Global Optimization for Shape Fitting

Victor Lempitsky
Department of Mathematics and Mechanics
Moscow State University
victorlempitsky@gmail.com

Yuri Boykov
Department of Computer Science
University of Western Ontario
yuri@csd.uwo.ca

Abstract

We propose a global optimization framework for 3D shape reconstruction from sparse noisy 3D measurements frequently encountered in range scanning, sparse feature-based stereo, and shape-from-X. In contrast to earlier local or banded optimization methods for shape fitting, we compute global optimum in the whole volume removing dependence on initial guess and sensitivity to numerous local minima. Our global method is based on two main ideas. First, we suggest a new regularization functional with a data alignment term that maximizes the number of (weakly-oriented) data points contained by a surface while allowing for some measurement errors. Second, we propose a touch-expand *algorithm for finding a minimum cut on a huge 3D grid using an automatically adjusted band. This overcomes prohibitively high memory cost of graph cuts when computing globally optimal surfaces at high-resolution. Our results for sparse or incomplete 3D data from laser scanning and passive multi-view stereo are robust to noise, outliers, missing parts, and varying sampling density.*

1. Introduction

This paper develops a global optimization approach to 3D reconstruction from sparse points containing noise, outliers, and gaps. We assume that each data point comes with some estimate of surface orientation - a vector that we use to softly constrain the (outward) normal of a surface that fits the point. Unlike many previous methods using estimated normals, our approach needs orientation only at observed (sparse) data points and it is relatively robust to orientation errors. Our experiments use only sensor/camera directions (e.g. two arrows in Fig. 1) available in most applications.

We propose a novel surface fitting functional directly enforcing geometric proximity to data points. Intuitively, it counts data points contained by a surface while allowing for localization errors. This data-fit functional does not bias a reconstructed surface to any particular shape but it can be combined with appropriate shape priors (smoothness/regularization), volumetric occupancy data, or other terms (e.g. photoconsistency). A wide class of such combined geometric surface functionals can be globally optimized via graph cuts [2, 13], avoiding the pitfalls of local optimization such as dependence on initial guess (Fig. 2,3).

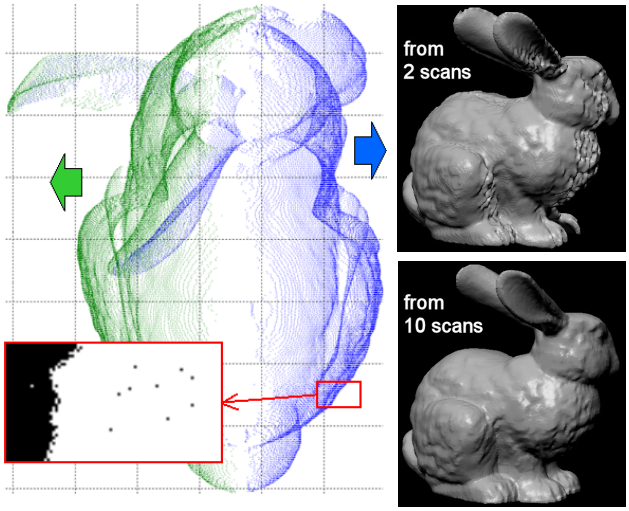
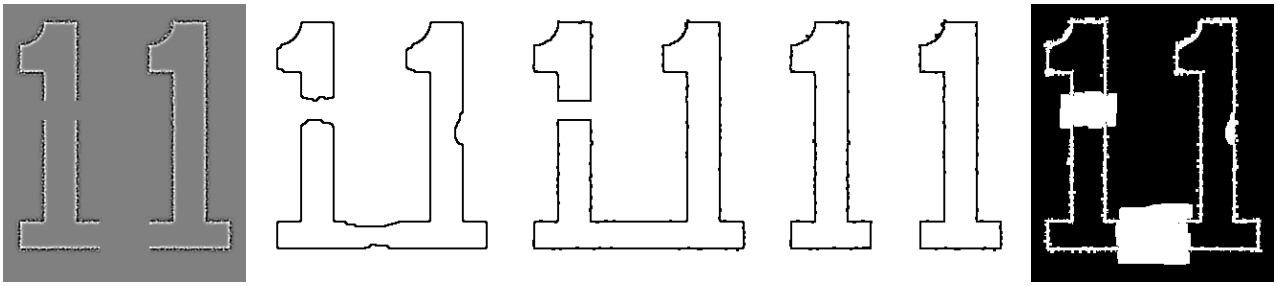


Figure 1. *Left* – an example of input data for our method showing 2 (out of 10) registered range scans (courtesy of Stanford's repository). Typical noise and outliers are shown in a close-up (from a similar dataset). The arrows show 2 scan directions which is often sufficient surface orientation information for our method. *Right* – surface fitting result for 2 opposite view scans (top) shows some artifacts in the areas tangential to both scans. In such areas there are relatively few noisy data points and scan directions are particularly bad orientation estimates for them. Adding points from other views gives good overall accuracy (bottom).

Graph cuts typically require a lot of memory for high resolution volumes. Previous hierarchical or narrow band techniques [16, 10] do not guarantee global minima and generally behave as local optimization methods, see Fig. 2,3. In order to do globally optimal surface fitting at high resolutions, we propose a new memory-efficient *touch-expand* algorithm that guarantees global minimum cut in the whole volume without performing computations on a full grid.

Global optimization of our new functional with the touch-expand algorithm produces high-resolution water-



a) Data term b) Initial shape c) Local optimization d) Global optimization e) Band (in white)

Figure 2. Global vs. local/banded optimization for 2D shape reconstruction. We minimize a functional similar to (1) or (11). The energy's data term is based on pixels' potentials (a): dark – negative, bright – positive, gray – zero. The regularization term is geometric length of the shape's boundary. (c) shows local optimization (level sets) results for initial solution (b). Global optimization (graph-cut on a full-size 2D grid) is in (d). Our method detects this global minima using a small band (e) automatically adjusted by our *touch-expand* algorithm.

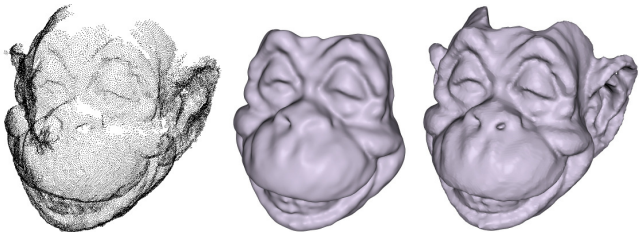


Figure 3. A cloud of data points (*left*). **Banded graph-cut** from unoriented points [10] (*middle*): incorrect estimate of initial shape may lead to significant reconstruction errors. **Global graph-cut** approach proposed in this work (*right*) avoids local minima problem. We also show that even rough estimates of data points' orientation may significantly improve alignment with the data.

tight surfaces closely fitting data points without overfitting to noise and outliers.

2. Related work on surface fitting

While a detailed review of related work on surface fitting can be found in [14], below we can mention only a fraction of existing methods putting stronger emphasis on regularization-based techniques which better deal with noise, outliers, and large gaps. The majority of standard methods can be grouped based on their numerical approach to surface representation. While explicit mesh-based representations are very common, e.g. [5], to avoid mesh-related numerical problems, many methods represent surfaces implicitly using level-sets [6, 21, 22, 18]. Both meshes and level-sets can be locally optimized via gradient descent.

We use a different approach where surfaces are represented as binary partitionings (s/t cuts) on discrete grids [2, 13]. This implicit representation technique allows global optimization of a large class of surface functionals [13] via network flow algorithms. Robust global graph-cut methods have already demonstrated a strong potential for many problems in computer vision but they have not been applied to surface fitting. One noticeable exception is a very recent method of Hornung and Kobbelt [10], but they compute an optimal cut in a fixed narrow band (a crust) whose shape directly affects the solution. Similarly to local or variational methods, their results depend on the initialization, *i.e*. on the particular shape of the constructed crust. Using banded (non-global) optimization may lead to reconstruction errors, while discarding orientation information may lead to over-smoothing (Fig. 3). Our graph-cut based algorithm is the first surface fitting method that guarantees global optima in the whole volume. We also designed a novel geometric surface-to-data fit functional suited for global optimization.

Regardless of their approach to surface representation, regularization-based surface fitting techniques can be compared by the energy. Below we review existing geometric functionals. To merge incomplete range scans obtained from multiple view points Whitaker [21] uses functional

$$E(S) = \int_{interior(S)} D(v)dv + \int_S ds \qquad (1)$$

where $D(v) = \sum_j D_j(v)$ is a potential function based on a combination of signed distance maps of independent scans obtained from multiple view points. Function $D(v)$ is very similar to the interior function used by Curless and Levoy [6]. Both [21] and [6] assume that each j-th scan is dense which allows the corresponding signed distance function to be defined in a straightforward fashion. In fact, minimizing the first term in (1) is equivalent to extracting a zero-level isosurface of $D(v)$ which is exactly the algorithm in [6].

Zhao *et al*. [22] proposed a different regularization-based surface fitting functional applicable to an arbitrarily sparse set of points P

$$E(S) = \int_S d_P^n(s) \cdot ds, \quad 1 \le n \le \infty \qquad (2)$$

where $d_P(s)$ is a (unsigned) distance from point s on surface S to the nearest data point in P. No estimates of surface orientation are required. As discussed in [22], a local minima of (2) is a surface composed of facets from Delaunay triangulation of points in P. Such local minima solution will strongly depend on initialization.

Hornung and Kobbelt [10] apply graph-cut surface optimization [2] to a sparse data fitting functional very similar to (2). However, (2) is not appropriate for global optimization since its global minima is a trivial (null) surface. Thus, [10] compute their solution in a (fixed) narrow band reducing graph-cut framework to local optimization essentially equivalent to the variational approach in [22].

Savadjiev *et al*. [18] formulate surface fitting as a problem of estimating a dense vector field $\{v_p\}$ of surface normals from sparse data. Then, a continuous surface can be recovered from a dense field $\{v_p\}$ by optimizing a *flux* functional

$$E(S) = \int_S \langle v_s, n_s \rangle \cdot ds \qquad (3)$$

where n_s is a normal of S. To estimate a dense field $\{v_p\}$ from sparse data points, they use *variational relaxation* framework. However, a lack of global surface orientation in [18] results in ± 180 degree ambiguity of the estimated vectors $\{v_p\}$. Then, they use some heuristics deviating from (3) and causing geometric artifacts (e.g. "thickness").

3. Our energy formulation

We skip attempts to accurately estimate a dense field of surface normals. Instead, we assume a sparse vector field

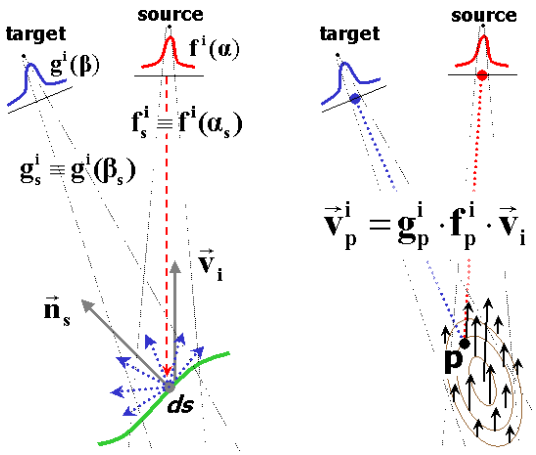
$$\{v_i |\ p_i \in P\}$$

where each vector v_i softly constrains the true surface normal n_i at data point p_i. In particular, we expect large positive values of the *cosine* of the angle between v_i and n_i. To account for localization errors we "blur" vectors v_i obtaining a semi-dense field as in Fig. 5. Flux of S with respect to such a vector field can be seen as counting points covered by S. Our approach to surface fitting is based on combining such flux with regularization (or shape prior) in a single geometric functional. The regularization term propagates information and resolves ambiguities while flux enforces alignment with data. Combining flux with regularization is commonly used in vision for reconstruction of elongated structures, narrow protrusions, and other fine details [19, 12, 13, 4]. Section 3.1 further motivates our data fit functional statistically and geometrically. Section 3.2 combines it with surface regularization (generic shape priors).

3.1. Quality of Fit to Data

Ideally, if all discrete data points have absolute precision then the quality of a surface fit can be measured by the number of data points that lie on it. Reconstruction of a continuous/dense surface would require imposing a shape prior (*e.g*. regularization) even for precise sparse data but this factor is separate from the quality of fit to data.

In practice, it is well known that even range scanning measurements are not exact, not to mention accuracy of



(a) Observing surface S (b) Vector field $\bar{v}^i_p = g^i_p f^i_p \bar{v}_i$

Figure 4. Data points from active or passive reconstruction techniques (*e.g*. range scanning or feature-based stereo) are not as discrete as they may appear. For example, the i-th data point from a laser scanner (a) corresponds to noisy non-deterministic measurements of an illuminated patch which could be smaller or larger depending on the beam's width and the patch's orientation. In laser scanning, "source" image $f^i(\alpha)$ is a probability of emitting a photon in direction α and "sensor" image $g^i(\beta)$ measures *radiance* of surface points viewed at angle β. Statistically motivated functionals like (6) suggest that sparse/discrete surface measurements can be represented via flux of continuous vector fields (b).

sparse stereo. How should the quality of a fit be measured in the presence of noisy data? In order to clarify this question it is very instructional to analyze the uncertainty of a single data point in laser scanning which is often considered to be the most accurate approach to surface reconstruction. Our basic analysis can be extended to other active or passive light acquisition methods. Laser scanning is chosen primarily as the simplest example illustrating our main approach.

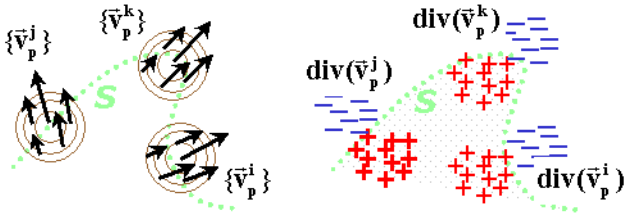
Consider the active light surface measurement process shown in Fig. 4 (a). The amount of light energy reaching surface patch ds from the source in a unit of time is

$$ds \cdot \qquad f^i_s \cdot \langle \bar{v}_i, \bar{n}_s \rangle \qquad (4)$$

where $\bar{n}_s$ is surface normal, $\bar{v}_i$ gives direction to the source, and $f^i_s := f^i(\alpha_s)$ for angle α_s at which the source views given point s in space. The Lambertian assumption about surface S implies that *radiance* of points on patch ds is the same in all angular directions. Since the image sensor measures the radiance of surface points then the total energy of light reflected by patch ds in a unit of time (e.g. see [7]) is

$$ds \cdot \qquad \pi \cdot g^i_s \qquad (5)$$

where $g^i_s := g^i(\beta_s)$ for angle β_s at which the sensor observes point s in space. There is a connection between (4) and (5): if surface *albedo* is ρ then $\pi \cdot g^i_s = \rho \cdot f^i_s \cdot \langle \bar{v}_i, \bar{n}_s \rangle$.



(a) Flux (b) Divergence

Figure 5. Maximization of flux through vector field $\bar{v}_p = \sum_{i=1}^{N} \bar{v}_p^i$ can be seen as maximization of the number of data points contained by a surface while allowing measurement uncertainty (a). This is equivalent to maximizing divergence of vector field $\{\bar{v}_p\}$ inside surface interior (b). Points with non-zero (positive or negative) values of scalar function $div(\bar{v}_p)$ are indicated by $+$ or $-$.

Functions (4) and (5) are defined for points s on any surface S. However, the linear relationship

$$g_s^i \propto f_s^i \cdot \langle \bar{v}_i, \bar{n}_s \rangle \quad \forall s \in S$$

only holds if points s and normals n_s belong to the (unknown) surface that reflected the light. It is natural to estimate surface S by maximizing some measure of similarity between the two functions on S. Maximization of the dot product

$$max \xleftarrow{S} \int_S g_s^i \cdot f_s^i \cdot \langle \bar{v}_i, \bar{n}_s \rangle \cdot ds \qquad (6)$$

is a reasonable approach (see also [14]) to align (4) and (5).

Note that the Lambertian assumption above is non-essential for our main argument as long as there is a correlation between the irradiance on patch ds and the reflected light radiance in the direction to the sensor.

In the case of multiple data points $1 \leq i \leq N$, our measure of surface fitness sums functionals (6) for all points i. This is analogous to counting the number of points contained by S in the case of absolutely precise data. We get functional

$$max \xleftarrow{S} \int_S \langle \bar{v}_s, \bar{n}_s \rangle \cdot ds \qquad (7)$$

which is *flux* for vector field $\bar{v}_p = \sum_{i=1}^{N} \bar{v}_p^i$ representing all (uncertain) data points, see Fig. 4(b) and Fig. 5(a). Note that the Gauss-Ostrogradsky (a.k.a. *divergence*) theorem

$$\int_S \langle \bar{v}_s, \bar{n}_s \rangle \cdot ds = \int_{inter(S)} div(\bar{v}_p) \cdot dp. \qquad (8)$$

implies that (7) is equivalent to maximizing the integral of vector field's *divergence* in the interior of S, see Fig. 5(b).

General idea of data fit functional (7) goes beyond laser scanning. In the majority of existing active or passive light methods for acquiring discrete/sparse surface measurements, each i-th data point naturally comes with direction to the source $\bar{v}_i$ or even better estimate of surface orientation. Moreover, instead of a single data point p_i most methods can report some distribution function $\rho_i(p)$ describing probability that i-th measurement corresponds to a surface patch located at point $p \in \mathcal{R}^3$. The specific form of $\rho_i(\cdot)$ depends on particularities of the specific acquisition method. In many cases this distribution is a Gaussian with a given mean (data point p_i) and some covariance matrix specific to each technique. In general, we suggest flux-based functional (7,8) as a generic surface-to-data fit quality measure where vector field $\{\bar{v}_p\}$ representing data is

$$\bar{v}_p = \sum_i \rho_i(p) \cdot \bar{v}_i \qquad \forall p \in \mathcal{R}^3. \qquad (9)$$

In the special case of laser scanning we had $\rho_i(p) = g_p^i \cdot f_p^i$ which in most cases is a Gaussian whose variance depends on the laser beam width. In feature-based sparse stereo, variance may depend on the size of the image features.

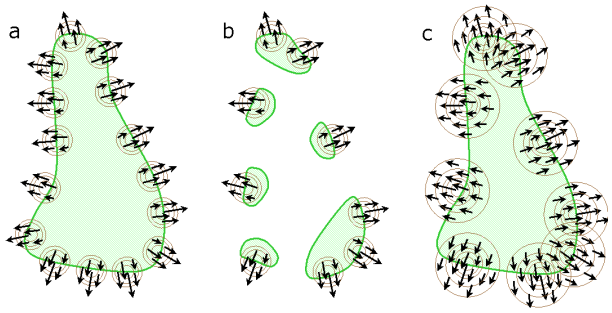


Figure 6. Surface fitting with different sampling density and priors. a) Optimization of the data-fit term in combination with a membrane prior produces consistent reconstruction for densely sampled points. b) The same fails when the sampling density is low. c) Widening the support of each point (ad-hoc "low curvature" prior) allows to obtain correct reconstruction.

3.2. Adding Shape Prior and/or Occupancy data

In the presence of large gaps and outliers in the data, imposing some shape prior is essential for surface reconstruction. One simple approach is to augment functional (7) with an area-based regularization term (*elastic membrane* prior) giving the following minimization problem:

$$\min \xleftarrow{S} \int_S \lambda \cdot ds \; - \int_S \langle \bar{v}, \bar{n}_s \rangle ds \qquad (10)$$

where $\{\bar{v}_p\}$ is a vector field representing data, as in (9). The exact value of λ determines the strength of the membrane prior and may be chosen according to the sampling density.

The optimization of the functional (10) maximizes the number of collected data points while minimizing the area, thus handling noise and outliers in the initial data, see Fig. 6(a). However, if data points are sampled at distances much sparser than each point's "support", see Fig. 6(b), then the membrane prior may produce inconsistent results for all choices of λ.

An alternative approach is to use a low-curvature (*stiff membrane*) prior. The corresponding variational models are known. We propose a simple heuristic for low curvature shapes. Note that each observed data point p_i with estimated orientation $\bar{v}_i$ may correspond to a small surface patch of certain size Δ. The smaller expected surface curvature is, the larger Δ gets. Then, the i-th data point support function $\rho_i(p)$ that represents the likelihood of point p to be on a surface with orientation $\bar{v}_i$ (see (9)) can be widened by Δ in the direction orthogonal to $\bar{v}_i$. Thus, the low curvature prior may justify support functions $\rho_i(p)$ wider than the span of the measurement error model. Fig. 6(c) demonstrates the effect of this ad-hoc low-curvature prior.

In many cases, there is also some information about the scene's geometry that comes in the form of *spatial occupancy*. It can be defined by a volumetric function $O(p)$ where positive values indicate that point p is likely to be inside the surface and vice versa, while absolute values of $O(p)$ correspond to the certainty. Such occupancy data usually comes from *line-of-sight* information (for each range scan, the space between the scan surface and the scanner is likely to be empty) or silhouette intersection and may be easily incorporated into functional (10). As a result, in the most general case the optimization problem has the form:

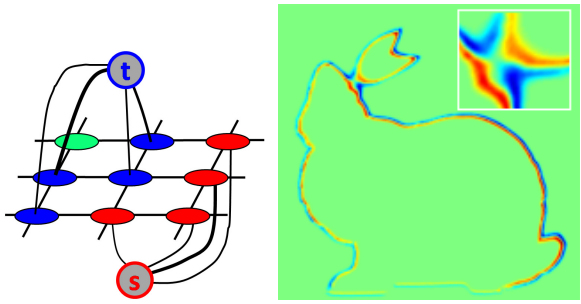
$$\min \overset{S}{\leftarrow} \int_S \lambda \, ds \; - \int_{inter(S)} U(p) \, dp \qquad (11)$$

where $U(p) = div(\bar{v}_p) \; + \; O(p)$ is a volumetric potential unifying divergence (flux) of field $\{\bar{v}_p\}$ and occupancy data.

4. Global Min-Cut via *Touch-Expand* Method

[2, 13] showed that global minima of geometric surface functionals like (11) can be computed via graph cut algorithms. A typical graph construction is shown in Fig. 7(a). A grid of nodes is embedded in a bounding box. Adjacent nodes are connected via *n-links*, whereas *t-links* connect nodes to the terminals s or t according to each node's potential $U(p)$. A surface S corresponds to an s/t-*cut* on the constructed graph. An appropriate choice of edge weights ensures that the cost of this s/t-cut closely approximates the value of functional (11) corresponding to surface S; the weights of severed t-links approximate the volumetric potential term, while the weights of severed n-links approximate the surface area [2, 13]. Consequently, a global minimum surface for (11) can be found by computing a minimal s/t-cut in the constructed graph.

The spatial resolution of the grid should be sufficiently high to match the resolution of the data. Memory requirements make existing mincut/maxflow algorithms practically infeasible (for global optimization in full volumes) due to the sheer size of the grids required for high-resolution data sets. However, it may be observed (Fig. 7(b)) that, for our application, most of the grid nodes (pixels/voxels) are located away from the data points where the minimal surface is likely to pass. It seems that such structural redundancy can be exploited via optimization in a narrow band near the data points, but this converts graph cut into a local optimization technique (like level-sets) and makes surface reconstruction sensitive to the initial guess (Fig. 2).



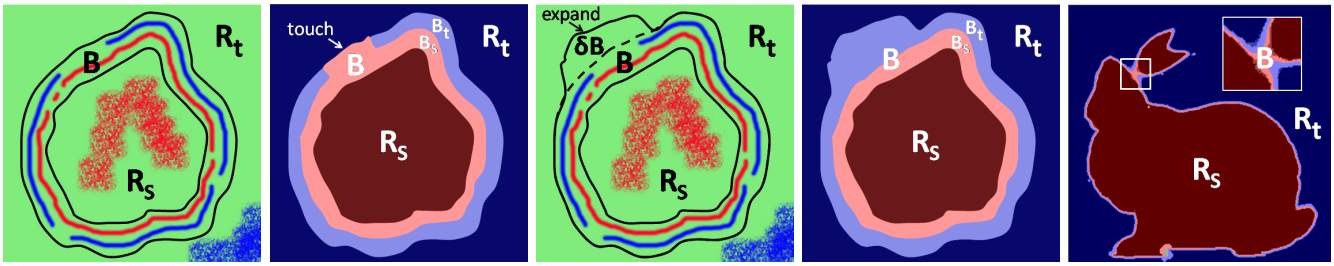
(a) Grid graph (2D) (b) Real data example

Figure 7. (a) A grid construction for optimizing functional (11) as in [2, 13]: neighboring nodes are connected via n-links representing regularization cost, the nodes are also connected to the terminals via t-links representing volumetric potentials. (b) T-links (or potentials) for a slice of 3D "bunny" data in Fig.1: red nodes, $U(p) > 0$, are connected to the source, and blue nodes, $U(p) < 0$, are connected to the sink. Green nodes, $U(p) = 0$, are not connected to the terminals. Our *touch-expand* method attempts to save computer memory required for the green area nodes where the minimum cut is less likely (but possible!) to pass.

We developed a more principled approach to address memory efficiency that preserves global optimality of graph cuts. Our *touch-expand* algorithm computes an exact minimum cut for the entire graph, while allocating memory only for a band (subgraph). The algorithm initializes some band around an initial guess, and then modifies the band in a series of *touch-expand* steps. Crucially, at convergence the minimum cut on a modified band is guaranteed to coincide with the minimum cut on the entire graph.

Consider a network graph G. Our algorithm maintains a subgraph (band) B including a subset of the nodes in G, the terminals s and t, and the edges connecting all these vertices. The rest of the graph is split into two subgraphs R_s and R_t (Fig. 8(a)). It is required that the nodes in R_s are not directly connected via t-links to t and that the nodes in R_t are not connected to s. Also, it is required that R_s does not "touch" R_t (no n-links from R_s to R_t). Given an initial guess in a form of some s/t-cut, such configuration is established by adding to B all nodes incident to the severed edges, all nodes in s-component connected to t, and all nodes in t-component connected to s. The remaining nodes in s- and t-components form R_s and R_t, correspondingly.

Each touch-expand iteration works as follows. Compute a minimal cut in a current band B and let B_s and B_t denote its s- and t-components. Identify the nodes in R_s adjacent to B_t and nodes in R_t adjacent to B_s (Fig. 8(b)).



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

Figure 8. *Touch-expand* algorithm computes minimum cut without allocating memory for the whole graph (unless necessary). It maintains band B and subgraphs R_s and R_t as in (a) where red pixels/nodes are t-linked to s and blue nodes are t-linked to t. If a min-cut for B results in a *touch* as in (b), the band is *expanded* (c). If a minimum cut for a new band B does not result in a "touch" as in (d), then it is guaranteed to coincide with the min-cut on the entire graph and the algorithm stops. (e) shows a band at convergence for a real problem.

Such nodes form a "touch" set δB, which is subtracted from R_s and R_t and added to the band B (expand operation – Fig. 8(c)). A min-cut can be computed in an expanded band and a sequence of touch-expand iterations continues until $\delta B = \emptyset$ (Fig. 8(d)). It can be proven that if $\delta B = \emptyset$ then the minimum cut on B severs the same edges as the minimum cut on the entire graph G. In other words, the minimum cut on G partitions the graph into $B_s \cup R_s$ and $B_t \cup R_t$. A simple proof using mincut/maxflow duality is given in [14].

The monotonic growth of B guarantees a finite number of iterations. It is also important that the maximum flow used to compute a min-cut for B at one iteration is feasible for an expanded band B at the next iteration. Thus, we can efficiently reuse flow between iterations, which is faster than computing the maximum flow from scratch.

Touch-expand algorithm is guaranteed to find a globally minimum cut. However, its efficiency (i.e. the final size of B and the number of iterations) may depend on the accuracy of initialization. For our problem, a good initialization may be easily obtained with a hierarchical approach where min-cut at a courser scale is used to initialize the touch-expand algorithm at a finer resolution. Using such initialization for the experiments in this paper resulted in final bands B that were typically an order of magnitude smaller than the entire graph (Fig. 8(e)). Also, the particular hierarchical algorithm detailed in [14] spends most of the time (e.g. 90%) on a single max-flow computation during the first touch-expand iteration at the finest resolution.

5. Experimental evaluation

Our implementation of touch-expand method was based on the max-flow algorithm [3] for computing minimum cuts. To extract an isosurface from binary segmented grids without aliasing artifacts, we used the modification of the method [20] (details are given in [14]). In general, our algorithm produced watertight meshes of correct genus but when the resolution was too high, the algorithm might give a few floating voxels. The algorithm did not require a lot of parameter tuning, *e.g.* all Stanford data sets (Fig. 9,10) were processed using the same parameters.

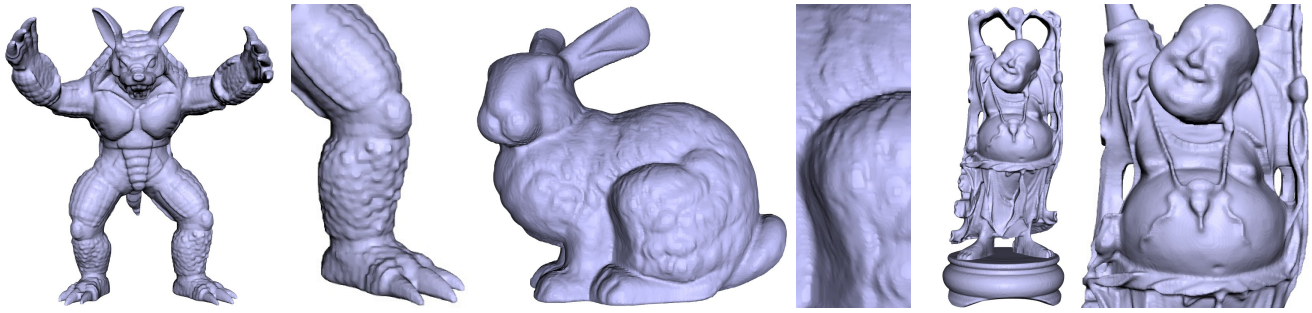
Scene	Grid size	Band size	Time
Bunny	551x544x428(6)	3.72%	61.7
Dragon	602x425x269(6)	9.61%	62.9
Dragon(VRIP)	901x637x403(6)	3.06%	47.5
Armadillo	601x635x501(6)	2.23%	70.7
Buddha	377x914x377(6)	10.75%	1269.9
Temple Full	452x704x332(6)	10.37%	359.5
Temple Ring	451x707x331(6)	11.91%	640.7
Temple Sparse	275x427x202(26)	21.92%	1116.1
Dino Full	252x300x253(26)	20.59%	629.3
Dino Ring	251x300x253(26)	21.31%	1038.2
Dino Sparse	252x300x253(26)	21.68%	667.3

Table 1. Touch-expand method statistics for our experiments. Neighborhood system size is shown in brackets. The fourth column contains the percentage of all nodes that were included in the subgraph B on convergence of touch-expand algorithm. Time (for 1.8 GHz processor) is given in seconds.

Numerical performance of touch-expand algorithm is summarized in Tab. 1. The peak RAM allocation size for all models was 1.5-2.5 GB. Note that most of these tests would demand prohibitively large amount of memory if standard mincut/maxflow methods run at the same high resolutions.

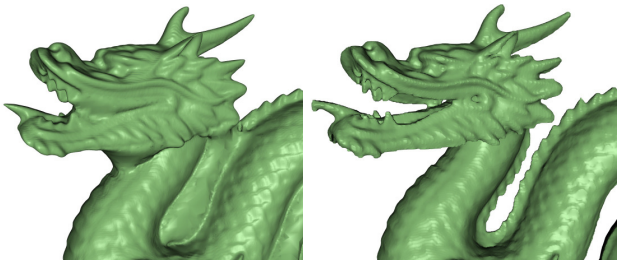
Fitting to laser-scanned data. We tested our method on the range scan datasets from Stanford 3D Scanning Repository, containing 10–112 registered laser scans. Importantly, we assigned the single orientation vector (corresponding to scan viewing direction) for all points in each scan, thus relying on the extremely coarse normal estimates. For Buddha dataset, the coarse line-of-sight information was used near the legs to ensure correct hole-filling. High-resolution results obtained with our method are shown on Fig. 9.

We also compared our method with Poisson surface reconstruction (author implementation). Being introduced in [11], it was demonstrated there to compare favorably to several other state-of-the-art methods. In our experiments, it was however unable to handle Stanford Dragon dataset with coarse orientation estimates, whereas our method produced more consistent results (Fig. 10). Note that the Dragon



Armadillo Bunny Buddha

Figure 9. Results of our method on the range scan datasets from the Stanford Repository. Coarse orientation estimates were used (one direction per scan). Touch-expand algorithm allowed to compute the results corresponding to global optima of our functional at high resolution.



Poisson Ours

Figure 10. Comparison of our method with Poisson surface reconstruction [11] on the Stanford Dragon dataset. As the points had very coarse orientation estimates, Poisson reconstruction lead to oversmoothing, whereas our method was able to produce more consistent results.

dataset with the vertices of VRIP-ed reconstruction and accurate normal estimates was handled equally well by both methods [14].

We also tested the ability of our approach to handle large variations in sampling density. To do that, we removed 98% of points from the right half of Armadillo and then applied our method. With the use of non-uniform Euclidean regularization $\lambda = \lambda(p)$ in (10) proportional to the local sampling density, our method was able to handle such 50-to-1 sampling difference gracefully (Fig. 11).

Fitting to passive-stereo data. Recently, a number of multiview reconstruction methods have been proposed that produce high-accuracy "quasi-dense" output, leaving holes where the stereo correspondence can not be established [15, 8, 9]. Our method can post-process such outputs turning them into complete watertight meshes, while preserving the high level of detailization.

To test this, we considered incomplete meshes produced by mutltiview reconstruction method [8] on Middlebury multiview stereo page datasets [17] kindly provided by the authors of [8]. For our algorithm, we treated the vertices as input points and estimated the normals from local mesh structure. In order to perform the hole filling in ambiguous cases, we used the occupancy information coming from the coarse silhouette intersection.

Figure 11. Result of our method (right) applied to the pruned Stanford Armadillo range scans data (left). 98% of points were removed from the left part, thus creating 50-to-1 difference in density challenging to most algorithms. Using spatially-varying regularization within our method allowed to produce a consistent reconstruction.

The results are summarized on Fig. 12. While our method was quite successful in producing watertight meshes from the input, the results on Dino reveal metrication artifacts, which are probably the major limitation of our method. For such problems where filling of large holes is required, the use of continuous maxflow [1] rather than discrete version may be considered (note that touch-expand method can be used for continuous maxflow as well).

6. Future Work

This work focused on global optimization of geometric surface fitting functionals that can be represented via binary submodular 2-nd order energies. Higher order interactions and/or supermodular terms could encode wider classes of shape priors (e.g. minimum total curvature) and geometric data-alignment constraints (e.g. for surface orientation).

Acknowledgements

Our research benefited from discussions with Alexander Hornung and Leif Kobbelt who shared their "monkey" dataset as well as their results. Michael Goesele kindly pro-

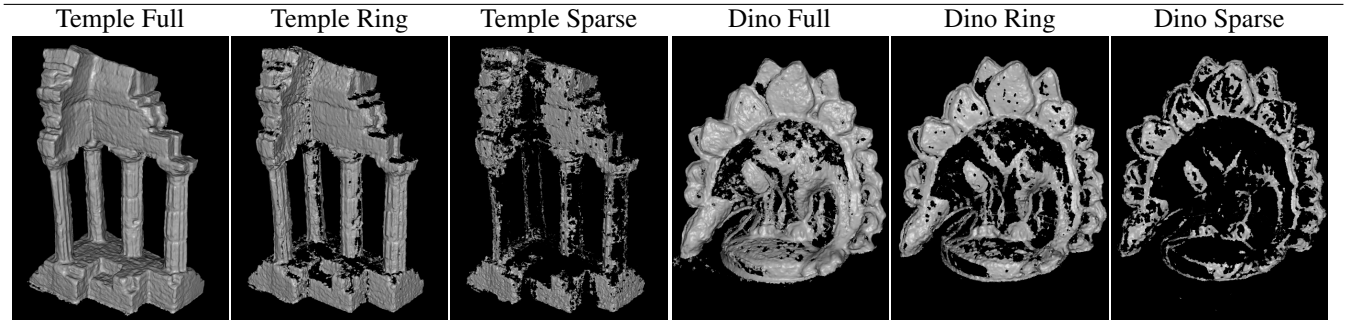


Figure 12. Shape fitting to incomplete meshes. Top – results of Goesele *et al* method [8], bottom – watertight shapes fitted to those results using our method. First numbers are the accuracy measures (90% of the surface lies within this distance from the ground truth). Second numbers are completeness measures (percentage of the ground truth that lies within 1.25 mm from the surface). Visual inspection and statistical measurements suggest that our method significantly improved the completeness of the meshes, while the accuracy of the reconstructions remained on a par with the input.

vided the results of [8]. Olivier Juan suggested a modification of [20], which we used to gain subpixel accuracy. Discussions with Andrew Delong helped to clarify many important points. We also thank Daniel Scharstein, Brian Curless, Steven Seitz, and other maintainers of [17] for the statistical evaluation of our results on Middlebury data and designing a custom-made web-page for them[1].

References

[1] B. Appleton and H. Talbot. Globally minimal surfaces by continuous maximal flows. *IEEE TPAMI*, 28(1):106–118, 2006.

[2] Y. Boykov and V. Kolmogorov. Computing geodesics and minimal surfaces via graph cuts. In *ICCV'03(I)*, page 26.

[3] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE TPAMI*, 26(9):1124–1137, 2004.

[4] Y. Boykov and V. Lempitsky. From photohulls to photoflux optimization. In *BMVC'06(III)*, pages 1149–1158.

[5] L. Cohen and I. Cohen. Finite element methods for active contour models and balloons from 2-D to 3-D. In *CVPR'92*, pages 592–598.

[6] B. Curless and M. Levoy. A volumetric method for building complex models from range images. In *SIGGRAPH'96*, pages 303–312.

[7] D. Forsyth and J. Ponce. *Computer Vision, a modern approach*. Prentice Hall, 2003.

[8] M. Goesele, B. Curless, and S. Seitz. Multiview stereo revisited. In *CVPR'06(II)*, pages 2402–2409.

[9] M. Habbecke and L. Kobbelt. Iterative multi-view plane fitting. In *VMV'06*, pages 73–80.

[10] A. Hornung and L. Kobbelt. Robust reconstruction of watertight 3d models from non-uniformly sampled point clouds without normal information. In *Eurographics Symposium on Geometry Processing'06*, pages 41–50.

[11] M. Kazhdan, M. Bolitho, and H. Hoppe. Poisson surface reconstruction. In *Eurographics Symposium on Geometry Processing'06*, pages 61–70.

[12] R. Kimmel and A. M. Bruckstein. Regularized Laplacian zero crossings as optimal edge integrators. *International Journal of Computer Vision*, 53(3):225–243, 2003.

[13] V. Kolmogorov and Y. Boykov. What metrics can be approximated by geo-cuts, or global optimization of length/area and flux. In *ICCV'05(I)*, pages 564–571.

[14] V. Lempitsky and Y. Boykov. Global optimization for shape fitting. In *UWO, Comp. Sci., Tech. Report 679, ISBN 978-0-7714-2592-9*, Feb. 2007.

[15] M. Lhuillier and L. Quan. Quasi-dense reconstruction from image sequence. In *ECCV*, pages 125–139, 2002.

[16] H. Lombaert, Y. Sun, L. Grady, and C. Xu. A multilevel banded graph cuts method for fast image segmentation. In *ICCV'05*, pages 259–265.

[17] Middlebury Multiview Stereo Page. http://vision.middlebury.edu/mview/.

[18] P. Savadjiev, F. P. Ferrie, and K. Siddiqi. Surface recovery from 3d point data using a combined parametric and geometric flow approach. In *EMMCVPR'03, pp. 325-340*.

[19] A. Vasilevskiy and K. Siddiqi. Flux maximizing geometric flows. *PAMI*, 24(12):1565–1578, December 2002.

[20] R. Whitaker. Reducing aliasing artifacts in iso-surfaces of binary volumes. pages 23–32.

[21] R. Whitaker. A level-set approach to 3d reconstruction from range data. *IJCV*, 29(3):203–231, 1998.

[22] H. Zhao, S. Osher, and R. Fedkiw. Fast surface reconstruction using the level set method. In *VLSM '01, p. 194*.

[1] http://vision.middlebury.edu/mview/2006-lempitsky/