

Oriented Visibility for Multiview Reconstruction

V. Lempitsky¹, Y. Boykov², and D. Ivanov¹

¹ Department of Mathematics,
Moscow State University,
Moscow, Russia
{vitya, dvi}@fit.com.ru

² Department of Computer Science,
The University of Western Ontario,
London, ON, Canada
yuri@csd.uwo.ca

Abstract. Visibility estimation is arguably the most difficult problem in dense 3D reconstruction from multiple arbitrary views. In this paper, we propose a simple new approach to estimating visibility based on position and orientation of local surface patches. Using our concept of *oriented visibility*, we present a new algorithm for multiview reconstruction based on exact global optimization of surface photoconsistency using graph cuts on a CW-complex. In contrast to many previous methods for 3D reconstruction from arbitrary views, our method does not depend on initialization and is robust to photometrically difficult situations.

Keywords: multiview reconstruction, image-based modeling, visibility, dense stereo, graph cuts, directed graphs, CW-complex, global optimization.

1 Introduction

A multiview reconstruction is a problem of inferring a 3D shape of a scene from a set of its 2D views. In a sequel, we assume that these views are registered within the global world coordinate system, i.e. for each point in the world space, it is possible to determine the coordinates of its projection onto each view.

Recent advances in multiview reconstruction are by far concerned with discrete optimization methods. Such methods (graph cuts [5], belief propagation [12], tree-reweighted message passing [19]) allow efficient minimization of a specific class of energies that can be associated with Markov random fields (MRF). These methods do not require initialization and converge to strong minima of the energy functionals; in particular, graph cuts are able to find a globally optimal labelling for a wide class of binary-labelled MRFs [10]. Due to all these benefits, approaches based on discrete optimization methods are now considered as state-of-the-art for several special cases of multiview reconstruction, namely, reconstruction from a stereo pair [8], from a set of views with similar viewing directions [9, 11], and from a set of views with controlled background [16].

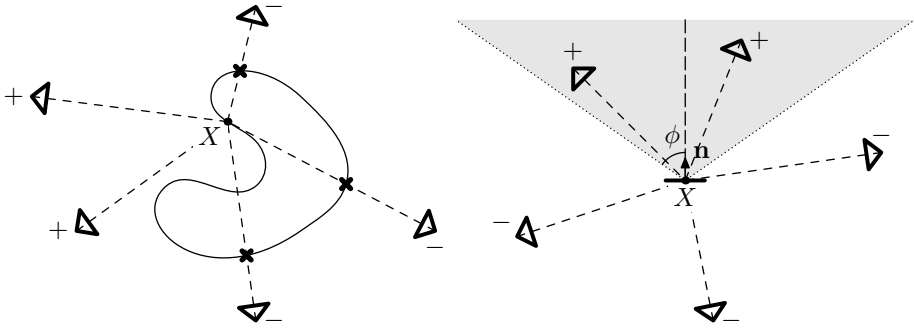


Fig. 1. Two approaches to visibility approximation. Left — *state-based visibility*: a current global scene configuration is used to estimate the visibility for a point X on its surface. Right — *oriented visibility*: a local patch (X, \mathbf{n}) is considered visible from the viewpoints within the predefined angle from the normal direction \mathbf{n} .

In this paper, we consider the most general case of multiview reconstruction, i.e. reconstruction from the views observing the scene from the arbitrarily distributed viewpoints. Only this case allows to infer the complete shape of an object and therefore is the most interesting for many applications. Unfortunately, this case is the most difficult, as any matching process between views has to reason explicitly about *visibility* of different scene parts in different views.

To estimate the true visibility of some surface element one needs to know the true scene geometry and vice versa. To solve this chicken-and-egg problem, it is necessary to use some approximation of visibility. Current approaches (space carving [7], level sets stereo [4]) reconstruct the scene geometry during iterative process, and at each moment of time, a point is considered visible from a viewpoint if it is not occluded with current scene configuration (Fig. 1-left). We call this approach *state-based visibility*, as the visibility is determined by the current state of the scene.

Using iterative optimization results in a significant problem: iterative updates are not guaranteed to converge to the globally optimal configuration. This convergence essentially depends on the initialization and/or on the threshold values. The problem with convergence is worsened by the fact that if the current scene state is far from the true state, state-based visibility approximates the true visibility with significant errors.

The convergence problem could be solved by the application of discrete optimization. Unfortunately, state-based visibility results in an energy function that models interaction between distant scene parts, and discrete optimization methods become really inefficient for such energies with long-range dependence. One possible way to get rid of long-range dependence and to apply discrete optimization is proposed in [15, 17, 18]. There, the reconstruction is initialized with some given-from-aside approximation, and the state-based visibility is calculated based on this approximation. As the true surface is assumed to be close to the initialization, the energy function ignores the visibility changes and is, therefore, short-range dependent and suitable for discrete optimization. Though

using discrete optimization, these approach still require a good initial approximation. Thus, state-based visibility prevents them from taking full advantage of the discrete optimization’ virtues.

In this paper, we propose a reconstruction method employing an alternative approach to visibility estimation. This approach called *oriented visibility* (section 2) is based on the geometric fact that the visibility of a local patch is to a large extent determined by its orientation. Therefore, oriented visibility estimates the visibility of a patch from a viewpoint based on the patch position and normal direction (Fig. 1-right).

This local approach to visibility is in contrast with the traditional global state-based visibility. The main benefit is that the energy functional based on this visibility estimate (formulated in section 3) is amenable for efficient discrete minimization with graph cuts (section 4), which yield its global minimum. From the application standpoint, the key advantage of our method is its ability to find the globally optimal (with respect to the reconstruction functional) scene configuration within given bounding volume without any initialization.

Both state-based visibility (global) and oriented visibility (local) are not exact. We argue, however, that in the situation when no good initialization is given, our reconstruction method based on oriented visibility is a good choice. At the very least, it can be used to supply an initial guess to any of the reconstruction methods relying on state-based visibility; due to their dependency on initialization, this would greatly promote their convergence to the correct scene state. We also briefly discuss an alternative iterative reconstruction method, which fuses state-based and oriented visibilities (section 5). The results of our approach on real and synthetic imagery are demonstrated in section 6, and the discussion of its perspectives in section 7 concludes the paper.

2 Oriented Visibility

To formalize the idea of oriented visibility, we need to introduce some notations. We assume, that the whole scene (or its part we are interested in) is located within some bounding volume \mathcal{B} . Each allowed scene configuration is characterized by some occupied subvolume $M \subset \mathcal{B}$ with the piecewise-smooth oriented boundary ∂M (the *scene surface*).

Let us assume that the scene is observed by N views, taken with pinhole cameras with viewpoints p_1, p_2, \dots, p_N . The positions, orientations, and intrinsic parameters of the cameras are assumed known. Consequently, each point $X \in \mathcal{B}$ can be projected onto each view. Then, let $c_1(X), c_2(X), \dots, c_N(X)$ be the colors of the projections (either grayscale intensities or RGB triples). Let also $\mathbf{v}_1(X), \mathbf{v}_2(X), \dots, \mathbf{v}_N(X)$ be the vectors representing normalized viewing directions from X to p_i :

$$\mathbf{v}_i(X) = \frac{p_i - X}{\|p_i - X\|}. \quad (1)$$

Let (X, \mathbf{n}) denote an infinitesimal *patch* located at point X and having outward looking normal \mathbf{n} . Then the scene surface ∂M can be regarded as a union

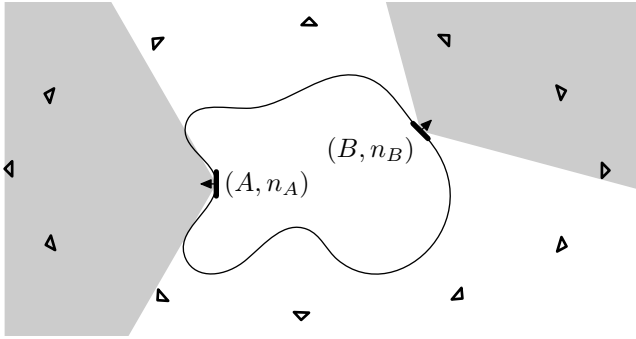


Fig. 2. Oriented visibility in action. Note, that the shown oriented visibility corresponding to the value $\phi = 60^\circ$ is correct for patch (A, \mathbf{n}_A) but underestimates visibility for patch (B, \mathbf{n}_B) .

of patches $(X, \mathbf{n}_{\partial M})$, where X lies on ∂M and $\mathbf{n}_{\partial M}$ is an outward normal to ∂M at X .

Let $\alpha_1(X, \mathbf{n}), \alpha_2(X, \mathbf{n}), \dots, \alpha_M(X, \mathbf{n})$ be the binary variables indicating the visibility of a patch (X, \mathbf{n}) in the corresponding view. Then, orientation-based α_i is calculated as

$$\alpha_i(X, \mathbf{n}) = \begin{cases} 1, & \text{if } \angle(\mathbf{v}_i(X), \mathbf{n}) < \phi \\ 0, & \text{otherwise} \end{cases}, \quad (2)$$

where ϕ is some acute angle (Fig. 1-right).

Thus, our approximation of visibility reflects the fact that the surface element will be always self-occluded for the viewpoints behind it and frequently self-occluded for the viewpoints observing it from oblique angles. The angle ϕ , therefore, determines the threshold of obliqueness, below which the observation from the viewpoint is considered unreliable. Setting ϕ small allows to estimate visibility correctly for concave parts of a scene, while setting it large allows to involve more cameras and hence to increase the discriminative power of our photoconsistency measure (Fig. 2). In our experiments, we found $\phi \approx 60^\circ$ to suit a large variety of scenes and adhered to this value.

3 Energetic Formulation

The goal of this section is to render a multiview reconstruction as an energy minimization problem by assigning an energy cost based on oriented visibility to each scene configuration. As do most other approaches, we assume that the surface of the unknown scene is nearly lambertian, i.e. the color of some small patch on its surface is independent on the viewpoint it is observed from.

Under this assumption, patches belonging to the true surface should have similar colors in the viewpoints observing them (be *photoconsistent*). Therefore, the energy cost $A(X, \mathbf{n})$ of a patch can be defined as:

$$A(X, \mathbf{n}) = \frac{1}{T} \sum_{\substack{i < j \\ \alpha_i(X, \mathbf{n})=1, \alpha_j(X, \mathbf{n})=1}} \|c_i(X) - c_j(X)\|^2, \quad (3)$$

where T is the number of items in the summation. Thus, this cost is the mean of pairwise squared differences between the colors of the projections onto the views observing the patch (L_2 -norm is used for RGB triples). The smaller is this cost, the more photoconsistent is the patch and the more likely does it belong to the surface of the scene.

The overall surface energy cost for the scene configuration $(M, \partial M)$ is then calculated by integrating the patches' costs over the scene surface:

$$\mathbf{E}_{\mathbf{I}}(\partial M) = \iint_{\partial M} A(X, \mathbf{n}_{\partial M}) dS. \quad (4)$$

Now, we have expressed the surface photoconsistency with an energy term $\mathbf{E}_{\mathbf{I}}(\partial M)$. Minimizing $\mathbf{E}_{\mathbf{I}}(\partial M)$ solely is, however, uninteresting as it has an obvious global minimum $(M, \partial M) = \emptyset$ that equals zero. In fact, it has been demonstrated in [7] that in the absence of noise, the scene configuration consistent with a given set of views is not unique, and there is a continuous family of such configurations. Therefore, reconstruction based solely on photoconsistency is an ill-posed problem.

To regularize it, we propose to augment the energy functional with a regularization term $\mathbf{E}_{\mathbf{R}}(M) = \iiint_M B(X) dV$. Here, $B(X)$ is some volume potential corresponding to the prior tendency for point X to belong or not to the reconstruction. E.g., constant negative $B(X)$ produces monotonic ballooning effect biasing the reconstruction process towards larger reconstructions. This simple potential can be used if no prior knowledge is available. It is also possible to introduce boundary conditions in the problem by setting $B(X)$ to large positive or negative values near the boundary. In our experiments, we used the combination of the ballooning potential and the potential encoding the boundary conditions (more details are given in Section 6). An interesting option is to construct $B(X)$ based on the information about background in a way analogous to [16]. Finally, $B(X)$ can encode some prior domain-specific knowledge about the scene geometry.

In conclusion of this section, let us write down the full energy functional guiding the reconstruction:

$$\mathbf{E}(M, \partial M) = \mathbf{E}_{\mathbf{I}}(\partial M) + \mathbf{E}_{\mathbf{R}}(M) = \iint_{\partial M} A(X, \mathbf{n}_{\partial M}) dS + \iiint_M B(X) dV. \quad (5)$$

4 Energy Minimization

4.1 Problem Discretization

To make the minimization of functional (5) tractable, we discretize our problem. The bounding volume \mathcal{B} is subdivided into polyhedral cells R_1, R_2, \dots, R_K

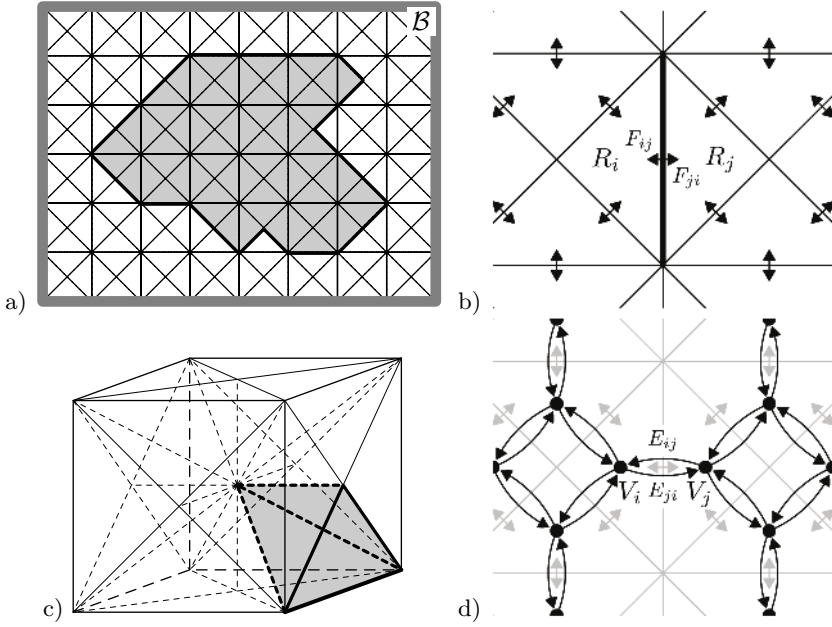


Fig. 3. Discrete optimization of our energy. **a)** A bounding volume \mathcal{B} is discretized into a complex \mathcal{C} . One of the \mathcal{C} -consistent configurations is shaded. **b)** The local structure of our complex. Two adjacent cells R_i and R_j are separated with a pair of oriented faces F_{ij} and F_{ji} . **c)** The cells of our three-dimensional complex produced by voxel subdivision. One of the cells is emphasized. **d)** Local structure of the graph \mathcal{G} dual to the complex from b). The vertices V_i and V_j dual to the cells R_i and R_j are connected with two directed n-links E_{ij} and E_{ji} dual to the faces F_{ij} and F_{ji} (t-links are not shown).

(Fig. 3a). Each pair of neighboring cells R_i and R_j is considered to be separated with a pair of oriented polygonal faces – F_{ij} separating R_j from R_i and F_{ji} separating R_i from R_j (Fig. 3b). We denote an outward looking normal to F_{ij} directed towards R_j as \mathbf{n}_{ij} . We will refer to our discretization structure $\{R_i, F_{ij}\}$ as *complex* (borrowing this term from algebraic geometry where similar structures are called ‘CW-complexes’).

Given a particular complex \mathcal{C} , we may introduce the notion of \mathcal{C} -consistent scene configuration (Fig. 3a). We call the scene configuration $(M, \partial M)$ *consistent with complex \mathcal{C}* (or, simply, *\mathcal{C} -consistent*) if M is composed from the cells of the complex:

$$M = \bigsqcup_{k=1..Q} R_{i_k} . \quad (6)$$

Then, the boundary ∂M consists of oriented faces, separating a cell not belonging to M from a cell within M :

$$\partial M = \bigsqcup_{\substack{i,j \\ R_i \subset M, R_j \subset \overline{M}}} F_{ij} . \quad (7)$$

Assuming that our complex has an appropriate resolution, we can restrict our optimization process to the set of \mathcal{C} -consistent scene configurations. The energy cost $\mathbf{E}(M, \partial M)$ for a \mathcal{C} -consistent configuration can be calculated as:

$$\begin{aligned} \mathbf{E}(M, \partial M) &= \iint_{\partial M} A(X, \mathbf{n}_{\partial M}) dS + \iiint_M B(X) dV = \\ &= \sum_{\substack{i,j \\ R_i \subset M, R_j \subset \overline{M}}} \iint_{F_{ij}} A(X, \mathbf{n}_{ij}) dS + \sum_{A_i \subset M} \iiint_{R_i} B(X) dV = \\ &= \sum_{\substack{i,j \\ R_i \subset M, R_j \subset \overline{M}}} w_{ij} + \sum_{A_i \subset M} w_i, \end{aligned}$$

where $w_{ij} = \iint_{F_{ij}} A(X, \mathbf{n}_{ij}) dS$ is the energy cost for including the oriented face F_{ij} into the scene surface ∂M (*face cost*), and $w_i = \iiint_{R_i} B(X) dV$ is the energy cost for including the cell R_i into the scene M (*cell cost*).

After the numeric computation of these costs, we have a purely discrete optimization problem: find the set of cells with the minimal sum of cell costs and boundary faces costs.

4.2 Graph Cuts Minimization

Although there are so many (2^K) possible \mathcal{C} -consistent configurations, the best \mathcal{C} -consistent configuration can be found in a low order polynomial on K time using graph cuts, recently employed for the optimization of energies in many vision problems (e.g. [1, 14, 2]; the last one having the most similar optimization scheme to ours). Under graph cuts, they usually mean the mincut/maxflow algorithm solving the following problem.

Consider a directed graph with two distinguished vertices (*terminals*) S and T . To each edge between non-terminal vertices (*n-link*), a nonnegative scalar weight is assigned. Edges going to and from terminal vertices (*t-links*) are attributed with arbitrary real weights. A *cut* is a partition of all vertices into two non-intersecting sets called *S-set* and *T-set*, such that the former contains terminal S and the latter contains terminal T . A weight of a cut is by definition the sum of the weights of all edges going from a vertex in *S-set* to a vertex in *T-set* (*cut edges*). Mincut algorithms are able to find the cut with the minimal possible weight (the *minimal cut*) in the time that is low order polynomial on the graph complexity.

To render our problem as a mincut/maxflow problem, we embed a dual graph \mathcal{G} into our complex \mathcal{C} (Fig. 3d). For each cell R_i from \mathcal{C} , \mathcal{G} contains a vertex V_i located in the center of a cell. For each oriented face F_{ij} from \mathcal{C} , \mathcal{G} contains an n-link E_{ij} going from V_i to V_j with the weight w_{ij} . Thus, the direction of this n-link is in accordance with the direction of an outward looking normal \mathbf{n}_{ij} . We also augment \mathcal{G} with two terminal vertices S and T and add t-links E_i^T going from V_i to T having weights w_i .

To any \mathcal{C} -consistent scene configuration $(M, \partial M)$, there corresponds the cut on \mathcal{G} with S -set including the vertices corresponding to the cells within M plus the terminal S .

$$M = \bigsqcup_{k=1..Q} R_{i_k} \quad \longleftrightarrow \quad S\text{-set} = \{V_{i_k} | k = 1..Q\} \cup \{S\}. \quad (8)$$

With such correspondence, any n-link is cut iff the corresponding oriented face is included in the boundary ∂M , and any t-link is cut iff the corresponding cell is included in the scene. Consequently, the weight of a cut always equals the energy cost of a corresponding \mathcal{C} -consistent configuration.

Due to this equality, the minimal cut corresponds to the \mathcal{C} -consistent scene configuration with the minimal energy. The multiview reconstruction is therefore performed as follows. First, construct a complex \mathcal{C} . Second, calculate face costs and cell costs. Third, embed a dual graph \mathcal{G} ; find a minimal cut on \mathcal{G} and the corresponding scene configuration.

4.3 Complex Construction

Let us now consider the choice of an exact structure of the complex \mathcal{C} . Assume that our bounding volume \mathcal{B} is a box, and let us qualitatively analyze the factors that should be taken into consideration while choosing the complex.

The choice of \mathcal{C} determines how "densely" \mathcal{C} -consistent configurations sample the set of all configurations and how close would be the minimal \mathcal{C} -consistent scene configuration to the global minimum over the whole set of scene configurations. Obviously, the smaller is the size of cells and faces, and the larger are their numbers, the richer is the set of \mathcal{C} -consistent configurations. The fineness of resolution is not, however, the only factor to be considered. Due to the dependence of $A(x, \mathbf{n})$ on the orientation, another important matter is how densely the orientations of complex faces sample the set of all possible orientations. Thus, a straightforward but not a proper choice for \mathcal{C} would be a commonly-used rectangular voxel grid. The deficiency of such grid is that it has the oriented faces of only six orientations irrespective of the resolution.

There can be several strategies in constructing complex better than rectangular grid. In our experiments, we first subdivide our bounding box into voxel cubes and then subdivide each cube with six planes, each passing through a pair of opposite cube edges (Fig. 3c). As a result, the voxel is split into 24 tetrahedral cells. This complex has an advantage of having oriented faces with as much as 18 different orientations.

Apart from the tripled number of orientations, another pleasant property of our complex is that the surfaces of scene configurations consistent with it are triangular meshes suitable for immediate storage and rendering. More than that, with such complex, typical reconstructed surfaces are not so jaggy as those composed from voxels and do not require additional smoothening quality-degrading postprocessing like marching cubes.

5 Semi-local Optimization

Oriented visibility gives incorrect visibility estimate in a situations when a part of a scene has an orientation visible from a viewpoint but is occluded by another distant part of a scene (*distant occlusion*). In many cases, however, such distantly occluded areas constitute a relatively small part of the scene surface.

When the accurate reconstruction of distantly occluded parts is required, the result obtained with our method can be used as a starting point for any other reconstruction method relying on iterative optimization and state-based visibility. Alternatively, we developed our own algorithm for a (semi-)local optimization, which searches for a globally optimal configuration within a band around the previous configuration.

Making use of the current state of the scene, this algorithm combines both state-based and oriented approaches to visibility, yielding the following visibility estimate for a patch (X, \mathbf{n}) from the i th viewpoint:

$$\alpha_i(X, \mathbf{n}) = \alpha_i^{or}(X, \mathbf{n}) \cdot \alpha_i^{state}(X), \quad (9)$$

where $\alpha_i^{or}(X, \mathbf{n})$ is the oriented visibility indicator defined in (2), and $\alpha_i^{state}(X)$ is a visibility computed based on the current state of the scene. Thus, the role of $\alpha_i^{state}(X)$ is to detect distant occlusions. In all other aspects, our semi-local optimization procedure is similar to the initial global optimization step.

With such visibility updates, the semi-local optimization can be reapplied several times, each time considering the band around a previous configuration as a novel bounding volume. Changing the thickness of a band allows to trade between the accuracy of distant occlusion estimate on one side and the speed of convergence and the robustness to trapping in local minima on the other. Typically, a few (< 10) iterations is enough to converge.

As the computations are restricted to a narrow band around current configuration, it is also possible to use finer resolution of the complex within the same amount of memory and computation time, thus obtaining more accurate results.

6 Experimental Results

In this section, we present the results of our method on three image sets (Fig. 4). The artificial *solids* setup contains virtual objects “hanging” in the air. Its main challenges are fine texture details paired with a uniform background. For many photoconsistency-based algorithms such combination results in numerous “floating” artifacts. The real *camel* and *matreshkas* setups contain objects placed on the ground table/plane. The position of this plane as well as camera parameters using structure-and-motion methods. The *matreshkas* setup is particularly difficult for reconstruction, as the varnished surfaces are highly non-lambertian.



Fig. 4. Samples of the source imagery in our experiments. Image sets comprised 16–20 views surrounding the scene from all accessible positions. Left — *solids* setup was rendered artificially using POV-Raytm [13]. Middle and right — *camel* and *matreshkas* setups were taken using consumer digital camera.

The following volume potentials expressing the prior geometric knowledge were used for reconstruction (B_1 for *solids*, B_2 for *camel* and *matreshkas*):

$$B_1(X) = \begin{cases} +\infty & \text{near the bounding box boundary ,} \\ \beta & \text{inside bounding box .} \end{cases}$$

$$B_2(X) = \begin{cases} -\infty & \text{below ground plane ,} \\ +\infty & \text{near the bounding box boundary above ground plane ,} \\ \beta & \text{inside bounding box .} \end{cases}$$

Here, β is some small negative value introducing a slight ballooning effect, which was kept constant throughout our experiments. Infinities in the volume potential ensure the closeness of the recovered scene for *solids* and the "object on the ground" topology for *camel* and *matreshkas*.

For the evaluation purposes, we implemented an improved version [3] of the popular space carving approach [7]. Since space carving is very sensitive to the selection of photoconsistency threshold, we did our best while selecting the optimal threshold for each setup. However, as demonstrates Fig. 5a, our setups were too photometrically difficult for space carving.

The results of our method are presented on the rest of Fig. 5. Note, that despite significant distant occlusions, global optimization based on the oriented visibility solely (Fig. 5-middle) produced generally correct reconstructions for *solids* and *camel* setups. This suggests that the ability to find a global minimum often justifies the use of inexact visibility estimate.

In our experiments, the complex \mathcal{C} comprised upto 20 million of cells resulting in the same number of vertices in the dual graph \mathcal{G} . Therefore the

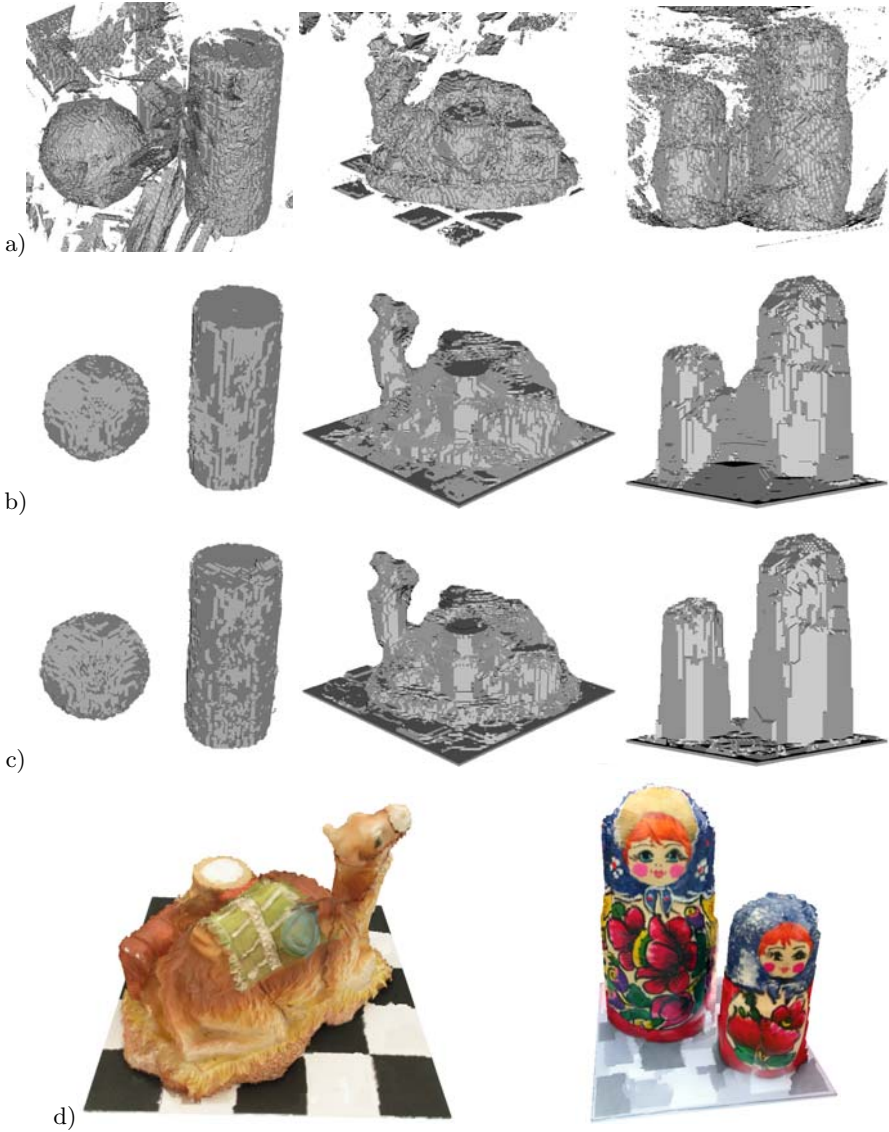


Fig. 5. The reconstruction results. **a)** The results of space carving algorithm. **b)** The results of the proposed algorithm after global optimization. **c)** The results after three steps of subsequent semi-local optimization. **d)** Renderings of image based textured models created from real imagery with our method.

mincut/maxflow algorithm was the computational bottleneck for our approach. Performing global reconstruction for such resolution demanded upto 20 minutes on a P4-2.6GHz computer. Subsequent semi-local updates took few minutes.

7 Discussion

In this paper, we have proposed a novel orientation-based approach to visibility estimation. Such purely local visibility estimate allows us to cast a multiview reconstruction problem as an optimization of a novel energy functional amenable for minimization with graph cuts.

Our main advantage over other methods, which rely on state-based visibility and on iterative updates, is the independence from the initialization due to the ability of our optimization to yield a global minimum of the energy functional. The result produced with our global optimization can be improved with our semi-local optimization combining oriented and state-based approaches to visibility. Alternatively, it can be used as a good starting point for any of the other reconstruction methods.

The main limitation of our method in its current implementation is its computational demands. To deal with this problem, one can consider complexes based on spatially non-uniform sampling of the reconstruction space. Non-uniform sampling can be driven by some domain-specific knowledge or the cues resulting from another reconstruction algorithm. In both cases, uncertain knowledge about position and orientation of the surface may be used to include faces in the complex at some positions and with some orientations more frequently than others. An interesting option for implementing this are random grids arranged in BSP-trees proposed for discrete minimal surface search in [6].

The second deficiency of our method is the difficulties it faces while recovering the protruding parts of objects (e.g. camel ears). This problem is, however, inherent to all minimal surface methods, since they minimize the integral of non-negative energy function over a scene surface.

Another prospect for future investigation is concerned with the fact that our patch energy cost $A(X, \mathbf{n})$ accounts for both position and orientation of a patch. This can allow to use different shading models in our method. Thus, we can consider the reconstruction based on non-lambertian reflectivity models (e.g. Phong model) or shape-from-shading reconstruction.

References

1. Y. Boykov and V. Kolmogorov. An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision. *PAMI*, vol. 26, no. 9, pp. 1124-1137, Sept. 2004.
2. C. Buehler, S. J. Gortler, M. Cohen, and L. McMillan. Minimal Surfaces for Stereo Vision. *Proc. ECCV 2002*, vol. 3, pp. 885-899.
3. W. Culbertson, T. Malzbender, and G. Slabaugh. Generalized Voxel Coloring. *Workshop on Vision Algorithms'1999*, pp. 100-115
4. O. Faugeras and R. Keriven. Complete Dense Stereovision Using Level Set Methods. *Proc. ECCV 1998*, vol. 1, pp. 379-393.
5. D. Greig, B. Porteous, and A. Seheult. Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society*, 51(2):271-279, 1989.
6. D. Kirsanov and S. J. Gortler. A Discrete Global Minimization Algorithm for Continuous Variational Problems. Harvard CS Technical Report TR-14-04, 2004.

7. K. N. Kutulakos and S. M. Seitz. A Theory of Shape by Space Carving. *IJCV*, 2000, 38(3), pp. 199–218.
8. Middlebury Stereo Vision Page. <http://cat.middlebury.edu/stereo/>
9. V. Kolmogorov and R. Zabih. Multi-camera Scene Reconstruction via Graph Cuts. *Proc. ECCV 2002*.
10. Vladimir Kolmogorov, Ramin Zabih. What Energy Functions Can Be Minimized via Graph Cuts? *Proc. ECCV'2002*, vol. 3, pp. 65–81
11. S. Paris, F. Sillion, and L. Quan. A Surface Reconstruction Method Using Global Graph Cut Optimization. *IJCV* (*to appear*).
12. J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of plausible Inference*. San Francisco, CA: Morgan Kaufmann, 1988.
13. POV-Raytm raytracer. <http://www.povray.org>
14. S. Roy and I. Cox. A maximum-flow formulation of the n-camera stereo correspondence problem. *Proc. ICCV 1998*, pp. 492–499.
15. S. Sinha and M. Pollefeys. Multi-View Reconstruction Using Photo-consistency and Exact Silhouette Constraints: A Maximum-Flow Formulation. *Proc. ICCV 2005*, pp. 349–356.
16. D. Snow, P. Viola, and R. Zabih. Exact voxel occupancy with graph cuts. *Proc. CVPR 2000*, vol. 1, pp. 345–352.
17. G. Vogiatzis, P. H. S. Torr, S. Seitz, and R. Cipolla. Reconstructing Relief Surfaces. *Proc. BMVC 2004*, pp. 117–126.
18. G. Vogiatzis, P. H. S. Torr, and R. Cipolla. Multi-view stereo via Volumetric Graph-cuts. *Proc. CVPR 2005*, pp. 391–398.
19. M. J. Wainwright, T. S. Jaakkola, A. S. Willsky. MAP estimation via agreement on trees: Message-passing and linear programming. *IEEE Trans. on Information Theory*, vol. 51, no. 11, 2005.