

Photoflux Maximizing Shapes

Yuri Boykov, Victor Lempitsky

Abstract

Our work was inspired by recent advances in image segmentation where flux-based functionals significantly improved alignment of object boundaries. We propose a novel *photoflux* functional for multi-view 3D reconstruction that is closely related to properties of photohulls. Our *photohull prior* can be combined with regularization. Thus, this work unifies two major groups of multiview stereo techniques: “space carving” and “deformable models”. Our approach combines benefits of both groups and allows to recover fine shape details without oversmoothing while robustly handling noise. Photoflux provides an intelligent ballooning force helping to segment thin structures or holes. We propose a number of different versions of photoflux based on global, local, or non-deterministic visibility models. Some forms of photoflux can be easily added into standard regularization techniques. For other forms we propose new optimization methods. We also show that photoflux maximizing shapes can be seen as regularized Laplacian zero-crossings.

1 Introduction

The reconstruction of 3D shape from a set of its views, *multiview reconstruction*, is one of the fundamental problems in computer vision. In this paper, we assume that all views are registered within the global coordinate system, such that for any 3D point $X \in \mathcal{R}^3$ it is possible to find its projection onto each of the views.

The majority of multiview shape-from-photoconsistency methods fall into two categories: greedy space carving approaches aiming at recovering maximal photoconsistent shapes, photohulls, and energy-based methods searching for some regularized surface with good photoconsistency (Figure 4b&c). Normally, energy-based methods achieve better surface quality by adding regularization into the reconstruction process removing noise. However, an explicit bias towards maximal photoconsistent surface, i.e. photohull, allows space carving to recover more complex geometries and to outperform energy-based methods on the scenes with thin, protruding, and/or textureless parts.

This work is in many respects inspired by the recent progress in image segmentation, which is another shape reconstruction problem in low-level vision sharing a lot in common with multiview reconstruction. Similarly to multiview reconstruction, a number of greedy methods such as thresholding or region growing that once dominated in the field of image segmentation, were replaced with regularized minimal/deformable surface approaches. However, standard minimum surface methods for image segmentation often oversmooth thin structures and undersegment blurred boundaries, which are similar to problems of current energy-based multiview reconstruction methods.

Recently, new energy-based approaches based on flux integration have been proposed for image segmentation [31] and significant improvement in segmentation of thin objects

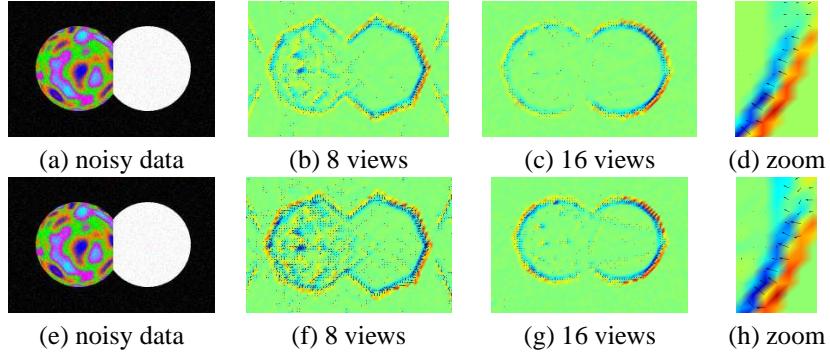


Figure 1: Two methods for computing *photoconsistency flow* vector fields (Sec.2.3) in 3D volumes. One camera view of a 3D scene is shown in (a) and (e). The spectrum of colors in (b-d,f-h) indicates divergence of vector fields in one slice of the volume. Green color corresponds to zero divergence. Pictures (b-d) illustrate vector fields based on average photoconsistency gradients (7) and pictures (f-h) show “expected” surface normals (9).

and in alignment of object boundaries has been reported. These approaches optimize the combination of a Riemannian area/length and the flux of a vector field normally computed from the gradients of a gray-scale image. The state of the art in two problems is summarized in the table shown below.

	Greedy methods	Minimal/deformable surfaces	Flux-based methods
Image segmentation	Thresholding [27] Region growing [1]	Snakes [12] Level-sets [23] Graph cuts [3]	Level-sets [31] Graph cuts [17]
Multiview reconstruction	Voxel coloring [28] Space carving [18]	Mesh-based [7] Level-sets [8, 30, 25] Graph cuts [29, 32, 20]	This work

This paper proposes flux optimization for multiview reconstruction. We also propose methods for calculating vector fields from the gradients of the photoconsistency function that allow to bias reconstructed surfaces towards photohull (Figure 1). We demonstrate that integrating flux of such fields, *photoflux*, into the reconstruction process allows to recover thin and textureless parts of objects. Photoflux can be easily integrated into regularization-based methods for multiview reconstruction. We propose a number of novel multiview reconstruction techniques combining photoflux and surface regularization. Such techniques can use both global and local optimization methods such as level-sets or graph cuts. We prove the concept of photoflux experimentally by comparing reconstruction results with and without photoflux.

2 From *photohulls* to *photoflux* optimization

The concept of *photohull* [18] is based on an “oracle” that makes “yes” or “no” decisions about photoconsistency of any given point on the current surface. The output of space

carving algorithm, *photohull*, will depend on a threshold for making hard decisions about photoconsistency. Like “region growing” is sensitive to “leaking” through weak spots on object boundary, space carving is not robust to noise or specularities on object surface that will often result in “leaking” and eventual cascading erosion of the shape.

However, space carving has one significant advantage over existing minimal surface methods for multiview stereo. Space carving can reconstruct thin objects or fine protruding details of the shape that previous energy based methods tend to oversmooth. Below we propose a surface functional for multiview reconstruction that avoids shrinking bias of standard regularization-based techniques. Our formulation of multiview reconstruction problem is motivated by the fact that a surface of a photohull should have large flux of photoconsistency gradients, that we also call *photoflux*.

Photoflux allows to integrate a bias to photohull into regularization methods for multiview stereo. Recent results in image segmentation suggest that incorporating flux-based functionals into local and global regularization methods such as level-sets [31, 13] and graph cuts [17] significantly improved alignment of segments with object boundaries.

2.1 Photoflux and global visibility

For now, assume a fixed surface \hat{S} . Let $\rho(X|\hat{S})$ be some *color variance* function measuring the deviation between the colors I_X^1, \dots, I_X^m of a point X in different views given visibility of X defined by the shape \hat{S} (for example, this could be a current solution in an iterative algorithm). Some examples of color variance measures are mentioned in Section 3.3. Under a Gaussian noise assumption, the likelihood of a surface at point X to be photoconsistent with the images is

$$P(X|\hat{S}) = \Pr(I_X^1, \dots, I_X^m | X \in \hat{S}) \propto \exp\left(-\frac{\rho(X|\hat{S})}{2\sigma^2}\right) \quad (1)$$

where σ is the noise variance. We will use this likelihood function throughout the paper. We will simply refer to (1) as a *photoconsistency function*.

Instead of (1) space carving algorithm considers a binary photoconsistency function

$$P(X|\hat{S}) = \begin{cases} 1 & \text{if } \rho(X|\hat{S}) \leq \sigma \\ 0 & \text{if } \rho(X|\hat{S}) > \sigma \end{cases}$$

where σ is a photoconsistency threshold. The algorithm can be regarded as a successive carving of inconsistent voxels X such that $P(X|\hat{S}) = 0$ and a simultaneous update of \hat{S} . Therefore, space carving algorithm converges to shape S (photohull) that is a boundary between points with $P(X|S) = 1$ and points with $P(X|S) = 0$.

In the context of our non-deterministic photoconsistency function (1), a boundary of a photohull is characterized by a large gradient of $P(X|S)$ directed from exterior of photohull to its interior. Based on this observation, we propose the following functional for energy-based multiview reconstruction

$$\text{PHOTOFLUX I :} \quad F(S) = - \int_S \langle \nabla P(X|S), N_X \rangle dS \quad (2)$$

where N_X is an outward looking unit normal of surface S at point X , and $\langle \cdot, \cdot \rangle$ is a scalar product that does not have to be necessarily Euclidean. Functional (2) corresponds to flux

of photoconsistency gradients through the surface. To be short, we call this functional a *photoflux* through an oriented surface S^1 .

Remember that reconstruction of photohulls using space carving algorithms does not include any regularization. Analogously, optimization of flux alone does not provide any regularization either [17]. Divergence theorem suggest that flux can be seen as an “intelligent” regional balloon force. To add regularization, we augment the functional (2) with an area surface term. This term can correspond either to Euclidean area $\int_S dS$ or to a photoconsistency-based area, used in other energy-based methods $\int_S \rho(X|S) dS$.

2.2 Photoflux and local visibility

In case of convex shapes we have $P(X|S) = P(X|N_{dS})$ where N_{dS} is an outward normal of a surface patch containing point X and

$$P(X|N) = \Pr(I_X^1, \dots, I_X^m | X, N) \propto \exp\left(-\frac{\rho(X|N)}{2\sigma^2}\right) \quad (3)$$

is a photoconsistency of point X based on its *local visibility* [20] defined by a given surface normal N . Clearly, photoconsistency $P(X|N)$ should be computed from cameras located in the half-space defined by a tangent plane at point X and its outward normal N .

Thus, convexity of shape S implies $\nabla P(X|S) = \nabla P(X|N_X)$ and photoflux of such shape can be written as

$$\text{PHOTOFUX II : } F(S) = - \int_S \langle \nabla P(X|N_X), N_X \rangle dS. \quad (4)$$

For general non-convex shapes S equation (4) is an approximation of the definition of photoflux in (2). In case of occlusions due to non-convexity of shape (e.g. in case of multiple objects) we have $P(X|S) \geq P(X|N_X)$ since some cameras used to compute photoconsistency based on local visibility contribute an error. Our experiments show even local visibility approximations like (4) handle occlusions gracefully.

2.3 Photoflux and non-deterministic visibility

Note that the earlier two definitions of *photoflux* are based on vector fields of gradients $\nabla P(X|S)$ or $\nabla P(X|N)$. These vectors are defined only at points X on some given surface S or on a given local patch with normal N . The surface or a patch must be fixed in order to estimate visibility of point X before photoconsistency $P(X|S)$ or $P(X|N)$ can be computed. We also assume that small perturbations of object shape near X do not change visibility so that the gradients of photoconsistency can be evaluated at X .

This section provides a more general definition of photoflux. Instead of computing visibility of point X deterministically for a given surface or a patch, we assume some probability distribution $\Pr(W|X)$ for possible values of visibility W at any given point X . Note that visibility $W = (w_1, \dots, w_m)$ is a vector of non-negative weights $w_k \geq 0$ describing correlation between intensity I_X^k of a pixel in camera k where point $X \in \mathcal{R}^3$ projects and a color of point X assuming it belongs to some surface.

¹The orientation of S is defined by a field of its “outward” normals.

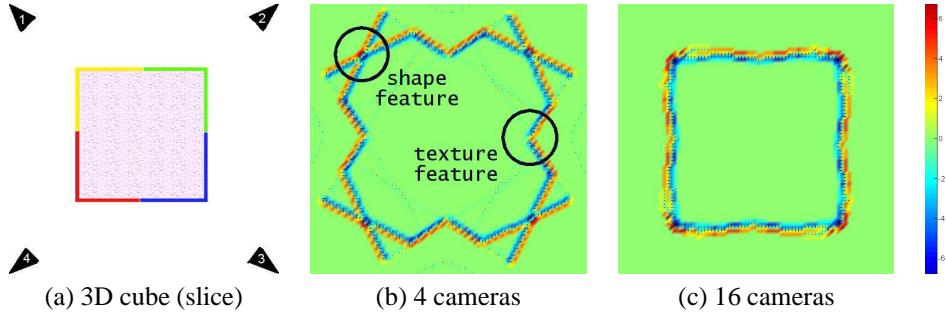


Figure 2: Reconstruction of a cube. Divergence of estimated photoconsistency gradients is shown in (b) and (c). The images of divergence demonstrate “apparent contours” of shape features (sharp corners) or surface texture features. Our method can be seen as regularization of Laplacian zero crossings of average photoconsistency (see Sec. 3.2).

The most general definition of photoflux in this paper is motivated by flux-based shape priors for image segmentation [17]. Our “photohull shape prior” is based on

$$\text{PHOTOFUX III :} \quad F(S) = - \int_S \langle v_X, N_X \rangle dS . \quad (5)$$

where a field of vectors $\{v_X\}$ describes *photoconsistency flow*, that is, direction and rate of increase in photoconsistency at every point X . One way to obtain such vectors is to compute expected gradients of photoconsistency by averaging unknown visibility out

$$v_X = \sum_W \nabla P(X|W) \cdot \Pr(W|X). \quad (6)$$

Functionals (2) and (4) are special cases of (5) when visibility of X is deterministic.

We also suggest a simpler but more practical version of photoconsistency flow

$$v_X = \sum_N \nabla P(X|N) \cdot \Pr(N|X) \quad (7)$$

where visibility is represented by a possible outward normal direction N . A possible heuristic for estimating probability $\Pr(N|X)$ in practice could be based on gradients of photoconsistency

$$\Pr(N|X) \propto \langle \nabla P(X|N), N \rangle^+ \quad (8)$$

where $\langle \cdot, \cdot \rangle^+$ is a positive part of the dot product. Figures 1(b-d) and 2 demonstrate vector fields computed according to equations (7) and (8).

Equation (8) requires a normalization multiplier to make it a true probability distribution. In practice, however, this may cause “division by zero” problems which can be avoided by truncating a vector field where the multiplier gets too small. This “division by zero” corresponds to points where surface is unlikely to pass and “visibility” is ill-defined. Practically speaking, equation (8) without any normalization generates very similar photoconsistency flows $\{v_X\}$. This heuristic is also easier to implement since no “division by zero” numerical problems arise.

In general, there are other options for computing vectors v_X for photohull prior (5). For example, we also experimented with vector fields

$$v_X = \sum_N N \cdot \Pr(N|X). \quad (9)$$

shown in examples of Figure 1)(e-h). Vector fields could also be obtained from approximate and/or sparse disparity maps. Such maps can be computed at each camera by evaluating correspondences with nearby cameras using narrow-base stereo methods or robust feature matching.

3 Properties of photoflux optimization

The definitions of photoflux in (2) and (4) use deterministic visibility and the corresponding vector fields are defined only on a surface of any fixed shape. More general form of photoflux (5) allows non-deterministic visibility and the corresponding vector field can be estimated at any point $X \in \mathcal{R}^3$. Despite these obvious distinctions, there are many important common properties of different forms of photoflux which we describe in this section. For simplicity, we discuss these properties in the context of equations (5) and (7).

3.1 Combining *photoflux* and regularization

Standard *space carving* methods for photohull reconstruction are analogous to adaptive region-growing techniques in image segmentation and they do not include any surface regularization. Section 2.1 “derived” photoflux functional from properties of photohulls. Thus, it should not be surprising that photoflux maximization itself does not imply any regularization. As discussed in [17], flux optimization is equivalent to thresholding a *divergence* of the corresponding vector field

$$\text{div}(v) = \nabla \cdot v = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}.$$

In the context of Figure 2, “thresholding” implies that all voxels of blue color (negative divergence values) should be selected as surface interior. It is clear that this approach would generate noisy and/or ambiguous results. Note that the only difference of space carving approach would be to carve away exterior voxels starting from a box containing the object. This exterior “region growing” should stop on boundary between the red and blue voxels (see Figure 2). Similar to “region growing” methods in image segmentation, space carving can leak through a single weak spots on an object boundary.

Unlike space carving, our approach to multiview reconstruction describes object boundary via a cost functional, photoflux. In contrast to photohulls, photoflux maximizing shapes can be regularized by combining photoflux (2) or (5) with any surface regularization functionals typically used in multiview reconstruction. For example, we can combine photoflux with an integral of photoconsistency over the whole surface obtaining a regularized photoflux functional

$$E(S) = - \int_S \langle \nabla P(X|N_X), N_X \rangle dS + \lambda \cdot \int_S \rho(X|N_X) dS. \quad (10)$$

Alternatively, it might also be possible to regularize photoflux via Euclidean area of the surface S . It is also possible to use non-Euclidean dot products in the definition of photoflux that may be based on photoconsistency.

Optimization of an energy combining photoflux and regularization as in (10) should reconstruct fine object details while handling noise. Photoflux can be seen as an intelligent balloon force making minimal surfaces better align with thin protrusions or cavities. Combining flux and regularization has similar effects in image segmentation [31, 13, 17].

3.2 Photoflux and regularized Laplacian zero-crossings

Integrating flux of image gradients into image segmentation energy results in better alignment with object boundaries. For example, [13] motivates flux-based approach to edge detection by connecting it with Laplacian zero-crossing filters which are image processing classics. Analogously, optimization of photoflux can give better alignment of 3D reconstruction results with an object surface (or its photohull).

This section shows that we compute regularized Laplacian zero-crossings of the likelihood function $\bar{P}(X) := \Pr(I_X^1, \dots, I_X^m | X)$ for observed intensities given that X is a point on a surface of an object. We have

$$\bar{P}(X) = \Pr(I_X^1, \dots, I_X^m | X) = \sum_N \Pr(I_X^1, \dots, I_X^m | X, N) \cdot \Pr(N | X) = \sum_N P(X | N) \cdot \Pr(N | X) \quad (11)$$

where $P(X | N)$ is the photoconsistency function in (3) and $\Pr(N | X)$ is a conditional probability function for a normal of a surface at point X given that the surface is passing through X . Generally speaking, averaging over N represents averaging over different possible visibilities of point X in the cameras observing some object from different views. Thus, $\Pr(N | X)$ can be interpreted as a probability of a given visibility for point X if it is known that it belongs to some surface.

Typically, global visibility of point X on surface S does not change much if the surface is slightly nudged at X . Similarly, local visibility of a surface patch dS should not change if it moves a bit while preserving its orientation. Therefore, it is natural to assume that $\nabla \Pr(N | X) \approx 0$ (differentiation over X). Then, the gradient of our likelihood function

$$\nabla \bar{P}(X) = \nabla \Pr(I_X^1, \dots, I_X^m | X) = \sum_N \nabla P(X | N) \cdot \Pr(N | X) = v_X$$

coincides with vector field v_X in (7). Ostrogradsky-Gauss theorem, also known as *divergence theorem*, implies that maximization of photoflux (5) is equivalent to thresholding regions with positive divergence $\text{div}(v_X)$ from regions with negative divergence. Since

$$\text{div}(v_X) = \nabla \cdot v_X = \nabla \cdot \nabla \bar{P}(X) = \Delta \bar{P}(X)$$

then combining photoflux optimization for vector field v_X in (7) with surface regularization as in Section 3.1 can be seen as computing regularized Laplacian zero-crossings of likelihood $\bar{P}(X)$.

3.3 Integrating *apparent contours* and texture features

This subsection demonstrates that photoflux based multi-view reconstruction implicitly relies on *apparent contours* and on detecting object textural features. To be specific,

assume that color variance is measured by function

$$\rho(X|N_X) = \sum_k w_k \cdot (I_X^k - \bar{I}_X)^2, \quad \bar{I}_X = \sum_k w_k \cdot I_X^k \quad (12)$$

or by

$$\rho(X|N_X) = \sum_{i,j} w_i \cdot w_j \cdot (I_X^i - I_X^j)^2 \quad (13)$$

where I_X^k is intensity of a pixel in camera k where point $X \in \mathcal{R}^3$ projects, \bar{I}_X is a weighted average of such intensities, and w_k is a non-negative weight describing visibility of point X in camera k . Typically, visibility w_k is large for a camera observing surface point X with normal N_X under a small angle and w_k approaches zero for angles near 90 degrees. Note that color consistency penalty as in (12) was proposed for 3D reconstruction in [30] and functionals like (13) were proposed in [25] (image-based regularization).

For both (12) and (13), differentiation of photoconsistency $P(X|N_X)$ in (1) gives

$$\nabla P(X|N_X) \propto P(X|N_X) \cdot \sum_k w_k \cdot (\bar{I}_X - I_X^k) \cdot \nabla I_X^k \quad (14)$$

where ∇I_X^k is a gradient of an image in camera k at a pixel where 3D point X projects. We treat ∇I_X^k as a 3D vector since k -th camera image plane orientation in \mathcal{R}^3 is known. Expressions analogous to (14) can be derived for other color variance models.

Equation (14) breaks down the gradient of photoconsistency into a linear combination of vectors contributed by individual cameras where point X is visible, that is $w_k > 0$. It is not surprising that each camera contribution is based on its image gradient ∇I_X^k . For example, a small number of cameras in Figure 2(b) allows to see straight “rays” of points with large gradients of photoconsistency. Each “ray” in Figure 2(b) is formed by points $X \in \mathcal{R}^3$ projecting onto the same pixel in one of the cameras where image gradient is large. Note that multiplication by a factor $P(X|N_X)$ in (14) discards gradients in all points $X \in \mathcal{R}^3$ that are not photoconsistent.

Computing a gradient of photoconsistency automatically obtains a “good” linear combination of image gradients (14) appropriate for estimating surface normals. Interestingly, each camera k individually does not know how to interpret its image gradients. In general, there are two possible reasons for a large image gradient at any given pixel; the camera could observe either a boundary between an object and a background (apparent contour) or a textural detail on a surface. Figure 2(b) demonstrates that rays of gradients from different cameras intersect at points that align with an object surface either at its *shape features* or at its *texture features*. Note that by computing gradients of photoconsistency we automatically detect both types of features and use them as surface “anchors”. Larger number of cameras, e.g. in Figure 2(c), gives tighter alignment of large photoconsistency gradients with such anchors producing a close outline of the true object surface.

4 Efficient algorithms

The photoflux optimization can be plugged into virtually any energy optimization framework. In particular, level-sets [8, 25] and banded graph cuts [32] reconstructions can include the photoflux optimization either in the deterministic form (2) or in the probabilistic form (5). In the former case, the current solution defines the shape estimate \hat{S} determining global visibility.

Local visibility photoflux in (4) allows to use global optimization methods using explicit graph cuts on a complex [14, 20]. However, the corresponding energy is not guaranteed to be submodular. Thus, it may be necessary either to truncate supermodular terms as in [26], or to use approximation methods like QPBO [2].

Non-deterministic approach to photoflux (5) can be implemented by global optimization methods based on either implicit surface representation [17] or complex-based graph cuts [14]. It is also interesting to investigate global methods that update estimate of visibility once parts of a surface are recovered in optimization methods that generate partial solutions (e.g., push-relabel [6], pseudo-flows [10], or active cuts [11]).

5 Experimental Results

To validate the general concept of photoflux in this paper, we present our results for *complex*-based multiview reconstruction [20] after integrating photoflux into it. This technique allows to obtain the global minimum of a discretized version of regularized photoflux functionals as in (10). In particular, we tested two forms of photoflux (4) and (5). We only briefly describe complex construction and an optimization algorithm presented in [20] focussing on the properties of our new regularized photoflux functional.

5.1 Estimating Photoconsistency Gradients

Our experimental results were obtained using vector fields of photoconsistency flow as in (4) and (5,7,8). The main component of these formulas is photoconsistency gradient $\nabla P(X|N)$ that can be estimated using finite differences. We tried the following approach.

Note that for any vector a we have $a \propto \sum_n n \cdot \langle n, a \rangle$ where unit vectors n are summed over all possible directions (or points on a unit sphere). Also, linear approximation of photoconsistency function $P(X|N)$ implies

$$\langle \nabla P(X|N), n \rangle \propto P(X + \varepsilon \cdot n|N) - P(X|N) \quad (15)$$

for sufficiently small constant ε . Therefore,

$$\nabla P(X|N) \propto \sum_n n \cdot [P(X + \varepsilon \cdot n|N) - P(X|N)].$$

In particular, for methods using vector fields as in Section 2.3 one can use

$$\langle \nabla P(X|N), N \rangle^+ \propto \left(\sum_n \langle n, N \rangle \cdot [P(X + \varepsilon \cdot n|N) - P(X|N)] \right)^+.$$

In practice, photoconsistency can be very shallow or thin around highly-textured surfaces. Then, photoconsistency function may be non-differentiable there just like function $|x|$ is not differentiable at $x = 0$. One possible solution is to compute “one-sided” derivatives. We found that “one-sided” estimate of gradient of photoconsistency $P(X|N)$ works somewhat more robustly in practice.

One-sided estimate of gradient can be computed as follows. For any vectors a and N we have

$$a \propto \sum_n n \cdot \langle n, a \rangle \approx \sum_{n: \langle n, N \rangle > 0} n \cdot \langle n, a \rangle + \sum_{n: \langle n, N \rangle < 0} n \cdot \langle n, a \rangle$$

$$\begin{aligned}
 &= \sum_{n: \langle n, N \rangle > 0} n \cdot \langle n, a \rangle + \sum_{n: \langle n, N \rangle < 0} (-n) \cdot \langle -n, a \rangle \\
 &= \sum_{n: \langle n, N \rangle > 0} n \cdot \langle n, a \rangle + \sum_{n: \langle n, N \rangle > 0} n \cdot \langle n, a \rangle
 \end{aligned}$$

that implies

$$a \propto \sum_{n: \langle n, N \rangle > 0} n \cdot \langle n, a \rangle$$

Then, (15) gives “one-sided” finite difference estimate of photoconsistency gradient

$$\nabla P(X|N) = \sum_{n: \langle n, N \rangle > 0} n \cdot [P(X + \varepsilon \cdot n|N) - P(X|N)] \quad (16)$$

and an estimate of a dot product in (8) becomes

$$\langle \nabla P(X|N), N \rangle^+ \propto \left(\sum_n \langle n, N \rangle^+ \cdot [P(X + \varepsilon \cdot n|N) - P(X|N)] \right)^+. \quad (17)$$

5.2 Tests with Complex-based Regularization of Photoflux

In the absence of any scene approximation, [20] treats visibility and photoconsistency $P(X|N)$ of a patch as functions of its position and orientation, only. In [20] all cameras with viewing directions lying within a certain angle from the normal contribute equally to the photoconsistency. In this work, we use photoconsistency measure similar to the image-based approach [30, 25] where visibility weights w_k in (12) are proportional to the area of the projection of a unit patch at given point with given orientation onto the image k . We also account for the sampling artifacts of photoconsistency function in a similar way to [19].

To find the globally optimal shape, we subdivide the approximate bounding box with 9 families of parallel planes. As a result, we obtain a *CW-complex* comprising polyhedral cells and the facets separating cells from each other. For each facet, we distinguish between two faces of this facet having opposite orientations. For every oriented facet F a regularized cost (10) is computed based on the facet’ area, position, and normal orientation. The minimum of a regularized cost (10) over all shapes comprised from the cells of the CW-complex can be efficiently found using a variant of a min-cut algorithm [4] applied to the graph, which is dual to the *CW-complex* [20].

Non-deterministic photoflux: Integrating functional (5) into complex-based regularization algorithm above is straightforward. One can estimate vector v_X at each facet and then introduce directed costs for an n-link corresponding to this facet as described in [17]. This work guarantees that the corresponding energy is submodular and that the global minima can be computed via max-flow/min-cut algorithms, e.g. [4].

In our experiments, we use complexes with 5–10 million of cells, with a typical running time of few minutes on a P4-2.8 architecture. The results are presented on the figures 4, 3, 5. For the benchmarking on Figure 4, we used similar photoconsistency measure for all three algorithms. The same method [20] without photoflux was used for the energy-based reconstruction in column c). We did our best to tune the parameters for all three methods. Empirically, photoflux-based optimization demonstrated the least sensitivity to

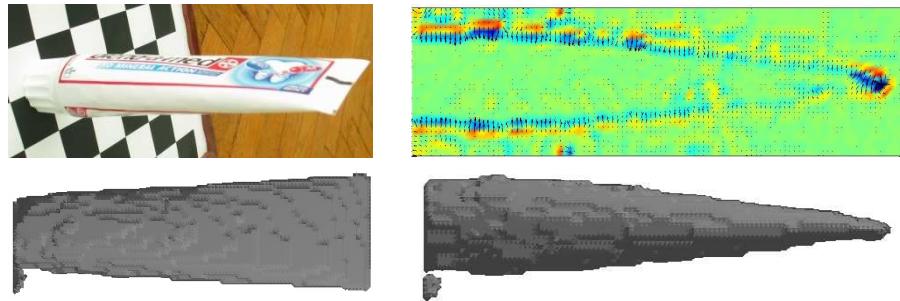


Figure 3: Results of our method for the *toothpaste* sequence. Note the sharp top of the tube. The surface also exhibits significant non-lambertian reflectivity effects. Top-right is the profile slice of a vector field. A strong photoflux vector field appears in the most difficult part (top) of the tube. Bottom row - two views of the reconstruction.

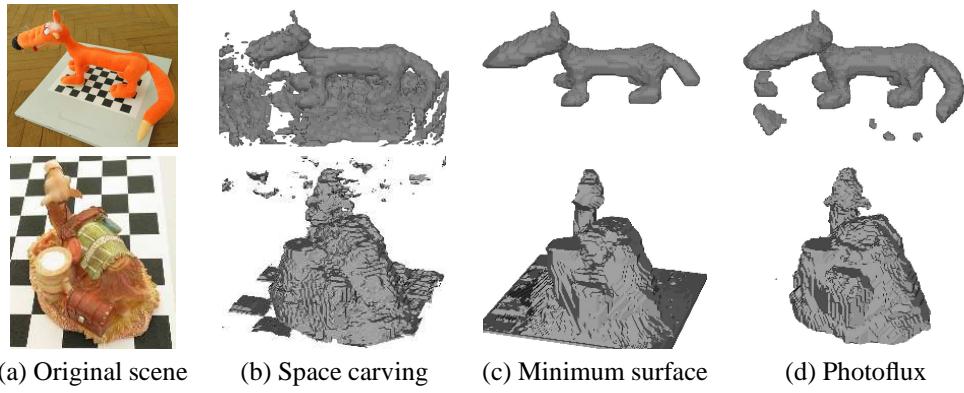


Figure 4: Comparison of different methods for multiview reconstruction on the fox and camel sequence. Typically, space carving generates noisy results while minimum surface methods oversmooth the shape. Adding maximization of photoflux into energy-based methods for 3D reconstruction allows to accurately restore thin protrusions and other fine details of the shape (ears of camel and a long tail of the fox).

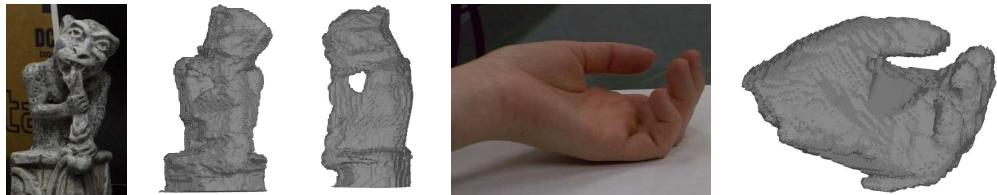


Figure 5: Results of our method on standard *gargoyle* and *hand* sequences. Note the complex topology of the gargoyle and the weak texture of the hand. Please, refer to other publications using these sequences for benchmarking.

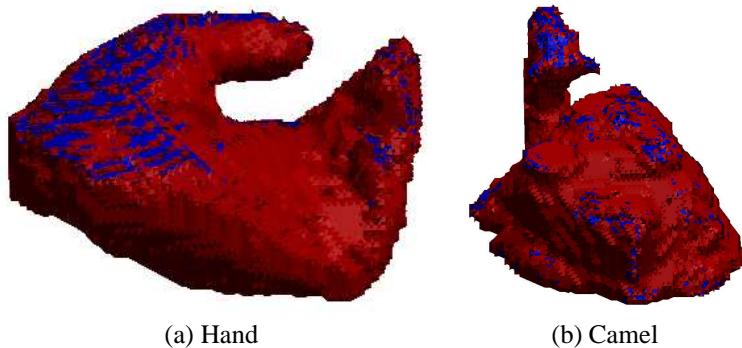


Figure 6: Regularized photoflux (4) results via global graph cuts optimization on a complex on standard *hand* and our own *camel* sequences. Blue facets correspond to non-submodular pairwise energy terms. Discretization demonstrates the actual resolution of the complex. See Figures 4 and 5 for original “hand” and “camel” images.

the parameter variation. Note, that the *gargoyle* and *hand* sequences (Figure 5) is a courtesy of Prof. Kutulakos and have been serving as testing datasets in a number of works (including [18, 19, 25]).

Local-visibility deterministic photoflux: Photoflux in (4) can also be integrated into complex based regularization framework since each facet on the complex has normal N and $\nabla P(X|N)$ can be evaluated when the complex is constructed. However, this approach separately evaluates two independent vectors $\nabla P(X|N)$ and $\nabla P(X|-N)$ for each facet since its outward normal can be either N or $-N$ and it is not known which one is correct a priori. Once the flux penalties $\langle \nabla P(X|N), N \rangle$ and $\langle \nabla P(X|-N), -N \rangle$ are computed for two possible orientations of the facet, we have to integrate them into pairwise interaction between two cells sharing this fixed facet. In this case, however, the corresponding energy is not guaranteed to be submodular as two penalties above are completely unrelated (computed from two different sets of cameras) and their relationship is unpredictable. The corresponding binary energy can be approached with a number of methods. We tested “truncation” approach as in [26] shown in Figure 6 where blue shows facets where non-submodular interactions were observed. It is very interesting to test performance of other optimization techniques for non-submodular energies like QPBO [2], tree-reweighted message passing [16], belief-propagation [24], and other methods.

6 Future Work

We are working on iterative estimation of visibility for photoflux which can be achieved via local improvement methods (e.g. level-sets [25], banded cuts [32], pde-cuts [5]), or via discrete algorithms demonstrating intermediate solutions (push-relabel [9], pseudo-flow [10], active cuts [11]). It is also interesting to compare our results with explicit graph cuts on complexes [20] with implementing regularized photoflux via implicit graph cuts techniques [3, 17]. In general, there are also other methods for computing vector fields that can estimate photoconsistency flow or surface normals defining our photohull shape

prior. Such dense or sparse fields could be computed from robust feature detection [22] or from approximate disparity maps. Our technique can also integrate additional reconstruction cues. In particular, it is clear how to integrate “silhouettes” if available (for example when background subtraction is feasible). For example, this can be done by adding a regional bias to visual hull [29]. There are obvious extensions to “shape-from-motion” with moving cameras and to dynamic applications with moving objects. In this case, efficient optimization methods using flow [15] or cut recycling [11] can be employed in the context of graph based optimization. 4D graph constructions for dynamic applications are easy to envision as well. Possible memory requirement issues can be addressed with hierarchical [21, 11], narrow-banded [32], and other local refinement techniques.

References

- [1] R. Adams and L. Bischof. Seeded region growing. *IEEE Trans. PAMI*, 16(6):641–647, 1994.
- [2] E. Boros and P. L. Hammer. Pseudo-boolean optimization. *Discrete Applied Mathematics*, 123(1-3):155–225, 2002.
- [3] Yuri Boykov and Vladimir Kolmogorov. Computing geodesics and minimal surfaces via graph cuts. In *ICCV '03*, page 26, 2003.
- [4] Yuri Boykov and Vladimir Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans. PAMI*, 26(9):1124–1137, 2004.
- [5] Yuri Boykov, Vladimir Kolmogorov, Daniel Cremers, and Andrew Delong. An integral solution to surface evolution PDEs via geo-cuts. In *European Conference on Computer Vision*, Graz, Austria, May 2006 (to appear).
- [6] Boris V. Cherkassky and Andrew V. Goldberg. On implementing push-relabel method for the maximum flow problem. In *Proceedings of the 4th International IPCO Conference*, pages 157–171, 1995.
- [7] Carlos Hernandez Esteban and Francis Schmitt. Silhouette and stereo fusion for 3d object modeling. *CVIU*, 96(3):367–392, 2004.
- [8] Olivier D. Faugeras and Renaud Keriven. Complete dense stereovision using level set methods. In *ECCV '98(1)*, pages 379–393, 1998.
- [9] Andrew V. Goldberg and Robert E. Tarjan. A new approach to the maximum-flow problem. *Journal of the Association for Computing Machinery*, 35(4):921–940, October 1988.
- [10] D. S. Hochbaum. The pseudoflow algorithm for the maximum flow problem. *Manuscript, UC Berkeley*, revised 2003, Extended abstract in: The pseudoflow algorithm and the pseudoflow-based simplex for the maximum flow problem. Proceedings of IPCO98, June 1998. *Lecture Notes in Computer Science*, Bixby, Boyd and Rios-Mercado (Eds.) 1412, Springer, 325-337.

- [11] Olivier Juan and Yuri Boykov. Active graph cuts. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (to appear)*, 2006.
- [12] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *IJCV*, 1(4):321–331, 1988.
- [13] R. Kimmel and A. M. Bruckstein. Regularized laplacian zero crossings as optimal edge integrators. *IJCV*, 53(3):225–243, 2003.
- [14] D. Kirasanov and S. J. Gortler. A discrete global minimization algorithm for continuous variational problems, 2004.
- [15] Pushmeet Kohli and Philip H.S. Torr. Efficiently solving dynamic markov random fields using graph cuts. In *International Conference on Computer Vision*, October 2005.
- [16] Vladimir Kolmogorov. Convergent tree-reweighted message passing for energy minimization. *IEEE transactions on Pattern Analysis and Pattern Recognition (PAMI)*, 2006 (to appear).
- [17] Vladimir Kolmogorov and Yuri Boykov. What metrics can be approximated by geo-cuts, or global optimization of length/area and flux. In *ICCV '05*, page 564, 2005.
- [18] K. N. Kutulakos and S. M. Seitz. A theory of shape by space carving. In *IJCV*, volume 38(3), pages 199–218, 2002.
- [19] Kiriakos N. Kutulakos. Approximate n-view stereo. In *ECCV (1)*, pages 67–83, 2000.
- [20] V. Lempitsky, Y. Boykov, and D. Ivanov. Oriented visibility for multiview reconstruction. In *ECCV'06 (to appear)*, 2006.
- [21] Herve Lombaert, Yiyong Sun, Leo Grady, and Chenyang Xu. A multilevel banded graph cuts method for fast image segmentation. In *International Conference on Computer Vision*, October 2005.
- [22] D. Lowe. Object recognition from local scale invariant features. In *International Conference on Computer Vision*, pages 1150–1157, Corfu, Greece, September 1999.
- [23] R. Malladi, J. A. Sethian, and B. C. Vemuri. Evolutionary fronts for topology-independent shape modeling and recovery. In *ECCV '94*, pages 3–13, 1994.
- [24] Judea Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1988.
- [25] Jean-Philippe Pons, Renaud Keriven, and Olivier Faugeras. Modelling dynamic scenes by registering multi-view image sequences. In *CVPR '05 (2)*, pages 822–827, 2005.
- [26] Carsten Rother, Sanjiv Kumar, Vladimir Kolmogorov, and Andrew Blake. Digital tapestry. In *CVPR '05 (1)*, 2005.

- [27] P. K. Sahoo, S. Soltani, A. K.C. Wong, and Y. C. Chen. A survey of thresholding techniques. *CVGIP*, 41(2):233–260, 1988.
- [28] Steven M. Seitz and Charles R. Dyer. Photorealistic scene reconstruction by voxel coloring. In *CVPR '97*, page 1067, 1997.
- [29] D. Snow, P. Viola, and R. Zabih. Exact voxel occupancy with graph cuts. In *CVPR'00*, pages 345–353, 2000.
- [30] Stefano Soatto, Anthony J. Yezzi, and Hailin Jin. Tales of shape and radiance in multi-view stereo. In *ICCV '03*, 2003.
- [31] Alexander Vasilevskiy and Kaleem Siddiqi. Flux maximizing geometric flows. *IEEE Trans. PAMI*, 24(12):1565–1578, 2002.
- [32] G. Vogiatzis, P. H. S. Torr, and R. Cipolla. Multi-view stereo via volumetric graph-cuts. In *CVPR '05(2)*, 2005.