International Journal of Computational Geometry & Applications

Estimating Surface Normals in Noisy Point Cloud Data*

NILOY J. MITRA
Stanford Graphics Laboratory, James H. Clark Center, Room S297,
318 Campus Drive, Stanford, CA 94305 USA
niloy@stanford.edu

AN NGUYEN
Stanford Graphics Laboratory, James H. Clark Center, Room S297,
318 Campus Drive, Stanford, CA 94305 USA
anguyen@cs.stanford.edu

LEONIDAS GUIBAS
Stanford Graphics Laboratory, James H. Clark Center, Room S293,
318 Campus Drive, Stanford, CA 94305 USA
guibas@cs.stanford.edu

Received 7 July 2003
Revised 20 October 2003
Communicated by ????

In this paper we describe and analyze a method based on local least square fitting for estimating the normals at all sample points of a point cloud data (PCD) set, in the presence of noise. We study the effects of neighborhood size, curvature, sampling density, and noise on the normal estimation when the PCD is sampled from a smooth curve in $\mathbb{R}^2$ or a smooth surface in $\mathbb{R}^3$ and noise is added. The analysis allows us to find the optimal neighborhood size using other local information from the PCD. Experimental results are also provided.

Keywords: normal estimation; noisy point cloud data; eigen analysis; neighborhood size estimation.

1. Introduction

Modern range sensing technology enables us to make detailed scans of complex objects generating point cloud data (PCD) consisting of millions of points. The data acquired is usually distorted by noise arising out of various physical measurement processes and limitations of the corresponding acquisition technologies.

*A preliminary version of this paper appeared in the Proc. of the 19^{th} ACM Symp. on Computational Geometry, 2003. The work was supported by NSF CARGO grant 0138456 and a Stanford Graduate Fellowship.

The traditional way to use PCD is to reconstruct the underlying surface model represented by the PCD, for example as a triangle mesh, and then apply well known methods on that underlying manifold model. However, when the size of the PCD is large, such methods may be expensive. To do surface reconstruction on a PCD, one would first need to filter out the noise from the PCD, usually by some smoothing filter [15]. Such a process may remove sharp features, however, which may be undesirable. A reconstruction algorithm such as those proposed by Amenta *et al.* [2,3] then computes a mesh that approximates the noise free PCD. Both the smoothing and the surface reconstruction processes may be computationally expensive. For certain applications like rendering or visualization, such a computation is often unnecessary and direct rendering of PCD has been investigated by the graphics community [17,19].

Alexa *et al.* [1] and Pauly *et al.* [19] have proposed to use PCD as a new modeling primitive. Algorithms for such a paradigm often require information about the normal at each of the points. For example, normals are used in rendering PCD, making visibility computation, answering inside–outside queries, etc. Also some curve (or surface) reconstruction algorithms [6,7] need to have the normal estimates as a part of the input data.

The normal estimation problem has been studied by various communities such as computer graphics, image processing, and mathematics, but mostly in the case of manifold representations of the surface. We would like to estimate the normal at each point in a PCD, given to us only as an unstructured set of points sampled from a smooth curve in $\mathbb{R}^2$ or a smooth surface in $\mathbb{R}^3$ and without any additional manifold structure.

Hoppe *et al.* [13] proposed an algorithm where the normal at each point is estimated as the normal to the fitting plane obtained by applying the total least square method to the k *nearest neighbors* of the point. This method is robust in the presence of noise due to the inherent low pass filtering. In this algorithm, the value of k is a parameter and is chosen manually based on visual inspection of the computed estimates of the normals, and different trial values of k may be needed before a good selection of k is found. Furthermore, the same value of k is used for normal estimation at all points in the PCD.

We note that the accuracy of the normal estimation using a total least square method depends on (1) the noise in the PCD, (2) the curvature of the underlying manifold, (3) the density and the distribution of the samples, and (4) the neighborhood size used in the estimation process. In this paper, we make precise such dependencies and study the contribution of each of these factors on the normal estimation process. This analysis allows us to find the optimal neighborhood size to be used in the method. The neighborhood size can be computed adaptively at each point based on its local information, given some estimates about the noise, the local sampling density, and bounds on the local curvature. The computational complexity of estimating all normals of a PCD with m points is only $O(m \log m)$.

1.1. *Related Work*

In this section, we summarize some of the previous works that are related to the computation of the normal vectors of a PCD. Many current surface reconstruction algorithms [2,3,10] can either compute the normal as part of the reconstruction, or the normal can be trivially approximated once the surface has been reconstructed. As the algorithms require that the input is noise free, a raw PCD with noise needs to go through a smoothing process before these algorithms can be applied.

The work of Hoppe *et al.* [13] for surface reconstruction suggests a method to compute the normals for the PCD. The normal estimate at each point is done by fitting a least square plane to its k nearest neighbors. The value of k is selected experimentally. The same approach has also been adopted by Pauly et al. [19] for local surface estimation. Higher order surfaces have been used by Welch *et al.* [18] for local parameterization. However, as pointed out by Amenta et al. [4] such algorithms can fail even in cases with arbitrarily dense set of samples. This problem can be resolved by assuming uniformly distributed samples which prevents errors resulting from biased fits. As noted before, all these algorithms work well even in presence of noise because of the inherent filtering effect. The success of these algorithms depends largely on selecting a suitable value for k, but usually little guidance is given on the selection of this crucial parameter.

Recently, Fleishman *et al.* [9] and Jones *et al.* [14] have independently proposed similar schemes that use bilateral filtering for fast feature-preserving mesh denoising. These algorithms assume that the connectivity information of the underlying manifold is provided in the form of the input mesh. Further, the techniques do not provide any guarantee on the quality of the smoothed mesh.

1.2. *Paper Overview*

In section 2, we study the normal estimation for PCD which are samplings of curves in $\mathbb{R}^2$, and the effects of different parameters on the error of that estimation process. In section 3, we derive similar results for PCD which come from surfaces in $\mathbb{R}^3$. In section 4, we provide some simulations to illustrate the results obtained in sections 2 and 3. We also provide an algorithm for using our theoretical results on practical data. We conclude in section 5.

2. Normal Estimation in $\mathbb{R}^2$

In this section, we consider the problem of approximating the normals to a point cloud in $\mathbb{R}^2$. Given a set of points, which are noisy samples of a smooth curve in $\mathbb{R}^2$, we can use the following method to estimate the normal to the curve at each of the sample points. For each point O, we find all the points of the PCD inside a circle of radius r centered at O, then compute the total least square line fitting those points. The normal to the fitting line gives us an approximation to the undirected normal of the curve at O. Note that the orientation of the normals is a global property of the PCD and thus cannot be computed locally. Once all the undirected normals are computed, a standard breadth first search algorithm [13] can be applied to obtain all the normal directions in a consistent way. For the rest of this paper, we only consider the computation of the undirected normals.

We analyze the error of the approximation when the noise is small and the sampling density is high enough around O. Under these assumptions, which we will make precise later, the computed normal approximates well the true normal. We observe that if r is large, the neighborhood of the point cannot be well approximated by a line in the presence of curvature in the data and we may incur large error. On the other hand, if r is small, the noise in the data can result in significant estimation error. We aim for the optimal r that strikes a balance between these opposing sources of error.

2.1. *Modeling*

Without loss of generality, we consider O the origin, and the y-axis to be along the normal to the curve at O. We assume that the points of the PCD in a disk of radius r around O come from a segment of the curve (a 1-D topological disk) of bounded curvature. Under this assumption, the segment of the curve near O is locally a graph of a single valued smooth function $y = g(x)$ defined over some interval $[-r, r]$ which we denote by R. Further, there exists a constant $\kappa > 0$ such that $|g''(x)| < \kappa, \; \forall x \in R$. For our purpose, it is enough to assume g to be C^2 continuous in R.

Let $\mathbf{p_i} = (x_i, y_i)$ for $1 \le i \le k$ be the points of the PCD that lie inside a circle of radius r centered at O. We assume the following probabilistic model for the points $\mathbf{p_i}$. Assume that x_i's are instances of a random variable X taking values in R, and $y_i = g(x_i) + n_i$, where the noise terms n_i are independent instances of a random variable N. X and N are assumed to be independent. We assume that the noise N has zero mean and standard deviation σ_n, and takes values in $[-n, n]$.

Using Taylor series, there are numbers $\psi_i, \; 1 \le i \le k$ such that $g(x_i) = g''(\psi_i)x_i^2/2$ with $|\psi_i| \le |x_i| \le r$. Let $\gamma_i = g''(\psi_i)$, then the bounded curvature assumption implies that $|\gamma_i| \le \kappa$.

Note that if κr is large, even when there is no noise in the PCD, the normal to the best fit line may not be a good approximation to the tangent as shown in Figure 1. Similarly, if σ_n/r is large and the noise is biased, this normal may not be a good approximation even if the original curve is a straight line, see Figure 2. In order to keep the normal approximation error low we assume *a priori* that κr and σ_n/r are sufficiently small.

We assume that the data is *evenly distributed*; there is a radius $r_0 > 0$ (possibly dependent on O) so that any neighborhood of size r_0 in R contains at least 2 points of the x_i's but no more than some small constant number of them. We observe that the number of points k inside any disk of radius r is bounded from above by $\Theta(1)r\rho$, and also is bounded from below by another $\Theta(1)r\rho$, where ρ is the sampling density

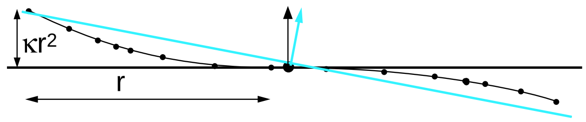


Fig. 1. Curvature causes error in the estimated normal

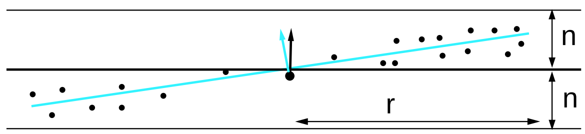


Fig. 2. Noise causes error in the estimated normal

of the point cloud. Here we use $\Theta(1)$ to denote some small positive constant, and for notational simplicity, different appearances of $\Theta(1)$ may denote different constants. We note that distributions satisfying the (ϵ, δ) sampling condition proposed by Dey *et al.* [8] are evenly distributed.

Under the above assumptions, we would like to bound the normal estimation error and study the effects of different parameters. The analysis involves probabilistic arguments to account for the random nature of the noise.

2.2. *Total Least Square Line*

This section describes the well-known total least square method in brief.

Given a set of points $\mathbf{p_i}$, $1 \le i \le k$, we would like to find the line $\mathbf{a}^T\mathbf{x} = c$, with $\mathbf{a}^T\mathbf{a} = 1$ such that the sum of square distances from the points $\mathbf{p_i}$'s to the line is minimized. Let $f(\mathbf{a}, c, \lambda) = \frac{1}{2k}\sum_{i=1}^{k}(\mathbf{a}^T\mathbf{p}_i - c)^2 + \lambda(1 - \mathbf{a}^T\mathbf{a})$, where λ is a Lagrangian multiplier. We would like to find $\mathbf{a}$, c and λ that minimize $f(\mathbf{a}, c, \lambda)$.

To solve this minimization problem, we solve the following equations, $\partial f(\mathbf{a}, c, \lambda)/\partial \mathbf{a} = 0$, $\partial f(\mathbf{a}, c, \lambda)/c = 0$, and $\partial f(\mathbf{a}, c, \lambda)/\partial \lambda = 0$. It follows that $\left(\frac{1}{k}\sum_{i=1}^{k}(\mathbf{p_i} - \bar{\mathbf{p}})(\mathbf{p_i} - \bar{\mathbf{p}})^T\right)\mathbf{a} = \lambda\mathbf{a}$, $c = \bar{\mathbf{p}}^T\mathbf{a}$, and $\mathbf{a}^T\mathbf{a} = 1$ where $\bar{\mathbf{p}} = \frac{1}{k}\sum_{i=1}^{k}\mathbf{p_i}$. These constraints put together results in $f(\mathbf{a}, c) = \frac{\lambda}{2}$. Thus λ is an eigenvalue of the covariance matrix $M = \frac{1}{k}\sum_{i=1}^{k}(\mathbf{p_i} - \bar{\mathbf{p}})(\mathbf{p_i} - \bar{\mathbf{p}})^T$ with $\mathbf{a}$ as the corresponding eigenvector of M. It is clear that to minimize $f(\mathbf{a}, c)$, λ has to be the minimum eigenvalue of M. The corresponding eigenvector $\mathbf{a}$ is the normal to the total least square line and is our normal estimate.

Note that this approach can be generalized to higher dimensional space. The normal to the total least square fitting plane (or hyperplane) of a set of k points $\mathbf{p_i}, 1 \le i \le k$ in $\mathbb{R}^d$ for $d \ge 2$ can be obtained by computing the eigenvector corresponding to the smallest eigenvalue of $M = \frac{1}{k}\sum_{i=1}^{k}(\mathbf{p_i} - \bar{\mathbf{p}})(\mathbf{p_i} - \bar{\mathbf{p}})^T$. We observe that the covariance matrix M is always symmetric positive semi-definite, and has non-negative eigenvalues and the diagonal is non-negative.

2.3. *Eigen-analysis of M*

We can write the 2×2 symmetric matrix M, as defined in the previous section, as $\begin{bmatrix} m_{11} & m_{12} \\ m_{12} & m_{22} \end{bmatrix}$. Note that in absence of noise and curvature, $m_{12} = m_{22} = 0$ which means 0 is the smallest eigenvalue of M with $[0\ 1]^T$ as the corresponding eigenvector. Under our assumption that the noise and the curvature are small, y_i's are small, and thus m_{12} and m_{22} are small. Let $\alpha = (|m_{12}| + m_{22})/m_{11}$. We would like to estimate the smallest eigenvalue of M and its corresponding eigenvector when α is small.

Using the *Gershgorin Circle Theorem* [11], there is an eigenvalue λ_1 such that $|m_{11} - \lambda_1| \le |m_{12}|$, and an eigenvalue λ_2 such that $|m_{22} - \lambda_2| \le |m_{12}|$. When $\alpha < 1/2$, we have that $\lambda_1 \ge m_{11} - |m_{12}| \ge m_{22} + |m_{12}| \ge \lambda_2$. It follows that the two eigenvalues are distinct, and λ_2 is the smallest eigenvalue of M. Let $[v\ 1]^T$ be the eigenvector corresponding to λ_2, then

$$\begin{bmatrix} m_{11} & m_{12} \\ m_{12} & m_{22} \end{bmatrix} \begin{bmatrix} v \\ 1 \end{bmatrix} = \lambda_2 \begin{bmatrix} v \\ 1 \end{bmatrix},$$
$$\begin{bmatrix} m_{11} - \lambda_2 \\ m_{12} \end{bmatrix} v = -\begin{bmatrix} m_{12} \\ m_{22} - \lambda_2 \end{bmatrix}.$$

Thus

$$\begin{aligned} v &= -\frac{(m_{11} - \lambda_2)m_{12} + m_{12}(m_{22} - \lambda_2)}{(m_{11} - \lambda_2)^2 + m_{12}^2}, \qquad (1) \\ |v| &\le \frac{|m_{12}|(m_{11} - \lambda_2 + |m_{12}|)}{(m_{11} - \lambda_2)^2}, \\ &\le \frac{\alpha(1+\alpha)}{(1-\alpha)^2}. \end{aligned}$$

Thus, the estimation error, which is the angle between the estimated normal and the true normal (which is $[0\ 1]^T$ in this case), is less than $\tan^{-1}(\alpha(1+\alpha)/(1-\alpha)^2) \approx \alpha$, for small α. Note that we could write the error explicitly in closed form, then bound it. Our approach is more complicated, though as we will show later, it can be extended to obtain the error bound for the 3D case. To bound the estimation error, we need to estimate α.

2.4. *Estimating Entries of M*

The assumption that the sample points are evenly distributed in the interval $[-r, r]$ implies that, given any number x in that interval, the number of points $\mathbf{p_i}$'s satisfying $|x_i - x| \ge r/4$ is at least $\Theta(1)k$. It follows easily that $m_{11} = \frac{1}{k}\sum_{i=1}^{k}(x_i - \bar{x})^2 \ge \Theta(1)r^2$. The constant $\Theta(1)$ depends only on the distribution of the random variable X.

For the entries m_{12} and m_{22}, we use $|x_i| \leq r$ and $|y_i| \leq \kappa r^2/2 + n$ to obtain the following trivial bound:

$$
\begin{aligned}
|m_{12}| &= \left| \frac{1}{k} \sum_{i=1}^{k} x_i y_i - \frac{1}{k^2} \sum_{i=1}^{k} x_i \sum_{i=1}^{k} y_i \right| \\
&\leq 2r(\kappa r^2/2 + n)\,, \\
m_{22} &\leq \frac{1}{k} \sum_{i=1}^{k} y_i^2 \\
&\leq 2((\kappa r^2/2)^2 + n^2)\,.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\alpha &\leq \Theta(1)\left(\kappa r + \frac{n}{r} + \kappa^2 r^2 + \frac{n^2}{r^2}\right) \\
&\leq \Theta(1)\left(\kappa r + \frac{n}{r}\right)\,. \qquad (2)
\end{aligned}
$$

This bound illustrates the effects of r, κ and n on the error. For large values of r, the error caused by the curvature κr dominates, while for a small neighborhood the term n/r is dominating. Nevertheless, the expression depends on the absolute bound n of the noise N. This bound n can be unnecessarily large or unbounded for many distribution models of N. We would like to use our assumption on the distribution of the noise N to improve our bound on α further.

Note that

$$
\begin{aligned}
|m_{12}| &= \left| \frac{1}{k} \sum_{i=1}^{k} x_i y_i - \frac{1}{k^2} \sum_{i=1}^{k} x_i \sum_{i=1}^{k} y_i \right| \\
&\leq \left| \frac{1}{k} \sum_{i=1}^{k} (\gamma_i x_i^3/2 + x_i n_i) \right| \\
&\quad + \left| \frac{1}{k^2} \sum_{i=1}^{k} x_i \sum_{i=1}^{k} (\gamma_i x_i^2/2 + n_i) \right| \\
&\leq \Theta(1)\kappa r^3 + \left| \frac{1}{k} \sum_{i=1}^{k} x_i n_i \right| \\
&\quad + \Theta(1) r \left(\kappa r^2 + \left| \frac{1}{k} \sum_{i=1}^{k} n_i \right| \right)\,.
\end{aligned}
$$

Furthermore, under the assumption that X and N are independent, we have $E[x_i n_i] = E[x_i]E[n_i] = 0$ since $E[n_i] = 0$ and $\mathrm{Var}(x_i n_i) = \Theta(1) r^2 \sigma_n^2$ since $\mathrm{Var}(n_i) = \sigma_n^2$. Let $\epsilon > 0$ be some small constant. Using the *Chebyshev Inequality* [16], we can show that the following bound on $|m_{12}|$ holds with probability at least $1 - \epsilon$:

$$
\begin{aligned}
|m_{12}| &\le \Theta(1)\kappa r^3 + \Theta(1)\sqrt{\frac{r^2\sigma_n^2}{\epsilon k}} + \Theta(1) r \sqrt{\frac{\sigma_n^2}{\epsilon k}} \\
&= \Theta(1)\kappa r^3 + \Theta(1)\sqrt{\frac{r^2\sigma_n^2}{\epsilon r \rho}} + \Theta(1) r \sqrt{\frac{\sigma_n^2}{\epsilon r \rho}} \\
&\le \Theta(1)\kappa r^3 + \Theta(1)\sigma_n \sqrt{\frac{r}{\epsilon \rho}}\,. \qquad (3)
\end{aligned}
$$

For reasonable noise models, we also have that

$$
\begin{aligned}
m_{22} &\le \frac{1}{k}\sum_{i=1}^{k} 2(\gamma_i^2 x_i^4/4 + n_i^2) \\
&\le \Theta(1)\kappa^2 r^4 + \Theta(1)\sigma_n^2\,.
\end{aligned}
$$

2.5. *Error Bound for the Estimated Normal*

From the estimations of the entries of M, we obtain the following bound on α, with probability at least $1-\epsilon$:

$$
\alpha \le \Theta(1)\kappa r + \Theta(1)\frac{\sigma_n}{\sqrt{\epsilon \rho r^3}} + \Theta(1)\frac{\sigma_n^2}{r^2}\,. \qquad (4)
$$

Note that this bound depends on the standard deviation σ_n of the noise N rather than its magnitude bound n.

For a given set of parameters κ, σ_n, ρ, and ϵ, we can find the optimal r that minimizes the right hand side of inequality 4. As this optimal value of r is not easily expressed in closed form, let us consider a few extreme cases.

- When there is no curvature ($\kappa = 0$) we can make the bound on α arbitrarily small by increasing r. For sufficiently large r, the bound is linear in σ_n and it decreases as $r^{-3/2}$.
- When there is no noise, we can make the error bound small by choosing r as small as possible, say $r = r_0$.
- When both noise and curvature are present, the error bound cannot be arbitrarily reduced. When the density ρ of the PCD is sufficiently high, $\alpha \le \Theta(1)\kappa r + \Theta(1)\sigma_n^2/r^2$. The error bound is minimized when $r = \Theta(1)\sigma_n^{2/3}\kappa^{-1/3}$, in which case $\alpha \le \Theta(1)\kappa^{2/3}\sigma_n^{2/3}$. The sufficiently high density condition on ρ can be shown to be $\rho > \Theta(1)\epsilon^{-1}\sigma_n^{-4/3}\kappa^{-1/3}$.
- When there are both noise and curvature, and the density ρ is sufficiently low, $\alpha \le \Theta(1)\kappa r + \Theta(1)\sigma_n/\sqrt{\epsilon\rho r^3}$. The bound is smallest when $r = \Theta(1)(\sigma_n^2/(\epsilon\rho\kappa^2))^{1/5}$, in which case, $\alpha \le \Theta(1)(\kappa^3\sigma_n^2/(\epsilon\rho))^{1/5}$. The sufficiently low condition on ρ can be expressed more specifically as $\rho < \Theta(1)\epsilon^{-1}\sigma_n^{-4/3}\kappa^{-1/3}$. We would like to point out that the constant hidden

in the $\Theta(1)$ notation in the "sufficiently low" condition is $3/4$ of that in the "sufficiently high" condition.

3. Normal Estimation in $\mathbb{R}^3$

We can extend the results obtained for curves in $\mathbb{R}^2$ to surfaces in $\mathbb{R}^3$. Given a point cloud obtained from a smooth 2-manifold in $\mathbb{R}^3$ and a point O on the surface, we can estimate the normal to the surface at O as follows: find all the points of the PCD inside a sphere of radius r centered at O, then compute the total least square plane fitting those points. The normal vector to the fitting plane is our estimate of the undirected normal at O.

Given a set of k points $\mathbf{p_i}$, $1 \leq i \leq k$, let $M = \frac{1}{k}\sum_{i=1}^{k}(\mathbf{p_i} - \bar{\mathbf{p}})(\mathbf{p_i} - \bar{\mathbf{p}})^T$ where $\bar{\mathbf{p}} = \frac{1}{k}\sum_{i=1}^{k}\mathbf{p_i}$. As pointed out in subsection 2.2, the normal to the total least square plane for this set of k points is the eigenvector corresponding to the minimum eigenvalue of M. We would like to bound the angle between this eigenvector and the true normal to the surface.

3.1. *Modeling*

We model the PCD in a similar fashion as in the $\mathbb{R}^2$ case. We assume that O is the origin, the z-axis is the normal to the surface at O, and that the points of the PCD in the sphere of radius r around O are samples of a topological disk on the underlying surface that has bounded curvature. Under these assumptions, we can locally represent the surface as the graph of a single valued function $z = g(\underline{x})$ where $\underline{x} = [x, y]^T$. Using Taylor Theorem, we can write $g(\underline{x}) = \frac{1}{2}\underline{x}^T H \underline{x}$ where H is the Hessian of g at *some* point ψ such that $|\psi| \leq |\underline{x}|$.

The assumption that the surface has bounded curvature in some neighborhood around O implies that there exists a $\kappa > 0$ such that the Hessian H of g satisfies $||H||_2 \leq \kappa$ in that neighborhood. For our purpose, the function g is C^2 continuous in the local neighborhood of O.

We write the points $\mathbf{p_i}$ as $\mathbf{p_i} = (x_i, y_i, z_i) = (\underline{x}_i, z_i)$. We assume that $z_i = g(\underline{x}_i) + n_i$, where the n_i's are independent instances of some random variable N with zero mean and standard deviation σ_n. We similarly assume that the points $\underline{x}_i$ are *evenly distributed* in the xy-plane on a disk D of radius r centered at O, i.e. there is a radius r_0 such that any disk of size r_0 inside D contains at least 3 points among the x_i's but no more than some small constant number of them. We also assume that the noise and the surface curvature are both small.

3.2. *Eigen-analysis in* $\mathbb{R}^3$

We write the analogous matrix $M = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{12} & m_{22} & m_{23} \\ m_{13} & m_{23} & m_{33} \end{bmatrix} = \begin{bmatrix} M_{11} & M_{13} \\ M_{13}^T & m_{33} \end{bmatrix}$. As pointed out in subsection 2.2, M is symmetric and positive semi-definite. Under the assumptions that the noise and the curvature are small, and that the points $\underline{x}_i$ are evenly distributed, m_{11} and m_{22} are the two dominant entries in M. We assume, without lost of generality, that $m_{11} \leq m_{22}$. Let $\alpha = (|m_{13}| + |m_{23}| + m_{33})/(m_{11} - |m_{12}|)$. As in the $\mathbb{R}^2$ case, we would like to bound the angle between the computed normal and the true normal to the point cloud in term of α.

Denote by $\lambda_1 \leq \lambda_2$ the eigenvalues of the 2×2 symmetric matrix M_{11}. Using again the Gershgorin Circle Theorem, it is easy to see that $m_{11} - |m_{12}| \leq \lambda_1, \lambda_2 \leq m_{22} + |m_{12}|$.

Let λ be the smallest eigenvalue of M. From the Gershgorin Circle Theorem we have $\lambda \leq |m_{13}|+|m_{23}|+m_{33} = \alpha(m_{11} - m_{12}) \leq \alpha\lambda_1$. Let $[\underline{\mathrm{v}}^T\ 1]^T$ be the eigenvector of M corresponding with λ. Then, as with Equation 1, we have that:

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

Note that

$$\begin{aligned}
&||(M_{11} - \lambda I)^{-2} M_{13}M_{13}^T||_2 \\
&\leq ||(M_{11} - \lambda I)^{-2}||_2||M_{13}||_2||M_{13}^T||_2 \\
&\leq (\lambda_1 - \lambda)^{-2}(m_{13}^2 + m_{23}^2) \\
&\leq (1 - \alpha)^{-2}\alpha^2\,.
\end{aligned}$$

Thus

$$\begin{aligned}
&\left|\left|\left(I + (M_{11} - \lambda I)^{-2} M_{13}M_{13}^T\right)^{-1}\right|\right|_2 \\
&\leq \frac{1}{1 - (1-\alpha)^{-2}\alpha^2} \leq \frac{(1-\alpha)^2}{1 - 2\alpha}\,.
\end{aligned}$$

It follows that

$$\begin{aligned}
||\underline{\mathrm{v}}||_2 &\leq \frac{1}{(1-\alpha)^2\lambda_1^2}\frac{(1-\alpha)^2}{1-2\alpha}\,(\lambda_2\alpha\lambda_1 + \alpha\lambda_1\alpha\lambda_1) \\
&\leq \frac{\alpha(1+\alpha)}{1-2\alpha}\,\frac{\lambda_2}{\lambda_1}\,.
\end{aligned}$$

When α is small, the right hand side is approximately $(\lambda_2/\lambda_1)\alpha$, and thus the angle between the computed normal and the true normal, $\tan^{-1}||\underline{\mathrm{v}}||_2$, is approximately bounded by $(\lambda_2/\lambda_1)\alpha \leq ((m_{22} + |m_{12}|)/(m_{11} - |m_{12}|))\alpha$,

3.3. *Estimation of the entries of M*

As in the $\mathbb{R}^2$ case, from the assumption that the samples are evenly distributed, we can show that $\Theta(1)r^2 \leq m_{11}, m_{22} \leq r^2$. We can also show that $m_{33} \leq \Theta(1)\kappa^2 r^4 + \Theta(1)\sigma_n^2$. Let ρ be the sampling density of the PCD at O, then $k = \Theta(1)\rho r^2$. Again, let $\epsilon > 0$ be some small positive number. Using the Chebyshev inequality, we can show that $m_{13}, m_{23} \leq \Theta(1)\kappa r^3 + \Theta(1)\sigma_n r/\sqrt{\epsilon k} \leq \Theta(1)\kappa r^3 + \Theta(1)\sigma_n/\sqrt{\epsilon\rho}$ with probability at least $1 - \epsilon$. For the term m_{12}, we note that $E[x_i y_i] = 0$ and $Var(x_i y_i) = \Theta(1)r^4$, and so, by the Chebyshev inequality, $m_{12} \leq \Theta(1)r/\sqrt{\epsilon\rho}$ with probability at least $1 - \epsilon$.

3.4. *Error Bound for the Estimated Normal*

Let $\beta = m_{12}/m_{11}$. We restrict our analysis to the cases when β is sufficiently less than 1, say $\beta < 1/2$. This restriction simply means that the points $\underline{x}_i$'s are not degenerate, i.e. not all of the points $\underline{x}_i$'s are lying on or near any given line on the xy-plane. With this restriction, it is clear that $(\lambda_2/\lambda_1)\alpha \leq (m_{22}/m_{11})((1+\beta)/(1-\beta))\alpha = \Theta(1)\alpha$.

From the estimations of the entries of M, we obtain the following bound with probability at least $1 - \epsilon$:

$$
\begin{aligned}
\frac{\lambda_2}{\lambda_1}\alpha &\leq \Theta(1)\kappa r + \Theta(1)\frac{\sigma_n}{r^2\sqrt{\epsilon\rho}} \\
&\quad \Theta(1)\kappa^2 r^2 + \Theta(1)\frac{\sigma_n^2}{r^2} \\
&\leq \Theta(1)\kappa r + \Theta(1)\frac{\sigma_n}{r^2\sqrt{\epsilon\rho}} + \Theta(1)\frac{\sigma_n^2}{r^2}
\end{aligned}
$$

This is an approximate bound on the angle between the estimated normal and the true normal. To minimize this error bound, it is clear that we should pick

$$
r = \left(\frac{1}{\kappa}\left(c_1\frac{\sigma_n}{\sqrt{\epsilon\rho}} + c_2\sigma_n^2\right)\right)^{1/3}, \tag{5}
$$

for some constants c_1, c_2. The constants c_1 and c_2 are small and they depend only on the distribution of the PCD.

We notice that from the above result, when there is no noise, we should pick the radius r to be as small as possible, say $r = r_0$. When there is no curvature, the radius r should be as large as possible. When the sampling density is high, the optimal value of r that minimizes the error bound is approximately $r = \Theta(1)(\sigma_n^2/\kappa)^{1/3}$. This result is similar to that for curves in $\mathbb{R}^2$, and it is not at all intuitive.

4. Experiments

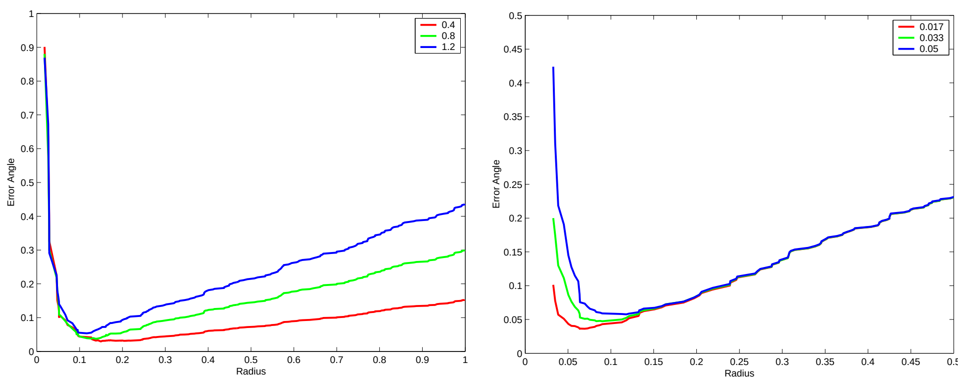
In this section, we discuss some simulations to validate our theoretical results. We then show how to use the results to obtaining a good neighborhood size for the normal computation with the least square method.

4.1. *Validation*

We considered a PCD whose points were noisy samples of the curves $(x, \kappa \ \mathrm{sgn}(x) \ x^2/2)$, for $x \in [-1, 1]$ for different choices of κ. We estimated the normals to the curves at the origin by applying the least square method on their corresponding PCD. As the y-axis is known to be the true normal to the curves, the angles between the computed normals and the y-axis gives the estimation errors.

To obtain the PCD in our experiments, we let the sampling density ρ be 100 points per unit length, and let x be uniformly distributed in the interval $[-1, 1]$. The y-components of the data were polluted with uniformly random noise in the interval $[-n, n]$, for some value n. The standard deviation σ_n of this noise is $n/\sqrt{3}$.

Figure 3(a) shows the error as a function of the neighborhood size r when $n = 0.05$ for 3 different values of κ, $\kappa = 0.4, 0.8$, and 1.2. As predicted by Equation 4 for large value of r, the error increases as r increases. In the experiments, it can be seen that the error is proportional to κr for $r > 0.2$. Note that the PCD we chose generates the worst case behavior of the error.



(a) Error due to curvature dominates for $r > 0.2$. The error–curves behave similarly for different choices of κ.

(b) Error due to noise dominates for $r <$ 0.6. The error–curves behave similarly under different amounts of noise.

Fig. 3. Effects of curvature and noise on estimation error for different choices of r.

Figure 3(b) shows the estimation error as a function of the neighborhood size r for small r when $\kappa = 1.2$ for 3 different values of n, $n = 0.017, 0.033$, and 0.05. We observe that the error tends to decrease as r increases for $r < 0.06$. This is expected as from Equation 4, the bound on the error is a decreasing function of r when r is small. To factor out the random effect of noise, the estimation error has been averaged out over 50 runs for each value of r.

4.2. *Estimating Neighborhood Size for the Normal Computation*

In this part, we used the results obtained in Section 3 to estimate the normals of a PCD. The data points in the PCD were assumed to be noisy samples of a smooth surface in $\mathbb{R}^3$. This is the case, for example, for PCD obtained by range scanners. To obtain the neighborhood size for the normal computation using the least square method, we would like to use Equation 5.

We assumed that the standard deviation σ_n of the noise was given to us as a part of the input. We estimated the other local parameters in Equation 5, then computed r. Note that this value of r minimizes the bound of the normal computation error, and there is no guarantee that this would minimize the error itself. The constants c_1 and c_2 depend on the sampling distribution of the PCD. While we could attempt to compute the exact values of c_1 and c_2, we tried to estimate the values of c_1 and c_2.

Algorithm 1 Estimates good normals for all the points of a PCD

estimate c_1, c_2
choose ϵ
for each point p **do**
  $k \leftarrow k_0$
  $count \leftarrow$ MaxCount
  repeat
    $r_{old} \approx$ distance from p to its k-th nearest neighbor
    $\rho \leftarrow k/\pi r_{old}^2$
    given k, compute κ locally (Gumhold et. al. [12])
    compute r_{new} using Equation 5
    $k \leftarrow \pi \rho r_{new}^2$
    if $k > k_{threshold}$ **then**
      break
    end if
    $count \leftarrow count - 1$
  until $count \neq 0$
  the normal to the least square fit plane to the k neareast neighbors of p gives a good estimate of the normal at p
end for

Given a PCD, we estimated the local sampling density as follows. For a given point p in the PCD, we used the approximate nearest neighbor library ANN [5] to find the distance s from p to its k-th nearest neighbor for some small number k_0. The local sampling density at p was then approximated as $\rho = k/(\pi s^2)$ samples per unit area.

To estimate the local curvature, we used the method proposed by Gumhold *et al.* [12]. Let p_j, $1 \leq j \leq k$ be the k nearest sample points around p, and let μ be the average distance from p to all the points p_j. We computed the best fit least square plane for those k points, and let d be the distance from p to that best fit plane. The local curvature at p can then be estimated as $\kappa = 2d/\mu^2$. This method gives an estimate of the local curvature without any guarantees on the approximation quality.

Once all the parameters were obtained, we computed the neighborhood size r using Equation 5. Note that the estimated value of r could be used to obtain a good value for k, which can to be used to re-estimate the local density and the local curvature. This suggests an iterative scheme in which we repeatedly estimate the local density, the local curvature, and the neighborhood size. In our experiments, we found that only a small number of iterations were enough to obtain good values for all the quantities. Algorithm 1 gives a summary of the iterative scheme. For the following experiments, k_0 was set to 15, and MAXCOUNT was set to 10. The value of ϵ was fixed at 0.1.

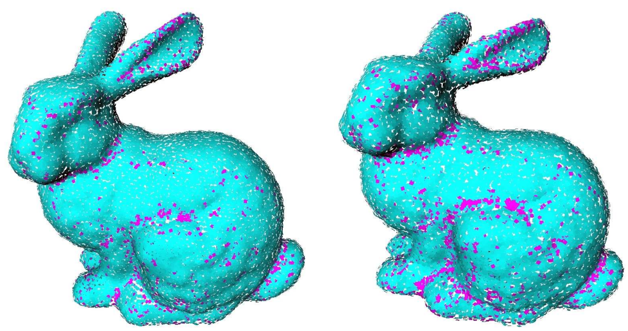
We still have problems with obtaining good estimates for the constants c_1 and c_2. Fortunately, we only have to estimate the constants once for a given PCD, and we can use the same constants for many PCD with a similar point distribution. We used Figure 6(a) for choosing c_1 and c_2. The PCD was created such that exact normals at all points (except those on the edges) were known. Estimation errors could be computed exactly at almost all the points and this information was used to choose values of the constants. We found that $c_1 = 1$, $c_2 = 4$ is a good pair of values and the same was used for all the other data sets.

(a) 1x noise level (b) 2x noise level

Fig. 4. Effects of curvature and noise on the choice of neighborhood size under different amounts of noise in the input data. At a few points, their corresponding neighbors as chosen by Algorithm 1 have been highlighted.

Noisy PCD used in our experiments were obtained by adding noise to the original data. The x, y, and z components of the noise were chosen independently and uniformly random. The magnitude of added noise was measured in a scale where the average spacing between neighboring points (the connectivity information from the mesh representation was used to determine neighbors) of the PCD was taken as one unit.

Figure 4 shows the effects of curvature and noise on the choice of neighborhood size. For a few points, their chosen neighboring points have been highlighted in the figures. Figure 4(a) demonstrates that bigger neighborhoods have been selected in flatter regions compared to the regions with more curvature. Figure 4(b) shows how increased noise leads to the choice of bigger neighborhoods.



(a) 1x noise level (b) 2x noise level

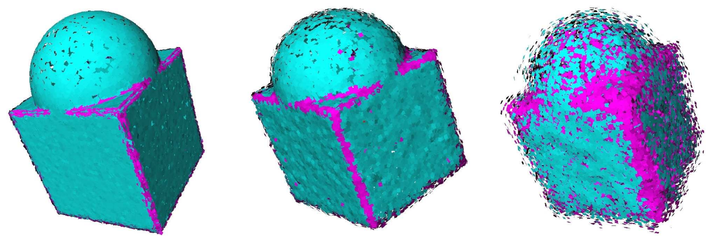
Fig. 5. Normal estimation errors for the bunny PCD with noise added. Points with more than 5° estimation error have been highlighted.

We computed the normals of the noisy PCD, and used the angles between those normals and the normals of the original PCD as estimates of the normal computation errors. The normals computed from the mesh representation of the data sets were used as the original normals. In Figure 5, we highlighted the points where the estimation error were more than 5° under two different amounts of noise.

Figure 6 shows the preformance of the algorithm under different noise conditions. In Figure 6(c), we observe that even in presence of significant noise, the algorithm performs well in flat faces of the object.

5. Conclusions

We have analyzed the method of least square in estimating the normals to a point cloud data derived either from a smooth curve in $\mathbb{R}^2$ or a smooth surface in $\mathbb{R}^3$, with noise added. In both cases, we provided theoretical bound on the maximum angle between the estimated normal and the true normal of the underlying manifold. This theoretical study allowed us to find an optimal neighborhood size to be used in the least square method.



(a) 1x noise (b) 2x noise (c) 4x noise

Fig. 6. Performance of the algorithm under various noise conditions. Points with more than 5° estimation error have been highlighted.

6. Acknowledgments

We are grateful to Tamal K. Dey for useful discussions and also for providing the software for evaluating normals using *cocone*. We also thank Marc Levoy, Ron Fedkiw for helpful discussions.

We also want to thank the numerous referees of the previous versions of this paper for their extremely useful suggestions.

References

1. Marc Alexa, Johannes Behr, Daniel Cohen-Or, Shachar Fleishman, David Levin, and Claudio T. Silva. Point set surfaces. In *Proceedings of the conference on Visualization '01*, pages 21–28. IEEE Computer Society, 2001.
2. N. Amenta, S. Choi, T. K. Dey, and N. Leekha. A simple algorithm for homeomorphic surface reconstruction. In *Proceedings of the sixteenth annual symposium on Computational geometry*, pages 213–222. ACM Press, 2000.
3. Nina Amenta, Marshall Bern, and Manolis Kamvysselis. A new Voronoi-based surface reconstruction algorithm. *Computer Graphics*, 32(Annual Conference Series):415–421, 1998.
4. Nina Amenta and Marshall W. Bern. Surface reconstruction by voronoi filtering. In *Symposium on Computational Geometry*, pages 39–48, 1998.
5. Sunil Arya, David M. Mount, Nathan S. Netanyahu, Ruth Silverman, and Angela Y. Wu. An optimal algorithm for approximate nearest neighbor searching fixed dimensions. *Journal of the ACM*, 45(6):891–923, 1998.
6. Jean-Daniel Boissonnat and Frederic Cazals. Smooth surface reconstruction via natural neighbour interpolation of distance functions. In *Symposium on Computational Geometry*, pages 223–232, 2000.
7. J. C. Carr, R. K. Beatson, J. B. Cherrie, T. J. Mitchell, W. R. Fright, B. C. McCallum, and T. R. Evans. Reconstruction and representation of 3d objects with radial basis functions. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 67–76. ACM Press, 2001.
8. T. K. Dey, J. Giesen, S. Goswami, and W. Zhao. Shape dimension and approximation from samples. In *Proc.* 13_{th} *ACM-SIAM Sympos, Discrete Algorithms*, pages 772–780,

2002.
9. Shachar Fleishman, Iddo Drori, and Daniel Cohen-Or. Bilateral mesh denoising. *ACM Transactions on Graphics (TOG)*, 22(3):950–953, 2003.
10. Stefan Funke and Edgar A. Ramos. Smooth-surface reconstruction in near-linear time. In *Proceedings of the thirteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 781–790. Society for Industrial and Applied Mathematics, 2002.
11. Gene H. Golub and Charles F. Van Loan. *Matrix computations (3rd ed.)*. Johns Hopkins University Press, 1996.
12. S. Gumhold, X. Wang, and R. MacLeod. Feature extraction from point clouds. In 10^{th} *International Meshing Roundtable, Sandia National Laboratories*, pages 293–305, October 2001.
13. Hugues Hoppe, Tony DeRose, Tom Duchamp, John McDonald, and Werner Stuetzle. Surface reconstruction from unorganized points. *Computer Graphics*, 26(2):71–78, 1992.
14. Thouis R. Jones, Frdo Durand, and Mathieu Desbrun. Non-iterative, feature-preserving mesh smoothing. *ACM Transactions on Graphics (TOG)*, 22(3):943–949, 2003.
15. In-Kwon Lee. Curve reconstruction from unorganized points. *Computer Aided Geometric Design*, 17(2):161–177, 2000.
16. A. Leon-Garcia. *Probability and Random Processes for Electrical Engineering*. Addison Wesley, 1994.
17. Szymon Rusinkiewicz and Marc Levoy. Qsplat: a multiresolution point rendering system for large meshes. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 343–352. ACM Press/Addison-Wesley Publishing Co., 2000.
18. William Welch and Andrew Witkin. Free-form shape design using triangulated surfaces. *Computer Graphics*, 28(Annual Conference Series):247–256, 1994.
19. Matthias Zwicker, Mark Pauly, Oliver Knoll, and Markus Gross. Pointshop 3d: an interactive system for point-based surface editing. In *Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 322–329. ACM Press, 2002.