

Contents

Contents 1

1 Convex Relaxation Techniques for Segmentation, Stereo and Multiview Reconstruction	2
1.1 Variational Methods, Partial Differential Equations and Convexity	2
1.2 Image Segmentation and Minimal Partitions	4
1.2.1 Classical Variational Approaches	4
1.2.2 A General Variational Formulation	5
1.2.3 Convex Representation	6
1.2.4 Convex Relaxation	8
1.2.5 Experimental Segmentation Results	9
1.3 Stereo Reconstruction	10
1.3.1 Experimental Stereo Results	11
1.4 Multiple View Reconstruction	13
1.4.1 Experimental Multiview Results	17
1.5 Summary and Conclusion	18

Bibliography 20

1

Convex Relaxation Techniques for Segmentation, Stereo and Multiview Reconstruction

Daniel Cremers, Technische Universität München, Germany

Thomas Pock, Graz University of Technology, Austria

Kalin Kolev, University of Bonn, Germany

Antonin Chambolle, Ecole Polytechnique, Paris, France

1.1 Variational Methods, Partial Differential Equations and Convexity

Digital images are discrete and hence it appears quite intuitive to revert to a spatially discrete representation and Markov random fields for modeling and solving problems in image processing and computer vision. Such discrete graphical representations of the computational domain have become very popular due to a multitude of highly efficient combinatorial algorithms for solving problems like shortest paths or minimum cuts on graphs that have been emerging since the 1950's.

Nevertheless, the world that is captured in digital images is not spatially discrete, and ideally algorithms to process images should be independent of the choice of the underlying grid on which they are sampled. Unfortunately, for many of the classical graph algorithms this is not the case. For example, when searching the shortest path from the lower left corner to the upper right corner of a unit square, the classical algorithm of Dijkstra for computing shortest paths will give a distance of 2 when applied to a regular 4-connected grid, although the Euclidean distance is clearly $\sqrt{2}$. The algorithm is not *consistent* in the sense that the numerical error does not go to zero when increasing the resolution of the graph. In fact, this concept of *continuum limit* is rarely considered in graph theoretic approaches.

This chapter is focused on describing recent developments in the theory of variational methods and partial differential equations which

make these a powerful alternative to Markov random field approaches for solving a variety of computer vision problems. While Markov random fields are inherently based on a discrete graph representations, variational methods are based on a representation of images as continuous-valued functions $I : \Omega \rightarrow \mathbb{R}^n$ on spatially continuous domains $\Omega \subset \mathbb{R}^d$. Similarly solutions to respective vision problems such as image segmentation or stereo and multiview reconstruction can be represented by respective functions on continuous domains.

A solution u to a given computer vision problem is determined by minimizing an appropriate functional $E(u)$. A necessary condition for optimality of E is given by the *Euler-Lagrange equation* which simply states that the variation of E with respect to u must vanish.

The last decades have brought about a number of break-throughs in the application of variational methods for computer vision, among others the variational approach of Horn and Schunck [12] for computing optical flow fields from pairs of images, the segmentation methods of Kass et al. [14] and of Mumford and Shah [23] with respective level set formulations of Caselles et al. [4], of Kichenassamy et al. [15] and of Chan and Vese [7], and level set formulations for 3D reconstruction from multiple views pioneered by Faugeras and Keriven [10].

Unfortunately, none of the above variational methods are based on convex functionals. As a consequence, solutions will merely correspond to *local* minima of the respective functional and typically depend on appropriate initializations. Since there exists no reliable procedure to compute configurations with any kind of optimality guarantee, the practical usefulness of such approaches is limited.

This chapter will provide a variety of recently developed convex relaxation techniques which allow to cast respective computer vision problems in terms of convex functionals. As a consequence, one can compute globally optimal solutions (or solutions with bounded optimality) that are independent of the initialization. Experimental comparison shows that the resulting PDE-based solutions typically require less memory, are substantially faster and provide more accurate solutions for the underlying vision problem than corresponding state-of-the-art graph-theoretic algorithms.

Central ideas presented in this book chapter were developed in various conference and journal papers, in particular [6, 22, 3, 21, 18, 5, 26, 30, 24]. The reader is referred to these works for further details.

1.2 Image Segmentation and Minimal Partitions

1.2.1 Classical Variational Approaches

One of the first areas of application for variational methods was that of image segmentation, where the goal is to partition the image plane into a set of meaningful regions. Among the most influential variational approaches were the approaches of Kass et al. [14], of Blake and Zisserman [1], and of Mumford and Shah [23]. They are complementary in the sense that the former (often called *edge-based* segmentation method) aims at identifying boundaries in the image that are supported by strong intensity gradients, whereas the latter two approaches (often called *region-based* segmentation methods) aims at identifying regions of smooth (or homogeneous) intensity.

Kass et al. [14] suggested to compute a segmentation of an image $I : \Omega \rightarrow \mathbb{R}$ on the domain $\Omega \subset \mathbb{R}^2$ in terms of a boundary C by minimizing a functional of the form

$$E(C) = \int_C -|\nabla I(C(s))|^2 + \alpha|C_s(s)|^2 + \beta|C_{ss}(s)|^2 ds \quad (1.1)$$

where the first term favors the boundary to lie in areas of strong intensity gradient, whereas the last two terms (weighted by parameters α and β) impose a certain regularity of the boundary.

Mumford and Shah [23] suggested to compute a piecewise smooth approximation u of the intensity function I by minimizing the functional

$$E(U, C) = \int_{\Omega} (u - I)^2 dx + \lambda \int_{\Omega - C} |\nabla u|^2 dx + \nu |C|. \quad (1.2)$$

While the first term imposes pointwise similarity of the approximation u with the input image I , the second term (weighted by λ) imposes smoothness of u everywhere except at the boundary C , the length of which is penalized with a weight ν . In the limit $\lambda \rightarrow \infty$ one obtains the case of a piecewise constant approximation.

One of the central algorithmic challenges addressed in this chapter is how to minimize such types of functionals. While parametric boundary representations [9] or implicit level set representations [7, 4, 15] typically only find locally optimal solutions with little or no optimality guarantees, recently developed convex relaxation schemes provide solutions which are either optimal or within a bound of the optimum. Some of these developments shall be detailed in the following.

1.2.2 A General Variational Formulation

Let $I : \Omega \rightarrow \mathbb{R}$ be a gray value input image on the domain $\Omega \subset \mathbb{R}^d$. Among a variety of variational approaches to image segmentation, let us consider the following rather general model:

$$\min_{\Omega_i} \left\{ \frac{1}{2} \sum_{i=0}^k \text{Per}_g(\Omega_i; \Omega) + \sum_{i=0}^k \int_{\Omega_i} f_i(x) dx \right\}, \quad (1.3)$$

such that $\bigcup_{i=0}^k \Omega_i = \Omega$, $\Omega_s \cap \Omega_t = \emptyset \forall s \neq t$,

Minimizing (1.3) partitions the domain $\Omega \subset \mathbb{R}^d$ into $k+1$ pairwise disjoint sets Ω_i . The first term imposes regularity of solutions. It measures the perimeter of the set Ω_i with respect to a metric defined by the nonnegative function $g(x)$.¹ The second term is the data term which is based on non-negative weight functions $f_i : \Omega \rightarrow \mathbb{R}^+$.

Model (1.3) includes as a special case the piecewise constant Mumford-Shah functional [23] discussed above, which arises when choosing

$$f_i(x) = (I(x) - c_i)^2,$$

which is the squared difference of the input image I to some mean intensity c_i . More generally, one can choose

$$f_i(x) = -\log P_i(I(x))$$

as the negative log likelihood for observing a certain intensity or color value [33, 16, 8]. Model (1.3) also includes as a special case edge-based segmentation approaches such as the geodesic active contours [4, 15] where a spatially inhomogeneous metric favors boundaries passing through areas of strong gradient by choosing for example

$$g(x) = \frac{1}{1 + |\nabla I(x)|}.$$

1. For simplicity we will only consider *isotropic* metrics. The approach is easily generalized to *anisotropic* metrics where boundaries favor certain orientations.

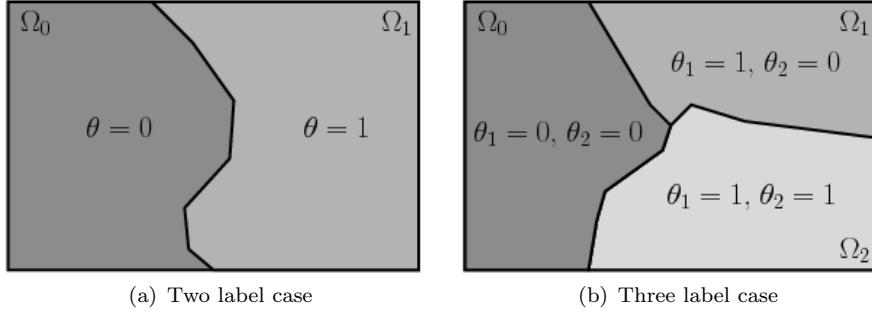


Figure 1.1 (a) One binary function θ is used to partition the image domain Ω into two regions. (b) Two binary functions $\theta_1 \geq \theta_2$ are used to partition the image domain Ω into three regions.

The discrete analogue of (1.3) is the *Potts model* [27] which is known to be NP hard. Several algorithms have been proposed to approximately minimize the Potts model. While the discrete problem can be tackled using iterated binary optimization via α -expansion [2] or roof duality relaxation [11, 28], such algorithms tend to exhibit a grid bias (metrication errors) in representing the continuous perimeters in (1.3). In the continuous domain, popular methods are based on level set methods [7, 4, 15] or parametric boundary representations [9]. The most crucial drawback of these methods is, that there is no guarantee to find globally optimal solutions.

1.2.3 Convex Representation

In the following, the $k + 1$ regions Ω_i in (1.3) are represented by a labeling function $u : \Omega \rightarrow \{0, \dots, k\}$, where $u(x) = l$ if and only if $x \in \Omega_i$. One can equivalently represent this multilabel function by k binary functions $\vec{\theta}(x) = (\theta_1(x), \dots, \theta_k(x))$ defined by

$$\theta_i(x) = \begin{cases} 1 & \text{if } u(x) \geq l \\ 0 & \text{otherwise} \end{cases}, \quad (1.4)$$

representing its upper level sets. In turn, the labeling function u can be recovered from these functions via the relation

$$u(x) = \sum_{i=1}^k \theta_i(x). \quad (1.5)$$

Figures 1.1(a) and 1.1(b) show examples of partitionings with one or two binary functions θ_i .

A one-to-one correspondence between multilabel functions $u(x)$ and vectors $\vec{\theta} = (\theta_1, \dots, \theta_k)$ of binary functions is guaranteed by constraining $\vec{\theta}$ to the ordered set

$$\mathcal{B} = \left\{ \vec{\theta} : \Omega \rightarrow \{0, 1\}^k, 1 \geq \theta_1(x) \geq \dots \geq \theta_k(x) \geq 0, \forall x \in \Omega \right\}. \quad (1.6)$$

How can one rewrite the optimization problem (1.3) in terms of the binary functions θ_i ? Let us start with the simple case of two regions ($k = 1$), where the region Ω_1 is defined by a single binary function θ_1 . The perimeter of the set Ω_1 is given by the weighted total variation of θ_1 :

$$\text{Per}_g(\Omega_1; \Omega) = \int g |D\theta_1| = \sup_{|\xi_1| \leq g} \int \xi_1 D\theta_1 = \sup_{|\xi_1| \leq g} - \int \theta_1 \operatorname{div} \xi_1 dx, \quad (1.7)$$

where $D\theta$ denotes the distributional derivative of θ . For differentiable functions θ it is simply given by $D\theta = \nabla \theta dx$. For binary-valued functions $\theta \in BV(\Omega)$ it is a surface measure supported on the boundary of the set $\{\theta = 1\}$. The second equality in (1.7) expresses this boundary length using the dual vector field $\xi_1 : \Omega \rightarrow \mathbb{R}^2$. These dual vector fields in fact play the role of the flow in the analogous graph cut approaches and Fenchel duality leads to a spatially continuous version of the Min Cut / Max Flow equivalence. We can now use standard LP-relaxation and let $\theta_1 \in [0, 1]$, then the *coarea formula* will ensure that the relaxed problem has the same solutions as the initial one.

The extension of this formulation to the case of multiple regions in (1.3) is not straight-forward. Simply summing the total variations of each function θ_i – as done in the analogous level set formulation of Chan and Vese [7] – would imply that certain boundaries are counted more than once: For the example shown in Figure 1.1(b), the boundary between Ω_0 and Ω_2 would be counted twice.

This ‘overcounting’ of boundaries can be elegantly suppressed by appropriate constraints that couple the dual variables $\vec{\xi} = (\xi_0, \dots, \xi_k)$.

Proposition 1.1

The optimization problem (1.3) is equivalent to the problem:

$$\min_{\vec{\theta} \in \mathcal{B}} \max_{\vec{\xi} \in \mathcal{K}} \left\{ \sum_{i=0}^k - \int_{\Omega} \theta_i \operatorname{div} \xi_i dx + \int_{\Omega} (\theta_i(x) - \theta_{i+1}(x)) f_i(x) dx \right\}, \quad (1.8)$$

with a set $\vec{\xi} = (\xi_0, \dots, \xi_k)$ of dual vector fields $\xi_i : \Omega \rightarrow \mathbb{R}^2$, constrained to the set

$$\mathcal{K} = \left\{ \vec{\xi} : \Omega \rightarrow \mathbb{R}^{d \times k}, \left| \sum_{i_1 \leq i \leq i_2} \xi_i(x) \right| \leq g(x), \forall x \in \Omega, 1 \leq i_1 \leq i_2 \leq k \right\}. \quad (1.9)$$

PROOF. For a proof the reader is referred to [5]. \square

The constraints on the dual variables $\xi_i(x)$ in (1.9) assure that each interface is counted exactly once. For the three-region case shown in Figure 1.1(b), for example, the above constraint implies that $|\xi_1(x) + \xi_2(x)| \leq 1$. This assures that the transition between Ω_0 and Ω_2 is counted once only. Interestingly this coupling constraint ties nicely into the subsequent convex optimization technique.

Proposition 1.2

The set \mathcal{K} defined in (1.9) is convex.

PROOF. For any two functions $\vec{\xi}, \vec{\xi}' \in \mathcal{K}$ and any $\alpha \in [0, 1]$ we have that $\alpha\vec{\xi} + (1 - \alpha)\vec{\xi}' \in \mathcal{K}$:

$$\left| \sum_{i_1 \leq i \leq i_2} \alpha \xi_i(x) + (1 - \alpha) \xi'_i(x) \right| \leq \alpha \left| \sum_{i_1 \leq i \leq i_2} \xi_i(x) \right| + (1 - \alpha) \left| \sum_{i_1 \leq i \leq i_2} \xi'_i(x) \right| \leq g(x). \quad (1.10)$$

\square

1.2.4 Convex Relaxation

Unfortunately, the overall optimization problem (1.8) is non-convex because the set \mathcal{B} defined in (1.6) is not convex. We therefore propose a convex relaxation which allows the functions θ_i to take on intermediate values between 0 and 1. To this end the set \mathcal{B} in the optimization problem (1.8) is replaced by the convex set:

$$\mathcal{R} = \left\{ \vec{\theta} : \Omega \rightarrow [0, 1]^k, 1 \geq \theta_1(x) \geq \dots \geq \theta_k(x) \geq 0, \forall x \in \Omega \right\}. \quad (1.11)$$

For $k = 1$ this formulation turns out to be equivalent to the two-region problem considered by Chan et al. [6] for which we have the following optimality guarantee:

Proposition 1.3

Let θ_1 be the optimum of the relaxed (convex) version of (1.8) for $k = 1$ with \mathcal{B} replaced by \mathcal{R} . Then threshold the solution θ_1 at any value $s \in (0, 1)$ will provide a global solution of the original non-convex labeling problem (1.8).

PROOF. For a proof, the reader is referred to [6]. \square

While this thresholding theorem does not extend to the general problem of more than two regions ($k > 1$), one can prove the following optimality bound.

Proposition 1.4

Let $\vec{\theta}^* \in \mathcal{R}$ be the solution of the relaxed version of (1.8) and let $\mathbf{1}_{\{\vec{\theta}^* \geq s\}} \in \mathcal{B}$ be a thresholded binary version for any $s \in [0, 1]$. Furthermore, let $\vec{\theta}' \in \mathcal{B}$ be the true global minimizer of the binary problem (1.8). Then one can provide the following bound on the energy of the computed solution.

$$|\mathcal{E}(\mathbf{1}_{\{\vec{\theta}^* \geq s\}}) - \mathcal{E}(\vec{\theta}')| \leq |\mathcal{E}(\vec{\theta}^*) - \mathcal{E}(\mathbf{1}_{\{\vec{\theta}^* \geq s\}})|. \quad (1.12)$$

PROOF. The bound follows directly, because in terms of their energies the optimal binary solution $\vec{\theta}'$ lies between the relaxed solution $\vec{\theta}^*$ and the thresholded one $\mathbf{1}_{\{\vec{\theta}^* \geq s\}}$:

$$\mathcal{E}(\vec{\theta}^*) \leq \mathcal{E}(\vec{\theta}') \leq \mathcal{E}(\mathbf{1}_{\{\vec{\theta}^* \geq s\}}). \quad (1.13)$$

\square

In many real world experiments this bound is actually zero or near zero, such that for these examples solutions are essentially optimal.

1.2.5 Experimental Segmentation Results

The convex relaxation framework for image segmentation introduced above is closely related to the theory of minimal surfaces. Figure 1.2 shows examples of minimal surface problems where the data term in (1.3) is switched off and the three or four colors are imposed as boundary constraints.

Figure 1.3 shows image segmentations computed with the proposed convex relaxation technique for model (1.3) with 4 and 8 labels, respectively.

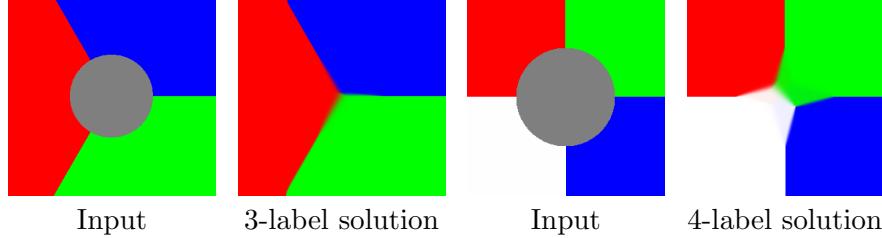


Figure 1.2 Three and four label solutions for predescribed boundary colors computed for the round area in the center with model (1.3) and no data term. The proposed relaxation scheme allows to accurately approximate triple junctions that are known to be the analytically optimal configurations of the respective minimal partition problems.

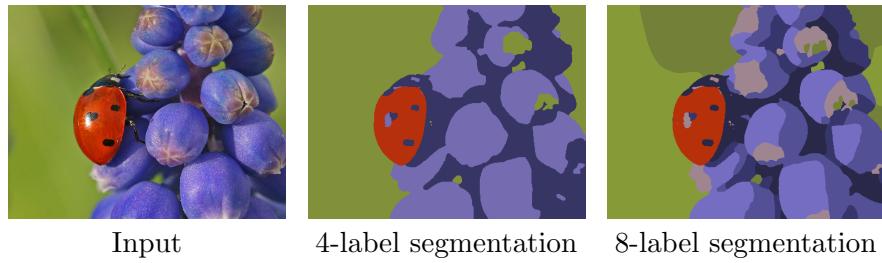


Figure 1.3 Image segmentations computed using the convex relaxation of model (1.3) with 4 and 8 labels.

1.3 Stereo Reconstruction

Characteristic for the segmentation model in (1.3) is that transitions in the labeling function $u : \Omega \rightarrow \{0, \dots, k\}$ are penalized in a manner that is independent of the size of transition. In stereo reconstruction, where $u(x)$ denotes the depth at a given image point $x \in \Omega$, one may want a penalty which favors spatially smooth depth fields. Moreover a data term $f(u, x)$ will favor different depth values for different points, for example based on the normalized cross-correlation of respective patches in either of the two input images. For a simple linear smoothness constraint, the resulting optimization problem is given by:

$$\min_{u: \Omega \rightarrow \{0, \dots, k\}} \int_{\Omega} f(u(x), x) dx + \int g(x) |Du|, \quad (1.14)$$

Due to the data term f , functional (1.14) is not convex. Yet, it turns out that one can replace this nonconvex approach with a corresponding convex one relying once more on the representation of the multi-label function u in terms of a vector of binary functions θ_i – see equation (1.4).

Proposition 1.5

The minimization problem (1.14) is equivalent to the minimization problem

$$\min_{\theta \in \mathcal{B}} \left\{ \sum_{i=0}^k \int_{\Omega} g|D\theta_i| + \int_{\Omega} f(i, x) (\theta_i(x) - \theta_{i+1}(x)) dx \right\}. \quad (1.15)$$

PROOF. For a complete proof the reader is referred to [5, 26]. \square

Moreover, one can compute optimal solutions to the original problem (1.14) by means of convex relaxation and thresholding.

Proposition 1.6

The minimization problem (1.15) can be solved globally by relaxation to a convex problem (replacing the domain \mathcal{B} by its convex hull \mathcal{R} defined in (1.11), solving the convex problem and thresholding the solution.

PROOF. For a complete proof the reader is referred to [5, 26]. \square

The above relaxation approach thus allows to optimally solve the original non-convex multi-label problem (1.14) by reducing it to the convex problem (1.15). A formulation with a continuous label space was developed in [26]. It shows that the proposed solution amounts to an anisotropic minimal surface problem. Using appropriate combinations of the convex constraints imposed in approaches (1.8) and (1.15) one can generalize this formulation to truncated linear potentials.

1.3.1 Experimental Stereo Results

The convex relaxation approach introduced above can be seen as the spatially continuous analogue of the discrete algorithm proposed by Ishikawa [13]. Figure 1.4 shows a comparison of stereo reconstructions computed with the proposed approach compared to respective results obtained with Ishikawa’s graph cut formulation obtained with 4- and 8-connected neighborhoods. Closeups show that the proposed spatially continuous solution does not exhibit any grid bias (metrication errors). Experimental comparisons shows that in comparison to discrete graph cut methods the proposed continuous shape optimization techniques

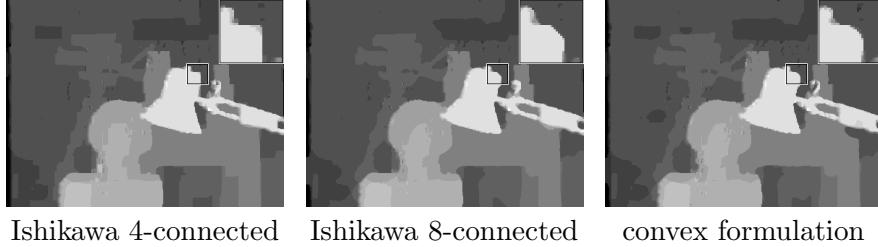


Figure 1.4 Absence of grid bias /metrication error in stereo reconstructions. In contrast to the spatially discrete approach of Ishikawa [13] (shown here for 4- and 8-connected neighborhood), the continuous solution based on convex relaxation does not favor solutions aligned to the underlying grid, such as 90 degree or 45 degree angles visible in the close-ups.

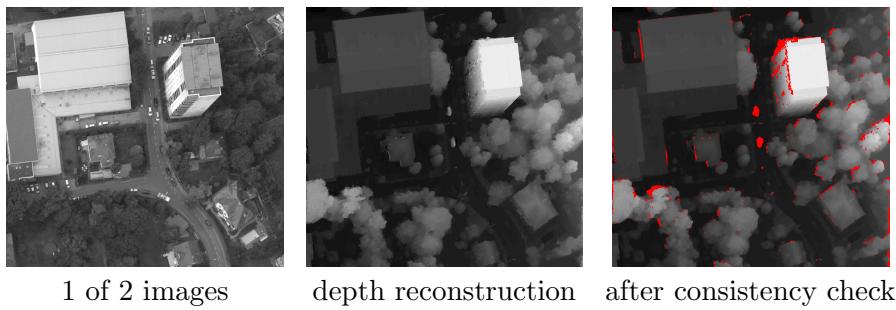


Figure 1.5 Depth reconstruction (brightness coded) compute from two aerial images of 1500×1400 pixels. While the center image shows the global minimum of the functional (1.14) computing using convex relaxation, the image on the right shows the same solution after consistency check which eliminates occlusion areas and areas of independently moving objects like the cars (colored in red).

typically provide more accurate solutions while requiring substantially less memory and lower computation times – see [17] for a detailed study.

Due to substantially reduced memory requirements, the algorithm can be applied to higher resolution image data. Figure 1.5 shows the depth reconstruction computed with the convex relaxation technique for a pair of aerial images of resolution 1500×1400 pixels. The brightness-encoded depth values clearly show fine scale details such as trees, cars, lamp posts and chimneys.

1.4 Multiple View Reconstruction

The reconstruction of three-dimensional shapes from a collection of calibrated images is among the classical challenges in computer vision. Rather than estimating point correspondences among pairs of images and triangulating these, a popular alternative for computing stereo-based reconstructions from multiple views pioneered by Faugeras and Keriven [10] is to directly compute a reconstruction as minimal weighted surfaces S in the volume $\Omega \subset \mathbb{R}^3$ based by solving the optimization problem

$$\min_S \int_S \rho(x) dA(x), \quad (1.16)$$

where $\rho : (\Omega \subset \mathbb{R}^3) \rightarrow [0, 1]$ is a *photoconsistency measure*. Based on the assumption of a Lambertian surface, $\rho(x)$ takes on small values if the projection of voxel x into pairs of images gives rise to similar color observations (i.e. the voxel is likely to be on the surface), while high values of $\rho(x)$ indicate that the colors observed in pairs of images are very different (i.e. the voxel x is likely not to be on the surface). Thus minimizing (1.16) gives rise to maximally photoconsistent surfaces.

The functional (1.16) has two important drawbacks: Firstly it is not convex such that computing good quality surfaces is not straightforward. Secondly, the global minimum of (1.16) is evidently the empty set which has zero energy while any other solution clearly has non-negative energy. In practice this latter drawback can be alleviated by either reverting to local optimization techniques such as the level set method [10] or by constraining the search space to some band around the visual hull [31].

Two alternative methods to remove the above problems shall be sketched in the following.

Solution 1: Volumetric photoconsistencies.

One can extend the functional by additional volumetric integrals over the interior $\text{int}(S)$ and exterior $\text{ext}(S)$ of the surface S :

$$\min_S \int_S \rho(x) dA(x) + \int_{\text{int}(S)} \rho_{int}(x) dx + \int_{\text{ext}(S)} \rho_{ext}(x) dx, \quad (1.17)$$



Figure 1.6 Stereo-based multiview reconstruction via convex relaxation. Two of 33 input images of resolution 1024×768 and three views of the reconstructed surface at volume resolution $216 \times 288 \times 324$. This solution corresponds to a global minimum of (1.18) obtained by convex relaxation and thresholding.

with appropriately defined regional photoconsistencies ρ_{int} and ρ_{ext} to model the log-likelihood that a voxel is inside or outside the surface. For details on the computation of these functions the reader is referred to [21].

Representing the surface S by a binary labeling function $\theta : \Omega \rightarrow \{0, 1\}$, where $\theta(x) = 1$ if $x \in \text{int}(S)$ and $\theta(x) = 0$ otherwise, the minimization problem (1.17) is equivalent to

$$\min_{\theta: \Omega \rightarrow \{0, 1\}} \int_{\Omega} \rho(x) |D\theta| + \int_{\Omega} \theta(x) (\rho_{int}(x) - \rho_{ext}(x)) dx \quad (1.18)$$

This turns out to be the two-region case of the model presented in Section 1.2.4. It can be solved optimally by minimizing the convex relaxation and thresholding of the minimizer.

Figure 1.6 shows reconstructions computed from multiple images of a bunny.

Solution 2: Imposing Silhouette Consistency.

What makes the empty set not a good reconstruction for a given image set is that its projection into each image does not match the observed silhouettes. Let us assume that we are also given the silhouettes $S_i : \Omega_i \rightarrow \{0, 1\}$ of the observed object in each image $\Omega_i \subset \mathbb{R}^2$ for $i = 1, \dots, n$, where $S_i(x) = 1$ if x is inside the object's silhouette. In reconstruction problems with known or homogeneous background these

are typically rather straight-forward to obtain beforehand. The following formulation allows to combine stereo and silhouette information via convex functionals over convex sets.

A silhouette consistent optimally photoconsistent reconstruction can then be computed by solving the constrained optimization problem

$$\begin{aligned} & \min_S \int_S \rho(x) dA(x), \\ \text{s.t. } & \pi_i(S) = S_i, \forall i = 1, \dots, n \end{aligned} \tag{1.19}$$

where π_i denotes the projection into image i .

As above, one can revert to an implicit representation of the surface with an indicator function $\theta : \Omega \rightarrow \{0, 1\}$. Problem (1.19) is equivalent to

$$\begin{aligned} & \min_{\theta: \Omega \rightarrow \{0, 1\}} \int_\Omega \rho(x) |D\theta(x)| \\ \text{s.t. } & \int_{R_{ij}} \theta(x) dR_{ij} \geq \delta \text{ if } j \in S_i, \quad \int_{R_{ij}} \theta(x) dR_{ij} = 0 \text{ if } j \notin S_i. \end{aligned} \tag{1.20}$$

Here the parameter $\delta > 0$ denotes a material-dependent constant corresponding to the thickness of material below which the object becomes translucent. In numerical implementations one can set $\delta = 1$. Thus the two constraints in (1.20) simply reflect the silhouette consistency constraint: For any pixel j which is part of the silhouette S_i observed in image i , the visual ray R_{ij} from the camera center through that pixel must cross the object in at least one voxel. On the other hand, for pixels j outside the silhouette S_i the ray R_{ij} through that pixel may not intersect with the object, i.e. the integral of θ along that ray must be zero. See Figure 1.7 for a schematic drawing.

Again, one can perform a relaxation of problem (1.20) by allowing θ to take on intermediate values between 0 and 1:

$$\min_{\theta \in D} \int_\Omega \rho(x) |D\theta(x)|, \tag{1.21}$$

with

$$\mathcal{D} := \left\{ \theta : V \rightarrow [0, 1] \mid \begin{array}{ll} \int_{R_{ij}} \theta(x) dR_{ij} \geq 1 & \text{if } j \in S_i \forall i, j \\ \int_{R_{ij}} \theta(x) dR_{ij} = 0 & \text{if } j \notin S_i \forall i, j \end{array} \right\} \tag{1.22}$$

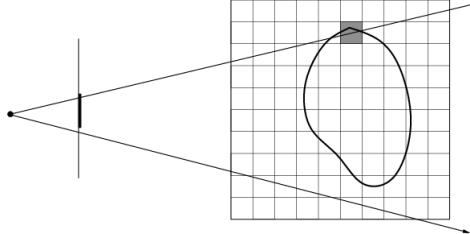


Figure 1.7 Schematic view of the silhouette consistency constraint: For a silhouette consistent reconstruction at least one voxel along each visual ray from the camera center through silhouette pixels (bold area) must be occupied, whereas all voxels along a ray through a non-silhouette pixel must be empty.

It turns out that in the implicit representation, the silhouette constraint nicely ties into the convex optimization framework.

Proposition 1.7

The set \mathcal{D} of all silhouette-consistent functions defined in (1.22) forms a convex set.

PROOF. Let $\theta_1, \theta_2 \in \mathcal{D}$ be two elements of \mathcal{D} . Then any convex combination $\theta = \alpha\theta_1 + (1 - \alpha)\theta_2$ with $\alpha \in [0, 1]$ is also an element in \mathcal{D} . In particular, $\theta(x) \in [0, 1]$ for all x . Moreover,

$$\int_{R_{ij}} \theta dR_{ij} = \alpha \int_{R_{ij}} \theta_1 dR_{ij} + (1 - \alpha) \int_{R_{ij}} \theta_2 dR_{ij} \geq 1 \text{ if } j \in S_i, \quad (1.23)$$

and similarly

$$\int_{R_{ij}} \theta dR_{ij} = \alpha \int_{R_{ij}} \theta_1 dR_{ij} + (1 - \alpha) \int_{R_{ij}} \theta_2 dR_{ij} = 0 \text{ if } j \notin S_i. \quad (1.24)$$

Thus $\theta \in \mathcal{D}$. \square

Since we are interested in minimizers of the non-convex binary labeling problem (1.20), a straightforward methodology is to threshold the solution of the convex problem (1.21) appropriately. Although this will not guarantee finding the global minimum of (1.20), the proposed strategy entails a series of advantages compared to classical local optimization techniques. Extending the set of admissible functions, computing the global minimum over this domain and subsequently projecting to the nearest point within the original set will provide a solution which is independent of initialization. Moreover, one can compute an

upper bound on the energetic deviation of the computed solution from the global minimum.

Proposition 1.8

Let θ^ be a minimizer of the relaxed problem (1.21) and let $\hat{\theta}^*$ be a projection onto the binary silhouette-consistent solutions. Then $\hat{\theta}^*$ is of bounded distance (in terms of its energy E) from the true solution θ' :*

$$|\mathcal{E}(\hat{\theta}^*) - \mathcal{E}(\theta')| \leq |\mathcal{E}(\theta^*) - \mathcal{E}(\hat{\theta}^*)|. \quad (1.25)$$

PROOF. The proof is analogous to that of proposition 1.4. \square

The projection $\hat{\theta}^*$ of a minimizer θ^* onto the silhouette-consistent binary functions can be computed by simple thresholding:

$$\hat{\theta}(x) = \begin{cases} 1, & \text{if } \theta^*(x) \geq \mu \\ 0, & \text{otherwise} \end{cases}, \quad (1.26)$$

where

$$\mu = \min \left\{ \left(\min_{i \in \{1, \dots, n\}, j \in S_i} \max_{x \in R_{ij}} \theta^*(x) \right), 0.5 \right\}. \quad (1.27)$$

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

Proposition 1.9

The computed binary solution exactly fulfills all silhouette constraints.

PROOF. For a proof the reader is referred to [18]. \square

1.4.1 Experimental Multiview Results

Figure 1.8 shows an experimental comparison of the two alternative multiview methods on the reconstruction of a metal head. Due to the strong reflections and highlights, the stereo information becomes unreliable and the resulting reconstruction is substantially deteriorated (left). Incorporating the silhouette constraint allows to substantially improve the reconstruction (right).

The minimal surface formulation suffers from a shrinking bias in the sense that small scale and elongated structures tend to get suppressed (as this leads to smaller surface energy). The silhouette constraint, on the other hand, allows to preserve many small scale structures. Figure 1.9 shows reconstructions from 24 images of a

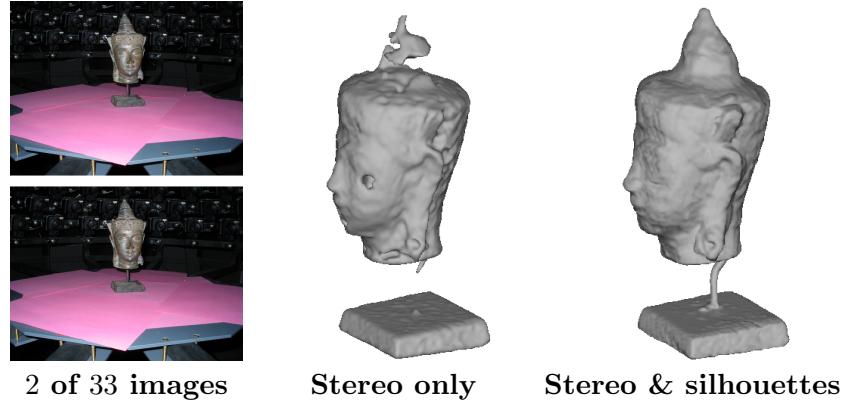


Figure 1.8 Silhouette and stereo integration. While the purely stereo-based reconstruction method (1.17) (middle) tends to remove thin structures and is heavily affected by specular reflections such as those of metal objects (left), the fusion of stereo and silhouette information using approach (1.20) allows to compute stereo-based reconstructions which are guaranteed to be silhouette consistent. As a consequence, concave areas (around the ears) as well as fine geometric details such as the pedestal are preserved in the reconstruction.

warrior statue, computed by solving the constrained optimization problem (1.21). The image data is courtesy of Yasutaka Furukawa – see <http://www.cs.washington.edu/homes/furukawa/research/mview/index.html>. Note that the silhouette constraint allows to preserve fine details like the hammer and the sword.

1.5 Summary and Conclusion

This chapter provided several convex relaxation techniques for central computer vision problems like image segmentation, stereo and multi-view reconstruction. Optimal solutions or solutions of bounded optimality are computed in a spatially continuous representation through the minimization of convex functionals and subsequent projection. In contrast to level set approaches, the presented convex relaxation schemes are independent of initialization and provide solutions with known optimality guarantees. In contrast to graph cut based solutions, the spatially continuous formulation does not suffer from metrification errors – see [17] for a detailed comparison. In addition physical constraints like the silhouette consistency of 3D reconstructions often give



Figure 1.9 Warrior sequence. 2 of 24 input images of resolution 1600×1600 and multiple views of the reconstructed surface. Note that thin structures such as hammer and sword as well as concavities – for example at the chest – are reconstructed accurately.

rise to convex constraints and can therefore be directly imposed in the optimization scheme.

Due to space limitations the numerical solution of the arising convex optimization problems were not detailed. There are many different algorithms, from simple gradient descent to more efficient algorithms based on fixed point iteration and successive over-relaxation [21] or primal-dual algorithms [26, 25]. These algorithms are typically straightforward to parallelize on the pixel or voxel grid. As a consequence, implementations on graphics hardware lead to substantial speed-ups over CPU implementations.

The concept of convex relaxation has become increasingly popular in computer vision over the last year. In addition to the approaches discussed in this chapter, there are further convex relaxation schemes to find optimal solutions to *anisotropic* minimal surface problems [26, 32, 29, 19] or to *ratio functionals* [20]. Recently, the first algorithm for minimizing convex relaxations of the *piecewise smooth Mumford-Shah functional* was proposed in [25]. All these developments indicate that the spatially continuous representation and pde-based optimization techniques provide a powerful alternative to graph-theoretic Markov random field approaches.

Bibliography

- [1] A. Blake and A. Zisserman. *Visual Reconstruction*. MIT Press, 1987.
- [2] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, 2001.
- [3] X. Bresson, S. Esedoglu, P. Vandergheynst, J. P. Thiran, and S. Osher. Fast global minimization of the active contour/snake model. *Journal of Mathematical Imaging and Vision*, 28(2):151–167, 2007.
- [4] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. In *Proc. IEEE Intl. Conf. on Comp. Vis.*, pages 694–699, Boston, USA, 1995.
- [5] A. Chambolle, D. Cremers, and T. Pock. A convex approach for computing minimal partitions. Technical report, Ecole Polytechnique, Paris, France, 2008.
- [6] T. Chan, S. Esedoglu, and M. Nikolova. Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM Journal of Applied Mathematics*, 66(5):1632–1648, 2006.
- [7] T.F. Chan and L.A. Vese. Active contours without edges. *IEEE Trans. Image Processing*, 10(2):266–277, 2001.
- [8] D. Cremers, M. Rousson, and R. Deriche. A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape. *Int. J. Computer Vision*, 72(2):195–215, April 2007.

- [9] D. Cremers, F. Tischhäuser, J. Weickert, and C. Schnörr. Diffusion Snakes: Introducing statistical shape knowledge into the Mumford–Shah functional. *Int. J. Computer Vision*, 50(3):295–313, 2002.
- [10] O. Faugeras and R. Keriven. Variational principles, surface evolution, PDE’s, level set methods, and the stereo problem. *IEEE Trans. Image Processing*, 7(3):336–344, March 1998.
- [11] P. L. Hammer, P. Hansen, and B. Simeone. Roof duality, complementation and persistency in quadratic 0-1 optimization. *Mathematical Programming*, 28(2):121–155, 1984.
- [12] B.K.P. Horn and B.G. Schunck. Determining optical flow. *J. Artificial Intelligence*, 17:185–203, 1981.
- [13] H. Ishikawa. Exact optimization for Markov random fields with convex priors. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25:1333–1336, October 2003.
- [14] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Int. J. Computer Vision*, 1(4):321–331, 1988.
- [15] S. Kichenassamy, A. Kumar, P. J. Olver, A. Tannenbaum, and A. J. Yezzi. Gradient flows and geometric active contour models. In *Proc. Int. Conf. on Computer Vision*, pages 810–815, 1995.
- [16] J. Kim, J. W. Fisher, A. Yezzi, M. Cetin, and A. Willsky. Nonparametric methods for image segmentation using information theory and curve evolution. In *Proc. Int. Conf. on Image Processing*, volume 3, pages 797–800, 2002.
- [17] M. Klodt, T. Schoenemann, K. Kolev, M. Schikora, and D. Cremers. An experimental comparison of discrete and continuous shape optimization methods. In *Proc. European Conf. Computer Vision*, October 2008.
- [18] K. Kolev and D. Cremers. Integration of multiview stereo and silhouettes via convex functionals on convex domains. In *Proc. European Conf. Computer Vision*, October 2008.
- [19] K. Kolev and D. Cremers. Anisotropic minimal surfaces for the integration of multiview stereo and normal information. Technical report, Department of Computer Science, University of Bonn, June 2009.

- [20] K. Kolev and D. Cremers. Continuous ratio optimization via convex relaxation with applications to multiview 3d reconstruction. In *Proc. Conf. Computer Vision and Pattern Recognition*, 2009.
- [21] K. Kolev, M. Klodt, T. Brox, and D. Cremers. Continuous global optimization in multiview 3d reconstruction. *Int. J. Computer Vision*, 2009.
- [22] K. Kolev, M. Klodt, T. Brox, S. Esedoglu, and D. Cremers. Continuous global optimization in multiview 3d reconstruction. In *Energy Minimization Methods for Computer Vision and Pattern Recognition*, 2009.
- [23] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.*, 42:577–685, 1989.
- [24] T. Pock, A. Chambolle, H. Bischof, and D. Cremers. A convex relaxation approach for computing minimal partitions. In *Proc. Conf. Computer Vision and Pattern Recognition*, 2009.
- [25] T. Pock, D. Cremers, H. Bischof, and A. Chambolle. An algorithm for minimizing the piecewise smooth mumford-shah functional. In *Proc. Int. Conf. on Computer Vision*, 2009.
- [26] T. Pock, T. Schoenemann, G. Graber, H. Bischof, and D. Cremers. A convex formulation of continuous multi-label problems. In *Proc. European Conf. Computer Vision*, October 2008.
- [27] R. B. Potts. Some generalized order-disorder transformations. *Proc. Camb. Phil. Soc.*, 48:106–109, 1952.
- [28] C. Rother, V. Kolmogorov, V. Lempitsky, and M. Szummer. Optimizing binary MRFs via extended roof duality. In *Proc. Conf. Computer Vision and Pattern Recognition*, 2007.
- [29] M. Unger, T. Mauthner, T. Pock, and H. Bischof. Tracking as segmentation of spatial-temporal volumes by anisotropic weighted TV. In *Energy Minimization Methods for Computer Vision and Pattern Recognition*, 2009.
- [30] M. Unger, T. Pock, D. Cremers, and H. Bischof. Tvseg - interactive total variation based image segmentation. In *Proc. British Machine Vision Conf.*, September 2008.

- [31] G. Vogiatzis, P. Torr, and R. Cippola. Multi-view stereo via volumetric graph-cuts. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 391–399, 2005.
- [32] C. Zach, M. Niethammer, and J.-M-Frahm. Continuous maximal flows and Wulff shapes: Application to MRFs. In *Proc. Conf. Computer Vision and Pattern Recognition*, 2009.
- [33] S. C. Zhu and A. Yuille. Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(9):884–900, 1996.

