Weighted Minimal Hypersurfaces and Their Applications in Computer Vision

Bastian Goldlücke and Marcus Magnor

Graphics - Optics - Vision
Max-Planck-Institut für Informatik
Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany
{bg,magnor}@mpii.de

Abstract. Many interesting problems in computer vision can be formulated as a minimization problem for an *energy functional.* If this functional is given as an integral of a scalar-valued weight function over an unknown hypersurface, then the minimal surface we are looking for can be determined as a solution of the functional's Euler-Lagrange equation. This paper deals with a general class of weight functions that may depend on the surface point and normal. By making use of a mathematical tool called *the method of the moving frame*, we are able to derive the Euler-Lagrange equation in arbitrary-dimensional space and without the need for any surface parameterization. Our work generalizes existing proofs, and we demonstrate that it yields the correct evolution equations for a variety of previous computer vision techniques which can be expressed in terms of our theoretical framework. In addition, problems involving minimal hypersurfaces in dimensions higher than three, which were previously impossible to solve in practice, can now be introduced and handled by generalized versions of existing algorithms. As one example, we sketch a novel idea how to reconstruct temporally coherent geometry from multiple video streams.

1 Introduction

A popular and successful way to treat many computer vision problems is to formulate their solution as a hypersurface which minimizes an energy functional given by a weighted area integral. In this paper, we want to expose, generalize and solve the mathematical problem which lies at the very heart of all of these methods. The aim is to find a k-dimensional regular hypersurface $\Sigma \subset \mathbb{R}^n$ which minimizes the energy functional

$$\mathcal{A}(\Sigma) := \int_{\Sigma} \Phi \, dA. \tag{1}$$

We will only investigate the case of codimension one, so throughout this text, $k = n - 1$. Such a surface is called a *weighted minimal hypersurface* with respect to the weight function. This function shall be as general as required in practice, so we allow it to depend on the surface point s as well as the surface normal $\mathbf{n}$.

In this paper, we derive a very elegant and short proof of the necessary minimality condition:

Theorem. A k-dimensional surface $\Sigma \subset \mathbb{R}^{k+1}$ which minimizes the functional $\mathcal{A}(\Sigma) := \int_\Sigma \Phi(s, \mathbf{n}(s))\ dA(s)$ satisfies the Euler-Lagrange equation

$$\langle \Phi_s, \mathbf{n} \rangle \ - \ \mathrm{Tr}\,(\mathbf{S})\,\Phi \ + \ \mathrm{div}_\Sigma(\Phi_\mathbf{n}) \ = \ 0, \tag{2}$$

where $\mathbf{S}$ is the shape operator of the surface, also known as the Weingarten map or second fundamental tensor. Using standard techniques, a local minimum can be obtained as a stationary solution to a corresponding surface evolution equation. Since this surface evolution can be implemented and solved in practice, the Theorem yields a generic solution to all problems of the form (1) for practical applications. In this work, we set aside the problems of convergence and local minima, as they are far from being solved yet. See e.g. [1] for a detailed analysis. This paper has thus two main contributions:

Unification: We unite a very general class of problems into a common mathematical framework. This kind of minimization problem arises in numerous contexts in computer vision, with dimension $n \leq 3$ and various choices of Φ. A few select examples are summarized in Sect. 5, among them the method of geodesic snakes for segmentation as well as a very general multi-view 3D reconstruction technique. Our theorem yields the correct surface evolution equations for all of them.

Generalization: Our result is valid in arbitrary dimension. We are not aware of a previously existing treatment in computer vision literature of this generality. Until now, the theorem has been proved separately in dimensions $k = 1$ and $k = 2$, using local coordinates on the surface [2]. The now freely selectable number of surface dimensions opens up new possibilities for future applications. As one example, we generalize the static 3D reconstruction of a surface towards a space-time reconstruction of an evolving surface, which can be viewed as a 3D volume in 4D space. The proposed method treats all frames of multiple video sequences simultaneously in order to provide a temporally coherent result.

In the special case that $\Phi = 1$ is constant, the problem of minimizing (1) is reduced to finding a standard minimal surface, which is defined to locally minimize area. As we deal with a generalization, it seems reasonable to adopt the same mathematical tools used in that context [3]. A brief review of this framework, known as the *method of the moving frame*, is given in Sect. 2. However, we are forced to assume that the reader has at least some familiarity with differential geometry, preferably of frame bundles. For improved readability, we organized the paper in such a way that Sect. 2 as well as Sect. 3, in which we prove our Theorem (2), can be skipped. The transition from the Euler-Lagrange equation to a surface and further to a level set evolution equation is reviewed in Sect. 4, where we also discuss some necessary implementation details. Applications are introduced in the last two sections. In Sect. 5, we summarize a few existing computer vision techniques in order to demonstrate how they fit into our framework. A novel idea for space-time consistent reconstruction is presented in Sect. 6.

2 Mathematical Framework

Our goal is to give a general proof that surfaces minimizing (1) can be obtained as a solution of the Euler-Lagrange equation (2) for the energy functional. The mathematical tool of choice is called the *method of the moving frame.* This section is intended to give a brief overview of this framework.

Any minimal surface Σ of the functional $\mathcal{A}$ is a *critical point* of the functional, i.e., in first order the value of the functional does not change under a small variation of the surface. This restriction is known as the functional's *Euler-Lagrange equation.* We are now going to give a, necessarily brief, overview of the mathematical framework in which this equation can be derived. For an excellent and thorough introduction, the reader is referred to [3].

We have to investigate how the functional behaves with respect to first order variations of the surface. To this end, let

$$X : \Sigma \times (-\epsilon, \epsilon) \to \mathbb{R}^n$$

be a variation of Σ with compact support, then for each $\tau \in (-\epsilon, \epsilon)$ a regular surface $\Sigma_\tau \in \mathbb{R}^n$ is given by $X(\Sigma, \tau)$. For each $(s, \tau) \in \Sigma \times (-\epsilon, \epsilon)$, let

$$\{\mathbf{e}_1(s,\tau), \dots, \mathbf{e}_n(s,\tau) =: \mathbf{n}(s,\tau)\}$$

be an orthonormal frame for the surface Σ_τ at s with $\mathbf{e}_n = \mathbf{n}$ normal to the tangent plane $T_s\Sigma_\tau$. The restrictions ω^i of the Maurer-Cartan forms of $\mathbb{R}^n$ to this frame are defined by

$$dX \;=\; \mathbf{e}_i \, \omega^i. \tag{3}$$

Throughout this text we use the Einstein convention for sums, which means that we implicitly compute the sum from 1 to n over all indices appearing twice on the same side of an equation. Because the frame is adapted to Σ_τ in the above sense, the forms ω^1 to ω^k are its usual dual forms on the surface. The *connection 1-forms* ω_i^j are defined by

$$d\mathbf{e}_i \;=\; \mathbf{e}_j \, \omega_i^j \tag{4}$$

and satisfy the *structure equations*

$$d\omega^i \;=\; -\omega_j^i \wedge \omega^j \qquad d\omega_j^i \;=\; \omega_k^i \wedge \omega_j^k, \tag{5}$$

which can be deduced by differentiating the definitions.

From the connection forms stems the true power of this method. They allow us to express derivatives of the frame, in particular of the normal, in terms of objects which are *part of the frame bundle themselves.* This is the one reason why we will never need local coordinates, because all necessary information about the embedding of the surface in space is encoded in the connection forms.

From the Euclidean structure on $\mathbb{R}^n$ it follows that the connection 1-forms are skew-symmetric, $\omega_i^j = -\omega_j^i$. The connection forms ω_i^n can be expressed in the base $\{\omega^1, \dots, \omega^k, d\tau\}$, courtesy of Cartan's Lemma [4]. To see this, first note that because of definition (3)

$$\omega^n \;=\; \langle dX, \mathbf{n} \rangle \;=\; \frac{\partial X}{\partial \tau}\, d\tau \;=:\; f\, d\tau. \tag{6}$$

Differentiating this equation yields together with (5)

$$df \wedge d\tau + \sum_{i=1}^{k} \omega_i^n \wedge \omega^i \;=\; 0,$$

therefore, by Cartan's Lemma, there exist functions h_{ij} such that

$$\begin{bmatrix} \omega_1^n \\ \vdots \\ \omega_k^n \\ df \end{bmatrix} = \begin{bmatrix} h_{11} & \dots & h_{1k} & f_1 \\ \vdots & \ddots & \vdots & \vdots \\ h_{k1} & \dots & h_{kk} & f_k \\ f_1 & \dots & f_k & f_n \end{bmatrix} \begin{bmatrix} \omega^1 \\ \vdots \\ \omega^k \\ d\tau \end{bmatrix}. \tag{7}$$

The top-left part $\mathbf{S} := (h_{ij})$ of this matrix is called the *shape operator*, and is closely related to the curvature of Σ_τ. In the lower dimensional cases, its entries are commonly known as follows:

- If $k = 1$, i.e. Σ_τ is a curve in $\mathbb{R}^2$, the sole coefficient h_{11} equals the scalar valued curvature usually denoted by κ.
- If on the other hand $k = 2$, i.e. Σ is a regular surface in $\mathbb{R}^3$, the entries of $\mathbf{S}$ are the coefficients of the second fundamental form of Σ_τ. More precisely,

$$II \;=\; \begin{bmatrix} \omega^1 & \omega^2 \end{bmatrix} \mathbf{S} \begin{bmatrix} \omega^1 \\ \omega^2 \end{bmatrix} \;=\; h_{11}(\omega^1)^2 + 2h_{12}\omega^1\omega^2 + h_{22}(\omega^2)^2.$$

Thus $H = \frac{1}{k}\mathrm{Tr}\,(\mathbf{S}) = \frac{1}{k}\sum_{i=1}^{k} h_{ii}$ is the mean curvature of the surface.

The f_i are just the directional derivatives of f in the directions of the $\mathbf{e}_i$. Using the structure equations (5), we immediately deduce an important relation for the area form dA on Σ_τ:

$$dA =: \omega_A = \omega^1 \wedge \ldots \wedge \omega^k \implies d\omega_A \;=\; -\mathrm{Tr}\,(\mathbf{S})\, \omega_A \wedge \omega^n, \tag{8}$$

We introduce the notation ω_A to remind the reader of the fact that the area element dA indeed is a differential form of degree k. Note that area in our sense does not imply "two-dimensional".

Finally, we need a notion of an 'integration by parts' for a surface integral. First, we generalize the usual operators from vector analysis to vector fields $\mathbf{v}$ and functions f on Σ:

$$\begin{aligned} \mathrm{div}_\Sigma(\mathbf{v}) \;&:=\; \sum_{i=1}^{k} \frac{\partial v^i}{\partial \mathbf{e}_i} \text{ with the expansion } \mathbf{v} \;=\; v^i\, \mathbf{e}_i, \text{ and} \\ \nabla_\Sigma f \;&:=\; \sum_{i=1}^{k} \frac{\partial f}{\partial \mathbf{e}_i}\, \mathbf{e}_i \;=\; \sum_{i=1}^{k} f_i e_i. \end{aligned}$$

Using the definitions and the product rule, we derive a generalization of an identity well-known from classical vector analysis,

$$\operatorname{div}_\Sigma(\mathbf{v}f) \;=\; \langle \mathbf{v}, \nabla_\Sigma f\rangle \;+\; \operatorname{div}_\Sigma(\mathbf{v})\, f, \tag{9}$$

which will be useful later as one possibility of shifting partial derivatives from one object to another. A second possibility is given by Gauss' Theorem for surfaces, which in our context reads

$$\int_\Sigma \operatorname{div}_\Sigma(\mathbf{v})\, dA \;=\; -\int_\Sigma \operatorname{Tr}(\mathbf{S})\, \langle \mathbf{v}, \mathbf{n}\rangle\, dA. \tag{10}$$

Note that $\mathbf{v}$ does not have to be tangential to Σ. Since we assume that all our surfaces are closed, the boundary term usually contributing to the formula has vanished.

We have now collected all the necessary tools to derive the Euler-Lagrange equation of $\mathcal{A}$, and do so in the next section. In Sect. 4, this will yield an evolution equation for the level sets of a function on $\mathbb{R}^n$.

3 Euler-Lagrange Equation

In this section we employ the mathematical framework to derive the Euler-Lagrange equation of the functional $\mathcal{A}$. The arguments can be followed just by abstract manipulation of symbols, without the need to understand all of the reasons which lead to the governing rules presented in Sect. 2.

The desired equation characterizes critical points of $\mathcal{A}$, and is given by the derivation of the functional with respect to τ at $\tau = 0$. We assume that $\Phi = \Phi(s, \mathbf{n})$ is a function of the surface point s and the normal $\mathbf{n}(s)$ at this point. Since Φ maps from $\mathbb{R}^n \times \mathbb{S}^n$, $\Phi_{\mathbf{n}}(s, \mathbf{n})$ is tangent to the unit sphere of $\mathbb{R}^n$ at $\mathbf{n}$, so we have the important relation

$$\langle \Phi_{\mathbf{n}}(s, \mathbf{n}), \mathbf{n}\rangle \;=\; 0. \tag{11}$$

This fact was overlooked in previous publications, which is the reason why our final equation is considerably simpler.

Let us now turn to the computation of the Euler-Lagrange equation. Using the Lie-derivative

$$\mathcal{L}_{\mathbf{v}}\omega \;=\; \mathbf{v} \rightharpoonup d\omega \;+\; d(\mathbf{v} \rightharpoonup \omega) \tag{12}$$

of a differential form ω in the direction of $\mathbf{v}$, we obtain

$$\begin{aligned}
&\left.\frac{d}{d\tau}\right|_{\tau=0} \mathcal{A}\,(\Sigma_\tau) \;\overset{(a)}{=}\; \int_\Sigma \mathcal{L}_{\frac{\partial}{\partial\tau}}(\Phi\,\omega_A) \;\overset{(b)}{=}\; \int_\Sigma \frac{\partial}{\partial\tau} \rightharpoonup d\,(\Phi\,\omega_A) \\
&\overset{(c)}{=}\; \int_\Sigma \frac{\partial}{\partial\tau} \rightharpoonup (d\Phi \wedge \omega_A \;+\; \Phi\, d\omega_A) \\
&\overset{(d)}{=}\; \int_\Sigma \frac{\partial}{\partial\tau} \rightharpoonup \big(\langle \Phi_s, \mathbf{e}_i\rangle\, \omega^i \wedge \omega_A \;+\; \Phi_{\mathbf{n}}\, d\mathbf{n} \wedge \omega_A \;-\; \operatorname{Tr}(\mathbf{S})\, \Phi\, \omega_A \wedge \omega^n\big) \\
&\overset{(e)}{=}\; \int_\Sigma \left[(\langle \Phi_s, \mathbf{n}\rangle - \;\operatorname{Tr}(\mathbf{S})\, \Phi)\; f\, \omega_A \;+\; \frac{\partial}{\partial\tau} \rightharpoonup (\Phi_{\mathbf{n}}\, d\mathbf{n} \wedge \omega_A) \right].
\end{aligned} \tag{13}$$

The five equalities above are justified by the following arguments:

a. A generalization of the 'Differentiation under the integral'-rule in classic calculus [3].
b. Cartan's rule (12) for expressing the Lie derivative and using the fact that $\omega^1(\mathbf{n}) = \cdots = \omega^k(\mathbf{n}) = 0$. Note that $\frac{\partial}{\partial\tau}$ is parallel to $\mathbf{n}$, so this equation also holds for $\frac{\partial}{\partial\tau}$.
c. Product rule for differential forms, note that Φ is a 0-form.
d. Expansion of $d\Phi = \Phi_s\, dX \,+\, \Phi_{\mathbf{n}}\, d\mathbf{n} = \langle \Phi_s, \mathbf{e}_i\rangle\, \omega^i \,+\, \Phi_{\mathbf{n}}\, d\mathbf{n}$. Here we inserted the definition (3) of the restrictions ω^i. The last term is due to (8).
e. Linearity of the left hook and again $\omega^1(\mathbf{n}) = \cdots = \omega^k(\mathbf{n}) = 0$. From (6), it follows that $\omega^n(\frac{\partial}{\partial\tau}) = f d\tau(\frac{\partial}{\partial\tau}) = f$.

We now turn our attention to the second term of the last integral. Inserting the definition (4) of the connection 1-forms and afterwards the expansion of the connection forms (7) due to Cartan's Lemma, we get

$$\begin{aligned}
&\frac{\partial}{\partial\tau} \rightharpoonup \left(\Phi_{\mathbf{n}}\, d\mathbf{n} \wedge \omega_A\right) \;=\; \frac{\partial}{\partial\tau} \rightharpoonup \big(\langle \Phi_{\mathbf{n}}, \mathbf{e}_j\rangle\; \omega_n^j \wedge \omega_A\big) \\
&= \frac{\partial}{\partial\tau} \rightharpoonup \left(-\left\langle \Phi_{\mathbf{n}}, \nabla_\Sigma f\right\rangle\, d\tau \wedge \omega_A\right) \;=\; -\left\langle \Phi_{\mathbf{n}}, \nabla_\Sigma f\right\rangle\, \omega_A \\
&= \mathrm{div}_\Sigma(\Phi_{\mathbf{n}})\, f\, \omega_A \;-\; \mathrm{div}_\Sigma\left(\Phi_{\mathbf{n}}\, f\right)\, \omega_A.
\end{aligned} \tag{14}$$

In the last equality, we have shifted derivatives using the product rule (9). We can finally compute the integral over the left term using Gauss' Theorem (10):

$$\int_\Sigma -\,\mathrm{div}_\Sigma\left(\Phi_{\mathbf{n}}\, f\right)\, dA \;=\; \int_\Sigma \mathrm{Tr}\,(\mathbf{S})\,\langle \Phi_{\mathbf{n}}, \mathbf{n}\rangle\; f\, dA \;=\; 0.$$

It vanishes due to (11). When we thus put equations (13) and (14) together, we see that we have derived

$$\left.\frac{d}{d\tau}\right|_{\tau=0} \mathcal{A}\left(\Sigma_\tau\right) = \int_\Sigma \left(\langle \Phi_s, \mathbf{n}\rangle -\; \mathrm{Tr}\,(\mathbf{S})\,\Phi + \mathrm{div}_\Sigma(\Phi_{\mathbf{n}})\right)\, f\, dA.$$

Since for a critical point this expression must be zero for any variation and hence for any f, we have arrived at the Euler-Lagrange equation of the functional

$$\langle \Phi_s, \mathbf{n}\rangle \;-\; \mathrm{Tr}\,(\mathbf{S})\,\Phi \;+\; \mathrm{div}_\Sigma(\Phi_{\mathbf{n}}) \;=\; 0, \tag{15}$$

and thus proved our Theorem (2).

4 Corresponding Level Set Equation

Level sets represent an efficient way to implement a surface evolution [5, 6], and are by now a well-established technique with a wide area of applications [7]. We will briefly review the transition from (15) to a surface evolution equation followed by one for a level set in this section. For the remainder of the text, let

$$\Psi := \langle \Phi_s, \mathbf{n}\rangle \;-\; \mathrm{Tr}\,(\mathbf{S})\,\Phi \;+\; \mathrm{div}_\Sigma(\Phi_{\mathbf{n}}).$$

A surface $\hat{\Sigma}$ which is a solution to the Euler-Lagrange equation $\Psi = 0$ is likewise a stationary solution to a surface evolution equation, where Ψ describes a force in the normal direction:

$$\frac{\partial}{\partial \tau} \Sigma_\tau \;=\; \Psi \mathbf{n}. \tag{16}$$

If we start with an initial surface Σ_0 and let the surface evolve using this equation, it will eventually converge to a local minimum of $\mathcal{A}$. Instead of implementing a surface evolution directly, we can make use of the level set idea. We express the surfaces Σ_τ for each parameter value $\tau \geq 0$ as the zero level sets of a regular function $u : \mathbb{R}^n \times \mathbb{R}^{\geq 0} \to \mathbb{R},\; u(s, \tau) = 0 \;\Leftrightarrow\; s \in \Sigma_\tau$.

We require $u(\cdot, \tau)$ to be positive in the volume enclosed by Σ_τ, thus if ∇ is the gradient operator for the spatial coordinates of u, we can compute the outer normal using

$$\mathbf{n} \;=\; -\frac{\nabla u}{|\nabla u|} \quad\Longrightarrow\quad |\nabla u| \;=\; -\left\langle \nabla u, \mathbf{n} \right\rangle .$$

Taking the derivative of $u(s, \tau) = 0$ with respect to τ and inserting (16), we deduce the evolution equation for u to be

$$\frac{\partial}{\partial \tau} u \;=\; -\left\langle \nabla u, \frac{\partial}{\partial \tau} \Sigma_\tau \right\rangle \;=\; -\left\langle \nabla u, \mathbf{n} \right\rangle \Psi \;=\; \Psi \left| \nabla u \right| . \tag{17}$$

Using the identity

$$\operatorname{Tr}(\mathbf{S}) \;=\; \operatorname{div}\left(\frac{\nabla u}{|\nabla u|} \right)$$

for the curvature of the level sets of u and the definition of Ψ, we arrive at the final reformulation of (16) in terms of a level set evolution:

$$\frac{\partial}{\partial \tau} u \;=\; \left[-\operatorname{div}\left(\Phi \cdot \frac{\nabla u}{|\nabla u|} \right) \;+\; \operatorname{div}_\Sigma(\Phi_{\mathbf{n}}) \right] \left| \nabla u \right| .$$

Note that the derivatives of Φ can be computed numerically. Thus, it is not necessary to compute an explicit expression for them manually, which would be very cumbersome for more difficult functionals. Instead, in an existing implementation of the evolution for a general function Φ, essentially any functional can be plugged in.

5 Known Applications in Computer Vision

Among the first variational methods which were successfully utilized for computer vision problems was the one now widely known as *Geodesic Active Contours* [8]. While originally designed for segmentation in 2D, it quickly became clear that it could be generalized to 3D [9], and also applied to other tasks. It is particularly attractive for modeling surfaces from point clouds [10, 11]. Geodesic contours were also employed for 2D detection and tracking of moving objects [12]. Also well analyzed in theory is how to employ minimal surfaces for 3D reconstruction of static objects from multiple views [13]. This technique was recently extended to simultaneously estimate the radiance of surfaces, and demonstrated to give good results in practice [14].

We will briefly review the above methods to demonstrate that all of them fit into our framework. In particular, our theorem applies to all of them and yields the correct surface evolution equations.

5.1 Segmentation via Geodesic Active Contours

Caselles, Kimmel and Sapiro realized that the energy which is minimized in the classical *snakes* approach [15] can be rewritten in terms of a geodesic computation in a Riemannian space by means of Maupertuis' Principle. The goal is to compute a contour curve $\mathcal{C}$ in an image I which is attracted by edges in the image while remaining reasonably smooth. Their final energy functional took the form

$$\begin{aligned} \mathcal{A}(\mathcal{C}) \; &:= \int_{\mathcal{C}} g \circ |\nabla I| \, ds, \\ &\text{where } g : \mathbb{R}^+ \to \mathbb{R}^+ \text{ is strictly decreasing with } \lim_{r \to 0} g(r) = 0. \end{aligned}$$

∇I acts as an edge detector, while g controls how image gradients are interpreted as energies. The main purpose of g is to act as a *stopping function*: The flow of the curve should cease when it arrives at object boundaries. Because the integral is minimized, the contour will move towards regions of high gradient. The smoothness requirement is enforced by the curvature term in equation (2). Note that $g \circ |\nabla I|$ depends only on the surface point and not on the normal, so the rightmost term in the Euler-Lagrange equation vanishes.

Essentially the same functional can be applied to 3D segmentation [9], where the source image I is replaced by a volumetric set of data, and the unknown curve $\mathcal{C}$ by an unknown 2D surface. Based on another derivation of the conformal length minimizing flow by Kichenassamy et al. [16], Zhao et al. [11, 17] chose an Euclidean distance function instead of an edge-based stopping potential for Φ to model surfaces from unstructured data sets.

5.2 Tracking

Paragios and Deriche combine geodesic active contours and a motion detection term in a single energy functional to track moving objects in a sequence of images [12]:

$$\mathcal{A}(\mathcal{C}) \; := \int_{\mathcal{C}} \gamma \, \underbrace{G_{\sigma_D} \circ I_D}_{Motion} + (1-\gamma) \, \underbrace{G_{\sigma_T} \circ |\nabla I|}_{Contours} \, ds,$$

where G_σ is a Gaussian with variance σ. The user-defined parameter γ weights the influence of the motion detection term against the boundary localization. The Gaussians play the same role as g in geodesic contours, their variances σ_T and σ_D are derived from the image statistics. The image I_D is designed to detect boundaries of moving regions in the current image I of the sequence, and constructed using a Bayesian model which takes into account the pixel differences to the previous frame.

5.3 3D Reconstruction

As a first step, Faugeras and Keriven [13] give a simple functional in dimension $n = 3$ for static 3D scene reconstruction which does not depend on the surface normal. It can be viewed as a space-carving approach [18] generalized from discrete voxels to a continuous surface model.

Let $C_1, \ldots, C_l$ be a number of cameras which project a scene in $\mathbb{R}^3$ onto images I_k via projections $\pi_k : \mathbb{R}^3 \to \mathbb{R}^2$. For each point $s \in \mathbb{R}^3$, let $\nu_k(s)$ denote whether or not s is visible in camera k in the presence of a surface Σ. $\nu_k(s)$ is defined to be one if s is visible, and zero otherwise. A measure of how good a surface Σ as a model of the scene geometry really is in accordance with a given set of images can be obtained as follows: Each surface point is projected into the set of images where it is visible, and the differences between the pixel colors for each pair of images are computed and summed up to get an error measure for the surface point. This error is integrated over the surface to get the total error. In mathematical notation,

$$\begin{aligned} \mathcal{A}(\Sigma) &:= \int_\Sigma \Phi^S dA, \text{ where} \\ \Phi^S(s) &:= \frac{1}{V_s(V_s-1)} \sum_{i,j=1}^{l} \nu_i(s)\nu_j(s) \cdot \left\| I_i \circ \pi_i(s) - I_j \circ \pi_j(s) \right\|_\infty . \end{aligned}$$

The number V_s of cameras able to see a point s is used to normalize the function.

Clearly, the above model is too simple to be of much use in multi-view reconstruction, since only single pixels with no regard to their neighborhoods are compared. A better functional was therefore suggested by Faugeras and Keriven, and can be applied using the results on how the evolution depends on the current normals. We present a slight modification of their original approach here. Our functional only depends on invariant surface properties and does not make use of geometric objects in the source camera views.

To each surface point s, we associate a small rectangle $\Box_{s,\mathbf{n}}$ in the tangent plane $T_s\Sigma$. In order to invariantly determine its orientation within the plane, we align the sides with the principal curvature directions. This rectangle is then projected into the images, and the normalized cross-correlation over the projected areas is computed. We choose the length of the rectangle sides to be inversely proportional to the curvature in the corresponding direction, up to a certain maximum, because the first order approximation of the surface by its tangent plane is valid over a larger region if the curvature is low. The corresponding functional can be written as

$$\begin{aligned} \mathcal{A}(\Sigma) &:= \int_\Sigma \Phi^C dA, \text{ where} \\ \Phi^C(s,\mathbf{n}) &:= -\frac{1}{V_s(V_s-1)} \sum_{i,j=1}^{l} \nu_i(s)\nu_j(s) \cdot \chi_{i,j}(s,\mathbf{n}) \text{ and} \\ \chi_{i,j}(s,\mathbf{n}) &:= \frac{1}{A(\Box_{s,\mathbf{n}})} \int_{\Box_{s,\mathbf{n}}} \left(I_i \circ \pi_i - \overline{I}_i^{s,\mathbf{n}} \right) \cdot \left(I_j \circ \pi_j - \overline{I}_j^{s,\mathbf{n}} \right) dA. \end{aligned}$$

The correlation integral has to be normalized using the area $A\left(\square_{s,\mathbf{n}}\right)$ of the square. The mean values are computed using

$$\overline{I}_i^{s,\mathbf{n}} := \frac{1}{A\left(\square_{s,\mathbf{n}}\right)} \int_{\square_{s,\mathbf{n}}} I_i \circ \pi_i \, dA.$$

When this functional is minimized, not only the position, but also the surface normal is adjusted to best match the images. This approach can also be employed to improve the normals for a known geometry approximation, i.e., the visual hull. When a segmentation of the images into background and foreground objects can be obtained, the visual hull also constitutes a good initial surface Σ_0 for the evolution equation (16), since it is by construction a conservative estimate of the object regions.

5.4 Reflectance Estimation

Jin, Soatto and Yezzi combine the reconstruction framework with a simultaneous reflectance estimation [14]. They use the functional

$$\mathcal{A}\left(\Sigma\right) := \int_{\Sigma} \left\| \tilde{R} - R \right\|_F^2 \, dA,$$

where the Frobenius norm $\|\cdot\|_F$ is employed to compute the difference of the measured radiance tensor field $\tilde{R}$ to an idealized R obtained from a reflection model, which depends on the surface Σ.

As claimed previously, all of the problems reviewed in this section are of the form required by the main theorem, and can thus be subsumed under the unifying framework presented in this paper.

6 A Novel Technique: Space-time 3D Reconstruction

In this section and for the first time, we are going to exploit the fact that we can now handle variational problems posed in higher-dimensional space. We present only one of many new applications one can think of. For instance, the additional degrees of freedom could also be used to optimize parameters defined in each surface point. We employ them to introduce a temporal dimension, which allows us to compute temporally coherent estimates for the scene geometry reconstruction from multiple video streams.

Since our equation is valid in arbitrary dimension, we can interpret it in dimension $n = 4$ as an evolution equation for a hypersurface in space-time. Its cross sections with planes of constant time $\{t = t_0\}$ yield the scene geometry at the time instant t_0. That way, a global optimum for the scene geometry, including the normals, *and its change over time* can be found which takes into account *every frame at every time instant* simultaneously. Moreover, the minimization problem is formulated and can be solved in an elegant mathematical context.

In order to distinguish normal surfaces in $\mathbb{R}^3$ from hypersurfaces, we will denote the latter by $\mathfrak{H}$. Points in space-time $\mathbb{R}^4$ are written as $x = (s, t) \in \mathbb{R}^3 \times \mathbb{R}$. A hypersurface $\mathfrak{H}$ gives rise to a family (Σ_t) of regular surfaces

$$\Sigma_t \; := \; \mathfrak{H} \cap (\mathbb{R}^3, t) \; \subset \; \mathbb{R}^3$$

for each time instant t. We also have to deal with time-dependent visibilities $\nu^t(s)$ depending on point s and the surface Σ_t at time-instant t, as well as time-dependent images I^t from each camera.

The resulting functional looks almost identical to the one in Sect. 5.3:

$$\begin{aligned}
\mathcal{A}(\mathfrak{H}) &:= \int_{\mathfrak{H}} \Phi^G dV, \text{ where} \\
\Phi^G(x, \mathbf{n}) &:= -\frac{1}{V_{s,t}(V_{s,t}-1)} \sum_{i,j=1}^{l} \nu_i^t(s)\nu_j^t(s) \cdot \chi_{i,j}^t(s, \mathbf{n}^t). \\
\chi_{i,j}^t(s, \mathbf{n}^t) &:= \frac{1}{A(\square_{s,\mathbf{n}^t})} \int_{\square_{s,\mathbf{n}^t}} \left(I_i^t \circ \pi_i - \overline{I}_i^{x,\mathbf{n}} \right) \cdot \left(I_j \circ \pi_j - \overline{I}_j^{x,\mathbf{n}} \right) \, dA.
\end{aligned}$$

The normal $\mathbf{n}^t$ to the surface Σ_t is the projection of the normal to the hypersurface $\mathfrak{H}$ onto the tangent space of Σ_t, which is a subspace of the tangent space of $\mathfrak{H}$. Note that the square $\square_{s,\mathbf{n}^t}$ lies inside the tangent plane of Σ_t. The mean values $\overline{I}_i^{x,\mathbf{n}}$ are of course also computed using the images at time t.

When this functional is minimized, two constraints are optimized simultaneously. First, each surface Σ_t together with its normals is selected to best match the images at that time instant. Second, a smooth change of the surfaces Σ_t with time is encouraged because of the curvature term in the Euler-Lagrange equation. Our experiments with real-world data using a parallel implementation of this scheme gave very promising results, which we are going to present in a future publication.

7 Conclusion

Using the mathematical tool of the *method of the moving frame*, we have derived the Euler-Lagrange equations for weighted minimal surfaces in arbitrary dimensions. We allowed for weight functions general enough to cover the variational problems encountered in computer vision research. Previously, existing proofs used local coordinates and were restricted to dimensions two or three, so our approach is more general. As demonstrated by several examples, weighted minimal surfaces lie at the heart of several well-established computer vision techniques. Our result for arbitrarily high dimensions paves the way for new, future research. In particular, we sketched a technique designed to achieve temporal coherence in 3D reconstruction from multiple video streams. In the near future, we are going to experimentally investigate its possibilities.

References

1. Chen, Y., Giga, Y., Goto, S.: Uniqueness and existence of viscosity solutions of generalized mean curvature flow. Journal of Differential Geometry **33** (1991) 749–786
2. Siddiqi, K., Lauziere, Y.B., Tannenbaum, A., Zucker, S.W.: Area and length minimizing flows for shape segmentation. IEEE Transactions on Image Processing **3** (1998) 433–443
3. Clelland, J.: MSRI Workshop on Lie groups and the method of moving frames. Lecture Notes. Department of Mathematics, University of Colorado (1999) *http://spot.Colorado.EDU/~jnc/MSRI.html*.
4. Sharpe, R.: Differential Geometry. Graduate Texts in Mathematics. Springer (1997)
5. Osher, S., Sethian, J.: Fronts propagating with curvature dependent speed: Algorithms based on the Hamilton-Jacobi formulation. Journal of Computational Physics **79** (1988) 12–49
6. Chop, D.: Computing minimal surfaces via level set curvature flow. Journal of Computational Physics **106** (1993) 77–91
7. Sethian, J.A.: Level Set Methods and Fast Marching Methods. 2nd edn. Monographs on Applied and Computational Mathematics. Cambridge University Press (1999)
8. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. In: Proc. International Conference on Computer Vision. (1995) 694–699
9. Caselles, V., Kimmel, R., Sapiro, G., Sbert, C.: Three dimensional object modeling via minimal surfaces. In: Proc. European Conference on Computer Vision. Volume 1., Springer (1996) 97–106
10. Caselles, V., Kimmel, R., Sapiro, G., Sbert, C.: Minimal surfaces based object segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence **19** (1997) 394–398
11. Zhao, H., Osher, S., Fedkiw, R.: Fast surface reconstruction using the level set method. 1st IEEE Workshop on Variational and Level Set Methods, 8th ICCV **80** (2001) 194–202
12. Paragios, N., Deriche, R.: Geodesic active contours and level sets for the detection and tracking of moving objects. IEEE Transactions on Pattern Analysis and Machine Intelligence **22** (2000) 266–280
13. Faugeras, O., Keriven, R.: Variational principles, surface evolution, PDE's, level set methods and the stereo problem. IEEE Transactions on Image Processing **3** (1998) 336–344
14. Jin, H., Soatto, S., Yezzi, A.: Multi-view stereo beyond Lambert. In: IEEE Conference on Computer Vision and Pattern Recognition. Volume I., Madison, Wisconsin, USA (2003) 171–178
15. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: active contour models. International Journal of Computer Vision **1** (1988) 321–331
16. Kichenassamy, S., Kumar, A., Olver, P.J., Tannenbaum, A., Yezzi, A.: Gradient flows and geometric active contour models. In: ICCV. (1995) 810–815
17. Zhao, H., Osher, S., Merriman, B., Kang, M.: Implicit and non-parametric shape reconstruction from unorganized points using variational level set method. In: Computer Vision and Image Understanding. (2000) 295–319
18. Kutukalos, K.N., Seitz, S.M.: A theory of shape by space carving. International Journal of Computer Vision **38** (2000) 197–216