Continuous Ratio Optimization via Convex Relaxation with Applications to Multiview 3D Reconstruction

Kalin Kolev and Daniel Cremers
Department of Computer Science
University of Bonn, Germany

Abstract

We introduce a convex relaxation framework to optimally minimize continuous surface ratios. The key idea is to minimize the continuous surface ratio by solving a sequence of convex optimization problems. We show that such minimal ratios are superior to traditionally used minimal surface formulations in that they do not suffer from a shrinking bias and no longer require the choice of a regularity parameter. The absence of a shrinking bias in the minimal ratio model is proven analytically. Furthermore we demonstrate that continuous ratio optimization can be applied to derive a new algorithm for reconstructing three-dimensional silhouette-consistent objects from multiple views. Experimental results confirm that our approach allows to accurately reconstruct deep concavities even without the specification of tuning parameters.

1. Introduction

Shape optimization is at the heart of fundamental Computer Vision problems like image segmentation and multiview 3D reconstruction. Over the years minimal surface formulations have become the established paradigm to solve such problems, either in a discrete or in a spatially continuous setting. Many of these models were originally introduced in the context of image segmentation [11, 2] and subsequently generalized to multiview 3D reconstruction where the reconstruction is computed by minimizing a photoconsistency-weighted minimal surface [8]. More recently, researchers have developed algorithms to globally optimize respective energies for multiview reconstruction in the discrete setting [20] and the continuous setting [13].

While minimal surface approaches were shown to provide good reconstructions and robustness to noise, they suffer from two important shortcomings:

• Minimal surfaces are known to exhibit a shrinking bias. In particular the global optimum of the minimal surface energy is the empty set. The empty set can be suppressed by constraining optimization to the vicinity of the visual hull [20], by introducing ballooning or flux terms [19, 1], by reverting to region-based terms using stereoscopic volume subdivision [9, 14], or by imposing silhouette consistency [13]. Nevertheless, the shrinking bias prevails in the sense that indentations and protrusions are energetically disfavored.

• The traditional variational approach of data term plus regularity requires choosing an appropriate regularity weight. As discussed in [17], this weight cannot be chosen automatically as it defines the spatial scale at which segmentations are to be computed. Formulated differently, the traditional shape optimization approach suffers from the fact that it is not scale invariant: For a given object, the quality of reconstruction highly depends on the spatial scale at which it is perceived. Consequently one cannot assure good reconstruction performance unless we know the scale of objects.

In this paper, we provide an exact mathematical characterization of both the shrinking bias and the scale dependency of traditional minimal surface approaches. Moreover,
we introduce a method which allows to optimally compute surfaces of minimal ratio in a spatially continuous setting. It is based on solving a sequence of convex optimization problems. In contrast to existing spatially discrete ratio optimization methods [10, 16], we derive a continuous counterpart, which entails a series of advantages such as straightforward parallelizability, an absence of metrication errors and substantially lower memory requirements [12].

We prove that the proposed minimal ratio solution does not suffer from the above problems: Firstly, it does not exhibit a shrinking bias and allows for better reconstructions of concavities and protrusions. In particular, we prove that any disjoint surface with the same energy can be added without affecting the overall energy costs. Secondly, the minimal ratio formulation does not have any tuning parameters and is shown to be scale invariant. Based on these observations, we developed a novel algorithm for reconstructing silhouette-consistent surfaces of minimal ratio from multiple views, based on photometric information. Experimental results confirm that it allows for accurate reconstructions that are independent of the geometric structure or the spatial scale of the object – see Figure 1.

The paper is organized as follows. In the next section, we present some related work. Section 3 briefly reviews a recent minimal surface formulation of multiview reconstruction [13]. In Section 4, we show how ratio optimization can be carried out in a spatially continuous setting and discuss the strengths of the model. Section 5 provides details on the numerical implementation. In Section 6, we show experimental results on real data sets. We conclude with a brief summary.

2. Related Work

Ratio optimization goes back to the 1960’s [6], where a generalization of the Newton’s method to fractional optimization was proposed. However, the potential of this technique for solving computer vision problems was recognized much later [4, 10]. Ratio minimization was successfully applied for image segmentation [4, 10, 16] and multiview 3D reconstruction [16]. However, all of these approaches were developed within a spatially discrete setting like minimal ratio cycles [10] or maxflow/mincut framework [16]. In contrast, the formulation proposed in this work is continuous and relies on convex relaxation. It generalizes previous convex minimization techniques [3, 15] and shows their applicability for ratio optimization. The continuous formulation entails advantages like parallelizability and reduction of metrication errors and memory requirements (see [12] for a discussion). Moreover, it allows the integration of convex constraints like silhouette consistency in the optimization process [13]. Our approach shares similarities to the continuous concave-convex procedure proposed in [21]. However, a crucial difference is that the method of [21] is suboptimal and generally does not give globally optimal solutions.

The concept of estimating stereoscopic regional terms specifying surface interior/exterior, required by our model, appeared in multiple previous works related to multiview 3D reconstruction [5, 9, 22, 14]. The key idea is to assign regional costs to each point in space based on the location of maximal photoconsistency along the viewing rays passing through it. While the approaches in [5, 9, 22] rely on pre-computed range images/disparity maps, the method of [14] is entirely volumetric and does not depend on the resolution of the input images. An alternative technique is to derive regional terms determining the surface interior as divergence of an estimated vectorfield, which specifies the shape orientation [1]. However, estimating surface normals based on multiview stereo is, in general, a challenging task, since it is very sensitive to image noise and mismatches. In our implementation, we used the approach of [14] due to its effectiveness and high accuracy.

3. Convex Integration of Silhouettes and Stereo

In this paragraph, we will briefly review the method of Kolev and Cremers [13] to impose silhouette-consistency as convex constraints.

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 projections 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, whose projection exactly coincides with the observed silhouettes, can be obtained according to the following minimal surface model:

\[
E(S) = \int_S \rho(x) \, dS, \quad (1)
\]

s. t. \( \pi_i(S) = S_i \quad \forall \ i = 1, \ldots, n \).

Representing the surface \( S \) implicitly by the characteristic function \( u : V \to \{0, 1\} \) of the surface interior \( S_{int} \) yields the following equivalent formulation:

\[
E(u) = \int_V \rho(x) |\nabla u(x)| \, d^3x
\]

s. t. \( u \in \{0, 1\} \)

\[
\int_{R_{ij}} u(x) \, dR_{ij} \geq 1 \quad \text{if} \quad j \in S_i
\]

\[
\int_{R_{ij}} u(x) \, dR_{ij} = 0 \quad \text{if} \quad j \notin S_i,
\]

where \( R_{ij} \) denotes the visual ray through pixel \( j \) of image \( i \). One can observe that the energy functional in (2) is convex, but it is optimized over a non-convex domain of binary functions. Relaxing the binary condition to \( u \in [0, 1] \) yields
4. Ratio Optimization via Relaxation

Let us assume, that a vectorfield $F : V \rightarrow \mathbb{R}^3$ is provided representing an estimate of the inward orientation of the desired shape. In Section 5 we will give more details on how such information can be obtained from the input images. With the notations introduced in Section 3, we consider the following constrained ratio optimization problem:

$$E(S) = \frac{\int_S \langle N_S(x), F(x) \rangle \, dS}{\int_S \rho(x) \, dS} \rightarrow \min$$

subject to

$$\pi_i(S) = S_i \quad \forall \, i = 1, \ldots, n,$$

where $N_S(x)$ denotes the outward normal of surface $S$ at point $x$. Minimization of (3) gives a silhouette-consistent shape, that optimally fulfills photoconsistency and normal field alignment criteria. Since during minimization the above energy functional becomes negative, optimal surfaces aim at maximizing the numerator magnitude and minimizing the denominator.

4.1. Absence of a Shrinking Bias

The shrinking bias of minimum surface methods has been extensively discussed in the literature. The following proposition is a precise statement that ratio optimization does not exhibit a shrinking bias.

**Proposition 1.** Let $S, S' \subset V$ be two arbitrary silhouette-consistent surfaces with the same ratio energy: $E(S) = E(S')$. Then adding the surface $S'$ to the surface $S$ does not affect the overall energy, i. e. $E(S \cup S') = E(S)$.

**Proof.** For simplicity, we will denote the numerator and denominator in (3) by $N(S) := \int_S \langle N_S(x), F(x) \rangle \, dS$ and $D(S) := \int_S \rho(x) \, dS$.

From the condition $E(S) = E(S')$ we can derive

$$N(S') = \frac{D(S') N(S)}{D(S)}$$

and hence

$$E(S \cup S') = \frac{N(S) + N(S')}{{D(S) + D(S')}} = \frac{N(S) + \frac{D(S') N(S)}{D(S)}}{D(S) + D(S')}$$

$$= \frac{N(S)}{D(S)} = E(S).$$

The above claim states, that one can iteratively expand a given surface by surface elements of the same ratio cost without affecting the total energy. Similarly removing areas of the same ratio cost does not decrease the energy, which implies that the model has no shrinking bias. This is in contrast to the minimum surface model in (1), where the energy would simply double (or halve) when adding (or removing) same cost surface elements.

In the context of multiview 3D reconstruction, the absence of shrinking bias implies that the ratio optimization does not have any inherent preference for a specific geometry. Hence, its accuracy does not depend on the presence of protrusions or indentations on the recovered surface.

4.2. Scale Invariance

Now, we will give a precise specification of scale invariance of the minimal ratio model (3), following observations in [10].

**Proposition 2.** For any arbitrary surface estimate $S$ and scaled version $S' = \gamma S$ with $\gamma > 0$, the ratio energy remains unchanged, i. e. $E(S) = E(S')$, provided that the data remains fixed:

$$F(x) = F(\gamma x)$$

$$\rho(x) = \rho(\gamma x).$$

**Proof.** Via change of variables, we obtain

$$E(S') = \frac{\int_{S'} \langle N_{S'}(x), F(x) \rangle \, dS'}{\gamma \int_{S'} \rho(x) \, dS'}$$

$$= \frac{\gamma \int_S \langle N_S(x), F(x) \rangle \, dS}{\gamma \int_S \rho(x) \, dS} = E(S).$$

The above property of the model is particularly useful when applying a multiresolution scheme. In that case, the scale invariance guarantees that a correct solution is computed at each resolution level without the cumbersome need to adjust respective parameters.
4.3. Continuous Optimization

While the above observations indicate that it is worthwhile studying minimal ratio formulations, the major computational challenge is to actually solve the optimization problem (3). In the following, we will propose an optimal solution which is based on sequential convex optimization.

Using the divergence theorem, we can derive the following equivalent formulation:

\[
E(S) = \frac{\int_{S_{int}} \text{div} F \, d^3x}{\int_S \rho(x) \, dS},
\]

s. t. \( \pi_i(S) = S_i \ \forall \ i = 1, \ldots, n, \)

where \( S_{int} \) denotes the interior of \( S \). Conversion to an implicit representation \( u = 1_{S_{int}} \), where \( 1_{S_{int}} \) denotes the characteristic function of \( S_{int} \), and relaxing the resulting binary condition yields:

\[
E(u) = \frac{\int_V \text{div} F \cdot u(x) \, d^3x}{\int_V \rho(x) |\nabla u(x)| \, d^3x},
\]

s. t. \( u \in [0, 1] \) \hspace{1cm} \hspace{1cm} (5)

\[
\int_{R_{ij}} u(x) \, dR_{ij} \geq 1 \text{ if } j \in S_i \text{ and } \int_{R_{ij}} u(x) \, dR_{ij} = 0 \text{ if } j \notin S_i,
\]

where \( R_{ij} \) denotes again the visual ray through pixel \( j \) of image \( i \). The constrained ratio minimization problem in (5) can be solved via the Dinkelbach’s method for fractional optimization [6]. It consists of sequentially minimizing

\[
G(u, \lambda) = \int_V \text{div} F \cdot u(x) \, d^3x - \lambda \int_V \rho(x) |\nabla u(x)| \, d^3x,
\]

s. t. \( u \in [0, 1] \)

\[
\int_{R_{ij}} u(x) \, dR_{ij} \geq 1 \text{ if } j \in S_i \text{ and } \int_{R_{ij}} u(x) \, dR_{ij} = 0 \text{ if } j \notin S_i
\]

(6)

for different values of \( \lambda \). Note that for a fixed parameter \( \lambda \) (6) exhibits a classical constrained convex optimization problem, i. e. a convex functional over a convex domain [13]. Hence, for each subproblem the global minimum can be computed efficiently.

In total, we have the following minimization procedure:

1. Initialize: pick \( u \) arbitrary and set \( \lambda = E(u) \).
2. Compute minimizer \( u^* \) of \( G(\cdot, \lambda) \).
3. Set \( u := u^* \) and \( \lambda := E(u^*) \).
4. If \( \lambda \) has decreased go to step (1), otherwise stop.

Now, we will prove the correctness of the algorithm.

**Proposition 3.** The above optimization procedure computes a (global) minimum of (5).

**Proof.** As before, we denote \( N(u) = \int_V \text{div} F \cdot u(x) \, d^3x \) and \( D(u) = \int_V \rho(x) |\nabla u(x)| \, d^3x \). Upon convergence, the method gives a solution \( u_{min} \) and a ratio \( \lambda_{min} \) such that \( \lambda_{min} = E(u_{min}) \). Convergence implies that \( G(u_{min}, \lambda_{min}) = 0 \) and \( u_{min} = \arg \min G(u, \lambda_{min}) \). From these two statements it follows

\[
0 = N(u_{min}) - \lambda_{min}D(u_{min}) \leq N(u) - \lambda_{min}D(u)
\]

for all \( u \) that fulfill the constraints in (5). Hence, we obtain

\[
\lambda_{min} \leq \frac{N(u)}{D(u)} = E(u)
\]

for all feasible functions \( u \), thus \( \lambda_{min} \) is the optimal ratio. \( \square \)

So far, we have discussed the minimization of the relaxed problem (5). Since we are interested in finding minimizers of the original non-convex problem (4), a straightforward methodology is to threshold the solution of the convex problem \( u_{min} \) appropriately in order to obtain a binary characteristic function \( \tilde{u} = 1_{S_{int}} \) for the surface interior \( S_{int} \):

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

(7)

where

\[
\mu = \min \left\{ \left( \min_{i \in \{1, \ldots, n\}} \max_{j \in S_i} u_{min}(x) \right), 0.5 \right\}
\]

(8)

This choice of threshold assures that the computed binary solution still fulfills exact silhouette consistency. Note that minimizing (4) is equivalent to minimizing the “binaryized” version of (5) (where \( u \in [0, 1] \) is replaced by \( u \in \{0, 1\} \)). Although this approach does not guarantee finding the global minimum of (4), it entails certain globality guarantees:

**Proposition 4.** Let \( u^* \) be a (global) minimum of the “binary” version of (5), \( \tilde{u} \) the computed solution and \( u_{min} \) a (global) minimum of (5). Then, a bound \( \gamma(u_{min}, \tilde{u}) \) exists such that

\[
E(\tilde{u}) - E(u^*) \leq \gamma(u_{min}, \tilde{u}).
\]
Since the binary functions are a subset of the real-valued functions, we have the relation

\[ E(u_{min}) \leq E(u') \leq E(\bar{u}) \]

As a consequence, we obtain the inequality

\[ E(\bar{u}) - E(u') \leq E(\bar{u}) - E(u_{min}) =: \gamma(u_{min}, \bar{u}). \]

\[ \square \]

5. Implementation

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

5.1. Data Terms

Following the formulation in (3), we need to define two data measures: a photoconsistency metric \( \rho : V \to [0, 1] \) and an inward normal field \( F : V \to \mathbb{R}^3 \).

The photoconsistency estimation that we used in our experiments is based on the voting scheme proposed in [7]. The choice of this technique was motivated by its robustness even without requiring explicit visibility information and increased accuracy compared to traditional methods. See [7] for more details.

Instead of directly estimating a normal field \( F \) representing the shape orientation, we compute a regional term \( \rho_{int} \) assigned to the interior of the surface, which can be interpreted as divergence of a corresponding vector field (see (4)). In order to obtain meaningful orientations, \( \rho_{int} \) should be defined only in a vicinity of the surface with negative values inside and positive values outside of it. A respective vector field \( F : V \to \mathbb{R}^3 \) with \( \text{div} F = \rho_{int} \) can be derived as

\[
F_1(x_1, x_2, x_3) = \frac{1}{3} \int_0^{x_1} \rho_{int}(x_1', x_2, x_3) \, dx_1' \\
F_2(x_1, x_2, x_3) = \frac{1}{3} \int_0^{x_2} \rho_{int}(x_1, x_2', x_3) \, dx_2' \\
F_3(x_1, x_2, x_3) = \frac{1}{3} \int_0^{x_3} \rho_{int}(x_1, x_2, x_3') \, dx_3'.
\]

where \( F = (F_1 \, F_2 \, F_3)^T \). Note that this definition is unique up to the addition of a divergence-free vector field. In our implementation, we used the approach of [14] to compute \( \rho_{int} \). The key idea is to assign regional costs to each point in space based on the location of maximal photoconsistency along the viewing rays passing through it. See [14] for more details.

A real example of estimated data volumes \( \rho \) and \( \rho_{int} \) is depicted in Fig. 2.

5.2. Numerics

As mentioned previously, the minimization of (5) poses a classical constrained convex optimization problem. 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 our implementation, we used a fixed-point iteration approach based on Successive Overrelaxation (SOR). See [15] for more details.

6. Experiments

We validate our approach on two real image sequences of weakly textured objects containing deep indentations, shown in Fig. 3 and 4. We compare the silhouette-constrained minimal surface model, introduced in [13], to the proposed ratio optimization. The first sequence is part of a recent multiview stereo evaluation project [18] and captures a textureless dinosaur figurine. Three of the input images and multiple views of the reconstructions with both models are shown in Fig. 3. The data set seems to be a very challenging test scenario for the minimal surface model, which produces clear oversmoothing effects by filling in deep concavities (e.g. at the legs). In contrast, the minimal ratio model, which is free of shrinking bias, accurately recovers the complete geometry. Note that none of the compared models uses a weighting parameter, which allows to control the amount of desired smoothing. On that condition, the minimal surface model strongly depends on the accuracy of the estimated photoconsistency measure and on the geometry of the recovered shape. In contrast, the success of its opponent is not affected by the particular structure of the reconstructed object or the spatial scale chosen. The above observations are confirmed by the second experiment, shown in Fig. 4. Once again, the minimal ratio model...
Figure 3. “dinoRing” sequence. 3 of 48 input images of resolution $480 \times 640$ and multiple views of the reconstructions with the silhouette-constrained minimal surface model, introduced in [13], and the proposed ratio optimization. Note that the minimal surface model produces clear oversmoothing effects by filling in deep concavities (e.g. at the legs). In contrast, the minimal ratio model, which is free of shrinking bias, accurately recovers the complete geometry.

Figure 4. Beethoven sequence. 2 of 33 input images of resolution $1024 \times 768$ and two views of the reconstructions with the silhouette-constrained minimal surface model, introduced in [13], and the proposed ratio optimization. Analogously, in contrast to the minimal ratio model, the minimal surface model tends to oversmooth concavities (e.g. at the cheeks or under the chin) due to the presence of shrinking bias.

achieves a higher grade of accuracy, especially at deep indentations (e.g. at the cheeks or under the chin).

The proposed ratio optimization procedure was implemented for a GPU by using a red-black strategy and evaluated on a PC with 2.8 GHz and 4 GB of main memory, equipped with a NVIDIA GeForce 280 GTX graphics card.

For both demonstrated experiments, we measured computational times in the range 2-3 minutes. Note that the computation of the data terms is not included in these runtimes.
7. Conclusion

We proposed a new approach for continuous ratio optimization based on relaxation and sequential convex optimization. The continuous formulation entails a series of advantages compared to discrete counterparts developed in the context of minimum ratio cycles or the parametric maximum flow framework like parallelizability, spatial consistency and reduced memory requirements. Based on the novel ratio optimization technique, we designed a new approach for multiview 3D reconstruction integrating multiview stereo and silhouette information. In a theoretical investigation, we proved that, in contrast to the classical minimal surface model, the minimal ratio model is free of shrinking bias while retaining properties like regularity and globality guarantees. They are confirmed by experiments on challenging real data sets capturing weakly textured objects of complex geometry.

Acknowledgments

We thank Thomas Schoenemann, Thomas Pock and Alan Yuille for fruitful discussions on ratio optimization. This work was supported by the German Research Foundation DFG CR-250/1-2 and DFG CR-250/4-1.

References