

2005

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

Jean-Francois Lalonde
Carnegie Mellon University

Ranjith Unnikrishnan
Carnegie Mellon University

Nicolas Vandapel
Carnegie Mellon University, vandapel@ri.cmu.edu

Martial Hebert
Carnegie Mellon University, hebert@ri.cmu.edu

Follow this and additional works at: <http://repository.cmu.edu/robotics>

 Part of the [Robotics Commons](#)

Published In

Fifth International Conference on 3-D Digital Imaging and Modeling (3DIM 2005), 285- 292.

This Conference Proceeding is brought to you for free and open access by the School of Computer Science at Research Showcase @ CMU. It has been accepted for inclusion in Robotics Institute by an authorized administrator of Research Showcase @ CMU. For more information, please contact research-showcase@andrew.cmu.edu.

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

Jean-François Lalonde, Ranjith Unnikrishnan, Nicolas Vandapel and Martial Hebert
Carnegie Mellon University
{jlalonde,ranjith,vandapel,hebert}@ri.cmu.edu

Abstract

Three-dimensional lidar data are commonly used to perform scene understanding for outdoor mobile robots, specifically in natural terrain. One effective method is to classify points using features based on local point cloud distribution into surfaces, linear structures or clutter volumes. But the local features are computed using 3-D points within a support-volume. Local and global point density variations and the presence of multiple manifolds make the problem of selecting the size of this support volume, or scale, challenging. In this paper we adopt an approach inspired by recent developments in computational geometry [5] and investigate the problem of automatic data-driven scale selection to improve point cloud classification. The approach is validated with results using data from different sensors in various environments classified into different terrain types (vegetation, solid surface and linear structure)¹.

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

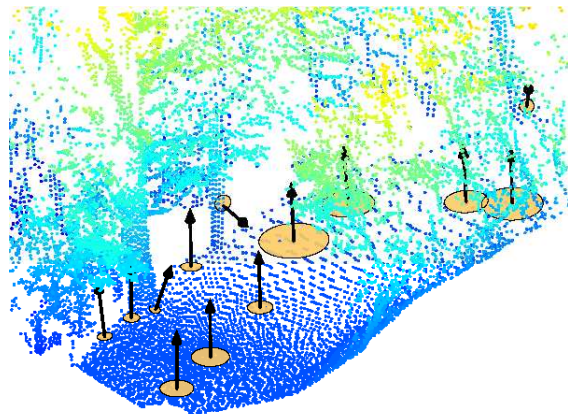


Figure 1. Support region sizes for selected points in outdoor vegetated terrain. Points are color-coded by height.

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. Sensor noise also confounds the feature computation process as a larger support region size may be needed to compensate for noise.

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 as illustrated in Figure 1. 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 exten-

¹Prepared 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.

sive experiments and detail our approach and its limitations in the sections that follow.

The next section presents various approaches followed previously to try to address the scale selection issue. The third section details the approach proposed for optimal scale selection to perform terrain classification. Results from various terrains sensed by three different lidar are presented in the next to last section. It is followed by concluding remarks and a discussion.

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.

Pioneering work by Lindeberg [3] 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 [2]. 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. However this body of work has focused solely on functions defined on a regular lattice and its applicability to unorganized point samples is unclear.

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 [7] 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 [1] uses k -neighborhoods to compensate for differences in sampling rate before computing eigenvalue-based features for detecting surfaces, creases and borders. However, no guarantee was given that a certain fixed choice of k would be representative of the underlying surface at all points.

Tang et al. [8] 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 piecewise lines and not applicable to classification. Work in [6] classifies points based on eigenvalues of the local covariance matrix in its n -neighborhood.

It defines a measure of deviation from planarity at a point that is a function of the eigen-values. The value of the scale (n) that maximizes the measure for 1-D sinusoidal signals is observed to be 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 were made regarding suitability of the proposed measure for 3-D surfaces or its optimality for classification.

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 [5][4] and compute a neighborhood size that minimizes an upper bound on expected angular error between the normal estimated at a point through 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.

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.

3. Approach

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 [5][4]. We start with a set of N_p points, $\mathbf{p}_i = [x_i \ y_i \ z_i]^T$, 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 [5], 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

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

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.

The remainder of the discussion makes the following assumptions:

(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.

(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 $[-n, n]$.

(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, we show a case where this assumption is violated and we propose a modification to address 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 .

Bounding entries of M 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 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} .

Also by definition:

$$|m_{12}| = \left| \frac{1}{k} \sum_{i=1}^k x_i y_i - \frac{1}{k^2} \sum_{i=1}^k x_i \sum_{i=1}^k y_i \right|$$

We assume that x_i and y_i are independently drawn instances of random variables X and Y respectively. So $E[XY] = E[X].E[Y]$, and hence $E[m_{12}] = 0$. Since x_i and y_i are bounded in R , we have a trivial upper bound on variance of m_{12} as $V[m_{12}] \leq \frac{1}{k} \theta_2 r^4$. It follows from Chebyshev's inequality that with probability $1 - \epsilon$:

$$|m_{12}| \leq \frac{\sqrt{\theta_2 r^4}}{\sqrt{\epsilon k}} = \frac{\sqrt{\theta_2 r^4}}{\sqrt{\epsilon p r^2}} = \sqrt{\theta_2} \frac{r}{\sqrt{\epsilon p}} \quad (3)$$

From the Taylor expansion of the surface $z = g(x, y)$ about $(0, 0)$, and the assumption of bounded curvature (A4), we have $\forall x_i, y_i \in \mathbb{R}$:

$$|z_i| = |g(x_i, y_i)| \leq \kappa \left(\frac{x_i^2}{2} + \frac{y_i^2}{2} \right) + n_i \quad (4)$$

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. (4), and since $|x_i| \leq r$ and $|y_i| \leq r$:

$$|m_{13}| \leq 2\kappa r^3 + \left| \frac{1}{k} \sum_{i=1}^k x_i n_i \right| + r \left| \frac{1}{k} \sum_{i=1}^k n_i \right|$$

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

$$|m_{13}| \leq 2\kappa r^3 + \theta_3 \frac{\sigma_n}{\sqrt{\epsilon p}} \quad (5)$$

$$\text{where } \theta_3 = \frac{\sqrt{V(x_i n_i)}}{r \sigma_n} + 1 \quad (6)$$

Symmetrically, the same procedure applies to m_{23} , by replacing x_i by y_i .

Finally for m_{33} ,

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

where θ_4 relates the noise variance σ_n^2 to the width, $2n$, of the noise range.

Eigen analysis We may write the covariance matrix M as

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{12} & m_{22} & m_{23} \\ m_{13} & m_{23} & m_{33} \end{bmatrix} \triangleq \begin{bmatrix} M_{11} & M_{13} \\ M_{13}^T & m_{33} \end{bmatrix} \quad (8)$$

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:

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

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 for \vec{v} and squaring gives the upper bound:

$$\begin{aligned} \|\vec{v}\|_2 &\leq \left(\|(M_{11} - \lambda I)^{-2}\|_2 \right. \\ &\quad \left. \|(\mathbb{I} + (M_{11} - \lambda I)^{-2} M_{13} M_{13}^T)^{-1}\|_2 \right) \\ &\quad \times \left(\|(M_{11} - \lambda I)\|_2 \|M_{13}\|_2 \right. \\ &\quad \left. + \|M_{13}\|_2 |(m_{33} - \lambda)| \right) \end{aligned} \quad (10)$$

It can be shown [5] that

$$\|\vec{v}\|_2 \leq \frac{\alpha(1 + \alpha)}{1 - 2\alpha} \frac{\lambda_2}{\lambda_1} \approx \left(\frac{\lambda_2}{\lambda_1} \right) \alpha \quad (11)$$

for small α .

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

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

Error bound for the estimated normals From Eqns. (2),(3),(5) and (7), we can replace each m_{ij} term in Eqn. (9) by its appropriate bound value to give:

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

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

$$\frac{\lambda_1}{\lambda_2} \alpha \leq \frac{m_{22} (1 + \beta)}{m_{11} (1 - \beta)} \alpha \leq K \alpha \quad (13)$$

we have from the previous lower bound

$$\frac{\lambda_1}{\lambda_2} \alpha \leq K \left(\frac{\theta_4}{\theta_1} \frac{\sigma_n^2}{r^2} + \frac{4}{\theta_1} \kappa r + \frac{2\theta_3}{\theta_1} \frac{\sigma_n}{r^2 \sqrt{\epsilon \rho}} \right) \quad (14)$$

Differentiating Eqn.(14) w.r.t r gives the required result:

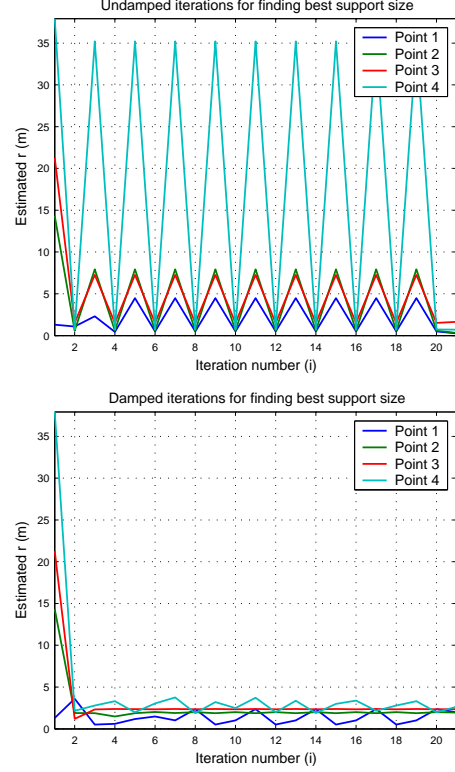


Figure 2. Plot of estimated support region size (r) at each iteration showing improvement with damped updates.

$$r = \left(\frac{1}{\kappa} \left(\theta_3 \frac{\sigma_n}{\sqrt{\epsilon \rho}} + \frac{\theta_4}{2} \sigma_n^2 \right) \right)^{\frac{1}{3}} \quad (15)$$

where the constants $d_1 = \theta_3$ and $d_2 = \theta_4/2$, as given in [5][4] 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. (5) and (7), θ_3 is related to $V(x_i n_i)$ while θ_4 is related to σ_n .

Estimating the optimal support region size The optimal r is estimated using an iterative procedure based on the suggestions in [5]. 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.(15). $k^{(i+1)}$ 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

[5][4] had poor convergence properties when assumption (A4) is broken. Figure 2 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 5. We modify the algorithm to perform damped updates to k using a learning rate γ with $\gamma \in [0, 1]$ as in:

$$k^{(i+1)} = \gamma k_{\text{computed}}^{(i+1)} + (1 - \gamma)k^{(i)} \quad (16)$$

This modification ensures controlled updates in each iteration 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 5 and in the region where the tree trunks meet the ground in Figure 1.

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 [9] 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.

Saliency features Our choice of features is inspired by the tensor voting framework in [7]. 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. 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 eigenvalues 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:

$$\begin{bmatrix} \text{point-ness} \\ \text{surface-ness} \\ \text{curve-ness} \end{bmatrix} \triangleq \begin{bmatrix} \lambda_1 \\ \lambda_2 - \lambda_1 \\ \lambda_3 - \lambda_2 \end{bmatrix} \quad (17)$$

Bayesian classification Using the features of Eqn. (17) 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.(17) belonging to class C_i is given by:

$$P(f(x)|C_i) = \sum_{j=1}^{n_i} \left(\frac{w_{(i,j)}}{(2\pi)^{3/2} |\Sigma_{(i,j)}|^{1/2}} \times e^{-\frac{1}{2}(f(x) - \mu_{(i,j)})^T \Sigma_{(i,j)}^{-1} (f(x) - \mu_{(i,j)})} \right) \quad (18)$$

The estimated class is the maximizer of the class posterior:

$$\begin{aligned} C_{\text{est}} &= \arg \max_i (P(C_i|f(x))) \\ &= \arg \max_i (P(f(x)|C_i)P(C_i)) \end{aligned} \quad (19)$$

where $P(C_i)$ represents the corresponding class prior.

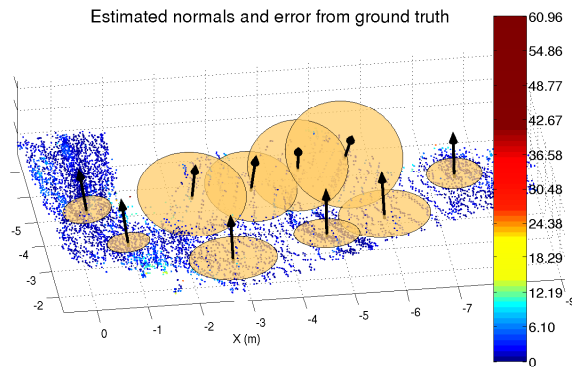
4. Experiments

4.1. Sensors and terrains

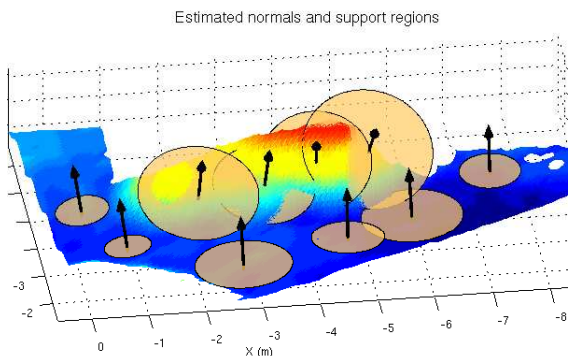
To validate the approach presented we used data collected with a Minolta scanner, an actuated SICK laser, a Zoller-Fröhlich high resolution scanner and the CMU autonomous helicopter. The Minolta Vivid 700 is a laser line striper that produces a 200×200 -pixel range image with 8 bits resolution. A SICK 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 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 [5] by testing it on geometric models for which the ground truth normals are known for each point.



(a)



(b)

Figure 3. 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.

Aerial lidar scan 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 long 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 use the Z+F data to produce the ground truth used to estimate the normal reconstruction error in the aerial data.

Figure 3-(a) shows the computed normals and the support regions for selected points in the aerial data. Figure 3-(b) shows the normal and support regions for the same points but overlaid on top of the high-resolution ground data. Points in Figure 3-(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.

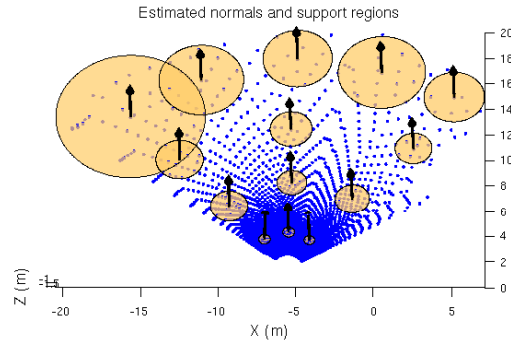


Figure 4. Plot of ground points with estimated support region size (r). Note the significant decrease in spatial density and corresponding increase in r with distance from the laser position

4.3. Validation of support regions

Outdoor ground scan In some situations, the density varies with the distance to the sensor and may become very low. This is illustrated in Figure 4, 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. The noise standard deviation σ_n is computed using calibration data, and depends on the distance to the sensor. 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 with the distance to the sensor. Moreover, the discontinuities located at the boundaries of the laser FOV represent a second important difference and break the assumptions 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.

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 that region. 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 5 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 significant error in normal estimation, especially around the

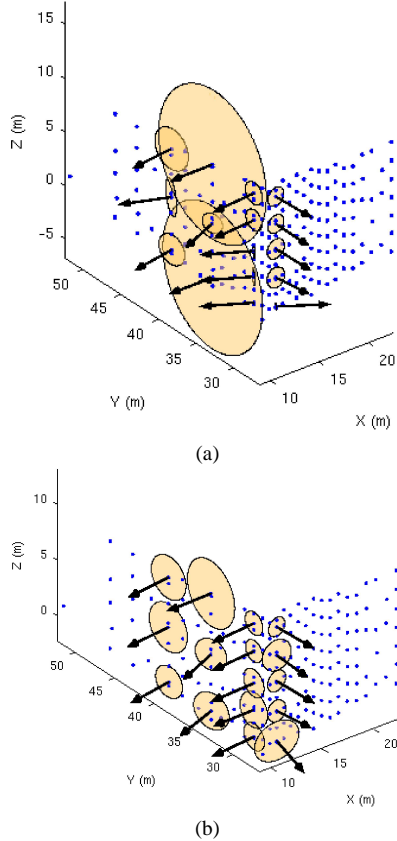


Figure 5. Estimate of support region size for wall corner (a) without and (b) with damped updates to k in each iterations.

discontinuity region. The results obtained with the learning rate $\gamma = 0.3$ (Figure 5) introduced in Section 3.1 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.

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, therefore the results are evaluated visually. Figure 1 shows the support region determined by our algorithm for different points 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.

4.4. Ground-based lidar classification of natural terrain

In this section, we use the algorithm with the classifier described in Section 3.2. The dataset is divided into cubic voxels with 10 cm edges. The classifier is then trained at scales ranging from 0.1 m to 2 m using manually labeled data. The best scale is chosen by applying our method and rounding the resulting support region radius to the nearest subdivision. Figure 6-(a) shows the classification results using the fixed support region size (radius of 40 cm, determined to be the best fixed scale experimentally) currently in use in our system. Obvious misclassification errors are made near the junction of the leftmost tree and the ground, and on the ground at a distance. Figure 6-(b) shows the improvement over the old strategy. We also present classification results on data collected with the Riegl scanner, as shown in Figure 6-(c/d). The data was manually labeled to produce ground truth classification. In performance evaluation, we only consider points that were misclassified using the fixed-scale strategy. The new strategy is found to give a 30% improvement in classification rate for the previously misclassified points.

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 6, 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 “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 6-(b/d) in the boundary regions of the dataset show that

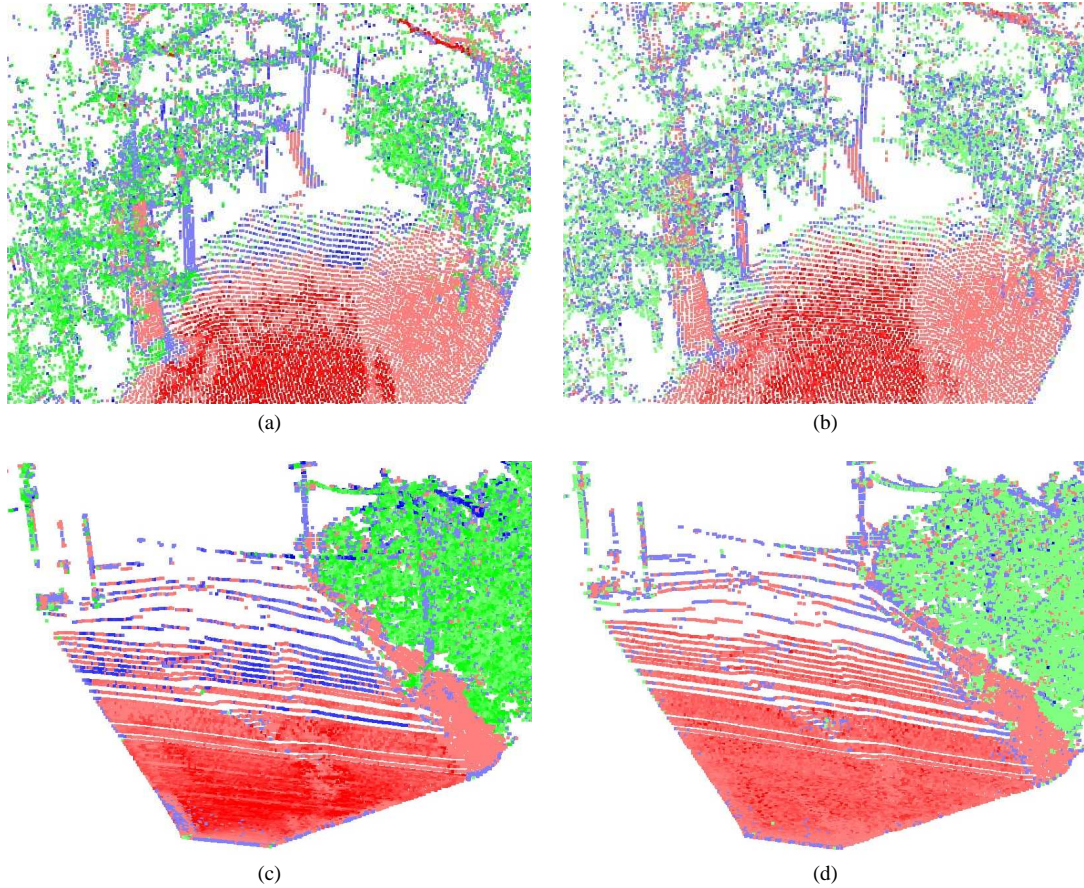


Figure 6. Outdoor terrain classification: (a-b) from the data set using in Figure 1 and (c-d) from data collected with the Riegl laser scanner. Points are colored green (vegetation), red (surface) or blue (linear structure). Darker shades indicate higher confidence in the estimated label. (a/c) Former strategy. (b/d) New strategy.

this hypothesis is violated. We attribute this to (1) the introduction of edge-effects in the chosen features (Eqn. (17)) 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. Gumhold, X. Wang, and R. Macleod. Feature extraction from point clouds. In *Intl. Meshing Roundtable*, 2001.
- [2] T. Lindeberg. Scale-space for discrete signals. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(3), 1990.
- [3] T. Lindeberg. Scale-space theory: A basic tool for analysing structures at different scales. *Journal of Applied Statistics*, 21(2), 1994.
- [4] N. Mitra and A. Nguyen. Estimating surface normals in noisy point cloud data. In *Symp. on Computational Geometry*, 2003.
- [5] N. Mitra, A. Nguyen, and L. Guibas. Estimating surface normals in noisy point cloud data. *Intl. Journal of Computational Geometry and Applications (to appear)*, 2005.
- [6] M. Pauly, R. Keiser, and M. Gross. Multi-scale feature extraction on point-sampled surfaces. In *Eurographics*, 2003.
- [7] 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 Trans. on Pattern Analysis and Machine Intelligence*, 26(5), 2004.
- [8] 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 Intl. Conf. on Robotics and Automation*, 2004.
- [9] N. Vandapel, D. Huber, A. Kapuria, and M. Hebert. Natural terrain classification using 3-d ladar data. In *IEEE Intl. Conf. on Robotics and Automation*, 2004.