

# Integration of Multiview Stereo and Silhouettes Via Convex Functionals on Convex Domains

Kalin Kolev and Daniel Cremers

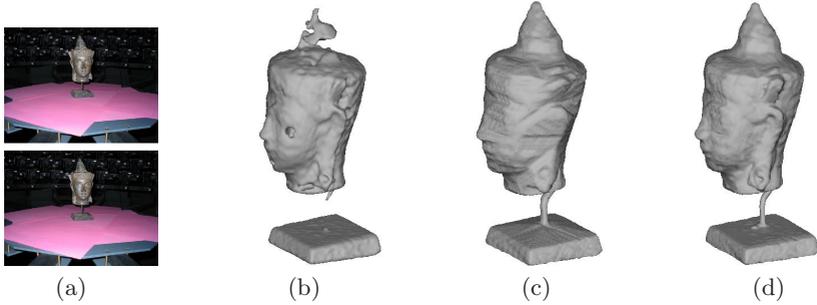
Department of Computer Science,  
University of Bonn, Germany  
{kolev,dcremers}@cs.uni-bonn.de

**Abstract.** We propose a convex framework for silhouette and stereo fusion in 3D reconstruction from multiple images. The key idea is to show that the reconstruction problem can be cast as one of minimizing a *convex* functional where the exact silhouette consistency is imposed as a convex constraint that restricts the domain of admissible functions. As a consequence, we can retain the original stereo-weighted surface area as a cost functional without heuristic modifications by balloon terms or other strategies, yet still obtain meaningful (nonempty) global minimizers. Compared to previous methods, the introduced approach does not depend on initialization and leads to a more robust numerical scheme by removing the bias near the visual hull boundary. We propose an efficient parallel implementation of this convex optimization problem on a graphics card. Based on a photoconsistency map and a set of image silhouettes we are therefore able to compute highly-accurate and silhouette-consistent reconstructions for challenging real-world data sets in less than one minute.

## 1 Introduction

Recovering three-dimensional geometrical structure from a series of calibrated images is among the fundamental problems in computer vision, with numerous applications in computer graphics, augmented reality, robot navigation and tracking. Among the multitude of existing methods for multiview reconstruction one can identify two major classes of approaches: shape from silhouettes and shape from stereo.

Historically, the first strategy for multiview 3D shape retrieval, dating back to the 1970's, has been to use the outlines of the imaged objects [1]. Most of these *shape from silhouettes* approaches aim at approximating the visual hull [2] of the observed solid. The visual hull is an outer approximation, constructed as the intersection of the visual cones associated with all image silhouettes. In the course of research, different shape representations have been proposed: volumetric [3], surface-based [4] and polyhedral [5]. Apart from shape representation, research has been also focused on the development of methods operating on raw image data instead of predetermined silhouettes. Most of them are based on an energy minimization framework allowing to impose regularization in the labeling process



**Fig. 1.** Silhouette and stereo integration. (a) Two of the input images. While stereo-based approaches (b) recover concavities but overcarve thin structures and areas of specular reflections, silhouette-based methods (c) reconstruct small-scale details captured by the silhouettes but fill indentations. In contrast, techniques for silhouette and stereo fusion (d) restore concave areas as well as fine geometric details.

[6]. The segmentation of each image is obtained through the evolution of a single surface in 3D rather than separate contours in 2D. As a result, such approaches exhibit considerable robustness to image noise and erroneous camera calibration.

The main drawback of silhouette-based approaches is their inability to reconstruct concavities, since these do not affect the silhouettes. *Stereo-based methods* capture such indentations by measuring photoconsistency of surface patches in space. The fundamental idea is that under the Lambertian assumption only points on the object's surface have a consistent appearance in the input images, while all other points project to incompatible image patches. The earliest algorithms use carving techniques to obtain a volumetric representation of the scene by repeatedly eroding inconsistent voxels [7]. They do not enforce the smoothness of the surface and this often results in rather noisy reconstructions. This drawback was overcome by energy minimization techniques [8,9,10], which typically aim at computing a weighted minimal surface, where the weights reflect the local photoconsistency.

Ideally, one would therefore like to combine both types of techniques in order to jointly achieve stereo and silhouette consistency (see Fig. 1). A simple strategy to fuse these complementary features is to use a visual hull (computed from silhouettes) as initialization for a stereo-based approach [11,12]. Firstly, this requires to constrain the solution space to avoid the empty set, secondly, the resulting reconstruction will generally not fulfill the silhouette constraint. Alternatively, one can unify both information sources in a single formulation. Two different techniques have been proposed to achieve this goal: One can integrate in the evolution silhouette-aligning forces [13,14,15,16], or one can use predetermined surface points [17,18,19] to impose exact silhouette constraints. Both strategies have their shortcomings. The first one could lead to a numerically unstable behavior and could introduce a bias near the visual hull boundary, while the second one requires premature decisions about voxel occupancy. To address these drawbacks [20] proposed a graph cut framework for silhouette and stereo

fusion. Unfortunately, the practical applicability of this method is limited due to its high memory requirements, which poses a severe restriction on the volume resolution. As a consequence, the development of robust and efficient schemes for silhouette and stereo integration remains an open challenge.

In this paper, we propose a novel mathematical framework for silhouette and stereo fusion in 3D reconstruction. The idea is to cast multiview stereovision as a *convex* variational problem where the exact silhouette consistency is imposed as a convex constraint that restricts the domain of admissible functions. Silhouette-consistent reconstructions are computed by convex relaxation, finding global minimizers of the relaxed problem and subsequent projection to the original non-convex set. Compared to existing fusion techniques, we thus compute guaranteed silhouette-consistent reconstructions without constraining the search space and without extending the original stereo-weighted cost functional by heuristic ballooning terms or more sophisticated balancing terms. Compared to classical local fusion techniques, the proposed formulation does not depend on initialization and leads to a more tractable numerical scheme by removing the bias near the visual hull boundary. In experiments on several challenging real data sets we show the advantages of silhouette-consistency in the reconstruction of small-scale structures which cannot be restored by state-of-the-art stereo algorithms.

In the next section, we will briefly review the formulation of stereo-based multiview reconstruction as a weighted minimal surface problem. In Section 3 we will show that the integration of stereo and silhouette constraints can be formulated as a problem of minimizing a convex functional over the convex set of silhouette-consistent functions. In Section 4 we provide details on the numerical implementation of the constrained convex optimization. In Section 5, we show experimental results on several real data sets which demonstrate the advantages of silhouette consistency for the reconstruction of fine-scale structures. We conclude with a brief summary.

## 2 3D Reconstruction as a Minimal Surface Problem

Let  $V \subset \mathbb{R}^3$  be a volume, which contains the scene of interest, and  $I_1, \dots, I_n : \Omega \rightarrow \mathbb{R}^3$  a collection of calibrated color images with perspective projections  $\pi_1, \dots, \pi_n$ . Let  $S_1, \dots, S_n \subset \Omega$  be the observed projections of the 3D object and  $\rho : V \rightarrow [0, 1]$  be a photoconsistency map measuring the discrepancy among various image projections. In particular, low values of  $\rho(x)$  indicate a strong agreement from different cameras on the observed image patches, indicating a high likelihood that the surface passes through the given point. More details on the computation of photoconsistency will be given in Section 4.1.

With the above definitions, multiview reconstruction can be done by minimizing the classical energy [8]:

$$E(S) = \int_S \rho(x) dS. \quad (1)$$

The reconstruction is therefore given by a minimal surface measured in a Riemannian metric that favors boundaries along photoconsistent locations. While

local optimization techniques (using coarse-to-fine strategies) provide useful reconstructions, there is little guarantee regarding optimality of the solutions. In fact, the question of optimality is somewhat meaningless, as the global minimum of (1) is obviously the empty set. A remedy to this problem is to either constrain the search space around the visual hull [12], or to add regional balancing terms to the cost functional using balloon forces [10] or heuristically constructed regional terms [21,22,23]. Global optima of respective cost functionals can then be computed either in a spatially discrete setting using graph cuts [12,10,23] or in a spatially continuous setting using convex relaxation techniques [22].

Nevertheless, two limitations of such methods are that firstly, the balancing regional terms are typically based on a number of somewhat heuristic assumptions and can often introduce a bias in the resulting segmentation. Secondly, the resulting reconstructions are not guaranteed to be silhouette-consistent in the sense that the projections of the surface do not necessarily coincide with the observed silhouettes.

### 3 Convex Integration of Silhouettes and Stereo

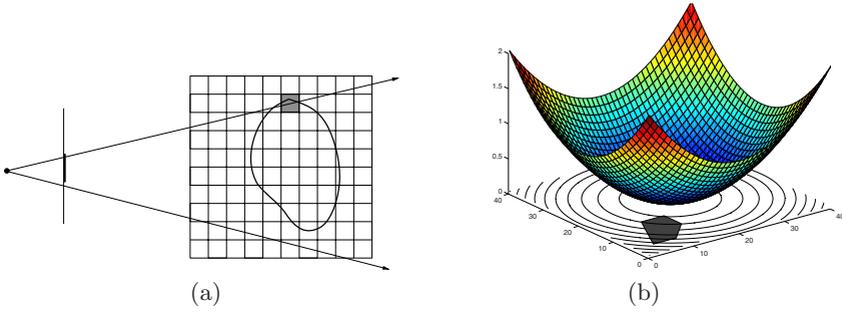
An alternative strategy to avoid trivial solution in the optimization of (1) is to impose silhouette alignment of the computed shape yielding the following constrained optimization problem:

$$\begin{aligned}
 E(S) &= \int_S \rho(x) dS, \\
 \text{s. t.} \quad \pi_i(S) &= S_i \quad \forall i = 1, \dots, n.
 \end{aligned}
 \tag{2}$$

In order to cast (2) as a convex optimization problem, the surface  $S$  is represented implicitly by the characteristic function  $u : V \rightarrow \{0, 1\}$  of the surface interior  $S_{int}$ . Hence, changes in the topology of  $S$  are handled automatically without reparametrization. With the implicit surface representation, we have the following constrained, convex energy functional equivalent to (2):

$$\begin{aligned}
 E(u) &= \int_V \rho(x) |\nabla u(x)| dx \\
 \text{s. t.} \quad u &\in \{0, 1\} \\
 \int_{R_{ij}} u(x) dR_{ij} &\geq \delta \text{ if } j \in S_i \\
 \int_{R_{ij}} u(x) dR_{ij} &= 0 \text{ if } j \notin S_i,
 \end{aligned}
 \tag{3}$$

where  $R_{ij}$  denotes the visual ray through pixel  $j$  of image  $i$  and  $\delta > 0$  denotes the thickness, below which the given material becomes translucent. In the following we will set  $\delta = 1$ . Here we have rewritten the constraint in (2) in following form: For a silhouette-consistent shape *at least one* of the voxels along a visual ray



**Fig. 2.** Silhouette constraints. (a) For a silhouette consistent shape at least one voxel along a visual ray through a silhouette pixel is occupied, whereas all voxels along a ray through a non-silhouette pixel are empty. The bold area on the image plane indicates the outlines of the observed object and the shaded voxel is an occupied one along the given viewing ray. (b) Multiview stereovision can be formulated as a convex optimization problem, where silhouette constraints determine a convex domain (shaded area) of admissible functions. Hence, global minimization is possible by using classical techniques like gradient descent.

through a silhouette pixel should be occupied, whereas *all* voxels along a ray determined by a non-silhouette pixel should be empty (see Fig. 2(a)).<sup>1</sup>

Due to the constraint that  $u$  is a binary-valued function, the minimization problem (3) is non-convex (because the space of binary functions is non-convex). By relaxing the binary constraint and allowing the function  $u$  to take on values in the interval  $[0, 1]$ , the optimization problem becomes that of minimizing a convex functional over a bounded convex set (see Fig. 2(b)):

$$\min_{u \in D} \int_V \rho(x) |\nabla u(x)| dx, \tag{4}$$

where

$$D := \left\{ u : V \rightarrow [0, 1] \mid \begin{cases} \int_{R_{ij}} u(x) dR_{ij} \geq 1 & \text{if } j \in S_i \forall i, j \\ \int_{R_{ij}} u(x) dR_{ij} = 0 & \text{if } j \notin S_i \forall i, j \end{cases} \right\} \tag{5}$$

is the set of continuous valued functions  $u$  which fulfill the silhouette constraints for all images  $i$  and all rays  $j$ .

**Proposition 1.** *The set  $D$  of all silhouette-consistent functions defined in (5) forms a compact and convex set.*

*Proof.*  $D$  is obviously compact, since it is determined by multiple restricting inequalities.

In order to show the convexity, let  $u_1, u_2 \in D$  be two elements of  $D$ . Then any convex combination  $u = \alpha u_1 + (1 - \alpha)u_2$  with  $\alpha \in [0, 1]$  is also an element in  $D$ . In particular,  $u(x) \in [0, 1]$  for all  $x$ . Moreover,

<sup>1</sup> Note that in case of imperfect silhouettes the above constraints can be applied only to the areas of high confidence.

$$\int_{R_{ij}} u \, dR_{ij} = \alpha \int_{R_{ij}} u_1 \, dR_{ij} + (1 - \alpha) \int_{R_{ij}} u_2 \, dR_{ij} \geq 1 \quad \text{if } j \in S_i,$$

and similarly

$$\int_{R_{ij}} u \, dR_{ij} = \alpha \int_{R_{ij}} u_1 \, dR_{ij} + (1 - \alpha) \int_{R_{ij}} u_2 \, dR_{ij} = 0 \quad \text{if } j \notin S_i.$$

Thus  $u \in D$ . □

The above statement implies that a global minimum  $u^*$  of the relaxed problem (4) exists and can be computed, for example by a simple gradient descent procedure or by more efficient numerical schemes. A necessary condition for a minimum of (4) is stated by the associated Euler-Lagrange equation

$$0 = \operatorname{div} \left( \rho \frac{\nabla u}{|\nabla u|} \right) = \rho \operatorname{div} \left( \frac{\nabla u}{|\nabla u|} \right) + \langle \nabla \rho, \frac{\nabla u}{|\nabla u|} \rangle. \tag{6}$$

A numerical solution to this partial differential equation within the domain of admissible functions specified by the convex constraints in (4) will be detailed in Section 4.

Since we are interested in minimizers of the non-convex binary labeling problem (3), a straightforward methodology is to threshold the solution of the convex problem appropriately. Although this will not guarantee finding the global minimum of (3), the proposed strategy entails a series of advantages compared to classical local optimization techniques. Intuitively, extending the set of admissible functions, computing the global minimum over this domain and subsequently projecting to the nearest point within the original set is expected to give a more accurate estimate than a simple gradient descent procedure for smooth functionals. In particular, this approach always gives an upper bound for the energetic deviation of the computed solution from the global minimum.

**Proposition 2.** *Let  $u^*$  be a minimizer of (4) and let  $D' \subset D$  be the set of binary silhouette-consistent functions. Let  $u' \in D'$  be the (global) minimum of (3) and  $\tilde{u}$  the solution obtained with the above procedure. Then, a bound  $\gamma(u^*, \tilde{u})$  exists such that*

$$E(\tilde{u}) - E(u') \leq \gamma(u^*, \tilde{u}).$$

*Proof.* The claim follows directly from the construction of the approach:

$$E(\tilde{u}) - E(u') \leq E(\tilde{u}) - E(u^*) =: \gamma(u^*, \tilde{u}).$$

The inequality is due to  $E(u^*) \leq E(u')$ , since  $D' \subset D$ . □

The projection  $\tilde{u} \in D'$  of a minimizer  $u^*$  onto  $D'$  can be computed by simple thresholding

$$\tilde{u}(x) = \begin{cases} 1, & \text{if } u^*(x) \geq \mu \\ 0, & \text{otherwise} \end{cases}, \tag{7}$$

where

$$\mu = \min \left\{ \left( \min_{i \in \{1, \dots, n\}, j \in S_i} \max_{x \in R_{ij}} u^*(x) \right), 0.5 \right\}. \quad (8)$$

This threshold  $\mu$  provides the closest silhouette-consistent binary function to the solution of the relaxed problem.

**Proposition 3.** *The reconstructed surface exactly fulfills all silhouette constraints, i.e.  $\tilde{u} \in D'$ .*

*Proof.* Let  $R_{pq}$  be a given ray. For  $q \notin S_p$  the silhouette constraint is fulfilled for any threshold  $\mu \in (0, 1)$ , since the labels  $\tilde{u}(x)$  of all voxels  $x$  along the respective ray are 0. For  $q \in S_p$ , we have:

$$\mu \leq \min_{i \in \{1, \dots, n\}, j \in S_i} \max_{x \in R_{ij}} u^*(x) \leq \max_{x \in R_{pq}} u^*(x).$$

This implies  $\exists x \in R_{pq} : u^*(x) \geq \mu$  and hence  $\exists x \in R_{pq} : \tilde{u}(x) = 1$ .  $\square$

Thus, the proposed methodology has the following advantages:

- It allows to incorporate *exact* silhouette constraints without making premature hard decisions about voxel occupancy along each viewing ray passing through a silhouette pixel.
- It does not depend on initialization, since the relaxed functional is optimized globally.
- It leads to a simple and tractable numerical scheme, which does not rely on a locally estimated surface orientation, and thus does not introduce a bias near the visual hull boundary.

All these benefits will be investigated in the experimental section.

## 4 Implementation

This section will give more details on the implementation of the proposed approach.

### 4.1 Photoconsistency Estimation

In this paper, we propose a novel strategy for integration of silhouette and stereo information. The presented method operates on a precomputed photoconsistency map  $\rho : V \rightarrow [0, 1]$  and is independent of its particular implementation. To validate its concept, we used the voting scheme proposed in [14] for photoconsistency computation. The choice of this technique was motivated by its robustness even without explicit visibility estimation and increased accuracy compared to traditional methods. See [14] for more details.

### 4.2 Constraint Realization

The minimization of (4) should be performed within the specified domain of admissible functions. To this end, one has to enforce the fulfillment of all constraints during the optimization process. A straightforward way to achieve this is to project the current estimate after each iteration to the next point in the restricted domain. For the first constraint in (4) this corresponds to just clipping values lying outside of the interval  $[0, 1]$ . The last constraint could be realized by starting with the visual hull as initialization (1 if the voxel is part of the visual hull and 0 otherwise) and keeping function values fixed outside of it. Since (4) is optimized globally, the initial guess has no impact on the result but only on the number of iterations needed until convergence. However, the second constraint in (4) requires more efforts. It states that the sum of the function values along each visual ray passing through a silhouette pixel should be at least 1. If this requirement is violated, the values of all voxels along the ray lying within the visual hull should be uniformly increased. Note that enforcing this constraint in a different order will generally produce a different result. However, when the evolution step is small enough, this issue is not crucial and can be ignored. In particular, the realization of this constraint can be done in parallel. In order to avoid computations of ray-volume intersections any time the constraint is checked, one can compute the set of relevant voxels to each viewing ray in a pre-processing step and store them in lists. However, the size of this data structure could grow significantly when the resolution of input images is high. For this reason, in our implementation we stored only the first and last voxel along each ray. Another important issue when using constraints is the frequency of enforcing them. In our implementation, we achieved a stable behavior when applying the first constraint after each optimization iteration and the silhouette constraints after each 10 iterations.

### 4.3 Linearization and Fixed-Point Iteration

In order to solve (6), we suggest to use a fixed point iteration scheme that transforms the nonlinear system into a sequence of linear systems. These can be efficiently solved with an iterative solver like successive over-relaxation (SOR).

The only source of nonlinearity in (6) is the diffusivity  $g := \frac{\rho}{|\nabla u|}$ . Starting with an initialization  $u^0$ , we can compute  $g$  and keep it constant. For constant  $g$ , (6) is linear and discretization yields a linear system of equations, which we solve with the SOR method. This means, we iteratively compute an update of  $u$  at voxel  $i$  by

$$u_i^{l,k+1} = (1 - \omega)u_i^{l,k} + \omega \frac{\sum_{j \in \mathcal{N}(i), j < i} g_{i \sim j}^l u_j^{l,k+1} + \sum_{j \in \mathcal{N}(i), j > i} g_{i \sim j}^l u_j^{l,k}}{\sum_{j \in \mathcal{N}(i)} g_{i \sim j}^l}, \tag{9}$$

where  $\mathcal{N}(i)$  denotes the 6-neighborhood of  $i$ . Finally,  $g_{i \sim j}$  denotes the diffusivity between voxel  $i$  and its neighbor  $j$ . It is defined as

$$g_{i \sim j}^l := \frac{g_i^l + g_j^l}{2}, \quad g_i^l := \frac{\rho_i}{\sqrt{|\nabla u_i^l|^2 + \epsilon^2}}, \quad (10)$$

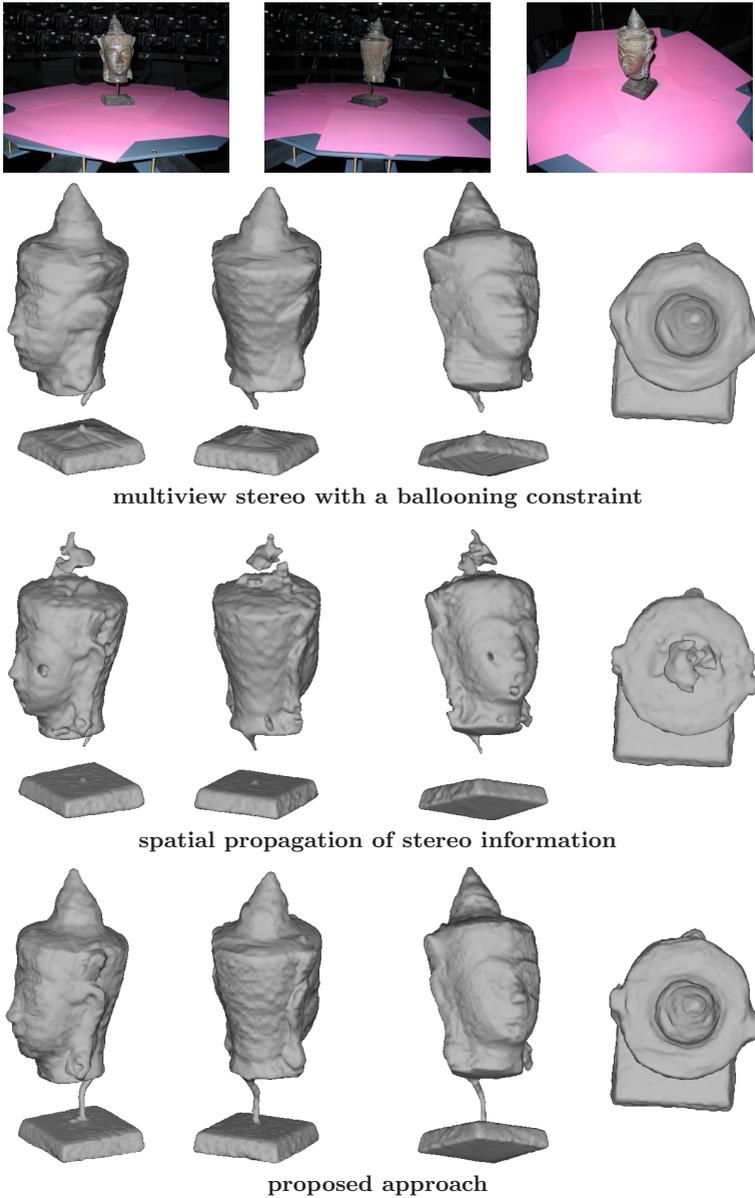
where  $\epsilon := 0.001$  is a small constant that prevents the diffusivity to become infinite when  $|\nabla u_i^l| = 0$  and  $|\nabla u_i^l|^2$  is approximated by standard central differences. The over-relaxation parameter  $\omega$  has to be chosen in the interval  $(0, 2)$  for the method to converge. The optimal value depends on the linear system to be solved. Empirically, for the specific problem at hand, we obtained a stable behavior for  $\omega = 1.5$ . After the linear solver yields a sufficiently good approximation (we iterated for  $k = 1, \dots, 10$ ), one can update the diffusivities and solve the next linear system. Iterations are stopped as soon as the energy decay in one iteration is in the area of number precision.

## 5 Experimental Results

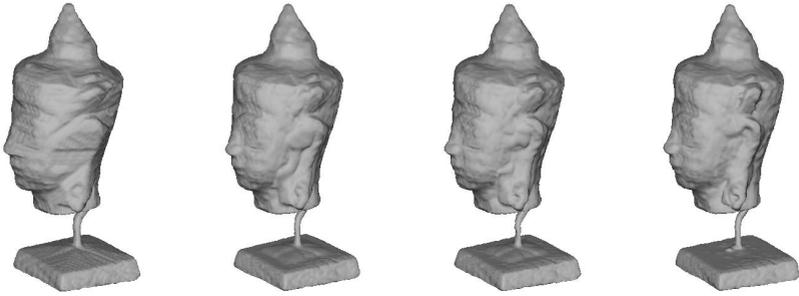
We validate the proposed approach on a scene of a head statue with complex reflection properties containing thin structures (the pedestal); see Fig. 3<sup>2</sup>. Scenes of this type are a known challenge for variational stereo-based methods due to the violation of the Lambertian assumption and the presence of a regularizer, which introduces a bias towards shapes with small area. In particular, we implemented two classical paradigms in multiview stereovision: a weighted minimal surface formulation with a ballooning constraint [12] and a stereo propagating scheme [21,22]. The first method produces clear oversmoothing effects at concavities and small-scale structures. The second approach retrieves shape indentations but also leads to erroneous carving at thin parts and specularities. In contrast, the introduced technique produces accurate reconstructions of thin structures (the pedestal) as well as concave areas by incorporating silhouette constraints in the optimization process. Note that all three models are based on a classical minimal surface formulation but use different methodologies to avoid the empty surface as a solution. Note also that all three methods use silhouette information to restrict the ballooning, for initialization or to constrain the domain of admissible functions. Fig. 4 shows intermediate steps in the evolution process of the proposed approach. Although global minimization is used, which makes the choice of the initial guess irrelevant, we initialized with the visual hull to emphasize the basic difference of our method to classical local optimization procedures. Usually, local minimization techniques use the surface orientation to identify locations to deform the current shape in order to minimize the resulting reprojection error. However, this could lead to instabilities and introduce a bias near the visual hull boundary by involving surface points beyond the contour generator. In contrast, the introduced method recovers shape indentations effortlessly, while retaining silhouette alignment during the optimization process.

Fig. 5 shows a comparison between the proposed method and the approach of [19] on an image sequence of a statue of a Greek goddess. Note the visible

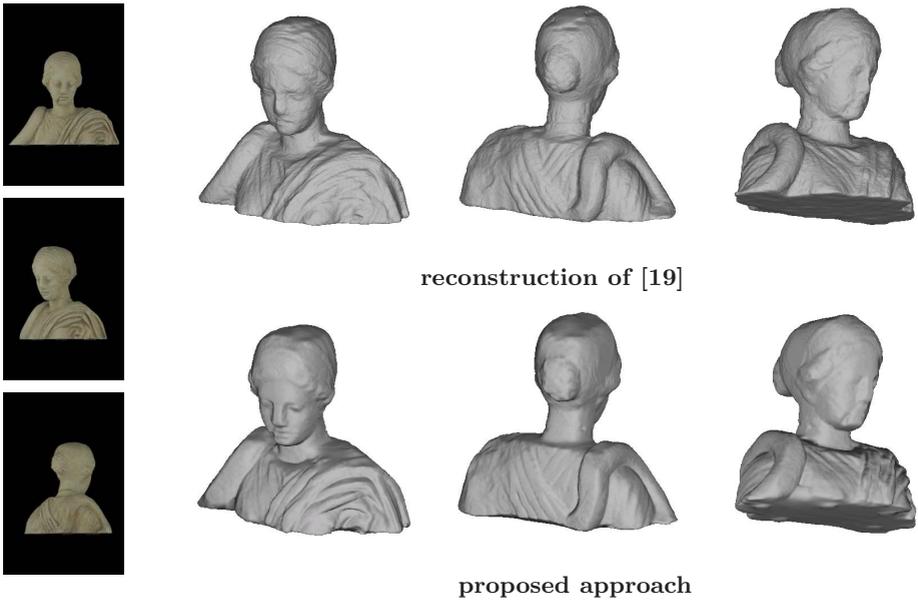
<sup>2</sup> The image sequences used in Figures 3 and 7 will be made available at <http://www-cvpr.iai.uni-bonn.de/data/>



**Fig. 3.** Head sequence. *First row:* 3 of 33 input images of resolution  $1024 \times 768$ . *Second row:* Multiple views of the reconstruction with a model based on the combination of multiview stereo with a ballooning constraint [12]. *Third row:* Multiple views of the reconstruction obtained with a more elaborate model based on spatial propagation of stereo information [21,22]. *Fourth row:* Multiple views of the reconstruction with the proposed approach. While both methods fail to recover the pedestal of the statue due to oversmoothing effects or erroneous carving, the proposed approach recovers accurately all relevant details.



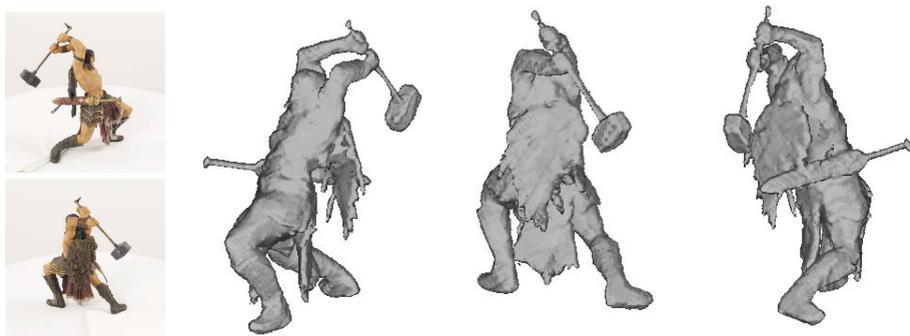
**Fig. 4.** Minimization process. Surface evolution starting from the visual hull, obtained by projecting the current estimate onto the original domain. Note that the presented method is able to generate accurate shapes starting from this initialization, since it does not take the local surface orientation into account.



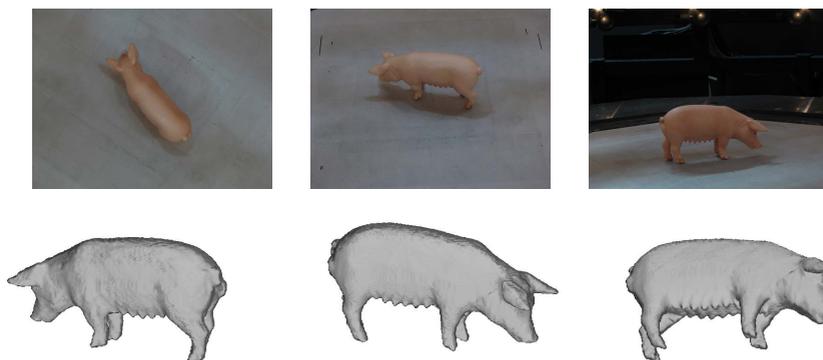
**Fig. 5.** Hygia sequence. 3 of 36 input images of resolution  $2088 \times 3040$  and multiple views of the reconstructed surface compared to the reconstruction of [19]. Our result exhibits a higher grade of smoothness, while recovering surface details more accurately (for example the face and the creases of the cloth). Note that even the legs of the statue are reconstructed.

improvements of our reconstruction in the area of the face and the creases of the cloth. Note that even the legs of the statue are recovered. On the other hand, however, image noise is suppressed by increasing the grade of surface smoothness.

Fig. 6 and 7 illustrate two additional challenging image sequences. The first one is publicly available and captures a figurine of an ancient warrior. See



**Fig. 6.** Warrior sequence. 2 of 24 input images of resolution  $1600 \times 1600$  and multiple views of the reconstructed surface. Note that thin structures (for example the handle of the hammer) as well as concavities (for example at the chest) are reconstructed accurately.



**Fig. 7.** Sow sequence. *First row:* 3 of 27 input images of resolution  $1024 \times 768$ . *Second row:* Multiple views of the reconstructed surface. Note the accurately reconstructed tits.

<http://www-cvr.ai.uiuc.edu/~yfurukaw/research/mview/index.html> for the data set including a reconstruction with the approach of [15]. Our result exhibits a high grade of smoothness, while preserving all fine geometric details. Similarly, the proposed approach generates a high-quality reconstruction of the sow figurine in Fig. 7 (see the accurately recovered tits). It is important to note that the absence of texture results in a minimal surface fulfilling silhouette consistency being generated, which allows to restore also homogeneous objects.

Apart from robustness, another crucial issue for a silhouette and stereo integration approach is the computational time needed. To this end, we used a GPU implementation of the presented method, where the SOR optimization in a red-black strategy as well as the imposed silhouette constraints run on the GPU. On a PC with 2.8 GHz and 4 GB of main memory, equipped with a NVIDIA

GeForce 8800 GTX graphics card, we measured computational times in the range of 30-60 seconds for all demonstrated experiments. Note that photoconsistency estimation is not included in these runtimes.

## 6 Conclusion

We proposed a novel framework for integrating silhouette and stereo information in 3D reconstruction from multiple images. The key idea is to cast multiview stereovision as a convex variational problem and to impose exact silhouette constraints by restricting the domain of feasible functions. Relaxation of the resulting formulation leads to the minimization of a convex functional over the convex set of silhouette-consistent functions, which can be performed in a globally optimal manner using classical techniques. A solution of the original problem is obtained by projecting the computed minimizer onto the corresponding restricted domain. In contrast to classical techniques for silhouette and stereo integration, it does lead to a more robust and tractable numerical scheme by avoiding hard decisions about voxel occupancy and removing the bias near the visual hull boundary. The proposed approach allows to compute accurate silhouette-consistent reconstructions for challenging real-world problems in less than one minute.

## Acknowledgments

This research was supported by the German Research Foundation, grant # CR250/1-2. We thank Carlos Hernandez and Yasutaka Furukawa for providing the data sets in Figures 5 and 6. We thank Sudipta Sinha for sharing his results for Figure 5.

## References

1. Baumgart, B.: Geometric modeling for computer vision. PhD thesis, Department of Computer Science, Stanford University, USA (1974)
2. Laurentini, A.: The visual hull concept for visual-based image understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16, 150–162 (1994)
3. Martin, W.N., Aggarwal, J.K.: Volumetric descriptions of objects from multiple views. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 5, 150–158 (1983)
4. Cipolla, R., Giblin, P.: Visual motion of curves and surfaces. Cambridge University Press, Cambridge (2000)
5. Franco, J.S., Boyer, E.: Exact polyhedral visual hulls. In: *Proceedings of the Fourteenth British Machine Vision Conference*, Norwich, UK, pp. 329–338 (2003)
6. Yezzi, A., Soatto, S.: Stereoscopic segmentation. *International Journal of Computer Vision* 53, 31–43 (2003)
7. Seitz, S., Dyer, C.: Photorealistic scene reconstruction by voxel coloring. In: *Proc. International Conference on Computer Vision and Pattern Recognition*, pp. 1067–1073 (1997)

8. Faugeras, O., Keriven, R.: Variational principles, surface evolution, PDE's, level set methods, and the stereo problem. *IEEE Transactions on Image Processing* 7, 336–344 (1998)
9. Duan, Y., Yang, L., Qin, H., Samaras, D.: Shape reconstruction from 3D and 2D data using PDE-based deformable surfaces. In: Pajdla, T., Matas, J(G.) (eds.) *ECCV 2004. LNCS, vol. 3024*, pp. 238–251. Springer, Heidelberg (2004)
10. Lempitsky, V., Boykov, Y., Ivanov, D.: Oriented visibility for multiview reconstruction. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) *ECCV 2006. LNCS, vol. 3953*, pp. 226–238. Springer, Heidelberg (2006)
11. Matsumoto, Y., Fujimura, K., Kitamura, T.: Shape-from-silhouette/stereo and its application to 3D digitizer. In: *Proceedings of Discrete Geometry for Computing Imagery*, pp. 177–190 (1999)
12. Vogiatzis, G., Torr, P., Cippola, R.: Multi-view stereo via volumetric graph-cuts. In: *Proc. International Conference on Computer Vision and Pattern Recognition*, pp. 391–399 (2005)
13. Cross, G., Zisserman, A.: Surface reconstruction from multiple views using apparent contours and surface texture. In: *Confluence of Computer Vision and Computer Graphics*, Norwell, MA, USA, pp. 25–47. Kluwer Academic Publishers, Dordrecht (2000)
14. Esteban, C.H., Schmitt, F.: Silhouette and stereo fusion for 3D object modeling. *Computer Vision and Image Understanding* 96, 367–392 (2004)
15. Furukawa, Y., Ponce, J.: Carved visual hulls for image-based modeling. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) *ECCV 2006. LNCS, vol. 3954*, pp. 564–577. Springer, Heidelberg (2006)
16. Gargallo, P., Prados, E., Sturm, P.: Minimizing the reprojection error in surface reconstruction from images. In: *Proceedings of the International Conference on Computer Vision, Janeiro, Brazil*. IEEE Computer Society Press, Los Alamitos (2007)
17. Isidoro, J., Sclaroff, S.: Stochastic refinement of the visual hull to satisfy photometric and silhouette consistency constraints. In: *Proc. International Conference on Computer Vision, Washington, DC, USA*, pp. 1335–1342 (2003)
18. Tran, S., Davis, L.: 3D surface reconstruction using graph cuts with surface constraints. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) *ECCV 2006. LNCS, vol. 3952*, pp. 219–231. Springer, Heidelberg (2006)
19. Sinha, S., Mordohai, P., Pollefeys, M.: Multiview stereo via graph cuts on the dual of an adaptive tetrahedral mesh. In: *Proc. International Conference on Computer Vision, Rio de Janeiro, Brazil* (2007)
20. Sinha, S., Pollefeys, M.: Multi-view reconstruction using photo-consistency and exact silhouette constraints: A maximum-flow formulation. In: *Proc. International Conference on Computer Vision, Washington, DC, USA*, pp. 349–356. IEEE Computer Society, Los Alamitos (2005)
21. Hernández, C., Vogiatzis, G., Cippola, R.: Probabilistic visibility for multi-view stereo. In: *Proc. International Conference on Computer Vision and Pattern Recognition, Minneapolis, Minnesota, USA*. IEEE Computer Society, Los Alamitos (2007)
22. Kolev, K., Klodt, M., Brox, T., Cremers, D.: Propagated photoconsistency and convexity in variational multiview 3D reconstruction. In: *Workshop on Photometric Analysis for Computer Vision, Rio de Janeiro, Brazil* (2007)
23. Boykov, Y., Lempitsky, V.: From photohulls to photoflux optimization. In: *Proc. British Machine Vision Conference, vol. 3*, pp. 1149–1158 (2006)