Photo-consistent surface reconstruction from noisy point clouds

Ehsan Aganj, Renaud Keriven, Jean-Philippe Pons

PHOTO-CONSISTENT SURFACE RECONSTRUCTION FROM NOISY POINT CLOUDS

Ehsan Aganj, Renaud Keriven

{aganj,keriven}@certis.enpc.fr
CERTIS - Ecole des Ponts - Universite Paris-Est
19, rue Alfred Nobel - Cite Descartes
77455 Marne-la-Vallee - France

Jean-Philippe Pons

jean-philippe.pons@cstb.fr
CSTB, 290 route des Lucioles,
BP 209, 06904 Sophia-Antipolis
Cedex, France

ABSTRACT

Existing algorithms for surface reconstruction from point sets are defeated by moderate amounts of noise and outliers, which makes them unapplicable to point clouds originating from multi-view image data. In this paper, we present a novel method which incorporates the input images in the surface reconstruction process for a better accuracy and robustness. Our approach is based on the medial axis transform of the scene, which our algorithm estimates through a global photo-consistency optimization by simulated annealing. A faithful polyhedral representation of the scene is then obtained by inversion of the medial axis transform.

***Index Terms*—** Surface reconstruction, point cloud, medial axis transform, stereovision.

1. INTRODUCTION

The problem of approximating a surface from a set of sample points has received a considerable interest in computational geometry, and more generally in mesh processing and in computer graphics. This problem, which is referred to as *surface reconstruction*, arises in many applications of science and engineering. Many approaches have been proposed [1, 2, 3], among which some offer theoretical guarantees on the output geometry and topology when the input sampling is sufficiently dense. Of particular interest in this work is the *Power Crust* algorithm [2], which builds a discrete approximation of the medial axis transform, then recovers a polyhedral surface as its inverse.

However, all the aforementioned algorithms assume that the sample is free of noise and outliers. This is an important restriction, since no real scanning device provides exact data. To handle this problem, some recent algorithms have been developed. In [4] theoretical guarantees have been provided by considering a noise model in which both the sampling density and the noise level depend on the local level of surface detail. A limitation of their algorithm is that it does not handle arbitrarily over-sampled datasets. This limitation has been overcome in a recent paper [5]. This work augments the *Power Crust* algorithm [2] with a greedy filtering process which discards parts of the medial axis transform originating from noise. However, the noise assumptions used in these two recent works do not hold in practice: the density and the noise level of samples produced by real scanning devices do not depend on the *local feature size* of the surface.

In this paper, we tackle the special case of point clouds extracted from calibrated multi-view image datasets. These point clouds typically feature higher levels of noise and higher proportions of outliers than those of active scanning devices. This discards most, if not all, standard surface reconstruction algorithms. Our work continues along the line of [5, 2]: it replaces the greedy estimation of the medial axis transform with a global photo-consistency optimization by simulated annealing, thus yielding improved accuracy and robustness.

The remainder of this paper is organized as follows. Section 2 presents some useful computational geometry concepts. Our method is described in Section 3 and is experimentally validated in Section 4.

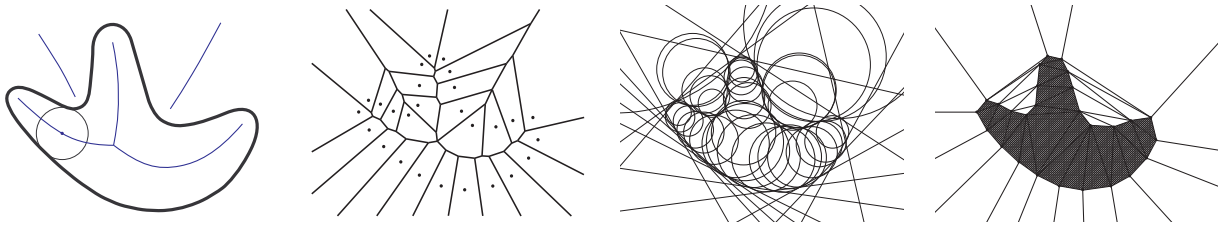
2. BACKGROUND

2.1. Medial axis transform

Since its introduction by Blum [6], the *medial axis* (and the *skeleton*, a closely related mathematical notion), has become a standard tool in shape analysis, recognition and classification. Formal definitions vary from author to author. Please refer to a recent review [7] for a thorough mathematical presentation.

Here, we define the medial axis of a closed bounded surface S as the closure of the set of points with at least two closest points on S. The *inner medial axis* (resp. the *outer medial axis*) is the subset of the medial axis inside (resp. outside) S. Figure 1(a) shows a two-dimensional example of the medial axis of a curve in the plane.

If we weight each point x of the medial axis with the radius $\rho(x)$ of the maximal ball centered at x whose interior does not intersect S, i.e. the distance from x to its closest points on S, we obtain the *medial axis transform* of the shape. This transformation is reversible, meaning that S can unambiguously be reconstructed from its medial axis transform.



(a) (b) (c) (d)

Fig. 1. *(a)* A two-dimensional curve in plane, its medial axis, and a maximal circle centered in a point of the medial axis. *(b)* The Voronoi diagram of a set of points in the plane. *(c)* The inner and outer polar balls. *(d)* The power diagram of the inner and outer polar balls (see text).

2.2. Voronoi diagram, poles, polar balls

We now consider the transposition of the above continuous notions to the discrete case. Of particular interest in surface reconstruction is the result by Amenta and Bern [1] that, given a sufficiently dense sample $P = \{p_1, \ldots, p_n\} \in \mathbb{R}^3$ of a surface S, the medial axis of S is approximated by a subset of the *Voronoi diagram* of P.

We recall that the *Voronoi cell* of a point p_i, denoted by $V(p_i)$, is the region of space that is closer from p_i than from all other points in P:

$$V(p_i) = \{x \in \mathbb{R}^3 \ : \ \forall j, \ \|x - p_i\| \leq \|x - p_j\|\} \ .$$

$V(p_i)$ is a convex polytope, possibly unbounded. The *Voronoi diagram* of P is roughly the set of cells, convex polygonal facets, edges and vertices induced by the Voronoi cells $V(p_i)$. For illustration purposes, Figure 1(b) shows the Voronoi diagram of a set of points in the plane.

Following [1], we call *poles* the vertices of the Voronoi diagram that approximate the medial axis. A pole is called an *inner pole* or an *outer pole* depending on whether it lies inside or outside S. We call *polar ball* a maximal ball centered at a pole whose interior does not intersect P. In other words, the radius of a polar ball is the distance from the pole to its closest points in P. Figure 1(c) shows the set of inner and outer polar balls of the Voronoi diagram 1(b).

The set of all inner and outer polar balls is the discrete counterpart of the medial axis transform. Similarly to the medial axis transform, the set of polar balls can be inverted: the *Power Crust* algorithm [2] builds a polyhedral surface that approximates the input point set P, by computing the *power diagram* of the set of polar balls.

2.3. Power diagram and power crust

We denote by $b_{c,\rho}$ a ball of radius ρ centered at c. The *power distance* between two balls b_{c_1,ρ_1} and b_{c_2,ρ_2} is defined by

$$d^2_{pow}(b_{c_1,\rho_1}, b_{c_2,\rho_2}) = \|c_1 - c_2\|^2 - \rho_1^2 - \rho_2^2 \ .$$

Using this distance, we can generalize the Voronoi diagram of a set of points to a set of balls $B = \{b_{c_1,\rho_1}, \ldots, b_{c_n,\rho_n}\}$: the *power cell* of a ball b_{c_i,ρ_i} is the region of space where a non-weighted point x is closer from b_{c_i,ρ_i}, in term of power distance, than from all other balls in B. The different power cells induce the *power diagram*. Like Voronoi diagrams, power diagrams have convex polytope cells. Figure 1(d) shows the power diagram cells of the inner and outer polar balls in Figure 1(c).

Consider the polar balls of a point sample P of a surface S. The *power crust* is the boundary between the power cells of inner polar balls and the power cells of outer polar balls. It is a polyhedral surface which approximates the surface under some sampling assumptions [2].

2.4. Surface reconstruction

We now focus on the problem of surface reconstruction: S is unknown. P may not fulfill the aforementioned sampling assumptions, and is corrupted by a certain amount of noise and outliers. The previous notions are then faced to significant difficulties: (i) the classification of poles as inner or outer is no more straightforward, (ii) due to noise and outliers, some poles may not approximate the medial axis.

In the original *Power Crust* algorithm [2], poles are labeled as inner or outer by a greedy propagation process driven by a priority queue. Unfortunately, this greedy approach has been shown to fail with a moderate amount of noise and outliers in [8].

3. APPROACH

In this work we propose to replace the greedy estimation of the medial axis transform with a global photo-consistency optimization by simulated annealing. A faithful polyhedral representation of the scene is then obtained by inversion of the medial axis transform.

3.1. Formulation

Given a sample set $P = \{p_1, \ldots, p_n\} \in \mathbb{R}^3$, we denote by $B = \{b_1, \cdots, b_m\}$ the set of m polar balls of the corresponding Voronoi diagram. We define $\vec{l} = \{l_1, \cdots, l_m\} \in \{-1, 0, +1\}^m$ a labeling vector describing the state of the

poles: a pole is labeled -1 if it is at the interior, or $+1$ if it is at the exterior of the surface. A pole labeled 0 is considered as noise and will be discarded in the estimation of the medial axis transform. We denote by $S_B(\vec{l})$ the inversion of the medial axis transform estimated by the set of labeled polar balls $(B, \vec{l})$. This is a polyhedral surface obtained by the power crust of the poles in B labeled inside (-1) or outside $(+1)$ as described before. In order to estimate correctly the medial axis transform we wish to minimize an energy functional dealing with photo-consistency on the surface of power crust,

$$\mathrm{E}(S_B(\vec{l})) = \sum_{f \in S_B(\vec{l})} \gamma(f) \tag{1}$$

where f is a face of the polyhedral surface $S_B(\vec{l})$. This measures how well the given surface matches the different input images in which it is seen. It is defined as the sum over the whole surface of some photo-consistency measure γ. The optimal surface is obtained by

$$S_{\mathrm{opt}} = S_B\left(\arg\min_{\vec{l}} \mathrm{E}(S_B(\vec{l}))\right) \tag{2}$$

3.2. Optimization

In order to minimize 1, we use the simulated annealing optimization algorithm [9]. The algorithm starts with an initial labeling vector $\vec{l}_0$. A global variable T is taken as the temperature of the optimization process, it is initialized to T_0. At each iteration, it takes a random pole b_i in B. Then, it changes its label l_i to a random new state and computes the energy variation δE. If δE is negative, it accepts the new labeling and goes to next iteration. Otherwise, it decides with a probability given by a function $\Pr(\delta E, T)$ if the new state should be accepted. In case it should not, it returns to the old configuration. The temperature T decreases during the process by a function which depends on the number of passed iterations. The algorithm stops when the energy is sufficiently low, or when a predefined time limit is over. The algorithm returns $S_B(\vec{l})$ as the final surface.

4. EXPERIMENTAL RESULTS

We have tested our method on a real dataset publicly available at `http://cvlab.epfl.ch/~strecha/multiview/`. All implementations have been done using *CGAL* library (`http://www.cgal.org/`)[10]. We have compared our method with a recent multi-view reconstruction method [11] robust to outliers. Our experiment uses the eight views Herz-Jesu-P8 dataset from [12]. Some images of this dataset are shown in Figure 3(c). 6950 points are generated by matching image keypoints. In order to reduce the computation time of our method we have initialized it with a labeling vector obtained by [11] and we have reduced the optimization space to the set of polar balls with radius smaller than the characteristic noise level of the dataset. The optimization process terminated after 40 minutes on a standard workstation. Two other methods have also been tested as sanity checks: the reconstruction algorithm [5] and the Laplacian smoothing algorithm which have been both initialized by the result of [11]. We have compared the cumulative error distribution of the four meshes with respect to the mesh obtained by the LIDAR technique provided by [12]. Table 1 shows the error distribution of the results provided by an automatic multi-view evaluation program [12]. For each method four cumulative error distributions have been computed. Column n represents the percentage of image pixels with an error less than $n\sigma$, where σ is the characteristic noise of LIDAR. More accurate results yield to higher scores in the table. Figure 2 shows the meshes obtained by four experimented methods. Our result compares favorably with the three other methods in terms of accuracy and robustness. In a second experiment we have initialized a variational multi-view stereovision method [13] with our mesh and with the output of [11]. Figures 3(a) and 3(b) show the results obtained by the two initializations. The results show that [13] provides more accurate and less noisy surface when it is initialized by our method.

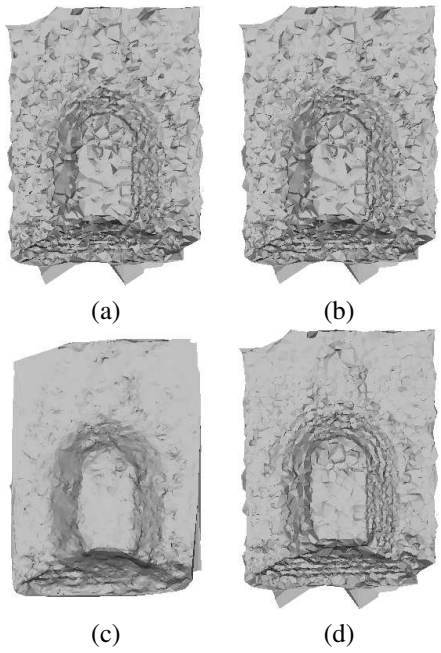
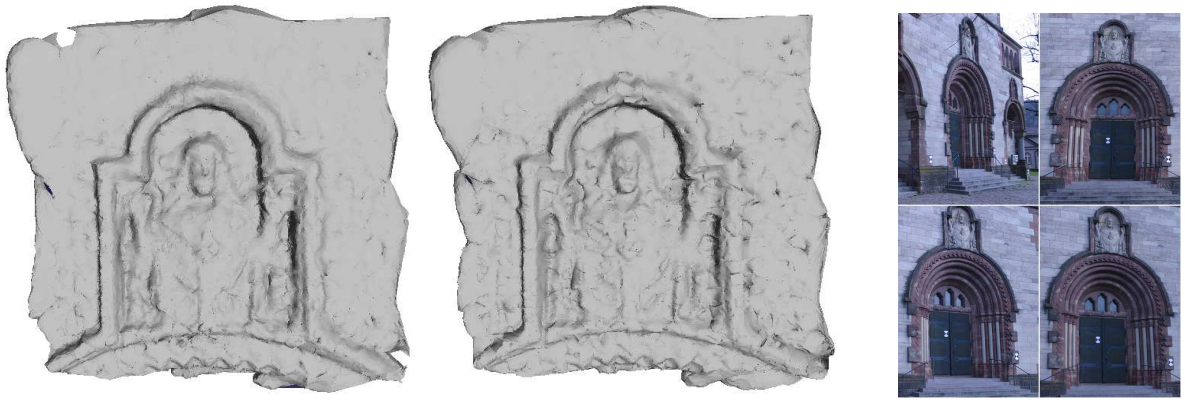


Fig. 2. Different meshes obtained by: a) The multi-view reconstruction method [11] b) The reconstruction algorithm [5] c) Laplacian Smooth d) our method.

5. DISCUSSION AND CONCLUSION

We have proposed a photo-consistent surface reconstruction method from noisy point clouds based on the estimation of the medial axis transform. Our work replaces the greedy algorithms of [5, 2] with a global photo-consistency optimization by simulated annealing. We have validated our method on real datasets. Our results compare favorably with state of



(a) (b) (c)

Fig. 3. Mesh obtained by the method of [13] initialized by the output of: a) our method b) The multi-view reconstruction method [11]. c) Some images of the Herz-Jesu-P8 dataset. We have used only a partial cut of the original images.

Experiment	1	3	5	7
[11]	2.37	10.81	16.80	20.89
Laplacian Smooth	2.21	10.18	15.80	20.14
[5]	2.38	10.84	16.87	20.98
Our method	3.60	15.10	21.79	25.72

Table 1. Quantitative results of our different numerical experiments. Column n represents the percentage of image pixels with an error less than $n\sigma$. More accurate result yields to a higher score in the table.

the art methods in terms of accuracy and robustness to noise. A drawback of our method is its computational complexity due to the optimization process. We hope that by modifying the form of the energy function we will be able to use more practical optimization algorithms such as graph-cuts.

6. REFERENCES

[1] N. Amenta and M. Bern, "Surface reconstruction by voronoi filtering," *Discrete and Computational Geometry*, vol. 22, pp. 481–504, 1999.

[2] N. Amenta, S. Choi, and R. K. Kolluri, "The power crust, unions of balls, and the medial axis transform," *Computational Geometry: Theory and Applications*, vol. 19, pp. 127–153, 2001.

[3] J.-D. Boissonnat and F. Cazals, "Smooth surface reconstruction via natural neighbour interpolation of distance functions," 2000, pp. 223–232.

[4] T. K. Dey and S. Goswami, "Provable surface reconstruction from noisy samples," *Comput. Geom.*, vol. 35, no. 1-2, pp. 124–141, 2006.

[5] B. Mederos, N. Amenta, L. Velho, and L.-H. de Figueiredo, "Surface reconstruction from noisy point clouds," in *SGP '05*, Aire-la-Ville, Switzerland, 2005, p. 53, Eurographics Association.

[6] H. Blum, "A Transformation for Extracting New Descriptors of Shape," in *Models for the Perception of Speech and Visual Form*, Weiant Wathen-Dunn, Ed., pp. 362–380. MIT Press, Cambridge, 1967.

[7] D. Attali, J.-D. Boissonnat, and H. Edelsbrunner, *Mathematical Foundations of Scientific Visualization, Computer Graphics, and Massive Data Exploration*, Springer-Verlag, Mathematics and Visualization, 2009.

[8] R. Kolluri, J.R. Shewchuk, and J.F. O'Brien, "Spectral surface reconstruction from noisy point clouds," in *SGP '04*, 2004, pp. 11–21.

[9] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of State Calculations by Fast Computing Machines," *J. Chem. Phys.*, vol. 21, pp. 1087–1092, 1953.

[10] J.-D. Boissonnat, O. Devillers, M. Teillaud, and M. Yvinec, "Triangulations in cgal (extended abstract)," in *SCG '00*, NY, USA, 2000, pp. 11–18, ACM.

[11] P. Labatut, J.-P. Pons, and R. Keriven, "Efficient multi-view reconstruction of large-scale scenes using interest points, delaunay triangulation and graph cuts," in *ICCV*, Rio de Janeiro, Brazil, Oct 2007.

[12] C. Strecha, W. von Hansen, L.J. Van Gool, and U. Thoennessen P. Fua, "On benchmarking camera calibration and multi-view stereo for high resolution imagery," 2008, pp. 1–8.

[13] J.-P. Pons, R. Keriven, and O. Faugeras, "Multi-view stereo reconstruction and scene flow estimation with a global image-based matching score," *IJCV*, vol. 72, no. 2, pp. 179–193, 2007.