Anisotropic Minimal Surfaces Integrating Photoconsistency and Normal Information for Multiview Stereo

Kalin Kolev[1], Thomas Pock[2], and Daniel Cremers[1]

[1] Department of Computer Science, TU München, Munich, Germany
[2] Institute for Computer Graphics and Vision, Graz University of Technology

Abstract. In this work the weighted minimal surface model traditionally used in multiview stereo is revisited. We propose to generalize the classical photoconsistency-weighted minimal surface approach by means of an anisotropic metric which allows to integrate a specified surface orientation into the optimization process. In contrast to the conventional isotropic case, where all spatial directions are treated equally, the anisotropic metric adaptively weights the regularization along different directions so as to favor certain surface orientations over others. We show that the proposed generalization preserves all properties and globality guarantees of continuous convex relaxation methods. We make use of a recently introduced efficient primal-dual algorithm to solve the arising saddle point problem. In multiple experiments on real image sequences we demonstrate that the proposed anisotropic generalization allows to overcome oversmoothing of small-scale surface details, giving rise to more precise reconstructions.

1 Introduction

Recovering 3D geometry of the observed scene from multiple calibrated cameras is one of the fundamental problems in Computer Vision. An established paradigm for solving this problem – often called *multiview stereo* – is to reconstruct the spatial structure in a way that maximizes the photoconsistency along the object surface, i. e. in a way that the projection of surface points into pairs of cameras gives rise to the same colors or local neighborhood structure. In contrast to other techniques like shape from shading or shape from silhouettes, multiview stereo does not require a controlled environment or additional user interaction. It is applicable to arbitrary Lambertian objects and forms the basis of the currently most competitive generic reconstruction methods.

However, the viability of the stereo paradigm strongly relies on the success of the matching process. When objects exhibit low texture or specular reflections and when the camera calibration is erroneous, this matching process may fail. In order to suppress the influence of different sources of error, a robust regularization scheme is therefore crucial. Among the pioneering approaches to this problem are the variational methods introduced by Faugeras and Keriven [1]

K. Daniilidis, P. Maragos, N. Paragios (Eds.): ECCV 2010, Part III, LNCS 6313, pp. 538–551, 2010.

Fig. 1. The figure depicts a challenging test scenario (1 out of 21 images is shown). The input sequence illustrates a bird figurine with a complex geometric structure comprising thin protrusions (e.g. the wings) and fine-scale details (e.g. the feathering). Note that classical photoconsistency is not able to capture the feathering of the wings due to their extremely low thickness. In this paper we propose a new energy minimization approach based on a robust and transparent integration of photoconsistency, silhouette and normal information, capable of accurately recovering such objects. Two views of the reconstructed surface are visualized.

and by Yezzi and Soatto [2]. The key idea in [1] is to specify the 3D geometry as a weighted minimal surface model, where the local metric is defined in terms of a photoconsistency measure reflecting the agreement of projected surface colors among pairs of images. While this approach successfully suppresses noise, it suffers from two important limitations. Firstly, the authors in [1] merely compute suboptimal local minima. In fact, the global minimum is actually the empty set. Secondly, the minimal surface formulation causes oversmoothing effects. As a result, protrusions and surface indentations tend to be suppressed in the reconstructions [3,4].

To overcome these limitations, researchers have devised a variety of strategies. Firstly, one can remove the trivial solution and the related shrinking bias by introducing constant ballooning terms so as to favor shapes of larger volume [5]. As an undesired side effect the resulting expansion force tends to fill in concavities or, respectively, cut off protrusions, depending on the strength of the inflating force. More elaborate methods were presented in [3,4], where the constant ballooning term is replaced by a data-aware volume subdivision. While these procedures help to drastically alleviate the shrinking bias of minimal surface models, they become unreliable in the presence of specular reflections or with very few input images and the reconstruction fails.

An alternative approach to avoid the empty set was recently proposed in [6]. The authors advocate to retain the weighted minimal surface model but restrict the optimization to the set of silhouette-consistent configurations. By imposing that along visual rays, passing silhouette pixels, the voxel occupancy must be at least one, the trivial solution is no longer feasible. Nevertheless, the shrinking bias is still present. Even within the set of silhouette-consistent solutions the obtained reconstruction may suffer from oversmoothing and suppression of indentations that are not captured by the silhouettes.

A different strategy to address the multiview stereo problem was suggested in [7,8,9]. The key idea is to specify additionally to the surface localization in terms of spatial photoconsistency also the local surface orientation. In [9] the gradient of the photoconsistency measure was used to approximate the orientation of the observed shape. [7,8] proposed to estimate the surface orientation directly via an optimization procedure over the local patch distortion. The optimal patch is determined based on the agreement of its projections onto the images, where it is visible. Directionally sensitive anisotropic metrics provide a powerful tool to integrate such shape normal information in the minimization process. Recent advances in this field within the context of discrete optimization [10,11] as well as continuous counterparts [12,13] provide the necessary machinery to compute a globally optimal solution.

Many of the discussed ideas have been developed independently with focus on particular weaknesses of previous methods for multiview reconstruction. For example, the fusion of multiview stereo and silhouettes aims at recovering thin protrusions while retaining concavities. Moreover, the normal information is known to capture high-frequency surface details and enhance reconstructions. This raises the straightforward question of how these different sources of information can be integrated in a unified framework in a robust and transparent manner. The ultimate goal is to combine their advantages and extend the range of applicability of established techniques. In particular, such an approach should be able to reconstruct fine-scale shape details on thin protruding structures, a test scenario for which most of the existing methods fail – see Fig. 1.

The contribution of this paper is to propose an energy minimization framework for multiple view reconstruction that allows to combine multiview photoconsistency, silhouette and normal information. We show that the reconstruction can be efficiently determined as an anisotropic minimal surface which favors not only locations of good photoconsistency but also orientations that are consistent with the specified normal field. By adaptively reducing the smoothing along the predetermined directions, the inherent shrinking bias of traditional minimal surface models is alleviated. Making use of convex relaxation techniques we pose the reconstruction problem as one of minimizing a convex energy functional. We show that globally optimal shapes which best fit the photoconsistency values and the specified normal field are obtained by thresholding the solution of the relaxed problem. In addition, exact silhouette consistency can be imposed by constraining the optimization to the convex set of silhouette consistent surfaces. It should be noted that the proposed approach is a portent of a wide range of applications involving normal field integration including range data fusion [14], shape from shading [15] and photometric stereo [16].

The paper is organized as follows. In the next section we introduce the concept of anisotropic minimal surfaces for multiview stereo. In Section 3 we derive corresponding specific energy functionals to integrate photoconsistency, surface orientation and silhouettes. In Section 4 we provide an efficient primal-dual algorithm for solving the arising saddle point problem. In Section 5 we show experimental results on challenging real data sets. We conclude with a brief summary.

2 Continuous Anisotropic Minimal Surfaces

This section introduces the main concept of anisotropic metrics preferring certain orientation selectivity by generalizing the traditionally used weighted minimal surface model. Similar anisotropic formulations have independently been developed in the context of binocular stereo [12] and image segmentation [13].

We start with some notations. Let $V \subset \mathbb{R}^3$ be a volume which contains the scene of interest and $I_1, \ldots, I_n : \Omega \to \mathbb{R}^3$ a collection of calibrated color images with perspective projections $\pi_1, \ldots, \pi_n$. Let $S_1, \ldots, S_n \subset \Omega$ be the observed silhouettes of the 3D object and $\rho : V \to [0, 1]$ be a photoconsistency map measuring the discrepancy among various image projections.

The most photoconsistent shape can be obtained according to the following weighted minimal surface model:

$$E(S) = \int_S \rho(s)\, ds. \tag{1}$$

The model encourages the surface to pass through points with high observation agreement. Its minimization identifies shapes with minimal overall costs according the local isotropic metric induced by ρ. However, the model does not explicitly affect the orientation of the estimated shape. To this end, in case of given surface orientation, a generalization has to be developed. This can be achieved by introducing a family of positive semidefinite anisotropic tensors $D(x) \in \mathbb{R}^{3\times3}, x \in V$ tolerating certain directional selectivity. Now, the minimal surface model (1) generalizes to:

$$E(S) = \int_S \sqrt{N_S(s)^T D(s) N_S(s)}\, ds, \tag{2}$$

where $N_S(s) \in S^2 \subset \mathbb{R}^3$ denotes the unit outward surface normal at point s. Obviously, the tensors $D(x)$ can be designed to energetically favor certain shape orientations while suppressing others. Note that D is defined pointwise. However, in the remainder of this section we will omit the argument for the sake of simplicity. The energy model (2) can sill be interpreted as a minimal surface formulation according to the Riemannian metric induced by D (i. e. $||v||_D = \sqrt{v^T D v}$). The Euclidean metric, which treats all spatial directions equally, appears as a special case with $D = I$, where $I \in \mathbb{R}^{3\times3}$ denotes the identity matrix. Examples of local distance maps in 2D of the Euclidean and the more general Riemannian metric are visualized in Fig. 2. The classical weighted minimal surface model (1) also appears as a special case for $D = \rho^2\, I$.

Now, we are confronted with the question of defining the family of anisotropic tensors D appropriately. Let us assume that a vectorfield $F : V \to \mathbb{R}^3$ is provided representing an estimate of the unit outward orientation of the desired shape. In practice, meaningful normal estimates can be computed only for points on the surface of the observed object. For all other points we can set $D = \rho^2\, I$, which corresponds to the conventional isotropic case. Thus, in the sequel we will assume $F : V \to S^2 \subset \mathbb{R}^3$, where S^2 denotes the unit sphere. In Section 4 we

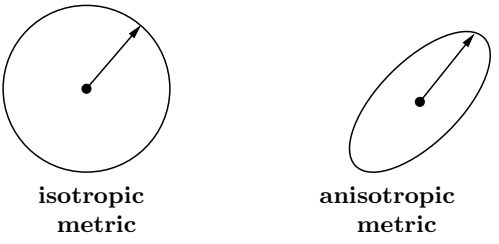

Fig. 2. Local distance maps. Examples of local distance maps: $\sqrt{v^T v} = 1$ (isotropic case) and $\sqrt{v^T D v} = 1$ (anisotropic case). While isotropic metrics treat all directions equally, anisotropic metrics possess directional selectivity.

will give more details on how orientation information can be obtained from the input images. Based on this data, we would like to suppress regularization along the corresponding normal and encourage the process along the tangent plane. Moreover, the photoconsistency map ρ should be taken into account. This can be achieved by setting

$$D = \rho^2 \left(\tau F F^T + \frac{3-\tau}{2} (I - F F^T) \right), \tag{3}$$

where $\tau \in [0, 1]$ is a weighting parameter that controls the distortion of the corresponding metric, i. e. the tolerance of the normal field F. In effect, the formulation in (3) realizes a basis transformation and subsequent scaling. The first term treats the component along F and scales it by τ, whereas the second one affects the tangential components. It is easy to verify that the choice $\tau = 1$ gives the original model (1). On the other hand, $\tau = 0$ will completely turn off smoothing along the vectorfield F. In our experiments we found out $\tau = 0.15$ to be a good compromise.

Next, we will show some favorable properties of the tensor D defined above.

Proposition 1. *For a normalized vector F, $\rho \geq 0$ and $\tau \in [0, 1]$ the matrix D defined in* (3) *is symmetric and positive semidefinite with $tr(D) = 3\rho^2$.*

Proof. The symmetry of D is obvious.
In order to show that D is positive semidefinite, we observe for $v \in \mathbb{R}^3$

$$v^T D v = \rho^2 \, \frac{3\tau - 3}{2} \, \left(v^T F \right)^2 + \rho^2 \, \frac{3-\tau}{2} \, ||v||^2 \geq \rho^2 \, \tau \, ||v||^2 \geq 0$$

due to $||F|| = 1$.
Finally, we obtain

$$tr(D) = \rho^2 \, \left(\frac{3\tau - 3}{2} \, ||F||^2 + \frac{3\,(3-\tau)}{2} \right) = 3\rho^2. \qquad \square$$

The condition that $tr(D)$ does not depend on the choice of the parameter τ assures that the overall smoothing remains fixed.

It should be noted that the inverse of the matrix in (3) as well as its square root can easily be computed as

$$\begin{aligned} D^{-1} &= \frac{1}{\rho^2}\left(\frac{1}{\tau}FF^T + \frac{2}{3-\tau}(I - FF^T)\right) \\ D^{1/2} &= \rho\left(\sqrt{\tau}FF^T + \sqrt{\frac{3-\tau}{2}}(I - FF^T)\right). \end{aligned} \tag{4}$$

This will be useful for optimization purposes (see Section 4.2).

3 Fusing Photoconsistency, Orientation and Silhouettes

In this section we will formulate specific energy functionals based on the anisotropic minimal surface model (2) which will serve as a data-aware smoothness term.

3.1 Adding Regional Terms

As mentioned previously, an undesired property of minimal surface models of the form (2) is that the empty set always exhibits a global minimum. One way to avoid this trivial solution is to derive additional information from the images, which gives a closer specification of the observed object. This can be achieved by introducing data terms $\rho_{int} : V \to [0,1]$ and $\rho_{ext} : V \to [0,1]$ defining costs for each point within the volume for being inside or outside the imaged shape, respectively. Now, we obtain the following energy model:

$$\begin{aligned} E(S) = \lambda &\left(\int_{int(S)} \rho_{int}(x)\, dx + \int_{ext(S)} \rho_{ext}(x)\, dx\right) \\ &+ \int_S \sqrt{N_S(s)^T D(s) N_S(s)}\, ds, \end{aligned} \tag{5}$$

where $int(S), ext(S) \subset V$ denote the surface interior and exterior, respectively, and $\lambda \in \mathbb{R}_{\geq 0}$ is a weighting parameter.

Next, we are confronted with the optimization of the functional (5). To this end, the first steps are a conversion to an implicit representation $u = \mathbf{1}_{int(S)}$, where $\mathbf{1}_{int(S)}$ denotes the characteristic function of $int(S)$, and subsequent relaxation:

$$E(u) = \lambda \int_V (\rho_{int}(x) - \rho_{ext}(x))\, u(x)\, dx + \int_V \sqrt{\nabla u(x)^T D(x) \nabla u(x)}\, dx, \tag{6}$$

where $u \in C_{rel} := \{\, \hat{u} \,|\, \hat{u} : V \to [0,1] \,\}$. Note that the "binary" version of (6), i. e. optimization over the set of binary functions $u \in C_{bin} := \{\, \hat{u} \,|\, \hat{u} : V \to \{0,1\} \,\}$, is equivalent to (5). Fortunately, the optimization of (5) turns out to be as simple as minimizing (6) which exhibits a constrained convex optimization problem. This is stated by the following

Theorem 1. *Let $u^* : V \to [0,1]$ be a global minimizer of the functional* (6)*. Then the characteristic functions of all upper level sets (i.e. thresholded versions)*

$$\Sigma_{\mu,u^*} = \{x \in V \,|\, u^*(x) > \mu\}, \qquad \mu \in (0,1), \tag{7}$$

of u^* are also global minimizers of (6)*.*

Proof. The claim follows directly from the layer cake representation of u^* (see [17]) and the anisotropic coarea formula (see [13] for a detailed derivation). □

The above theorem implies that we can obtain a global minimum of (5) by solving the constrained convex optimization problem (6) and thresholding the result by some $\mu \in (0,1)$.

3.2 Incorporating Silhouette Constraints

In certain practical scenarios obtaining reliable volume subdivision terms may be a challenging task. In such cases a reasonable alternative could be to retain the original minimal surface model (2) but to restrict the domain of feasible shapes in order to exclude the trivial solution. The object silhouettes serve as a useful tool that could provide such constraints:

$$\begin{aligned} E(S) &= \int_S \sqrt{N_S(s)^T D(s) N_S(s)}\, ds, \\ \text{s. t.} \qquad & \pi_i(S) = S_i \quad \forall\, i = 1, \ldots, n. \end{aligned} \tag{8}$$

Note that for $D = \rho^2 I$ the above formulation boils down to the model proposed in [6].

Unfortunately, global optimization of (8) is not a trivial task. Nevertheless, a global minimum can be obtained up to an energetic upper bound. Reverting to an implicit representation and subsequent relaxation yields:

$$\begin{aligned} E(u) &= \int_V \sqrt{\nabla u(x)^T D(x) \nabla u(x)}\, dx, \\ &\text{s. t.} \qquad u \in [0,1] \\ &\int_{R_{ij}} u(x)\, dR_{ij} \geq 1 \text{ if } j \in S_i \\ &\int_{R_{ij}} u(x)\, dR_{ij} = 0 \text{ if } j \notin S_i, \end{aligned} \tag{9}$$

where R_{ij} denotes the visual ray through pixel j of image i. It can be verified that (9) exhibits a constrained convex optimization problem for which the global minimum can be obtained. Since we are interested in finding minimizers of the original non-convex problem (8), we threshold the solution of the convex problem u_{min} appropriately:

$$\tilde{u}(x) = \begin{cases} 1, & \text{if } u_{min}(x) \geq \mu \\ 0, & \text{otherwise} \end{cases}, \tag{10}$$

where

$$\mu = \min \left\{ \left(\min_{i \in \{1,\ldots,n\}, j \in S_i} \max_{x \in R_{ij}} u_{min}(x) \right) \ , \ 0.5 \right\}. \tag{11}$$

The threshold is estimated such that the computed binary solution is the closest one that still fulfills exact silhouette consistency. Note that minimizing (8) is equivalent to minimizing the "binarized" version of (9) (where $u \in [0,1]$ is replaced by $u \in \{0,1\}$). Although this approach does not assure finding the global minimum of (8), it entails certain globality guarantees.

Proposition 2. *Let u' be a (global) minimum of the "binary" version of* (9)*, $\tilde{u}$ the computed solution and u_{min} a (global) minimum of* (9)*. Then, a bound $\gamma(u_{min}, \tilde{u})$ exists such that*

$$E(\tilde{u}) - E(u') \leq \gamma(u_{min}, \tilde{u}).$$

Proof. Since the binary functions are a subset of the real-valued functions, we have the relation

$$E(u_{min}) \leq E(u') \leq E(\tilde{u})$$

As a consequence, we obtain the inequality

$$E(\tilde{u}) - E(u') \leq E(\tilde{u}) - E(u_{min}) =: \gamma(u_{min}, \tilde{u}). \qquad \square$$

Generally, we used the energy model in (5) in our experiments, since it does not require silhouette information to be provided, and switched to (8) in cases where computing accurate regional terms was not feasible.

4 Implementation and Numerics

This section will give more details on the particular choice of data terms and the numerical implementation of the proposed approach.

4.1 Data Term Computation

Following the formulation in (3), we need to define multiple data measures: a photoconsistency map ρ, regional subdivision costs ρ_{int}, ρ_{ext} and an outward normal field F.

The photoconsistency estimation that we used in our experiments is based on the voting scheme proposed in [19]. Moreover, we used the propagating approach in [4] to derive volumetric assignment costs for object interior/exterior.

In order to obtain an estimate of a normal field F representing the surface orientation, we assume a sparse oriented point cloud

$$P = \{ \, p_i \mid p_i \in V \, \}$$
$$O = \left\{ \, v_i \mid v_i \in S^2 \, \right\},$$

where S^2 denotes the unit sphere. Such data can be obtained via an optimization procedure over the local photometric consistency (see Fig. 3). In our experiments,

(a) (b) (c)

Fig. 3. Surface normal estimation. (a) The orientation of a point in space is obtained, based on a local planar patch. The optimal orientation is given by the maximal photometric agreement of the projections of the patch onto the images, where it is visible. (b) A point cloud for the data set in Fig. 4, generated with the software at [18]. (c) Corresponding color-coded normal vectors.

we used the approach of [7], an implementation of which is publicly available at [18]. See [7] for more details. A sample oriented point cloud for a real image sequence, obtained with the above procedure, is visualized in Fig. 3 (b), (c). Based on this data, we define the vectorfield F as

$$F(x) = \begin{cases} v_i, & \text{if } x = p_i \\ 0, & \text{otherwise} \end{cases} . \tag{12}$$

In practice, we replace F with a semi-dense blurred version $\tilde{F}$ in order to account for inaccuracies due to image noise.

4.2 Efficient Primal-Dual Optimization

As mentioned previously, the minimization of (6) and (9) poses classical constrained convex optimization problems. Hence, any iterative local optimization procedure will provide the global minimum. However, the particular choice of minimization method will affect the speed of convergence. In the current work, we adopt the primal-dual method proposed in [20].

First, we observe that that the energy functionals in (6) and (9) are both in the form

$$E(u) = \int_V \sqrt{\nabla u^T D \nabla u}\, dx + \int_V f u\, dx, \tag{13}$$

where $f : V \to \mathbb{R}$ summarizes the constant part not dependent on u. We proceed by splitting D as $D = D^{1/2} D^{1/2}$ and introducing a dual variable $\xi : V \to \mathbb{R}^3$, which allows for the following conversion:

$$\begin{aligned} E(u) &= \int_V \sqrt{\nabla u^T D \nabla u}\, dx + \int_V f u\, dx \\ &= \int_V ||D^{1/2} \nabla u||\, dx + \int_V f u\, dx \\ &= \max_{||\xi|| \leq 1} \int_V \langle \xi, D^{1/2} \nabla u \rangle\, dx + \int_V f u\, dx. \end{aligned} \tag{14}$$

Now, we obtain a new functional

$$E(u,\xi) = \int_V \langle \xi, D^{1/2}\nabla u\rangle \, dx + \int_V f u \, dx, \tag{15}$$

that should be minimized with respect to u and maximized with respect to ξ under the constraint $||\xi|| \leq 1$. This states a typical saddle point problem that can be solved by a projected gradient descent/ascent strategy. Denoting by $K := \{ \xi \in \mathbb{R}^3 \mid ||\xi|| \leq 1 \}$ the unit ball, our optimization scheme can be described as follows. We choose $(u^0, \xi^0) \in C_{rel} \times K$ and let $\bar{u}^0 = u^0$. We choose two time-steps $\tau, \sigma > 0$. Then, we iterate for $n \geq 0$

$$\begin{aligned} \xi^{n+1} &= \Pi_K(\xi^n + \sigma(D^{1/2}\nabla \bar{u}^n)) \\ u^{n+1} &= \Pi_{C_{rel}}(u^n + \tau(\mathrm{div}(D^{1/2}\xi^{n+1}) - f)) \\ \bar{u}^{n+1} &= 2u^{n+1} - u^n, \end{aligned} \tag{16}$$

where Π_K and $\Pi_{C_{rel}}$ denote projections onto the corresponding sets. Both projections are realized by simple normalization and clipping, respectively. For projection onto the set of silhouette-consistent solutions, imposed in (9), we refer to [21]. Note that the matrix square root $D^{1/2}$ can easily be computed according to the construction (see (4)). Note however that $D^{1/2}$ is, in general, spatially varying.

For sufficiently small time-step parameters convergence of the above iterative procedure can be proved [20]. In our experiments we observed stable behavior for $\tau = \sigma = 0.1$.

5 Experiments

To motivate the exploration of anisotropic minimal surface models allowing to integrate normal information, we start with a challenging image sequence illustrated in Fig. 4. Depicted are 3 out of 21 input images and multiple views of the reconstructions with the classical weighted minimal surface model [6] and the proposed anisotropic generalization (8). The data set is quite inconvenient due to the complex geometry of the imaged object comprising multiple thin structures (e.g. the wings or the legs). While the isotropic minimal surface model accurately recovers all elongated structures, it completely fails at small-scale surface details (e.g. the feathering) in contrast to the proposed anisotropic approach which clearly enhances the reconstruction. In Fig. 3 the utilized normal field is illustrated. It should be noted that for very thin geometric structures no meaningful photoconsistency can be derived. In this case the weighted minimal surface model boils down to Euclidean minimal surface model and produces the smoothest silhouette-consistent shape.

Although the Middlebury benchmark [23] is essentially exhausted and no longer provides a major challenge for multiview stereo approaches, it remains one of the most established benchmarks. Fig. 5 shows multiple views of the reconstructions obtained with the model in (5) on the well-known "dinoRing" and

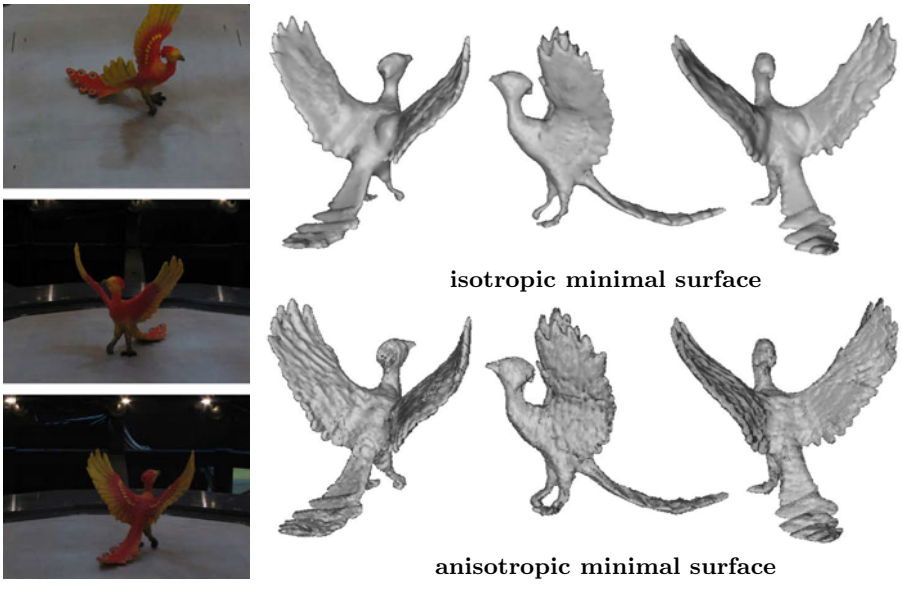

Fig. 4. Bird sequence. 3 out of 21 input images of resolution 1024×768 and multiple views of the reconstructions with the classical isotropic minimal surface model (used for example in [6]) and the proposed anisotropic generalization (8). Note that small-scale structures like the feathering are clearly oversmoothed by the isotropic model in contrast to the proposed anisotropic approach.

Fig. 5. Middlebury data sets. Two of the input images and two views of the reconstructions obtained with the model in (5) on the well-known "dinoRing" (48 images of resolution 640×480) and "templeRing" (47 images of resolution 640×480) data sets. See Table 1 for a quantitative evaluation and [22] for a comparison to other approaches.

"templeRing" data sets. We refer to Table 1 for a quantitative evaluation and a comparison to the results reported in [4] for an isotropic version of (5). Moreover, we refer to [22] for a comparison to alternative approaches. Surprisingly, despite the already very low errors, experiments show that the proposed anisotropic formulation leads to a further reduction of accuracy and completeness scores. It should be recalled that the proposed method operates on a discrete volume grid, which poses a restriction on the precision of the recovered 3D meshes. Additional

Table 1. Quantitative comparison between the anisotropic model in (5) and the isotropic version reported in [4] on the Middlebury data sets (see Fig. 5). The numbers give accuracy (in mm) and completeness (in %). The completeness score measures the percentage of points in the ground truth model that are within $1.25mm$ of the reconstructed model. The accuracy metric shown is the distance d that brings 90% of the reconstructed surface within d from some point on the ground truth.

	isotropic model	anisotropic model
dinoSparseRing	0.53mm / 98.3%	**0.48mm / 98.6%**
dinoRing	0.43mm / 99.4%	**0.42mm / 99.5%**
templeSparseRing	1.04mm / 91.8%	**0.97mm / 92.7%**
templeRing	0.72mm / 97.8%	**0.7mm / 98.3%**

Fig. 6. Niobe sequence. 2 out of 38 input images of resolution 2048×3072 and multiple views of the reconstruction obtained by optimizing the energy model in (5). Note the accurate reconstruction of fine-scale surface details.

post-processing refinement of the generated triangle meshes could be included to further increase the scale of accuracy and obtain better evaluation results but this is beyond the scope of the current work.

Finally, we conclude with an experiment on an image sequence capturing a Greek statue (Niobe, reproduction from the 19th century, 2m high). Two of the input images and multiple views of the high-quality reconstruction obtained by optimizing the energy model in (5) are depicted in Fig. 6. Despite the fixed volumetric discretization most of the relevant fine-scale surface details like body parts and creases of the clothing are recovered accurately. Note also the severe brightness variations that make a robust optimization scheme indispensable. It is interesting to mention that the input photographs were acquired by a hand-held camera and calibrated with the Bundler software [24].

All demonstrated reconstructions were computed on volumetric grids consisting of $18-40$ million voxels. On a consumer PC we measured runtimes of up to a couple of hours, whereas most of the time was spent on data term computations.

6 Conclusion

In this paper, we proposed to integrate multiview stereo information and surface orientation estimates by means of anisotropic minimal surfaces. The key idea is that the local photoconsistency enters as the overall weight of the metric, while the additional normal information induces a local anisotropy of the metric which favors discontinuities of the labeling function in directions of the prescribed normal field. We prove that optimal anisotropic minimal surfaces can be computed using convex relaxation and thresholding techniques. In addition, exact silhouette consistency can be imposed by constraining the optimization to the convex set of silhouette-consistent solutions. In experiments on real-world data we demonstrate that stereo-based reconstruction results can be enhanced both qualitatively and quantitatively by incorporating normal information.

Acknowledgments

This research was supported by the German Research Foundation, grant # CR250/1-2 and # CR250/4-1. We thank Svetlana Matiouk for assistance in the image acquisition for the experiment in Fig. 6. We thank the Bundler team for providing their calibration software.

References

1. 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)
2. Yezzi, A., Soatto, S.: Stereoscopic segmentation. International Journal of Computer Vision 53, 31–43 (2003)
3. Hernández, C., Vogiatzis, G., Cipolla, R.: Probabilistic visibility for multi-view stereo. In: Proc. International Conference on Computer Vision and Pattern Recognition, Minneapolis, Minnesota, USA, IEEE Computer Society Press, Los Alamitos (2007)
4. Kolev, K., Klodt, M., Brox, T., Cremers, D.: Continuous global optimization in multiview 3D reconstruction. International Journal of Computer Vision 84, 80–96 (2009)
5. 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)
6. Kolev, K., Cremers, D.: Integration of multiview stereo and silhouettes via convex functionals on convex domains. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part I. LNCS, vol. 5302, pp. 752–765. Springer, Heidelberg (2008)
7. Furukawa, Y., Ponce, J.: Accurate, dense, and robust multi-view stereopsis. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Minneapolis, Minnesota, USA (2007)
8. Habbecke, M., Kobbelt, L.: A surface-growing approach to multi-view stereo reconstruction. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Minneapolis, Minnesota, USA, pp. 1–8 (2007)

9. Boykov, Y., Lempitsky, V.: From photohulls to photoflux optimization. In: Proc. British Machine Vision Conference, vol. 3, pp. 1149–1158 (2006)
10. Kolmogorov, V., Boykov, Y.: What metrics can be approximated by geo-cuts, or global optimization of length/area and flux. In: Proc. International Conference on Computer Vision, Beijing, China, pp. 564–571. IEEE Computer Society, Los Alamitos (2005)
11. Kirsanov, D., Gortler, S.: A discrete global minimization algotithm for continuous variational problems. In: Harvard Computer Science Technical Report: TR-14-04 (2004)
12. Zach, C., Niethammer, M., Frahm, J.M.: Continuous maximal flows and Wulff shapes: Application to MRFs. In: Proc. International Conference on Computer Vision and Pattern Recognition, Miami, FL (2009)
13. Olsson, C., Byröd, M., Overgaard, N.C., Kahl, F.: Extending continuous cuts: Anisotropic metrics and expansion moves. In: International Conference on Computer Vision (2009)
14. Lempitsky, V., Boykov, Y.: Global optimization for shape fitting. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Minneapolis, Minnesota, USA (2007)
15. Horn, B., Brooks, M.: Shape from shading. MIT Press, Cambridge (1989)
16. Woodham, R.J.: Photometric method for determining surface orientation from multiple images. Optical Engineerings 19, 139–144 (1980)
17. Chan, T., Esedoḡlu, S., Nikolova, M.: Algorithms for finding global minimizers of image segmentation and denoising models. SIAM Journal on Applied Mathematics 66, 1632–1648 (2006)
18. PMVS. `http://www.cs.washington.edu/homes/furukawa/research/pmvs/`
19. Hernandez, C., Schmitt, F.: Silhouette and stereo fusion for 3D object modeling. Computer Vision and Image Understanding 96, 367–392 (2004)
20. Pock, T., Cremers, D., Bischof, H., Chambolle, A.: An algorithm for minimizing the piecewise smooth Mumford-Shah functional. In: IEEE on International Conference,Computer Vision (ICCV), Kyoto, Japan (2009)
21. Dykstra, R.: An algorithm for restricted least squares regression. Journal of the American Statistical Association 78, 837–842 (1983)
22. Middlebury, `http://vision.middlebury.edu/mview/`
23. Seitz, S., Curless, B., Diebel, J., Scharstein, D., Szeliski, R.: A comparison and evaluation of multi-view stereo reconstruction algorithms. In: Proc. International Conference on Computer Vision and Pattern Recognition, pp. 519–528 (2006)
24. Bundler, `http://phototour.cs.washington.edu/bundler/`