Computers & Graphics 35 (2011) 904–915

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Technical Section

Shape classification and normal estimation for non-uniformly sampled, noisy point data

Cindy Grimm *, William D. Smart

Campus Box 1045, Department of Computer Science and Engineering, Washington University in St. Louis, St. Louis, MO 63130, United States

ARTICLE INFO

Article history:
Received 8 October 2010
Received in revised form
11 March 2011
Accepted 17 March 2011
Available online 13 April 2011

Keywords:
Surface reconstruction
Normal estimation
Shape classification
Unlabeled data
Point data
One-ring

ABSTRACT

We present an algorithm for robustly analyzing point data arising from sampling a 2D surface embedded in 3D, even in the presence of noise and non-uniform sampling. The algorithm outputs, for each data point, a surface normal, a local surface approximation in the form of a one-ring, the local shape (flat, ridge, bowl, saddle, sharp edge, corner, boundary), the feature size, and a confidence value that can be used to determine areas where the sampling is poor or not surface-like.

We show that the normal estimation out-performs traditional fitting approaches, especially when the data points are non-uniformly sampled and in areas of high curvature. We demonstrate surface reconstruction, parameterization, and smoothing using the one-ring neighborhood at each point as an approximation of the full mesh structure.

1. Introduction

We present an algorithm for estimating surface normals and local shape from point data that is sampled from a 2D surface embedded in 3D space. We show that our algorithm is robust in the presence of both noise and non-uniform sampling. For each data point the algorithm produces a surface normal, a local surface approximation in the form of a one-ring, a local estimate of shape (flat, bowl, saddle, ridge, edge, corner, or boundary), a measure of the feature size, and a confidence value. This value reflects both the quality of the local sampling and noise in the function and can be used to detect places where the points do not represent a surface. The one-ring neighborhood can be used for further mesh processing such as reconstruction, smoothing or spectral mesh processing.

Traditional normal estimation approaches usually rely on fitting a local surface approximation to some subset of the k-nearest neighbors [1]. This approach works well most of the time, but it has several limitations (see Fig. 1). First, if the data are non-uniformly distributed, the fitting error becomes biased—the classic example of this is contour data, where the best planar fit can be *perpendicular* to the surface. Second, the local surface approximation may not have enough flexibility to match the surface, which introduces additional error and increases the sensitivity to non-uniformly distributed data. Increasing the flexibility of the representation, unfortunately, can lead to over-fitting. Third, there is no method for distinguishing between noisy data and poor local fit, or determining if the local samples even represent a surface.

Our fundamental idea is to build three *different* representations of the local surface—a surface normal, a good-quality one-ring, and a local shape model—and cross-validate them. If all three representations are mutually consistent with each other *and* the data, then we can be reasonably confident that they are correct. In this we are closer in spirit to approaches that use robust statistics [2,3]. We also gain a lot of information about the samples, namely how much noise is present, how even the sampling is, if there are sharp features or boundaries, the local feature size, and a plausible graph structure. There are several approaches that use one of these pieces of information to *derive* another—for example, normals from the graph structure [4]—but to our knowledge no-one else has used cross-validation to improve normal estimation and surface analysis.

We show that cross-validation produces better-quality normals than standard fitting approaches, particularly in areas with uneven sampling and high curvature. Not only are the averages better, but the variance is narrower, meaning we are less likely to return a "wrong" result. This is particularly true in saddle and ridge areas. Our approach also behaves well near boundaries and sharp features, and can explicitly identify them.

We make use of three observations. The first is that, locally, smooth surfaces are either flat (zero curvature), ridges (zero curvature in one direction), bowls (positive curvature), or saddles

* Corresponding author.
E-mail addresses: cmg@wustl.edu (C. Grimm), wds@wustl.edu (W.D. Smart).

0097-8493/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cag.2011.03.036

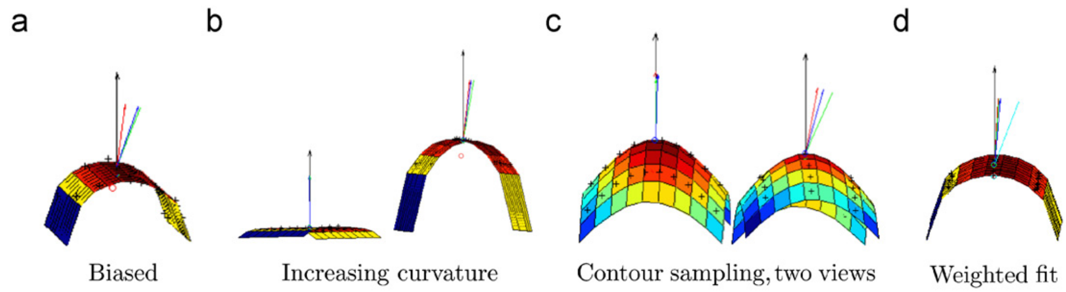

Fig. 1. Black is true normal, red is plane fitting, green is quadratic, and blue cubic: (a) more samples on one side than the other pull the normal estimation in that direction; (b) increasing the curvature (same samples) results in error in the normal; (c) two views, contour sampling—more samples on the right contour pull the normal in that direction; and (d) weighted fit (cyan arrow) can exacerbate the problem. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Comparison to known normals. Bunny is a sparse (2002) downsampling of the original data set, 2-holed mug is a contoured, undersampled smooth surface, dragon is a Marching Cubes mesh, gargoyle is a full laser scan data set, figure is an evenly sampled, C^4 surface with known normals, and radius is a contour data set from a CT SCan. Given is the average and standard deviation of the dot product of the calculated normal with the known one. Best is in bold.

	Bunny	Mug	Dragon	Gargoyle	Figure	Radius
Ours	**0.981, 0.057**	0.909, 0.215	**0.996, 0.033**	**0.999, 0.063**	**0.984, 0.063**	**0.989, 0.057**
SVD	0.935, 0.149	0.758, 0.348	0.990, 0.053	0.995, 0.019	0.974, 0.086	0.973, 0.116
Quad	0.880, 0.204	0.747, 0.321	0.968, 0.102	0.974, 0.073	0.928, 0.147	0.961, 0.115
Cubic	0.860, 0.230	0.678, 0.341	0.968, 0.105	0.977, 0.071	0.924, 0.162	0.956, 0.131
Del	0.952, 0.121	**0.951, 0.102**	0.985, 0.032	0.991, 0.022	0.974, 0.080	0.979, 0.083

(negative curvature). Given a known surface normal, we evaluate whether or not the data can be plausibly explained by one of these models (Section 5). If it can't, then either the surface normal is wrong or the data points do not, locally, form a surface. Second, a good-quality one-ring approximation of the surface—one that has roughly equal-sized, equilateral triangles—is the best method for estimating the normal because it does not suffer from inadvertent smoothing or over-fitting. (Smoothing can always be applied in a post-processing step if desired—see Section 8.) Third, the one-ring and the normal should be mutually consistent—projecting the data points onto the tangent plane should yield the same one-ring (Section 4). This is particularly important around ridges with high curvature, where the range of valid normals is small.

We use a combination of optimization and validation to find the surface normal, one-ring, and local surface model that are mutually consistent. We define evaluation scores for the shape models (Section 5) and for the one-ring (Section 4) along with cross-validation criteria (Section 3). Note that the search space is discrete—there are an enumerable number of one-rings. We use this fact to develop a heuristic algorithm that generates valid one-ring candidates and then optimizes them (Section 3).

We evaluate our normal construction and shape classification using both real and test data (Section 7). We demonstrate the usefulness of our one-rings for both surface reconstruction and smoothing (Section 8). Our contributions are:

- Robust normals even in the presence of noise in the data and non-uniform sampling.
- Construction of a well-behaved, minimal one-ring neighborhood from the k-nearest neighbors.
- Local shape estimation (flat, bowl, saddle, ridge, corner, sharp edge, or boundary).
- An optimization algorithm to find the above that cross-validates the results.
- Identification of outliers and poorly reconstructed areas.
- Two novel surface reconstruction algorithms based on the constructed one-rings.

Source code is available at https://sourceforge.net/projects/meshprocessing/.

2. Related work

One approach to normal estimation is surface reconstruction, either local [2] or global [5,6]. In the case of noise-free and dense sampling, several global, Delaunay-based techniques exist for accurate normal and feature size approximations [6–8]. In the presence of noise, however, these reconstructions can be incorrect. Recent work [9] extends this approach to noisy data by using an adaptive threshold to cull Delaunay balls that arise due to noise. We compare our local approach to this one and show that, particularly for unevenly sampled data, our normal reconstruction is more accurate (Table 1). However, the global approach can be more accurate in cases where the between sample distance is less than the between surface distance (Section 7.2).

The most common approach to normal estimation is plane fitting. Mitra and Nguyen [10] provide a formula for estimating the best number of neighbors, k, to use based on estimates of the noise and local curvature. They then calculate the normal by plane fitting and show that adaptively choosing k increases the accuracy of the normal estimation. The fitting approach was extended to quadratic and cubic surfaces with normal-based weights [11] which provide a better approximation in the presence of noise. A recent survey [1] also found quadratic fitting more accurate for moderate noise, although for very high noise planar fitting was better. We compare our approach to both planar, quadratic, and cubic fitting and show that, particularly for areas of high curvature and uneven sampling, our approach outperforms these surface fitting approaches (Section 7). This is because irregular sampling can easily "pull" the fitted surface away from the average, due to the nature of linear regression. This effect is worse when the underlying shape, such as a corner, cannot be approximated by the fitted surface (plane, quadric, or cubic).