically homogeneous (*e.g.*, globally planar segments are locally planar), and thus shape-distance estimates how unlikely it is that two ellipsoids belong to the same region. A threshold on the shape-distance determines which ellipsoids are too dissimilar to be merged.

Planar ellipsoids are given a distance based on how out-of-plane and misaligned they are. Cylindrical ellipsoids have a distance based on how off-axis and misaligned they are. Finally, spherical ellipsoids are given a distance related only to the Euclidean distance between the centroids (since there is no preferred orientation or position). Most ellipsoids have a shape identity somewhere between these extremes, so we weight each of the shape-specific component measures by how planar, cylindrical and spherical the ellipsoids are respectively, and normalise with respect to the relevant eigenvalue. *Shape-distance* D_{shape} is computed as

$$D_{shape} = D_{planar} + D_{cylindrical} + D_{spherical} \tag{7}$$

where

$$\begin{aligned}
D_{planar} &= \frac{||v_{3a}^T(\Delta\mu)||^2}{e_{3a}} \frac{n_a}{n_a+n_b} + \\
&\quad \frac{||v_{3b}^T(\Delta\mu)||^2}{e_{3b}} \frac{n_b}{n_a+n_b} + \\
&\quad (1 - ||v_{3a}^T v_{3b}||^2) \frac{\frac{p_a n_a + p_b n_b}{n_a+n_b}}{\frac{e_{3a}}{e_{2a}} + \frac{e_{3b}}{e_{2b}}} \\
D_{cylindrical} &= \frac{(1-p_a)||v_{2a}^T(\Delta\mu)||^2}{e_{2a}} \frac{n_a}{n_a+n_b} + \\
&\quad \frac{(1-p_b)||v_{2b}^T(\Delta\mu)||^2}{e_{2b}} \frac{n_b}{n_a+n_b} + \\
&\quad (1 - ||v_{1a}^T v_{1b}||^2) \frac{\frac{c_a n_a + c_b n_b}{n_a+n_b}}{\frac{e_{2a}}{e_{1a}} + \frac{e_{2b}}{e_{1b}}} \\
D_{spherical} &= \frac{(s_a)||v_{1a}^T(\Delta\mu)||^2}{e_{3a}} \frac{n_a}{n_a+n_b} + \\
&\quad \frac{(s_b)||v_{1b}^T(\Delta\mu)||^2}{e_{3a}} \frac{n_b}{n_a+n_b}
\end{aligned}$$

where, for distributions a and b, $\Delta\mu$ is the vector between the centroids, $v_{ia,ib}$ and $e_{ia,ib}$ are the i^{th} eigenvector and i^{th} eigenvalue of the covariance, $n_{a,b}$ are the sample counts, and $p_{a,b}, c_{a,b}, s_{a,b}$ are the planarity, cylindricality, and sphericity measures.

3 Evaluation Using AIC

Akaike's Information Criterion (AIC) is an information theoretic criterion developed by Hirotsugu Akaike [1974] for ranking statistical models based on how they balance bias (the number of model parameters) and variance (modelling error). AIC does not require the existence of ground-truth, and scores candidate models on a relative scale which allows for comparisons between models.

Given a particular data set and a selection of models which attempt to explain the data, AIC assigns a score to each model as

$$AIC = -2\log(L) + 2K \tag{8}$$

where $\log(L)$ is the is log of the maximised value of the likelihood function for the estimated model, and K is the number of free parameters in the model. For an ellipsoid we have $K = 9$; *i.e.*, 3 for the mean, and 6 for the covariance. The model with the lowest AIC score is the one which best explains the data with a minimum number of free parameters, and may be considered the best model of the data (from the given selection). We use a standard variant of AIC, known as AICc, which corrects for low bias estimates when there are a relatively large number of estimated parameters compared to the number of samples:

$$AICc = -2log(L) + 2K\left(\frac{n}{n-K-1}\right) \tag{9}$$

where n is the number of samples [Burnham and Anderson, 2001]. AICc converges asymptotically to AIC as n grows. We use AICc since many of the ellipsoids have relatively small sample counts compared to the number of parameters. In practice we see some degenerate ellipsoids with fewer than $K + 2$ points so, to prevent the bias term from going to infinity or negative we set a lower limit on the denominator term as $max(1, n - K - 1)$.

AIC must be evaluated with respect to the the Maximum Likelihood (ML) model parameters. Since the modelling ellipsoids are effectively multivariate Gaussian distributions then the ML parameters for our model are simply the mean and covariance of the segments, which we have already computed.

Using the formulation given by Bozdogan [1986], the variance term in Equation 9 for a multivariate Gaussian distribution is:

$$-2\log(L) = n\left(d\log(2\pi) + \log|\Sigma| + d\right) \tag{10}$$

where $d = 3$ is the number of dimensions. Thus AICc can be written as

$$AICc = n\left(d\log(2\pi) + \log|\Sigma| + d\right) + 2K\left(\frac{n}{n-K-1}\right) \tag{11}$$

Many features of interest (such as planes and cylinders) have relatively uniform sample distributions along one or more principal axes, and this invalidates the assumption of normal error distribution. However, these residuals are not of great importance since (for example) we would typically care only about the out-of-plane residuals for planar features, and the off-axis residuals for thin cylindrical features. Hence we modulate the variance along each eigenvector according to the shape of the distribution (again utilising p, c, s). This is achieved by first decomposing Σ into principal components S and V with

$$\Sigma = VSV^T \tag{12}$$

where V is the matrix of eigenvectors of Σ, and S is the diagonal matrix of eigenvalues of Σ (in descending order of size). We then modulate the eigenvalues in S by the shape of the ellipsoid, and impose a lower bound of 0.01 (the square of our measurement noise estimate) on the eigenvalues to prevent degeneracy when $p = 1$ or $c = 1$:

$$S_{mod} = diag\left(e_1 s, e_2\left(1-p\right), e_3\right) \tag{13}$$

where e_1, e_2 and e_3 are the descending sorted eigenvalues, and $s = 1 - p - c$. The covariance matrix is finally recomposed as

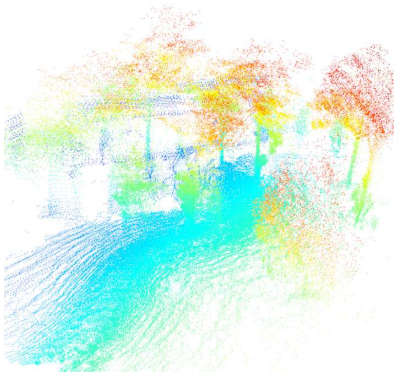
$$\Sigma_{mod} = VS_{mod}V^T \tag{14}$$

To see how the shape modulation works, consider the cases where $p = 1$, $c = 1$ and $s = 1$ (*i.e.*, $p = 0$, $c = 0$). For the $p = 1$ case, only the out of plane variance (along e_3) is significant, for the $c = 1$ case the variance has e_2 and e_3 components, perpendicular to the major axis, and for the $s = 1$ case all three components are utilised.

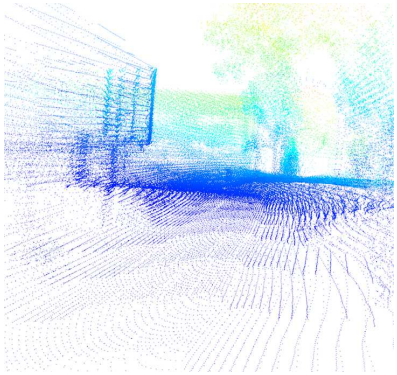
To calculate the overall AICc score for a modeled scene, the AICc score for each ellipsoid is computed separately, and these scores are summed. When comparing algorithms, the algorithm which generates the lowest AICc score for a given scene can be considered to have the best trade-off between fit and bias.

4 Experimental Results

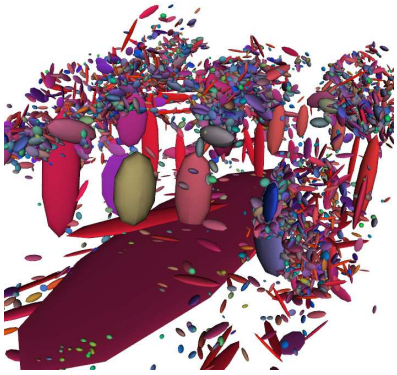
The segmentation technique was applied to point clouds from a number of different environments, including light industrial and off-road lightly wooded scenes (as shown in Figure 6). For each environment, 25 seconds of scan data were segmented with varying distance thresholds, using both distance measures. The AICc score was computed for the resulting models in each case. The optimal threshold for each distance measure (on each scene) was selected as that which minimised the AICc score of the produced model. Some example plots of AICc vs. threshold for the light-industrial scene are shown in Figure 4, where the minima are clearly seen. The optimal threshold selected by AICc for *shape-distance* appears to be quite stable across different environments, as indicated by the approximately coincident minima in Figure 7. This desirable property may negate any requirement to re-tune the threshold for each local point cloud, although preliminary experiments indicate that different



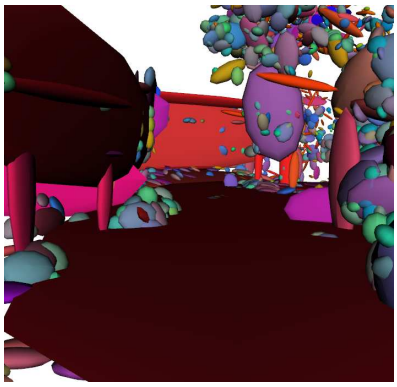
(a) Point Cloud - Woodland - 115,828 points



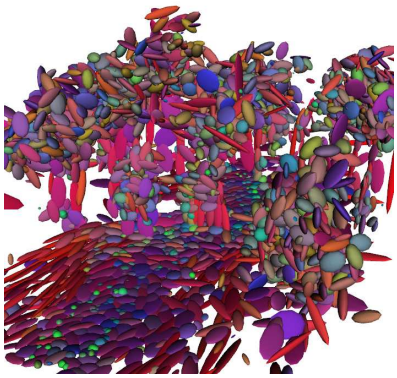
(b) Point Cloud - Light Industrial - 150,250 points



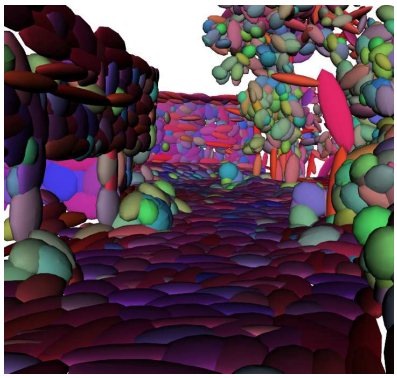
(c) Shape Distance - Minimum AICc Segmentation - 12,391 Ellipsoids



(d) Shape Distance - Minimum AICc Segmentation - 7,173 Ellipsoids



(e) Density Distance - Minimum AICc Segmentation - 4,057 Ellipsoids



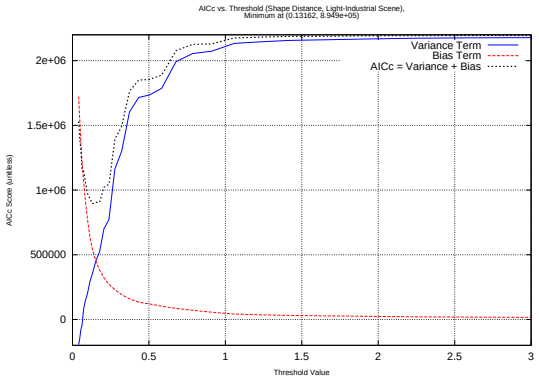
(f) Density Distance - Minimum AICc Segmentation - 5,061 Ellipsoids

Figure 6: Comparison of segmentations produced by using *shape-distance* and *density-distance* at their minimum AICc (optimal) threshold values for two different environments. The left column is an off-road woodland scene, and the right column is a light industrial scene. See Figure 1 for photographs.

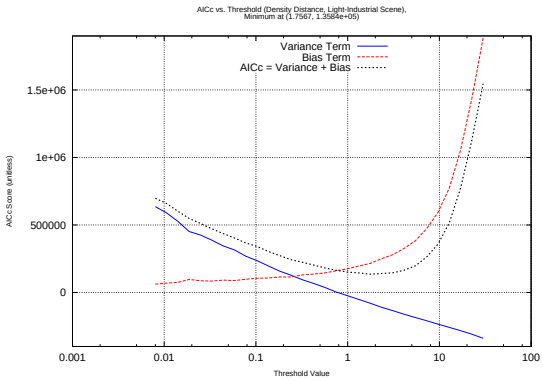
sensor arrangements have different optima. While we have seen similar results with other distance measures, further investigation is required to identify whether this outcome extends to all distance measures in general.

Plots of AICc vs. merge completion (Figure 5) reveal that AICc minima occur very far into the merge process (generally 80% completion or higher). This reflects AIC's preference for low complexity models that explain the most data.

Representative Segmentations Figure 6 shows the minimum AICc threshold segmentations for each dis-

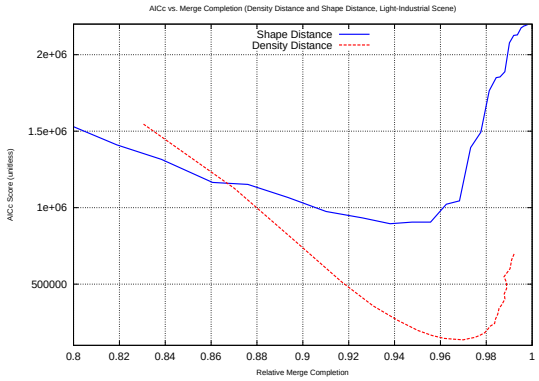


(a) AICc vs. Threshold - Shape Distance, Industrial Scene the minimum AICc occurs at a threshold value near 0.13

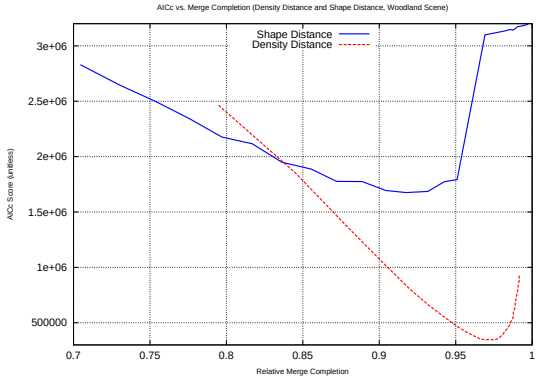


(b) AICc vs. Threshold - Density Distance, Industrial Scene the minimum occurs at a threshold value of around 1.7

Figure 4: AICc score vs. threshold parameter plots used for selecting optimal threshold values in each case, AICc minima are indicated. Segmentations at the optimal threshold values appear in Figure 6. Note that the measures are more easily compared using the relative merge completion in Figure 5.



(a) AICc vs. Relative Merge Completion - Shape Distance and Density Distance, Forest Scene



(b) AICc vs. Relative Merge Completion - Shape Distance and Density Distance, Industrial Scene

Figure 5: Plots of AICc score versus relative merge completion ratio (number of points/number of merges) show that the AICc minima occur very late in the merge process, later in the case of density distance and with a lower AICc score. This effect is due to AIC's preference for low-complexity models.

tance measure in two representative environments, corresponding to the photos in Figure 1. The left-hand column in Figure 6 shows an off-road woodland scene composed of a ground-plane, cylindrical tree trunks, and diffuse vegetation. The right-hand column shows a light industrial scene composed primarily of planar building façade and road surfaces, as well as some trees on the right hand side of the road. The *shape-distance* measure shows good generalisation and tends to extracts large-scale features, however this comes at the expense of higher variance, and a higher number of ellipsoids; whereas *density-distance* maintains a tight fit with fewer ellipsoids at the expense of large-scale structure. *Density distance* typically produces lower AICc segmentations with fewer ellipsoids as a result of the better fit to the data and the lack of geometric constrains on segment growth. The colouration of the ellipsoids is based on their shape according to Figure 3. *Shape-distance* leads to intuitively appealing segmentations by emphasising dominant structures, particularly in well-structured environments such as the industrial scene. *Density-distance* leads to subjective over-segmentation, but gives better data compression by providing a well fitted model with fewer parameters. Thus the choice of measure to use in a given situation depends on whether an accurate model is more important than a good segmentation. Figure 5 compares the measures based on AICc scores versus model simplicity for each scene. The *density-distance* measure clearly generates the most parsimonious models.

Woodland Scene This scene is dominated by a diffuse tree canopy and well-connected ground-plane. *Density-distance* appears to deal with the diffuse regions better than *shape-distance* and leads to a simpler model

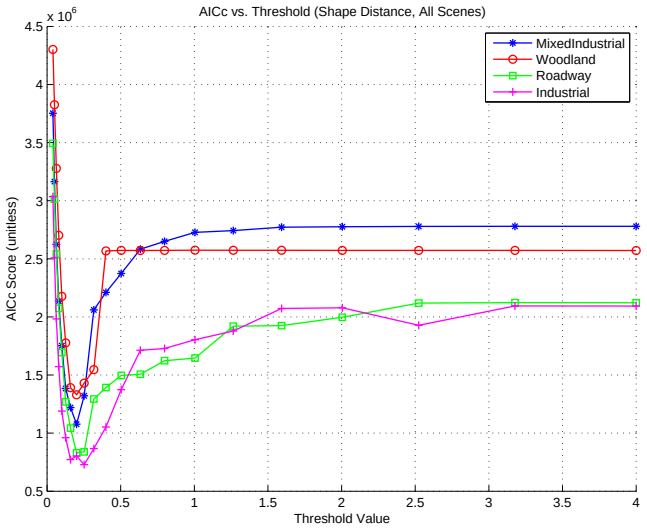


Figure 7: The optimal threshold parameter for *shape-distance* selected by AICc is stable across very different types of environments, as evidenced by the similar AICc minima. In addition to the example industrial and woodland scenes, we also include results from a road scene and a more structured industrial scene.

of such regions in general. The ground plane is composed of tessellated ellipsoids due to point-cloud density variations. This effect could be suppressed by re-sampling the point cloud to a more uniform density, but so far we have tried to limit pre-processing of the data as much as possible. *Shape-distance* performs extremely well on planar surfaces and this is evident as the ground-plane has been extracted as a single segment. Adjacent tree-trunks tend to be merged by *shape-distance* in this scene (although some of the effects are due to smearing from scan mis-registration). The cylinder merging effect can be reduced (if desired) by increasing the cylindrical offset weighting in the measure (see Equation 8).

Industrial Scene This scene is typical of light industrial building complexes and is dominated by the planar features of the roadway and building walls, with a moderate amount of vegetation in garden areas. *Shape-distance* generates a very simple model of the dominant features in such environments and even does a fair job of clustering compact canopy structures into spheroidal segments. Many small ellipsoids remain unmerged (not rendered) and the total model complexity remains higher than *density distance*. By contrast, the *density-distance* measure leads to a better model of the data with fewer parameters, but does not reveal large-scale structure.

Issues Diffuse objects generally have no good model (except globally perhaps) and are arbitrarily segmented based on the emerging shapes of the seed regions. This property makes *shape-distance* a worse choice than *density-distance* for scenes composed of many such distributions. Curved objects and surfaces tend to be over-segmented or represented by a single ellipsoid with high variance. Initial ellipsoids are not good models of sharp features and tend toward more spherical shapes that can affect the resulting model in various (detrimental) ways.

5 Conclusion and Future Work

In this paper we sought to show that a segmentation/model which was generated by manipulating intermediate models of the data can be evaluated on a consistent basis by evaluating the model (rather than the produced segments). AICc has been demonstrated to be a powerful and consistent method of model selection and parameter tuning. We also presented a novel *shape-distance* measure which generates a dissimilarity between a pair of ellipsoids in Euclidean space.

The shape-distance measure was shown to give more intuitively pleasing segmentations than density-distance at the cost of comparatively poor model fit and high AICc scores. Density-distance maintains a good fit to the data, but subjectively over-segments. The results illustrate the benefit of using AICc for parameter tuning by selecting sensible threshold values.

The sampling characteristics of our current sensor restrict us to fairly coarse global models at this stage, and we are looking at using a different sensor with higher resolution to help generate finer-grained models with more discriminative power. We will also investigate the effects of additional pre-processing of the point cloud to further remove noise, and perhaps generate more robust initial ellipsoids, before segmenting the data. Post-processing the results might also achieve a more desirable model; for example, by removing small ellipsoids to reduce local over-segmentation.

Future work will attempt to further verify the validity of the AICc threshold tuning method theoretically.

One issue that can confound the segmentation algorithm is the interaction of multiple overlapping objects. This problem is particularly apparent in the woodland environments we investigated, where a sparse shrub appears at the base of a tree. In some cases—for instance where one component is more porous, or one passes through another—additional analysis may be able to recover the independent components, thereby allowing us to separate these objects into multiple overlapping segments (*e.g.*, a cylindrical tree trunk within a spherical shrub). Another issue is that relatively small holes in planar surfaces, such as windows and open doorways, may also be ignored in the model in favour of fitting a larger wall segment to the data. This problem is challenging since, in the case of a hole due to occlusion, it

would be preferable to fill in such a gap. This trade-off can be considered when designing an appropriate distance measure, and depends on the relative importance of model compactness and accuracy.

As these algorithms are presently implemented in *Matlab*, they are not yet suitable for online application. However, we envision that this goal could be achieved by re-implementation in a more efficient programming language.

Overall, the modeling and segmentation techniques developed appear to form a good basis for future work in object classification and change detection in the context of outdoor autonomous robotics applications. Another potential area of application is in data association for SLAM. The current approach we have taken uses the raw ellipsoids extracted from the point cloud data [Bosse and Zlot, 2009a], and we plan to further investigate whether the merged ellipsoids are sufficiently stable such that they can be used to improve data association efficiency.

References

[Akaike, 1974] H. Akaike. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716–723, December 1974. original paper published in 1971.

[Bosse and Zlot, 2009a] M. Bosse and R. Zlot. Continuous 3D scan-matching with a spinning 2D laser. In *IEEE International Conference on Robotics and Automation*, 2009.

[Bosse and Zlot, 2009b] M. Bosse and R. Zlot. Place recognition using regional point descriptors for 3D mapping. In *Field and Service Robotics*, 2009.

[Bozdogan, 1986] H. Bozdogan. Multi-sample cluster analysis as an alternative to multiple comparison procedures. *Bulletin of Informatics and Cybernetics*, 22(1):95–130, 1986.

[Burnham and Anderson, 2001] K. P. Burnham and D. R. Anderson. Kullback-Lebler information as a basis for strong inference in ecological studies. *Wildlife Research*, 28(2):111–119, May 2001.

[Cole *et al.*, 2005] D. Cole, A. Harrison, and P. Newman. Using naturally salient regions for SLAM with 3D laser data. In *International Conference on Robotics and Automation, SLAM Workshop*, 2005.

[Fruh and Zakhor, 2001] C. Fruh and A. Zakhor. 3D model generation for cities using aerial photographs and ground level laser scans. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, 2001.

[Haala and Brenner, 1999] N. Haala and C. Brenner. Extraction of buildings and trees in urban environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2-3):130–137, 1999.

[Huang and Menq, 2001] J. Huang and C.-H. Menq. Automatic data segmentation for geometric feature extraction from unorganized 3-D coordinate points. *IEEE Transactions on Robotics and Automation*, 17(3):268–279, June 2001.

[Klasing *et al.*, 2008] K. Klasing, D. Wollherr, and M. Buss. A clustering method for efficient segmentation of 3D laser data. In *IEEE International Conference on Robotics and Automation*, 2008.

[Melkumyan, 2009] N. Melkumyan. Surface-based Synthesis of 3D Maps for Outdoor Unstructured Environments. 2009.

[Rabbania *et al.*, 2006] T. Rabbania, F. A. van den Heuvelb, and G. Vosselman. Segmentation of point clouds using smoothness constraint. In *ISPRS Commission V Symposium 'Image Engineering and Vision Metrology', Dresden, Germany*, September 2006.

[Rousseeuw and Zomeren, 1990] P. J. Rousseeuw and B. C. van Zomeren. Unmasking multivariate outliers and leverage points. *Journal of the American Statistical Association*, 85(411):633–639, 1990.

[Rusu *et al.*, 2008] R. B. Rusu, Z. C. Marton, N. Blodow, M. Dolha, and M. Beetz. Towards 3D Point cloud based object maps for household environments. *Robotics and Autonomous Systems*, 56(11):927–941, 2008.

[Slob and Hack, 2004] S. Slob and R. Hack. 3D terrestrial laser scanning as a new field measurement and monitoring technique. *Lecture Notes in Earth Sciences*, pages 179–189, 2004.

[Stamos and Allen, 2002] I. Stamos and P. K. Allen. Geometry and texture recovery of scenes of large scale. *Computer Vision and Image Understanding*, 88:94–118, November 2002.

[Thrun *et al.*, 2000] S. Thrun, W. Burgard, and D. Fox. A real-time algorithm for mobile robot mapping with applications to multi-robot and 3D mapping. In *IEEE International Conference on Robotics and Automation*, 2000.

[Unnikrishnan and Hebert, 2003] R. Unnikrishnan and M. Hebert. Robust extraction of multiple structures from non-uniformly sampled data. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2003.

[Vandapel *et al.*, 2004] N. Vandapel, D. F. Huber, A. Kapuria, and M. Hebert. Natural terrain classification using 3-D ladar data. In *IEEE International Conference on Robotics and Automation*, 2004.