

Scale Selection for Classification of Point-sampled 3-D Surfaces

Jean-François Lalonde, Ranjith Unnikrishnan,
Nicolas Vandapel and Martial Hebert

CMU-RI-TR-05-01

Updated: July 2005
Original publication: January 2005

Robotics Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

© Carnegie Mellon University

Abstract

This document is the extended version of the work published in [11]. Laser-based range sensors are commonly used on-board autonomous mobile robots for obstacle detection and scene understanding. A popular methodology for analyzing point cloud data from these sensors is to train Bayesian classifiers using locally computed features on labeled data and use them to compute class posteriors on-line at testing time. However, data from range sensors present a unique challenge for feature computation in the form of significant variation in spatial density of points, both across the field-of-view as well as within structures of interest. In particular, this poses the problem of choosing a scale for analysis and a support-region size for computing meaningful features reliably. While scale theory has been rigorously developed for 2-D images, no equivalent exists for unorganized 3-D point data. Choosing a satisfactory fixed scale over the entire dataset makes feature extraction sensitive to the presence of different manifolds in the data and varying data density. We adopt an approach inspired by recent developments in computational geometry [17] and investigate the problem of automatic data-driven scale selection to improve point cloud classification. The approach is validated with results using real data from different sensors in various environments (indoor, urban outdoor and natural outdoor) classified into different terrain types (vegetation, solid surface and linear structure).

Prepared in part through collaborative participation in the Robotics Consortium sponsored by the U.S Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-209912 and in part by the National Science Foundation under the grant IIS-0102272.

History

January 2005	Initial submission
February 2005	Revision from reviewer's comments
July 2005	Final revision, from conference attendees comments

Contents

1	Introduction	1
2	Related work	2
3	Approach	3
3.1	Normal Estimation in 3-D	4
3.1.1	Bounding entries of M	5
3.1.2	Eigen analysis	8
3.1.3	Error bound for the estimated normals	9
3.1.4	d_1 and d_2 estimation	10
3.1.5	Estimating the optimal support region size	11
3.1.6	Complexity analysis	12
3.2	Terrain Classification	13
3.2.1	Saliency features	13
3.2.2	Bayesian classification	14
4	Experiments	14
4.1	Sensors and terrains	14
4.2	Validation of computed normals	15
4.2.1	3-D models	15
4.2.2	Aerial ladar scan	16
4.3	Validation of support regions	16
4.3.1	Outdoor ground scan	16
4.3.2	Scan of wall corner	18
4.3.3	Outdoor natural terrain	18
4.4	Ground-based ladar classification of natural terrain	20
4.5	Comparison with multi-scale approach	20
5	Conclusions and Discussion	21

1 Introduction

Autonomous navigation in vegetated terrain remains a challenging problem in robotics due to the difficulties in modeling the high variability of outdoor environments. In this effort, laser range-finders have proven to be invaluable due to their high speed and direct sensing of depth information in the form of unorganized 3-D point clouds from objects in the scene. Depth cues allow more natural modeling of smooth, porous and linear surfaces as 3-D textures. Labeled data can then be used to compute 3-D features and train classifiers for distinguishing load-bearing surfaces, vegetation and linear structures respectively.

However the perspective sensing geometry of laser-range finders introduces significant variation in spatial density of observed points, both over the field-of-view as well as within the objects of interest. This poses the question of how to select the size of the support region, or *scale* of observation, for computing 3-D features that are representative of the local geometry. Scale theory has a rich literature for 2-D and 3-D images but no equivalent exists for unorganized point-sampled data. One method to circumvent this problem is to use a fixed scale that is satisfactory over the entire dataset. This however compromises feature computation both in regions where data is sparse as well as near the spatial boundaries between neighboring data belonging to two different classes. Another approach is to consider multiple scales at the same time. This approach clearly introduces a computational burden as it increases the dimension of the data. Sensor noise also confounds the feature computation process as a larger support region size may be needed to compensate for noise.

In Figure 1 we illustrates two problems associated with scale selection: the presence of multiple manifolds in the support region and the variable density of the data. In Figure 1-(a) we can see that the tree trunk separates into two large branches. The junction area is classified as surfaces as the support region encompass the tree trunk and both branches. Similarly in Figure 1-(b) we can see that the ground data become sparse as the distance to the sensor increase. In that case a fixed scale classification scheme will misclassify semantically those far away points as linear even though the point inside the area of interest define a line.

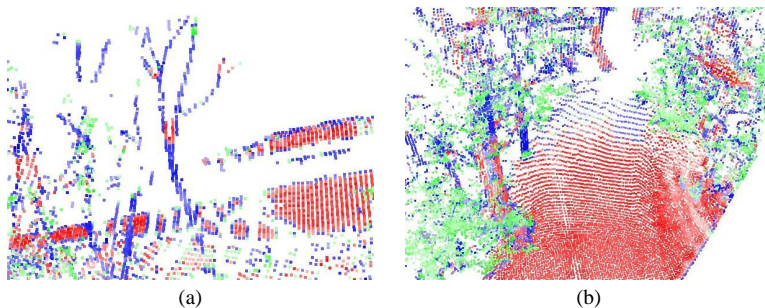


Figure 1: Issues associated with scale selection. (a) Junctions. (b) Density.

This paper presents a technique for determining the scale of observation of point-sampled data by computing the optimal size of the support region for computing surface

normals. Spatial features are then computed at this support size and used in a Bayes classifier for 3-D data segmentation. The method implicitly assumes that the scale that is representative of local geometry at a point is also the one that best discriminates its true class in feature space. We validate this assumption through extensive experiments and detail our approach and its limitations in the sections that follow.

Section 2 presents the related work on scale selection in 3-D. Section 3 details the estimation of the support size and our proposed algorithm. In Section 4 we present classification results on real outdoor data, and then summarize the contributions in Section 5.

2 Related work

It is widely accepted that real world objects appear as meaningful entities at different scales of observation. This has driven the need for rigorous, data-driven formalisms to identify representative scales in data, both for data representation as well as identification.

In [7] the authors emphasize the difference between change in an image (distance scaling) and change in the human perception of that image (information scaling), as the distance to the visual pattern varies.

Pioneering work by Lindeberg [14] equated analysis of continuous signals at successive scales to the suppression of local extrema, and showed that successive smoothing of the signal by gaussian convolution satisfied this property. By this principle, the scale at which the signal response to a normalized differential operator achieves local extrema is a characteristic length of the structure in the signal. This methodology has been extended to discrete signals in 1-D, 2-D and N-D lattices [13, 12]. The scale-invariance property has since been exploited extensively in computer vision as a technique to extract regions with sizes that accommodate scaling of the image and from which invariant features can be computed [4]. In [10] Karid is looking at the relationship between scale, saliency and scene description in 2D images for correspondance and matching problems. Saud [22] used shape token (scale, location and orientation) to describe objects.

However this body of work has focused solely on functions defined on a regular lattice and its applicability to unorganized point samples is unclear.

Some problems are scale or resolution dependent but the best scale/resolution cannot be determined explicitly. The use of multiple scales/resolutions simultaneously is then the best option. The resolution can be considered simultaneously but independently such as in [8] with histograms or the statistical relationship between signature at different scale learned like in [2].

In [19] the authors determines the relationship , at multiple scale, between intensity images of natural images and their corresponding depth images. The goal is to be able to performed scene inference on partial range images. Similarly in [26] the author look at such relationships but to analyze the human perception system. Such research was initiated by Mumford and co-authors by looking at range image statistics [9].

In the domain of point sampled data, efforts have been made to address the problem of scale for surface reconstruction and feature extraction. The tensor voting framework

in [23] equated scale to the region of influence of each tensor, and used it for fine-to-coarse analysis for surface reconstruction. However, no direct relation could be drawn between a choice of region size for tensor voting and that for computing a representative feature for classification. Work in [6] uses k -neighborhoods to compensate for differences in sampling rate before computing eigenvalue-based features for detecting surfaces, creases and borders. There was no guarantee presented that a certain fixed choice of k would be representative of the underlying surface at all points.

Tang et al. [24] use a Kalman filter-based discontinuity preserving line-smoother to detect junctions in 2-D scans. Successive iterations of the smoothing algorithm defined increasing scales of data. However, the method was focused to data modeled as piece-wise lines and not applicable to classification. Work in [18] classifies points based on eigen-values of the local covariance matrix in its n -neighborhood. They define a measure of deviation from planarity at a point that is a function of the eigen-values. It is observed that the value of the scale (n) that maximizes the measure for 1D sinusoidal signals is related to the wavelength of the signal. The scale corresponding to the maximum value is then chosen for computing the feature. However no theoretical guarantees are made regarding suitability of the proposed measure for 3D surfaces or its optimality for classification.

Sara, in [21], used a bottom-up approach to recover the scene geometry from 3-D data generated from a multi-head stereo camera system. Local oriented primitive are extracted at different scale, support region size, then connected.

Finally we would like to mention the work on analysis of galaxy distribution from [20, 3]. The authors are interested in comparing cosmological model with observations by the mean of statistical analysis of the shape distribution and the morphology of the two data sets. The technique, very time consuming and not applicable in our context, rely on using 3-D wavelet transform.

In contrast, this paper proposes to use a neighborhood size consistent with the estimate of local geometry at a point. We make use of recent work in computational geometry [17, 16] and compute a neighborhood size that minimizes an upper bound on expected angular error between the normal estimated at a point through Principal Component Analysis (PCA) and the true normal. The quality of this estimate is improved with knowledge of sensor geometry and error characteristics. A by-product of this process is an estimate of the local covariance matrix that is most consistent with the surface geometry. The eigen-values of this covariance matrix are used in a Bayes classifier to perform point-wise classification of the scene.

3 Approach

The core of this section is based on the work of Mitra et al presented in [17, 16], but we depart from those papers in several original ways: 1) we propose an approach to the estimation of the two critical parameters d_1 and d_2 (Section 3.1.4), 2) we introduce a modification of algorithm to estimate the optimal support region size that is robust to the presence of multiple manifolds (Section 3.1.5), 3) we evaluate the complexity of the approach with in mind implementation on-board a mobile robot (Section 3.1.6), and finally 4) we put the problem in the context of classification (Section 3.2). The reader

will realize that we detail fully the equations and keep the key parameters (density, curvature) visible to be able to intuit their relative influence.

3.1 Normal Estimation in 3-D

This section details the analysis of normal estimation on surfaces in 3-D point cloud data (PCD) as summarized in [17, 16]. We start with a set of N_p points, $\mathbf{p}_i = [x_i \ y_i \ z_i]^\top$, drawn at random from a surface in \mathbb{R}^3 . The goal is to compute the normal at each point of a point cloud with greatest accuracy. This is done by choosing a spatial neighborhood size r that minimizes the *expected* angular deviation of the computed normal at a point from its true normal. In contrast to the analysis in [17], we express the unknown parameters *explicitly* in terms of data dependent quantities and record their dependence on the data distribution and sensor model.

The total least-squares (TLS) estimate of the normal to a set of k points \mathbf{p}_i is given by the eigen-vector corresponding to the smallest eigen-value of the covariance matrix

$$M = \frac{1}{k} \sum_{i=1}^k (\mathbf{p}_i - \bar{\mathbf{p}})(\mathbf{p}_i - \bar{\mathbf{p}})^\top = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{12} & m_{22} & m_{23} \\ m_{13} & m_{23} & m_{33} \end{bmatrix} \quad (1)$$

where $\bar{\mathbf{p}} = \frac{1}{k} \sum_{i=1}^k \mathbf{p}_i$. Note that M is always symmetric positive semi-definite ($M \succeq 0$) and thus has non-negative eigenvalues.

We now review the assumptions made by Mitra et al in [17, 16] and discuss their validity in the context of our framework:

- (A1) Centered data: Without loss of generality, the dataset is centered about the origin O which is the point of interest. The z-axis is the normal to the surface at O and the points of the PCD in the sphere of radius r around O are i.i.d samples of a topological disk R on the underlying surface. We may then model the surface as a function $z = g(x, y)$ that is \mathbb{C}^2 continuous over the r -disk.
- (A2) Spatial density: There exists an $r_0 < \gamma$ such that a sphere of radius r_0 anywhere in R contains at least $k_0 > 0$ points. This implies that data has no holes and has spatial density $\rho > \rho_0 > 0$ everywhere. This assumption holds for full 3-D models. However, it will break when scanning large scale natural scenes, because there will be holes caused by range shadows, and discontinuities at the boundaries of the scans. Moreover, as it is shown in Section 4.3.1, the point density decreases as distance from point to sensor increases.
- (A3) Term z_i is observed with i.i.d. noise $n_i \sim N$ that is identically distributed over the interval R with zero mean, variance σ_n^2 and lies in the range $[-n, n]$. The datasets used in Section 4 satisfy this assumption. The novelty here is that the noise variance σ_n^2 depends on the distance of the point to the sensor. It is determined by sensor calibration results (c.f. Section 4.2.2).
- (A4) Bounded curvature in some neighborhood around the interest point: There exists a positive constant κ such that the Hessian H of g satisfies $\|H\|_2 \leq \kappa$ in the

r -neighborhood. In Section 3.1.5, we show a case where this assumption is violated and we propose a modification to account for it.

(A5) Noise σ_n and curvature κ are small: This in turn implies that m_{11} and m_{22} are the two dominant entries in M .

We proceed by computing bounds on the values in M and then use them to compute a bound on the angular error in the estimated normal at O .

3.1.1 Bounding entries of M

m_{11} and m_{22} By definition $m_{11} = \frac{1}{k} \sum_{i=1}^k (x_i - \bar{x})^2$. The assumption of the points being evenly distributed in the xy -plane bounds m_{11} in the following interval:

$$\theta_1 r^2 \leq m_{11} \leq r^2 \quad (2)$$

where $\theta_1 \in [0, 1]$. Symmetrically, the same applies for m_{22} .

m_{12} By definition,

$$\begin{aligned} |m_{12}| &= \left| \frac{1}{k} \sum_{i=1}^k (x_i - \bar{x})(y_i - \bar{y}) \right| \\ &= \left| \frac{1}{k} \sum_{i=1}^k x_i y_i - \frac{1}{k^2} \sum_{i=1}^k x_i \sum_{i=1}^k y_i \right| \end{aligned}$$

Let's first rewrite some elements of basic probabilities. Suppose we have a random variable X s.t. its instances are x_1, x_2, \dots, x_n . Then, an estimator for the mean is

$$\hat{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

The *bias* of the estimator is simply its expectation:

$$\mathbb{E}[\hat{x}] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[x_i] = \mathbb{E}[X]$$

Similarly, its variance is

$$\mathbb{V}[\hat{x}] = \frac{1}{n^2} \sum_{i=1}^n \mathbb{V}[x_i] = \frac{1}{n} \mathbb{V}[X] \quad (3)$$

Now, let's compute the expectation of m_{12} :

$$\begin{aligned} \mathbb{E}[m_{12}] &= \mathbb{E}\left[\frac{1}{k} \sum_{i=1}^k x_i y_i\right] - \mathbb{E}\left[\frac{1}{k^2} \sum_{i=1}^k x_i \sum_{i=1}^k y_i\right] \\ &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \end{aligned}$$

We assume X and Y to be independent. Therefore, $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ and $\mathbb{E}[m_{12}] = 0$. Now, using the assumption that $\mathbb{V}(x_i y_i) = \theta_2 r^4$ and using equation (3), the variance of m_{12} is expressed as:

$$\mathbb{V}[m_{12}] = \frac{1}{k^2} \sum_{i=1}^k \mathbb{V}[x_i y_i] + \frac{1}{k^4} \sum_{i=1}^k \mathbb{V}[x_i] \sum_{i=1}^k \mathbb{V}[y_i]$$

Because this second term can be made arbitrarily large, we derive the following upper bound on the variance of m_{12} :

$$\mathbb{V}[m_{12}] \leq \frac{1}{k} \theta_2 r^4$$

Chebyshev's Inequality. Let $\mu = \mathbb{E}(X)$ and $\sigma^2 = \mathbb{V}(X)$. Then

$$\mathbb{P}\left(|X - \mu| \leq \frac{\sigma}{\sqrt{\epsilon}}\right) = 1 - \epsilon \quad (4)$$

It follows from Chebyshev's Inequality that with probability $1 - \epsilon$:

$$\begin{aligned} |m_{12}| &\leq \frac{\sqrt{\theta_2 r^4}}{\sqrt{\epsilon k}} = \frac{\sqrt{\theta_2 r^4}}{\sqrt{\epsilon \rho r^2}} = \sqrt{\theta_2} \frac{r}{\sqrt{\epsilon \rho}} \\ &\leq \sqrt{\theta_2} \frac{r}{\sqrt{\epsilon \rho}} \end{aligned}$$

m_{13} and m_{23} We know from the Taylor expansion of $g(x, y)$ that:

$$\begin{aligned} g(x_i, y_i) &= g(0, 0) + x_i g_x(x_i, y_i) + y_i g_y(x_i, y_i) + \\ &\quad \frac{1}{2} \{x_i^2 g_{xx}(x_i, y_i) + 2x_i y_i g_{xy}(x_i, y_i) + y_i^2 g_{yy}(x_i, y_i)\} + \dots \\ &= g(0, 0) + x_i g_x(x_i, y_i) + y_i g_y(x_i, y_i) + \\ &\quad \frac{1}{2} \{x_i^2 g_{xx}(\psi_i, \varphi_i) + 2x_i y_i g_{xy}(\psi_i, \varphi_i) + y_i^2 g_{yy}(\psi_i, \varphi_i)\} \\ &= \frac{1}{2} \{x_i^2 g_{xx}(\psi_i, \varphi_i) + 2x_i y_i g_{xy}(\psi_i, \varphi_i) + y_i^2 g_{yy}(\psi_i, \varphi_i)\} \end{aligned}$$

for some $\psi_i \in [0, x_i]$ and $\varphi_i \in [0, y_i]$.

If we assume that the surface is rotationally invariant, then $g_{xy}(\psi_i, \varphi_i) = 0$ and the expression for $g(x_i, y_i)$ simplifies to:

$$g(x_i, y_i) = \frac{x_i^2}{2} g_{xx}(\psi_i, \varphi_i) + \frac{y_i^2}{2} g_{yy}(\psi_i, \varphi_i)$$

Furthermore, from assumption (A4), $|g_{xx}(x, y)| < \kappa$, $|g_{yy}(x, y)| < \kappa \forall x, y \in R$, we have:

$$\begin{aligned} |z_i| &= |g(x_i, y_i)| \\ &\leq \kappa \left(\frac{x_i^2}{2} + \frac{y_i^2}{2} \right) + n_i \quad \forall x_i, y_i \in R \end{aligned} \quad (5)$$

From the definition of m_{13} :

$$\begin{aligned} |m_{13}| &= \left| \frac{1}{k} \sum_{i=1}^k x_i z_i - \frac{1}{k^2} \sum_{i=1}^k x_i \sum_{i=1}^k z_i \right| \\ &\leq \left| \frac{1}{k} \sum_{i=1}^k x_i z_i \right| + \left| \frac{1}{k^2} \sum_{i=1}^k x_i \sum_{i=1}^k z_i \right| \end{aligned}$$

Substituting Eqn.5, and since $|x_i| \leq r$ and $|y_i| \leq r$:

$$\begin{aligned} |m_{13}| &\leq \left| \frac{1}{k} \sum_{i=1}^k \left(x_i \kappa \left(\frac{x_i^2}{2} + \frac{y_i^2}{2} \right) + n_i \right) \right| \\ &\quad + \left| \frac{1}{k^2} \sum_{i=1}^k x_i \sum_{i=1}^k \left(\kappa \left(\frac{x_i^2}{2} + \frac{y_i^2}{2} \right) + n_i \right) \right| \\ &\leq 2\kappa r^3 + \left| \frac{1}{k} \sum_{i=1}^k x_i n_i \right| + r \left| \frac{1}{k} \sum_{i=1}^k n_i \right| \end{aligned}$$

Under the assumption that X and N are independent, we note that $\mathbb{E}[x_i n_i] = \mathbb{E}[x_i] \mathbb{E}[n_i] = 0$ since $\mathbb{E}[n_i] = 0$ and $\mathbb{V}(x_i n_i) = C r^2 \sigma_n^2$. Using Chebyshev's inequality, we have that with probability $1 - \epsilon$:

$$\begin{aligned} |m_{13}| &\leq 2\kappa r^3 + C \sqrt{\frac{r^2 \sigma_n^2}{\epsilon k}} + r \sqrt{\frac{\sigma_n^2}{\epsilon k}} \\ &\leq 2\kappa r^3 + \theta_3 \frac{\sigma_n}{\sqrt{\epsilon \rho}} \end{aligned} \quad (6)$$

where $\theta_3 = C + 1$. Symmetrically, the same procedure applies to m_{23} , by replacing x_i by y_i .

m_{33} Finally for m_{33} ,

$$\begin{aligned}
m_{33} &= \frac{1}{k} \sum_{i=1}^k z_i^2 - \frac{1}{k^2} \left(\sum_{i=1}^k z_i \right)^2 \leq \frac{1}{k} \sum_{i=1}^k z_i^2 \\
&\leq \frac{1}{k} \sum_{i=1}^k (\kappa^2 r^4 + 2\kappa n_i r^2 + n_i^2) \\
&\leq \frac{1}{k} \sum_{i=1}^k 2(\kappa^2 r^4 + n_i^2) \\
&\leq 2\kappa^2 r^4 + \theta_4 \sigma_n^2
\end{aligned} \tag{7}$$

To summarize, we have defined the following bounds for each entry of M :

$\theta_1 r^2 \leq m_{11} \leq r^2$	$ m_{12} \sqrt{\theta_2} \frac{r}{\sqrt{\epsilon \rho}}$	$ m_{13} \leq 2\kappa r^3 + \theta_3 \frac{\sigma_n}{\sqrt{\epsilon \rho}}$
-	$\theta_1 r^2 \leq m_{22} \leq r^2$	$ m_{23} \leq 2\kappa r^3 + \theta_3 \frac{\sigma_n}{\sqrt{\epsilon \rho}}$
-	-	$m_{33} \leq 2\kappa^2 r^4 + \theta_4 \sigma_n^2$

3.1.2 Eigen analysis

We may write the covariance matrix M as

$$M = \left[\begin{array}{cc|c} m_{11} & m_{12} & m_{13} \\ m_{12} & m_{22} & m_{23} \\ \hline m_{13} & m_{23} & m_{33} \end{array} \right] \triangleq \begin{bmatrix} M_{11} & M_{13} \\ M_{13}^T & m_{33} \end{bmatrix} \tag{8}$$

Gershgorin Circle Theorem. For an $n \times n$ matrix, define

$$R_i = \sum_{\substack{j \\ i \neq j}}^n |M_{ij}|$$

Then each eigen-value of M is in at least one of the discs

$$\{z : |z - M_{ii}| \leq R_i\}$$

Let $\lambda_1 \leq \lambda_2$ be the eigen-values of M_{11} . Using the Gershgorin Circle Theorem (GCT), we have that $m_{11} - |m_{12}| \leq \lambda_1 \leq \lambda_2 \leq m_{22} - |m_{12}|$

Let us define a new dimensionless quantity α as:

$$\alpha \triangleq \frac{|m_{13}| + |m_{23}| + m_{33}}{m_{11} - |m_{12}|} \tag{9}$$

Let λ be the smallest eigen-value of M . Using GCT again gives $\lambda \leq |m_{13}| + |m_{23}| + m_{33} = \alpha(m_{11} - |m_{12}|) \leq \alpha \lambda_1$. If we take the eigen-vector corresponding to the minimum eigen-value of M as $[\vec{v}^T, 1]^T$, then

$$\begin{bmatrix} M_{11} & M_{13} \\ M_{13}^T & m_{33} \end{bmatrix} \begin{bmatrix} \vec{v} \\ 1 \end{bmatrix} = \lambda \begin{bmatrix} \vec{v} \\ 1 \end{bmatrix}$$

Expanding to solve the individual equations gives

$$\begin{aligned} \vec{v} = & -(M_{11} - \lambda\mathbb{I})^{-2} [\mathbb{I} + (M_{11} - \lambda\mathbb{I})^{-2} M_{13} M_{13}^T]^{-1} \\ & \times [(M_{11} - \lambda\mathbb{I}) M_{13} + M_{13} (m_{33} - \lambda)] \end{aligned} \quad (10)$$

$$\begin{aligned} \|\vec{v}\|_2 \leq & \left(\|(M_{11} - \lambda\mathbb{I})^{-2}\|_2 \right. \\ & \left. \|(\mathbb{I} + (M_{11} - \lambda\mathbb{I})^{-2} M_{13} M_{13}^T)^{-1}\|_2 \right) \\ & \times \left(\|(M_{11} - \lambda\mathbb{I})\|_2 \|M_{13}\|_2 \right. \\ & \left. + \|M_{13}\|_2 |(m_{33} - \lambda)| \right) \end{aligned} \quad (11)$$

It can be shown that

$$\|(M_{11} - \lambda\mathbb{I})^{-2} M_{13} M_{13}^T\|_2 \leq (1 - \alpha)^{-2} \alpha^2 \quad (12)$$

and hence

$$\|(\mathbb{I} + (M_{11} - \lambda\mathbb{I})^{-2} M_{13} M_{13}^T)^{-1}\|_2 \leq \frac{(1 - \alpha)^2}{1 - 2\alpha} \quad (13)$$

It then follows that:

$$\begin{aligned} \|\vec{v}\|_2 & \leq \frac{1}{(1 - \alpha)^2 \lambda_1^2} \frac{(1 - \alpha)^2}{1 - 2\alpha} (\lambda_2 \alpha \lambda_1 + (\alpha \lambda_1)^2) \\ & \leq \frac{\alpha(1 + \alpha)}{1 - 2\alpha} \frac{\lambda_2}{\lambda_1} \approx \left(\frac{\lambda_2}{\lambda_1} \right) \alpha \end{aligned} \quad (14)$$

for small α .

Hence the angle between the computed normal and the true normal is bounded from above by

$$\tan^{-1} \|\vec{v}\|_2 \leq \frac{\lambda_2}{\lambda_1} \alpha \leq \frac{(m_{22} + |m_{12}|)}{(m_{11} - |m_{12}|)} \alpha \quad (15)$$

3.1.3 Error bound for the estimated normals

From Eqns. (2),(3.1.1),(6) and (7), we can replace each m_{ij} term in Eqn (9) by its appropriate bound value to give:

$$\alpha \leq \frac{2\kappa^2 r^4}{\theta_1 r^2} + \frac{\theta_4 \sigma_n^2}{\theta_1 r^2} + \frac{2 \left| 2\kappa r^3 + \frac{\theta_3 \sigma_n}{\sqrt{\epsilon \rho}} \right|}{\theta_1 r^2}$$

Since the values κ , r , σ_n and ρ are always positive, by simplifying and re-arranging, we get:

$$\alpha \leq \frac{2}{\theta_1} \kappa^2 r^2 + \frac{\theta_4}{\theta_1} \frac{\sigma_n^2}{r^2} + \frac{4}{\theta_1} \kappa r + \frac{2\theta_3}{\theta_1} \frac{\sigma_n}{r^2 \sqrt{\epsilon \rho}}$$

Let us define $\beta \triangleq m_{12}/m_{11}$ and consider cases where $\beta < 1/2$. Since we have

$$\frac{\lambda_1}{\lambda_2} \alpha \leq \frac{m_{22} (1 + \beta)}{m_{11} (1 - \beta)} \alpha \leq K \alpha \quad (16)$$

we have from the previous lower bound

$$\begin{aligned} \frac{\lambda_1}{\lambda_2} \alpha &\leq K \left(\frac{2}{\theta_1} \kappa^2 r^2 + \frac{\theta_4}{\theta_1} \frac{\sigma_n^2}{r^2} + \frac{4}{\theta_1} \kappa r + \frac{2\theta_3}{\theta_1} \frac{\sigma_n}{r^2 \sqrt{\epsilon \rho}} \right) \\ &\leq K \left(\frac{\theta_4}{\theta_1} \frac{\sigma_n^2}{r^2} + \frac{4}{\theta_1} \kappa r + \frac{2\theta_3}{\theta_1} \frac{\sigma_n}{r^2 \sqrt{\epsilon \rho}} \right) \end{aligned} \quad (17)$$

Differentiating Eqn.(17) w.r.t r gives the required result:

$$r = \left(\frac{1}{\kappa} \left(\theta_3 \frac{\sigma_n}{\sqrt{\epsilon \rho}} + \frac{\theta_4}{2} \sigma_n^2 \right) \right)^{\frac{1}{3}} \quad (18)$$

where the constants $d_1 = \theta_3$ and $d_2 = \theta_4/2$, as given in [17, 16] are to be determined experimentally. Note that d_1 and d_2 depend only on the distribution of the PCD, since, as shown in Eqns (6) and (7), θ_3 is related to $\mathbb{V}(x_i n_i)$ while θ_4 is related to σ_n .

3.1.4 d_1 and d_2 estimation

The constants of d_1 and d_2 of Equation 18 needs to be estimated from the dataset. In [16], they mention that those constants were chosen by trial and error, and that they picked the one resulting in visually good results. In this section, we explore different ways of estimating those constants.

By simply re-arranging Equation 18, we see that it is now linear in d_1 and d_2 :

$$r^3 = \frac{1}{\kappa} \left(d_1 \frac{\sigma_n}{\sqrt{\epsilon \rho}} + d_2 \sigma_n^2 \right)$$

We want to minimize the following equation:

$$\min \sum_{i=0}^n \left\| \frac{1}{\kappa_i} \left(d_1 \frac{\sigma_n}{\sqrt{\epsilon \rho_i}} + d_2 \sigma_n^2 \right) - r_i^3 \right\|^2$$

This is done the following way.

1. Fix ϵ , k_0 and compute σ_n (either from sensor or from synthetic data).
2. For every point x_i :
 - (a) Find r_i that minimizes angular separation between computed and true normals at that point
 - (b) compute the local density $\rho_i = \frac{k_0}{\pi s^2}$ where s is the distance from the k_0 th nearest-neighbor to the point x_i .
 - (c) compute the local maximum curvature $\kappa_i = \frac{2d}{\mu}$, where d is the distance from the x_i to the TLS plane fitted on the neighborhood defined by r_i , and μ is the average distance from the point x_i to all its neighbors.

(d) build the linear system

$$\begin{bmatrix} \frac{\sigma_n}{\kappa_0 \sqrt{\epsilon \rho_0}} & \frac{\sigma_n^2}{\kappa_0} \\ \vdots & \vdots \\ \frac{\sigma_n}{\kappa_n \sqrt{\epsilon \rho_n}} & \frac{\sigma_n^2}{\kappa_n} \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} r_0^3 \\ \vdots \\ r_n^3 \end{bmatrix} \quad (19)$$

3. Solve the system in Eq. 19 by linear minimization using SVD or pseudo-inverse.

However, this method doesn't yield satisfying results because outliers are numerous, and least-squares method are typically sensible to them. A way to improve this is by using RANSAC. The linear system in Eq. 19 is first built the exact same way as before. Then, the first m rows in the matrices are chosen, and d_1 and d_2 are estimated using linear least-square applied on the resulting matrices. The score of the result is then estimated by counting the number of inliers, i.e. points for which the computed constants yield a good approximation of the desired radius. This process is repeated several times, and the constants that have the best score are kept.

The main drawback of trying to estimate d_1 and d_2 is that it requires knowledge of ground truth normals to estimate the best support region for every point. In real world scenes (Section 4.4), we have no such information and we can only evaluate results visually. Therefore, a calibration scene with known normals at each point would be necessary to perform such estimation. For example, the dataset in Section 4.3.2 might be appropriate since all points lie on surfaces of known relative orientation.

3.1.5 Estimating the optimal support region size

The optimal r is estimated using an iterative procedure based on the suggestions in [17]. An initial value of $k = k^{(i)}$ is used to compute a starting value of curvature $\kappa^{(i)}$ and $r^{(i)}$ is taken as the distance to the k -th nearest neighbor. An estimate of density $\rho^{(i)}$ is also obtained from $k = k^{(i)}$. The value of σ^2 is taken from the sensor model as a fixed function of the distance of the point from the laser. The value of $r^{(i+1)}$ for the $(i + 1)$ -th iteration is then computed using Eqn.(18). $k^{(i+1)}$ is then computed as the number of points in a neighborhood sized $r^{(i+1)}$ and the process is continued.

We observed that the iterative procedure suggested in [17, 16] had poor convergence properties when assumption (A4) is broken. As shown in Section 4.3.2, this can happen when two manifolds are located in the region of interest. Figure 2(a)-(c) shows the computed values of $r^{(i)}$ oscillating for points selected near regions of higher curvature, as in the case of intersecting walls in Figure 7. We modify the algorithm to perform damped updates to k using an infinite impulse-response (IIR) filter of the form:

$$k^{(i+1)} = \gamma k_{\text{computed}}^{(i+1)} + (1 - \gamma)k^{(i)} \quad (20)$$

The parameter γ defines how much importance is given to $k^{(i+1)}$ versus $k^{(i)}$. Figure 2(b)-(d) shows that the values of r after convergence do not depend on the initial $k^{(0)}$. It also illustrates that the IIR filter ensures controlled updates in each iteration

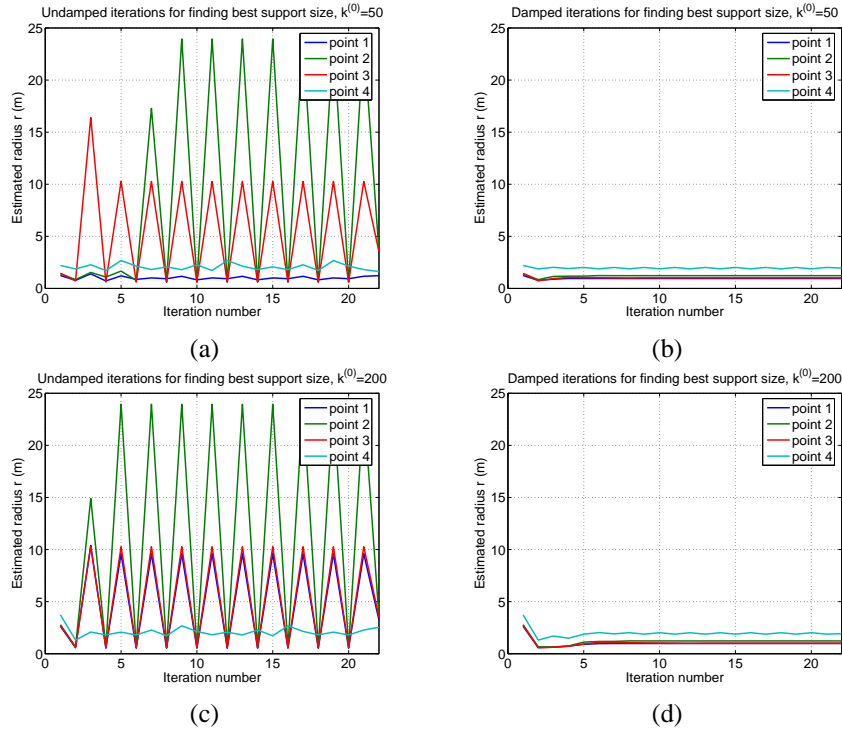


Figure 2: Plot of estimated support region size (r) at each iteration showing improvement with damped updates using (a)-(b) an initial value of $k^{(0)} = 50$, and (c)-(d) $k^{(0)} = 200$. The support region sizes converge to the same value, independently of $k^{(0)}$.

and assures sensible values of r near intersections of manifolds. This is reflected in the smaller support-region size near the intersections of the two walls in Figure 7 and in the region where the tree trunks meet the ground in Figure 8.

Figure 3 illustrates the progression of the damped and undamped algorithm with different values of $k^{(0)}$ for a typical point in the scene of Figure 7. In this example, the damped algorithm converges to a scale at which the normal estimation error is very low. Without dampening however, convergence is never reached and a high error is maintained. The final scale determined by the algorithm does not depend on the initial $k^{(0)}$ for 70% of points with the dampening, as opposed to only 28% without it. Therefore, the IIR filter reduces error in normal estimation, and makes the process less dependent on the initial $k^{(0)}$.

3.1.6 Complexity analysis

The most costly operation of the algorithm is the nearest-neighbor search. It is needed to compute the TLS plane and approximate the local curvature and density at each point. [1] has shown that an approximate nearest-neighbor search can be done in two

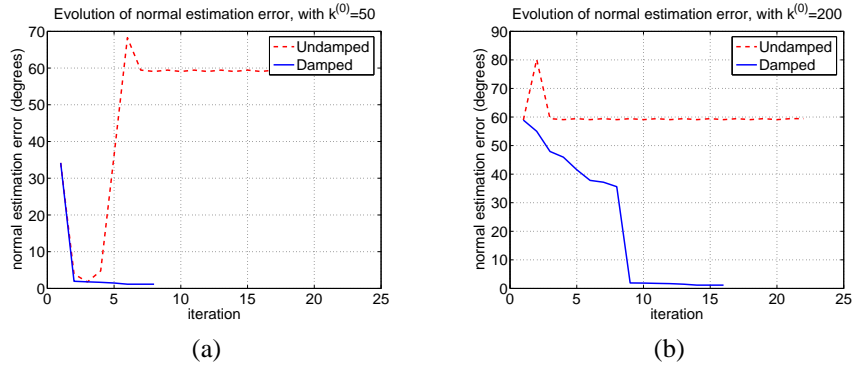


Figure 3: Comparison between evolution of normal estimation error for the undamped and damped versions of the algorithm, with (a) $k^{(0)} = 50$ and (b) $k^{(0)} = 200$. In the two cases, the damped algorithm converges to the same low error value in less than 20 iterations.

steps. First, it requires a data preprocessing stage that can be done in $O(dn \log n)$ time. In our case, $d = 3$, so preprocessing is of complexity $O(n \log n)$. Then, finding the k -nearest neighbors requires $O(kd \log n)$, which simplifies to $O(k \log n)$ in our case.

Since the k -nearest neighbor search is done for the n points and the number of iterations needed to reach convergence is bounded by a constant `maxCount`, the total complexity of the algorithm is then $O(n \log n) + O(nk \log n)$. Since k is bounded by the constant $k_{threshold}$, by choosing $k_{threshold} \ll n$ and since $k_{threshold}$ doesn't depend on n , we can approximate the complexity by $O(n \log n)$.

3.2 Terrain Classification

We focus on segmentation of ladar data into 3 classes – *clutter* to represent vegetation, *linear structures* to represent thin objects like wires and tree branches, and *surface* to capture ground, rock and tree-trunk surfaces. Our approach for classification is based on computing saliency features [25] that capture the local geometry at a point in terms of spatial distribution of points in its neighborhood. The distribution of saliency features is learned using a Gaussian mixture Model (GMM) automatically using the Expectation-Maximization (EM) algorithm. Given the distribution learned off-line, we can classify new data online using a Bayes classifier.

3.2.1 Saliency features

Our choice of features is inspired by the tensor voting framework in [23]. However, instead of looking at the distribution of surface normals in a neighborhood, we directly inspect the local distribution of 3-D points. This is done by computing the covariance matrix M (Eqn. (1)) corresponding to the scatter of the points in a local neighborhood, the support region.

The size of the support region defines the *scale* of the feature and is chosen to be the

radius r computed in Section 3.1.5. Note that M is computed in the intermediate steps while estimating r , and is representative of the local geometry of the neighborhood. Let $\lambda_1 \leq \lambda_2 \leq \lambda_3$ be the eigen-values of M corresponding to eigen-vectors m_1, m_2, m_3 respectively. In case of clutter, $\lambda_1 \approx \lambda_2 \approx \lambda_3$ and there is no dominant direction. For points on surfaces, $\lambda_3, \lambda_2 \gg \lambda_1$ and e_3, e_2 span the local plane of observations. For linear structures $\lambda_3 \gg \lambda_2, \lambda_1$ and e_3 is the dominant direction locally. Our saliency feature is defined as a linear combination of eigen-values in the 3-vector:

$$\begin{bmatrix} \text{point-ness} \\ \text{surface-ness} \\ \text{curve-ness} \end{bmatrix} \triangleq \begin{bmatrix} \lambda_1 \\ \lambda_2 - \lambda_1 \\ \lambda_3 - \lambda_2 \end{bmatrix} \quad (21)$$

3.2.2 Bayesian classification

Using the features of Eqn. (21) and a dataset labeled into the 3 classes, we train a GMM using the EM algorithm. Let the n_i components of the Gaussian mixture in the i -th class be specified by the set of weights, means and covariances as $C_i = \{(w_{(i,j)}, \mu_{(i,j)}, \Sigma_{(i,j)})_{j=1 \dots n_i}\}$ for $i = 1, 2, 3$. The likelihood of a new point x with feature $f(x) \in \mathbb{R}^3$ computed with Eqn.(21) belonging to class C_i is given by:

$$\begin{aligned} \mathbb{P}(f(x)|C_i) &= \sum_{j=1}^{n_i} \left(\frac{w_{(i,j)}}{(2\pi)^{3/2} |\Sigma_{(i,j)}|^{1/2}} \right. \\ &\quad \left. \times e^{-\frac{1}{2}(f(x)-\mu_{(i,j)})^T \Sigma_{(i,j)}^{-1} (f(x)-\mu_{(i,j)})} \right) \end{aligned} \quad (22)$$

The estimated class is the maximizer of the class posterior:

$$\begin{aligned} C_{est} &= \arg \max_i (\mathbb{P}(C_i|f(x))) \\ &= \arg \max_i (\mathbb{P}(f(x)|C_i)\mathbb{P}(C_i)) \end{aligned} \quad (23)$$

where $\mathbb{P}(C_i)$ represents the corresponding class prior.

4 Experiments

4.1 Sensors and terrains

To validate the approach presented we used data collected with a Minolta scanner, an actuated SICK laser, a Zoller-Fröhlich high resolution scanner [5] and the CMU autonomous helicopter [15]. The Minolta Vivid 700 is a laser line striper that produces a 200×200 -pixel range image with 8 bits resolution. A SICK LMS-291 is attached to a custom made scanning mount. The laser collects 60,000 points per scan. The angular separation between laser beams is $\frac{1}{4}$ degree over 100 degrees field of view. The angular separation between laser sweeps is $\frac{2}{3}$ of a degree over 115 degrees. The Zoller-Fröhlich (Z+F) LARA 21400 has a $360^\circ \times \pm 35^\circ$ FOV, producing 8000×1400 pixels range and reflectance images of the environment up to 21.4 m. The CMU autonomous helicopter

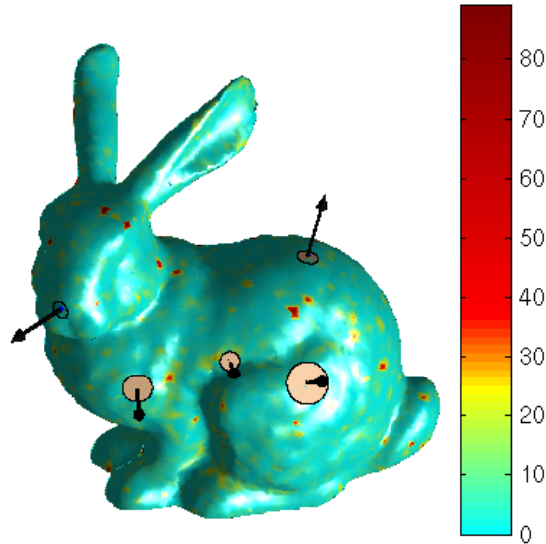


Figure 4: Normal estimation error for the bunny model and normals at selected points.

is equipped with a modified Riegl laser range finder that is capable of collecting 3D color data, with 10 cm accuracy.

We used these sensors to collect data from outdoor environments in urban settings, in natural open space and in a forest.

4.2 Validation of computed normals

In this section, we validate our implementation of the algorithm proposed by [17] by testing it on simple models for which the ground truth normals are known at each point. We first try on full 3-D models that satisfy the assumptions made in Section 3.1, then move on to data collected from the sensors previously described.

4.2.1 3-D models

We tested using the bunny model with the same parameters as in [17]. More specifically, we used $d_1 = 1$, $d_2 = 4$, $k_0 = 15$, $\epsilon = 0.1$, $k_{threshold} = 300$ and $maxCount = 10$. Ground truth normals are computed from the mesh. Figure 4 shows that the normal estimation error is generally very low, except on regions of high curvature. The error is illustrated by a color code, shown on the right of the figure. Points with highest error are colored in red, whereas light blue indicates low error. The error is the angular difference between the true and estimated normals and is expressed in degrees. The support regions are also shown for various points, along with the normals computed from those regions. The similarity of our results compared to those obtained in [17] confirm the expected behavior of the algorithm on full 3-D models.

4.2.2 Aerial lidar scan

A major difference between real datasets and full 3-D models, such as the bunny, is the distance of the scene to the sensor during the scanning process. With small objects, the sensor is close to the object, therefore acquisition noise is small. However, outdoor natural scenes have dimensions that make scanning at close range impossible, thus increasing noise.

We tested this approach using data from an open space natural environment containing a 1.5 m high pile of gravel surrounded by short cut and uncut grass. We collected high resolution, high density data with the Z+F laser. We also collected low-resolution aerial data for the same scene with the CMU autonomous helicopter. The two data sets are co-registered. We triangulate the Z+F data to produce the ground truth used to estimate the normal reconstruction error in the aerial data. The parameters used are the same as before, except that we increased `maxCount` to 20 to allow more time for convergence.

Figure 5 shows the results obtained. Figure 5-(a) shows the computed normals and the support regions for selected points in the aerial data. Figure 5-(b) shows the normal and support regions for the same points but overlaid on top of the high-resolution ground data. Points in Figure 5-(a) are color-coded by the difference between the error in estimated normals and the lowest possible error obtainable for any choice of support region in the aerial data. This lower bound on the error is computed by fitting a least-square plane on the k -nearest neighbors, with k ranging from k_0 to $k_{threshold}$, and retaining the smallest angular difference between the normal to the plane and the ground truth normal. In that example, the lower bound on the error averaged over all points is 5.1 degrees, while the method gives an average error of 9.9 degrees. Even with considerable noise, the algorithm behaves well by giving results close to the lowest possible error.

4.3 Validation of support regions

In this section we analyze the influence of diverse factors on the choice of support region size.

4.3.1 Outdoor ground scan

Another important difference between datasets used in [17] and our experiments lies in the greater size of real scenes. The former allows very high and constant density throughout the model, whereas in our case, the density varies with the distance to the sensor and may become very low. The effect of this important difference is illustrated in figure 6, which shows a scan of the ground taken by the SICK laser. The ground truth is defined as normals pointing along the positive z -axis.

Moreover, because the distance from scene points to the sensor may vary considerably over the dataset, the noise on each data point vary accordingly. To account for this effect, the noise standard deviation σ_n is computed using calibration data. Its value ranges from $\sigma_{min} = 0.0037$ at 1 meter, and $\sigma_{max} = 0.0125$ at 60 meters. As expected, the support region size grows as points are further away from the sensor. Moreover, the

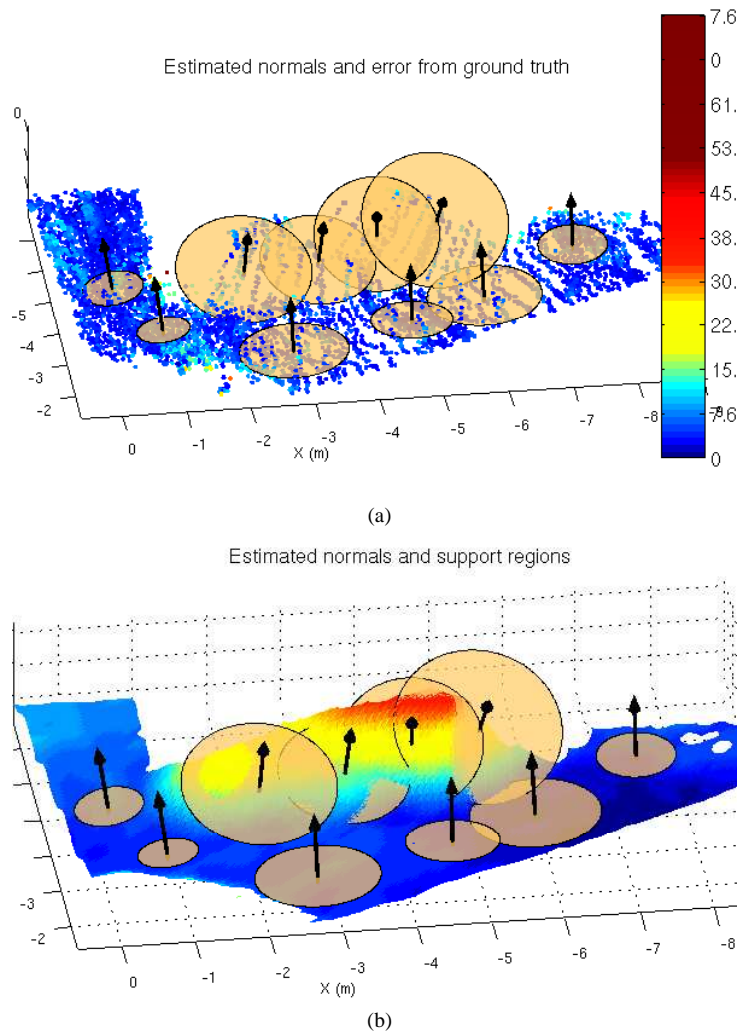


Figure 5: Normal estimation for the aerial data. Normal estimation and corresponding support region for selected points overlaid on top of the aerial data (a) (see text for explanation of the color coding) and ground data (b) with the elevation color coded.

discontinuities located at the boundaries of the laser FOV represent another important difference and break the assumption (A2) stated in Section 3.1. In this case, it doesn't affect the performance of the algorithm because all the points lie in the same plane.

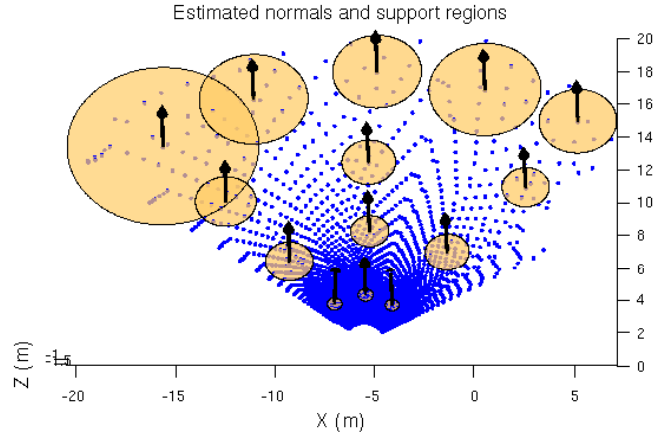


Figure 6: Plot of ground points with estimated support region size (r). Note the significant decrease in spatial density and corresponding increase in r with distance from the laser position (origin)

4.3.2 Scan of wall corner

This dataset is a scan of walls made using the SICK laser, and the sensor was placed at a distance of approximately 30 meters from the scene. Again, σ_n is computed using calibration data, and the same parameters are used in the algorithm. The scene presents a sharp change in curvature at the junction of both walls. This implies the presence of two different manifolds in the neighborhood of points located in the vicinity of that region. We note that the assumption (A5) in Section 3.1 is broken. Intuitively, we would expect that the support region should be relatively small near the junction, as not to include points lying on a different manifold. However, Figure 7 shows that it is not the case with the original algorithm. Undamped iterations cause the algorithm to stop at arbitrary values after a fixed number of iterations, with no guarantee of convergence. This results in badly estimated normals, especially around the discontinuity region. The results obtained with the IIR filter with $\gamma = 0.5$ (Figure 7) introduced in Section 3.1.5 corresponds to what we expected. The normal estimation is much better for the points lying near the corner, and is still as good for the other points. We obtain an average improvement of 10° in normal estimation using the dampening.

4.3.3 Outdoor natural terrain

This dataset was obtained using the SICK scanner and by looking at outdoor natural terrain, comprised of ground, trees and vegetation. Again, we would expect the support region to be small near sharp angles in the geometry of the scene, and larger if the scene is flat, or if the density is small. For this dataset, no ground truth is available, so the results are evaluated visually.

Figure 8 shows the support region determined by our algorithm for different points

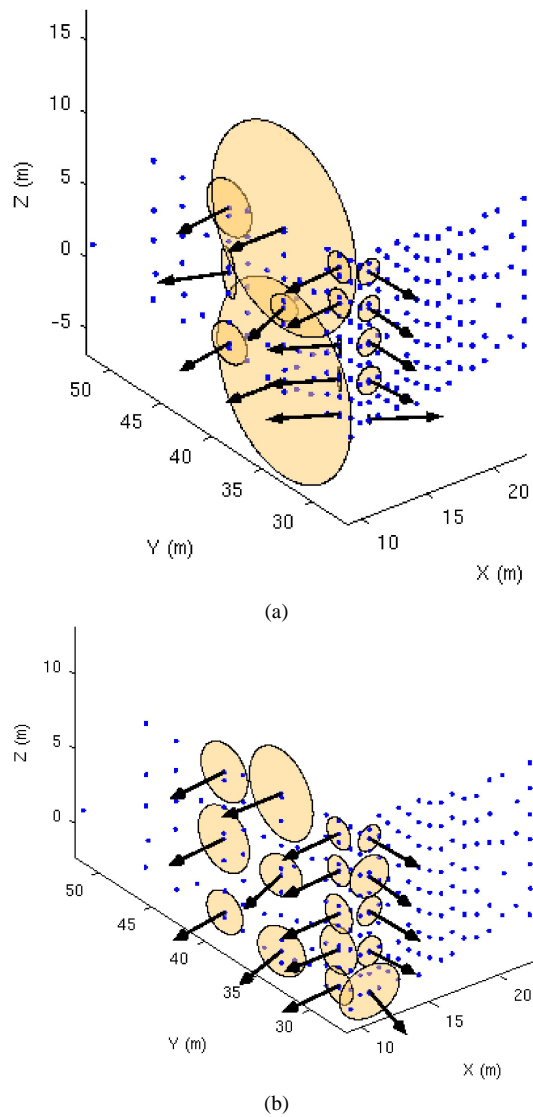


Figure 7: Estimate of support region size for wall corner (a) without and (b) with IIR filter to k in each iterations.

chosen at interesting locations in the scene. For example, the support regions of points located near the boundary of tree trunks and ground are much smaller than those in the center of the ground. This corresponds to the expected behavior.

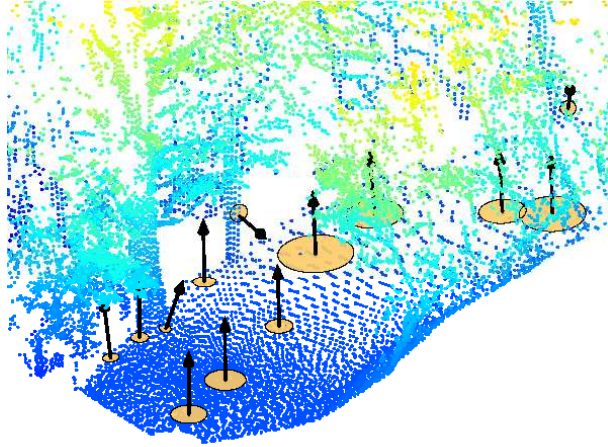


Figure 8: Support region sizes for selected points in outdoor vegetated terrain. Points are color-coded by height.

4.4 Ground-based lidar classification of natural terrain

In this section, we apply the algorithm to the classifier previously described in Section 3.2.2. The dataset is divided into cubic voxels with 10 cm edges. The classifier is then trained at scales ranging from 0.1 cm to 2 m using manually labeled data. The best scale is chosen by applying the method and rounding the given support region to the nearest subdivision. Figure 9-(a) shows the classification results using fixed support region size (radius of 40 cm). Obvious misclassification errors are made near the junction of the leftmost tree and the ground, and on the ground at a distance. Figure 9-(b) shows the improvement over the old strategy.

We manually labeled the data to produce ground truth classification. Over the whole dataset, 9575 points are labeled as surfaces. Using the old strategy, 1918 points are mis-labeled and identified either as clutter or linear structures. The new strategy is able to reduce this number to 1343 mis-classified points, an improvement of approximately 30 percent. On the other hand, of the $9575 - 1918 = 7657$ correctly classified points, only 172, or 2.25 percent are corrupted by the new method.

4.5 Comparison with multi-scale approach

A naïve alternative to the proposed algorithm is to train a different classifier for each scale in the set of considered scales, evaluate a test point on all the classifiers, and simply assign it the label returned with most confidence (highest posterior probability). However, when applying this strategy (with scales ranging from 0.1 m to 2 m) to outdoor natural terrain such as the one shown in Figure 9, we obtain a mere 45% of correctly classified points, as opposed to 84% with the method presented in this paper. As expected, the naïve strategy incorrectly favors very large support regions that include a large number of out-of-class points to give the commonly incorrect label of

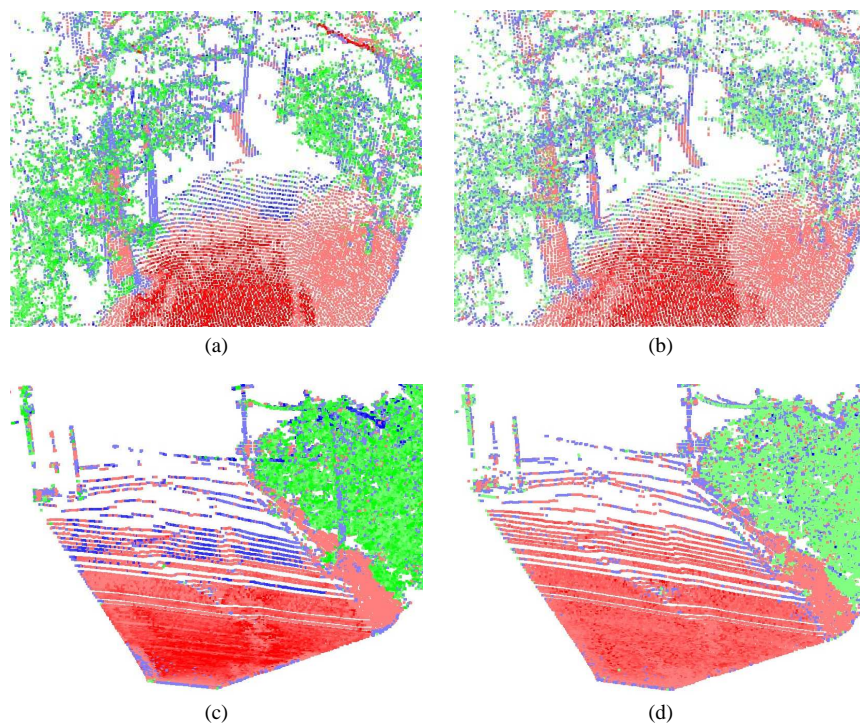


Figure 9: Outdoor terrain classification: (a-b) from the data set using in Figure 8 and (c-d) from data collected with the Riegl laser scanner. Points are colored green (vegetation), red (surface) or blue (linear structure). Darker shades indicate higher confidence in the estimated label. (a/c) Former strategy. (b/d) New strategy.

“vegetation” with high confidence.

5 Conclusions and Discussion

This paper presented a geometry-driven approach for choosing the scale of observation for classifying point-sampled surfaces in outdoor range data. Extensive experiments with outdoor and synthetic datasets confirm our hypothesis that feature computation at scales that are optimal in terms of inferred local geometry improve the quality of classification.

One implicit hypothesis of the proposed approach is that there exists at least one scale at which the data is classified correctly. Closer analysis of points misclassified in Figure 9-(b) in the boundary regions of the dataset show that this hypothesis is violated. We attribute this to (1) the introduction of edge-effects in the chosen features (Eqn. 21) causing them to be unrepresentative of the local geometry, and (2) the possibly poor discriminative ability of the classifier. The assumption of an underlying surface of bounded curvature at each point is also violated for scattered point clouds. In some

regions this results in a reduction of confidence for the vegetation class. The design of more representative shape features as well as eigen-analysis for curved and porous geometry is the subject of our current research.

References

- [1] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. We. An optimal algorithm for approximate nearest neighbor searching fixed dimensions. *Journal of the ACM*, 46(6):891–923, 2000.
- [2] J. D. Bonet, P. Viola, and J. Fisher. Flexible histograms: a multiresolution target discrimination model. In *Proceedings of SPIE*, volume 3370, 1998.
- [3] D. Donoho, O. Levi, J. Stack, and V. Martinez. Multiscale geometric analysis for 3-d catalogues. In *Multiscale Geometric Analysis: Theory, Tools, and Applications*, 2003.
- [4] Y. Dufournaud, C. Schmid, and R. Horaud. Matching image with different resolutions. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2000.
- [5] D. L. et al. Imaging lidar for 3-d surveying and cad modeling of real world environments. *International Journal of Robotics Research*, 19(11), 2000.
- [6] S. Gumhold, X. Wang, and R. Macleod. Feature extraction from point clouds. In *Intl. Meshing Roundtable*, 2001.
- [7] C.-E. Guo, Y. Wu, and S. Zhu. Information scaling laws in natural scenes. In *2nd Workshop on Generative Model Based Vision*, 2004.
- [8] E. Hadjidemetriou, M. Grossberg, and S. Nayar. Spatial information in multiresolution histograms. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2001.
- [9] J. Huang, A. Lee, and D. Mumford. Statistics of range images. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2000.
- [10] T. Karid and M. Brady. Saliency, scale and image description. *International Journal of Computer Vision*, 45(2), 2001.
- [11] J.-F. Lalonde, R. Unnikrishnan, N. Vandapel, and M. Hebert. Scale selection for classification of point-sampled 3-d surfaces. In *International Conference on 3-D Digital Imaging and Modeling*, June 2005.
- [12] J.-Y. Lim and H. S. Stiehl. A generalized discrete scale-space formulation for 2-d and 3-d signals. In *International Conference on Scale-Space Theories in Computer Vision*, 2003.
- [13] T. Lindeberg. Scale-space for discrete signals. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(3), 1990.
- [14] T. Lindeberg. Scale-space theory: A basic tool for analysing structures at different scales. *Journal of Applied Statistics*, 21(2), 1994.
- [15] J. R. Miller. *A 3D Color Terrain Modeling System for Small Autonomous Helicopters*. PhD thesis, Carnegie Mellon University, 2002.

-
- [16] N. Mitra and A. Nguyen. Estimating surface normals in noisy point cloud data. In *Symp. on Computational Geometry*, 2003.
- [17] N. Mitra, A. Nguyen, and L. Guibas. Estimating surface normals in noisy point cloud data. *Special Issue of Intl. Journal of Computational Geometry and Applications*, 14(4-5):261–276, 2004.
- [18] M. Pauly, R. Keiser, and M. Gross. Multi-scale feature extraction on point-sampled surfaces. In *Eurographics*, 2003.
- [19] B. Potetz and T. Lee. Scaling in natural scenes and the inference on 3d shape. In *Conference on Neural Information Processing Systems*, 2004.
- [20] P. Querre, J. L. Starck, and V. J. Martinez. Analysis of the galaxy distribution using multiscale methods. In *SPIE Conference on Astronomical Data Analysis*, 2002.
- [21] R. Sara and R. Bajcsy. Fish-scales: Representing fuzzy manifolds. In *IEEE International Conference on Computer Vision*, 2004.
- [22] E. Saund. Symbolic construction of a 2-d scale image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(8), 1990.
- [23] C. Tang, G. Medioni, P. Mordohai, and W. Tong. First order augmentations to tensor voting for boundary inference and multiscale analysis in 3-d. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(5), 2004.
- [24] F. Tang, M. Adams, J. Ibanez-Guzman, and W. Wijesoma. Pose invariant, robust feature extraction from range data with a modified scale space approach. In *IEEE International Conference on Robotics and Automation*, 2004.
- [25] N. Vandapel, D. Huber, A. Kapuria, and M. Hebert. Natural terrain classification using 3-d ladar data. In *IEEE International Conference on Robotics and Automation*, 2004.
- [26] Z. Yang and D. Purves. Image/source statistics of surface in natural scenes. *Network: Computation in Neural Systems*, 14:371–390, 2003.