

Variational Principles, Surface Evolution, PDE's, Level Set Methods and the Stereo Problem

Olivier Faugeras, Renaud Keriven

▶ **To cite this version:**

Olivier Faugeras, Renaud Keriven. Variational Principles, Surface Evolution, PDE's, Level Set Methods and the Stereo Problem. RR-3021, INRIA. 1996. inria-00073673

HAL Id: inria-00073673
https://inria.hal.science/inria-00073673

Submitted on 24 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Variational principles, Surface Evolution, PDE's, level set methods and the Stereo Problem

Olivier Faugeras
INRIA
MIT AI-Lab

Renaud Keriven
CERMICS / ENPC

N° 3021

26 Octobre 1996

THÈME 3



ISSN 0249-6399

INRIA
SOPHIA ANTIPOLIS

Variational principles, Surface Evolution, PDE's, level set methods and the Stereo Problem

Olivier Faugeras
INRIA
MIT AI-Lab

Renaud Keriven
CERMICS / ENPC

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projet Robotvis

Rapport de recherche n° 3021 — 26 Octobre 1996 — 38 pages

Abstract: We present a novel geometric approach for solving the stereo problem for an arbitrary number of images (greater than or equal to 2). It is based upon the definition of a variational principle that must be satisfied by the surfaces of the objects in the scene and their images. The Euler-Lagrange equations which are deduced from the variational principle provide a set of PDE's which are used to deform an initial set of surfaces which then move towards the objects to be detected. The level set implementation of these PDE's potentially provides an efficient and robust way of achieving the surface evolution and to deal automatically with changes in the surface topology during the deformation, i.e. to deal with multiple objects. Results of a two dimensional implementation of our theory are presented on synthetic and real images.

Key-words: Variational principles, PDE's, Geometry-Driven Evolution, Stereo.

(Résumé : tsvp)

Ce travail a été en partie financé par le contrat Esprit Improofs No. 23515, LTR. This work was partially supported by the Esprit grant Improofs, Nr. 23515, LTR domain.

Unité de recherche INRIA Sophia-Antipolis
2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex (France)
Téléphone : (33) 93 65 77 77 – Télécopie : (33) 93 65 77 65

Principes variationnels, évolution de surfaces, EDP's, méthodes de niveaux et le problème de la stéréoscopie

Résumé : Ce rapport expose une nouvelle méthode géométrique de résolution du problème de la stéréoscopie à partir d'un nombre quelconque d'images (plus grand ou égal à deux). Elle est basée sur un principe variationnel que doivent satisfaire les surfaces des objets de la scène ainsi que leurs images. Les équations d'Euler-Lagrange déduites de ce principe variationnel fournissent un ensemble d'EDP's qu'on utilise pour déformer un ensemble de surfaces initiales qui vont alors se déplacer vers les objets à détecter. La résolution de ce système d'EDP's par surfaces de niveaux permet potentiellement de réaliser de manière efficace et robuste le processus d'évolution des surfaces tout en prenant en compte automatiquement les problèmes de changement de topologie durant la déformation ce qui permet de traiter le cas d'objets multiples. Les résultats d'une implémentation bidimensionnelle de notre théorie sont présentés sur des images synthétiques et réelles.

Mots-clé : Principes variationnels, EDP's, évolutions géométriques, Stéréoscopie

1 Introduction and preliminaries

The idea that is put forward in this paper is that the methods of curve and surface evolutions which have been developed in computer vision under the name of snakes [14] and then reformulated by Caselles, Kimmel and Sapiro [1] and Kichenassamy et al. [15] in the context of PDE driven evolving curves can be used effectively for solving 3D vision problems such as stereo and motion analysis.

As a first step in this direction we present a mathematical analysis of the stereo problem in this context as well as a partial implementation.

The problem of curve evolution driven by a PDE has been recently studied both from the theoretical standpoint [8, 9, 19] and from the viewpoint of implementation [16, 21, 22] with the development of level set methods that can efficiently and robustly solve those PDE's. A nice recent exposition of the level set methods and of many of their applications can be found in [20].

The problem of surface evolution has been less touched upon even though some preliminary results have been obtained [22, 2].

The path we will follow to attack the stereo problem from that angle is, not surprisingly, a variational one. In a nutshell, we will describe the stereo problem (to be defined more precisely later) as the minimisation of a functional (we will explore several such functionals) with respect to some parameters (describing the geometry of the scene); we will compute the Euler-Lagrange equations of this functional, thereby obtaining a set of necessary conditions, in effect a set of partial differential equations, which we will solve as a time evolution problem by a level set method.

Stereo is a problem that has received considerable attention for decades in the psychophysical, neurophysiological and, more recently, in the computer vision literatures. It is impossible to cite all the published work here, we will simply refer the reader to some basic books on the subject [13, 10, 11, 12, 7]. To explain the problem of stereo from the computational standpoint, we will refer the reader to figure 1. Two, may be more, images of the world are taken simultaneously. The problem is, given those images, to recover the geometry of the scene. Given the fact that the relative positions and orientations and the internal parameters of the cameras are known which we will assume in this article (the cameras are then said to be calibrated [7]), the problem is essentially (but not only) one of establishing correspondences between the views: one talks about the matching problem. The matching problem is usually solved by setting up a matching functional for which one then tries to find extrema. Once a pixel in view i has been identified as being the image of the same scene point as another pixel in view j, the 3D point can then be reconstructed by intersecting the corresponding optical rays (see figure 1 again).

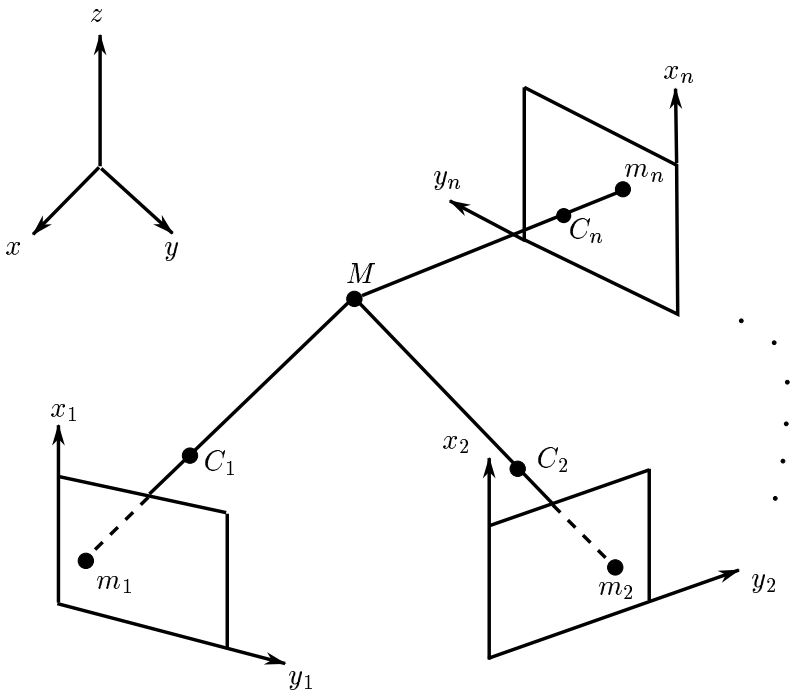


Figure 1: The multicamera stereo vision problem is, given a pixel m_1 in image 1, to find the corresponding pixel m_2 in image 2, ..., the corresponding pixel m_n in image n, i.e. the ones which are the images of the same 3D point M. Once such a correspondence has been established, the point M can be reconstructed by intersecting the optical rays $\langle m_i, C_i \rangle$, $i = 1, \cdots, n$.

In order to go any further, we need to be a little more specific about the process of image formation. We will assume here that the cameras perform a perpective projection of the 3D world on the retinal plane as shown in figure 2. The optical center, noted C in the figure, is the center of projection and the image of the 3D point M is the pixel m at the intersection of the optical ray $\langle C, m \rangle$ and the retinal plane $\mathcal{R}$. As described in many recent papers in computer vision, this operation can be conveniently described in projective geometry by a matrix operation. The projective coordinates of the pixel m (a 3×1 vector) are obtained by applying a 3×4 matrix $\mathbf{P}_1$ to the projective coordinates of the 3D point M (a 4×1 vector). This matrix is called the perspective projection matrix. If we express the matrix $\mathbf{P}_1$ in

the coordinate system (C, x, y, z) shown in the figure 2, it then takes a very simple form:

$$\mathbf{P}_1 = [\mathbf{I}_3\, \mathbf{0}]$$

where $\mathbf{I}_3$ is the 3×3 identity matrix. If we now move the camera by applying to it a rigid transformation described by the rotation matrix $\mathbf{R}$ and the translation vector $\mathbf{t}$, the expression of the matrix $\mathbf{P}$ changes accordingly and becomes:

$$\mathbf{P}_2 = [\mathbf{R}^T \; -\mathbf{R}^T\mathbf{t}]$$

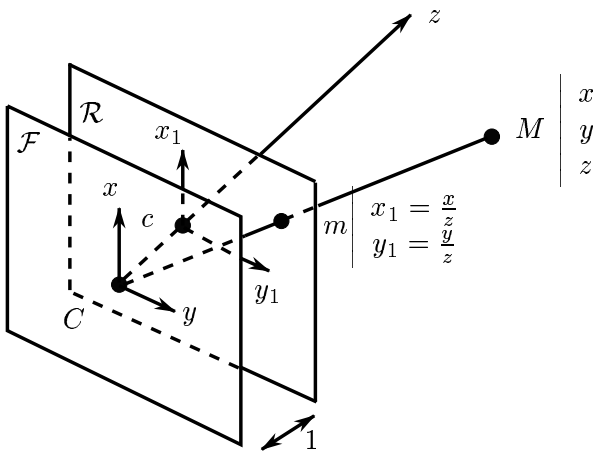


Figure 2: The focal plane (x, y) is parallel to the retinal plane (x_1, y_1) and at a distance of 1 from it.

With these preliminaries in mind we are ready to proceed with our program which we will do by progressing along two related axes. The first axis is that of object complexity, the second axis is that of matching functional complexity. They are related in the sense that an increase along one axis usually implies a corresponding increase along the other.

In the first two sections we will consider a simple object model which is well adapted to the binocular stereo case where it is natural to consider that the objects in the scene can be considered mathematically as forming the graph of an unknown smooth function (the depth function in the language of computer vision). In section 2 we consider an extremely simplified matching criterion which will allow us to convey

to the reader the flavor of the ideas that we are trying to push here. We then move in section 3 to a more sophisticated albeit classical matching criterion which is at the heart of the techniques known in computer vision as correlation based methods. Within the framework of this model we study two related shape models. The first model assumes that at every point in the scene the tangent plane is parallel to the retinal plane of one of the cameras (the so-called fronto parallel hypothesis). The second model relaxes this hypothesis by introducing a general tangent plane at every point. In the section 4 we introduce a more general shape model in which we do not assume anymore that the objects are the graph of a function and model them as a set of general smooth surfaces in three space. The next step would of course be to relax the smoothness assumption but we will postpone this to a future paper.

Let us decide on some definitions and notations. Images are denoted by I_k, k taking some integer values which indicate the camera with which the image has been acquired. They are considered as smooth (i.e. C^2, twice continuously differentiable) functions of pixels m_k whose coordinates are defined in some orthonormal image coordinate systems (x_k, y_k) which are assumed to be known. We note $I_k(m_k)$ or $I_k(x_k, y_k)$ the intensity value in image k at pixel m_k. We will use the first and second order derivatives of these functions, i.e. the gradient ∇I_k, a 2×1 vector equal to $[\frac{\partial I_k}{\partial x_k}, \frac{\partial I_k}{\partial y_k}]^T$, and the Hessian $\mathbf{H}_k$, a 2×2 symmetric matrix.

The pixels in the images are considered as functions of the 3D geometry of the scene, i.e. of some 3D point M on the surface of an object in the scene, and of the unit normal vector $\mathbf{N}$ to this surface.

Vectors and matrixes will generally be indicated in boldfaces, e.g. $\mathbf{x}$. The dot or inner product of two vectors $\mathbf{x}$ and $\mathbf{y}$ is denoted by $\mathbf{x} \cdot \mathbf{y}$. The cross-product of two 3×1 vectors $\mathbf{x}$ and $\mathbf{y}$ is noted $\mathbf{x} \times \mathbf{y}$ or $[\mathbf{x}]_\times \mathbf{y}$, where $[\mathbf{x}]_\times$ is an antisymmetric matrix.

We will make extensive use of differential calculus and of the chain rule for computing the derivatives of the composition of functions. We remind the reader that the derivative of a scalar function with respect to an $n \times 1$ vector is a $1 \times n$ vector, i.e. a linear form on $\mathbf{R}^n$, the second order derivative of a scalar function with respect to this vector is also called its Hessian, it is a symmetric bilinear form which is represented by a symmetric $n \times n$ matrix. The derivative of an $n \times 1$ vector with respect to a $p \times 1$ vector is a $n \times p$ matrix. We will also need derivatives of matrices with respect to vectors and matrices which are tensors but we will be able to avoid using tensor calculus altogether. Partial derivatives will be indicated either using the ∂ symbol, e.g. $\frac{\partial f}{\partial \mathbf{x}}$, or as a lower index, e.g. $f_\mathbf{x}$.

Our approach is an extension of previous work by Robert et al. and Robert and Deriche, [18, 17], where the idea of using a variational approach for solving the stereo problem was proposed first in the classical Tikhonov regularization framework and then by using regularization functions more proper to preserve discontinuities. We differ from this work because we do not assume that the depth is the graph of a function defined in the focal plane of the first camera which allows us to deal with an arbitrary number of cameras, because we take into account the projective deformation due to the orientation of the tangent plane to the object, as in [5], because we use a deformable surface approach based on a measurement criterion which is intrinsic, i.e. does not depend upon the parametrization of the objects, and therefore we can cope automatically with discontinuities, e.g. with multiple objects. Our work can be seen as a 3D extension of the approach proposed in [4] where we limit ourselves to the binocular case, to finding cross-sections of the objects with a fixed plane, and do not take into account the orientation of the tangent plane to the object.

2 A simple object and matching model

This section introduces in a simplified framework some of the basic ideas of this paper. We assume, and it is the first important assumption, that the objects which are being imaged by the stereo rig (a binocular stereo system) are modelled as the graph of an unknown smooth function $z = f(x, y)$ defined in the first retinal plane which we are trying to estimate. A point M of coordinates $[x, y, f(x, y)]^T$ is seen as two pixels m_1 and m_2 whose coordinates $(g_i(x, y), h_i(x, y)), i = 1, 2$, can be easily computed as functions of $x, y, f(x, y)$ and the coefficients of the perspective projection matrices $\mathbf{P}_1$ and $\mathbf{P}_2$. Let I_1 and I_2 be the intensities of the two images. Assuming, and it is the second important assumption, that the objects are perfectly Lambertian, we must have $I_1(m_1) = I_2(m_2)$ for all pixels in correspondence, i.e. which are the images of the same 3D point.

This reasoning immediately leads to the variational problem of finding a suitable function f defined, to be rigorous, over an open subset of the focal plane of the first camera which minimizes the following integral:

$$C_1(f) = \int\int (I_1(m_1(x, y)) - I_2(m_2(x, y))^2 dxdy = \int\int {}_1\Phi(f, x, y)dxdy, \tag{1}$$

computed over the previous open subset. Our first variational problem is thus to find a function f in some suitable functional space that minimizes the error measure

$C_1(f)$. The corresponding Euler-Lagrange equation is readily obtained:

$$(I_1 - I_2)(\nabla I_1 \cdot \frac{\partial \mathbf{m}_1}{\partial f} - \nabla I_2 \cdot \frac{\partial \mathbf{m}_2}{\partial f}) = 0 \tag{2}$$

The values of $\frac{\partial \mathbf{m}_1}{\partial f}$ and $\frac{\partial \mathbf{m}_2}{\partial f}$ are functions of f which are easily computed. The terms involving I_1 and I_2 are computed from the images. In order to solve (2) one can adopt a number of strategies.

One standard strategy is to consider that the function f is also a function $f(x, y, t)$ of time and to solve the following PDE:

$$f_t = \varphi(f)$$

where $\varphi(f)$ is equal to the left hand side of equation (2), with some initial condition $f(x, y, 0) = f_0(x, y)$. We thus see appear for the first time the idea that the shape of the objects in the scene, described by the function f, is obtained by allowing a surface of equation $z = f(x, y, t)$ to evolve over time, starting from some initial configuration $z = f(x, y, 0)$, according to some PDE, to hopefully converge toward the real shape of the objects in the scene when time goes to infinity. This convergence is driven by the data, i.e. the images, as expressed by the error criterion (1) or the Euler-Lagrange term $\varphi(f)$. It is known that if care is not taken, for example by adding a regularizing term to (1), the solution f is likely not to be smooth and therefore any noise in the images may cause the solution to differ widely from the real objects. This is more or less the approach taken in [18, 17]. We will postpone the solution of this problem until section 4 and in fact solve it differently from the usual way which consists of adding a regularization term to $C_1(f)$.

Another strategy is to apply the level set idea [16, 20]. Consider the family of surfaces S defined by $\mathbf{S}(x, y, t) = [x, y, f(x, y, t)]^T$. The parameters x and y are used to parameterize the surface, t is the time. The unit normal to this surface is the vector $\mathbf{N} = \pm \frac{1}{\sqrt{1+|\nabla f|^2}} [\nabla f^T, 1]^T$, the velocity vector is $\mathbf{S}_t = [0, 0, f_t]^T$ and hence the evolution of the surface can be written

$$\mathbf{S}_t = \frac{\varphi(f)}{\sqrt{1+ \mid \nabla f \mid^2}} \mathbf{N} \tag{3}$$

This expression of the evolution of the surface directly leads to a straightforward application of the level set methods. Consider a function $u(x, y, z, t)$ whose zero level set is the surface S, i.e. at each time instant t, the set of points (x, y, z) such that $u(x, y, z, t) = 0$ is identical to the surface S. Note that the function u can be

considered a temporal sequence of volumetric images. The next question is, given the fact that the time evolution of S is given by (3), what should the evolution of u be? this question has been answered in [16] and the answer is:

$$u_t = \frac{\varphi(f)}{\sqrt{1+\mid \nabla f \mid^2}} \mid \nabla u \mid$$

where ∇u is the gradient of u with respect to the first three variables. There are a couple of subtle points here. The first is that the level set methods have been designed for closed manifolds (curves or surfaces, say) but here the surface S is not closed in general, being a graph. This problem can be solved, as described for example in [3, 20]. The second point is that the coefficient of the term $\mid \nabla u \mid$ in the previous equation is defined only on the surface S and not in the whole (x, y, z) volume. But this term is needed at all points to solve for u.

We will not delve further into the last issue because it will be solved as we proceed toward better models.

3 A better functional for matching

It is clear that the error measure (1) is a bit simple for practical applications. We can extend in at least two ways. The first is to replace the difference of intensities by a measure of correlation, the hypothesis being that the scene is made of fronto parallel planes. The second is to relax this hypothesis and to take into account the orientation of the tangent plane to the surface of the object. In the first case we move along the matching criterion axis, in the second we move both along the shape and matching criterion complexity axes.

We explore those two avenues in the next sections.

3.1 Fronto parallel correlation functional

To each pair of values (x, y), corresponds a 3D point M, $\mathbf{M} = [x, y, f(x, y)]^T$ which defines two image points m_1 and m_2 as in the previous section. We can then classically define the unnormalized cross-correlation between I_1 and I_2 at the pixels m_1 and m_2. We note this cross-correlation $\langle I_1, I_2 \rangle(f, x, y)$ to acknowledge its analogy with an inner product and the fact that it depends on M:

$$\langle I_1, I_2 \rangle(f, x, y) = \frac{1}{4pq} \int_{-p}^{+p} \int_{-q}^{+q} (I_1(m_1 + m) - \overline{I_1}(m_1))(I_2(m_2 + m) - \overline{I_2}(m_2)) dm \quad (4)$$

equation where the averages $\overline{I_1}$ and $\overline{I_2}$ are classically defined as:

$$\overline{I_k}(m_k) = \frac{1}{4pq}\int_{-p}^{+p}\int_{-q}^{+q} I_k(m_k + m')dm' \quad k = 1, 2 \tag{5}$$

Finally, we note $\mid I \mid^2$ the quantity $\langle I, I\rangle$.

Note that $\langle I_1, I_2\rangle = \langle I_2, I_1\rangle$.

To simplify notations we write $\int^*$ instead of $\frac{1}{4pq}\int_{-p}^{+p}\int_{-q}^{+q}$ and define a matching functional which is the integral with respect to x and y of minus the normalized cross-correlation score $-\frac{\langle I_1, I_2\rangle}{|I_1|\cdot|I_2|}$:

$$C_2(f) = -\int\int \frac{\langle I_1, I_2\rangle}{\mid I_1 \mid \cdot \mid I_2 \mid} dxdy = \int\int {}_2\Phi(f, x, y)dxdy \tag{6}$$

the integral being computed, as in the previous section, over an open set of the focal plane of the first camera. The functional ${}_2\Phi$ is $-\frac{\langle I_1, I_2\rangle}{|I_1|\cdot|I_2|}(f, x, y)$. This quantity varies between -1 and +1, -1 indicating the maximum correlation. We have to compute its derivative with respect to f in order to obtain the Euler-Lagrange equation of the problem. The computations are simple but a little fastidious. It can first be seen that ${}_2\Phi_f$ is a sum of two terms:

$$\frac{(\mid I_1 \mid \cdot \mid I_2 \mid)_f}{\mid I_1 \mid^2 \cdot \mid I_2 \mid^2}\langle I_1, I_2\rangle - \frac{1}{\mid I_1 \mid \cdot \mid I_2 \mid}\langle I_1, I_2\rangle_f$$

The results needed to compute $\langle I_1, I_2\rangle_f$ and $(\mid I_1 \mid \cdot \mid I_2 \mid)_f$ are summarized in the following lemma which is proved in Appendix A.

Lemma 1 *The partial derivatives* $\langle I_1, I_2\rangle_f$ *and* $\mid I_k \mid_f \quad k = 1, 2$ *are given by the following formulas:*

$$\langle I_1, I_2\rangle_f = \frac{\partial \mathbf{m}_1}{\partial f}\cdot\langle\nabla I_1, I_2\rangle + \frac{\partial \mathbf{m}_2}{\partial f}\cdot\langle I_1, \nabla I_2\rangle \tag{7}$$

$$\mid I_k \mid \cdot \mid I_k \mid_f = \frac{\partial \mathbf{m}_k}{\partial f}\cdot\langle\nabla I_k, I_k\rangle \quad k = 1, 2 \tag{8}$$

where the values of the quantities $\langle\nabla I_1, I_2\rangle$, $\langle I_1, \nabla I_2\rangle$ *and* $\langle\nabla I_k, I_k\rangle$, $k = 1, 2$ *can be found in Appendix A.*

We can then proceed to solve the Euler-Lagrange equation as described in the previous section. But we will not pursue this task and explore rather a better functional.

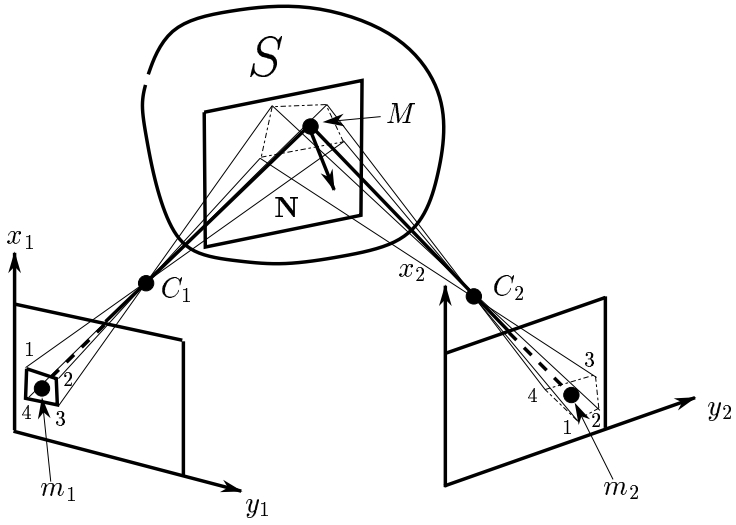


Figure 3: The square window $(1234)_1$ in the first image is back projected onto the tangent plane to the object S at point M and reprojected in the retinal plane of the second camera where it is generally not square. The observation is that the distortion between $(1234)_1$ and $(1234)_2$ can be described by a collineation which is function of M and the normal $\mathbf{N}$ to the surface of the object.

3.2 Taking into account the tangent plane to the object

We now take into account the fact that the rectangular window centered at m_2 is not rectangular but is the image in the second retina of the backprojection on the tangent plane to the object at the point $M = (x, y, f(x, y))$ of the rectangular window centered at m_1 (see figure 3). In esssence, we approximate the object S in a neighbourhood of M by its tangent plane but without assuming, as in the previous section, that this plane is fronto parallel, and in fact also that the retinal planes of the two cameras are identical. Let us first study the correspondence induced by this plane between the two images.

3.2.1 Image correspondences induced by a plane

Let us consider a plane of equation $\mathbf{N}^T\mathbf{M} - d = 0$ in the coordinate system of the first camera. d is the algebraic distance of the origin of coordinates to that plane and $\mathbf{N}$ is a unit vector normal to the plane. This plane induces a projective transformation

between the two image planes. This correspondence plays an essential role in the sequel.

To see why we obtain a projective transformation, let M be a 3D point in that plane, $\mathbf{M}_1$ and $\mathbf{M}_2$ be the two 3D vectors representing this point in the coordinate systems attached to the first and second cameras, respectively. These two 3×1 vectors are actually coordinate vectors of the two pixels m_1 and m_2 seen as projective points (see section 1). Furthermore, they are related by the following equation:

$$\mathbf{M}_2 = \mathbf{R}^T(\mathbf{M}_1 - \mathbf{t})$$

Since M belongs to the plane, $\mathbf{N}^T\mathbf{M}_1 = d$, and we have:

$$\mathbf{M}_2 = (\mathbf{R}^T - \frac{\mathbf{R}^T\mathbf{t}\mathbf{N}^T}{d})\mathbf{M}_1$$

which precisely expresses the fact that the two pixels m_1 and m_2 are related by a collineation, or projective transformation K. The 3×3 matrix representing this collineation is $(\mathbf{R}^T - \frac{\mathbf{R}^T\mathbf{t}\mathbf{N}^T}{d})$. This transformation is one to one except when the plane goes through one of the two optical centers when it becomes degenerate. We will assume that it does not go through either one of those two points and since the matrix of K is only defined up to a scale factor we might as well take it equal to:

$$\mathbf{K} = d\mathbf{R}^T - \mathbf{R}^T\mathbf{t}\mathbf{N}^T \tag{9}$$

3.2.2 The new criterion and its Euler-Lagrange equations

We just saw that a plane induces a collineation between the two retinal planes. This is the basis of the method proposed in [4] although for a very different purpose. The window alluded to in the introduction to this section is therefore the image by the collineation induced by the tangent plane of the rectangular window in image 1. This collineation is a function of the point M and of the normal to the object at M. It is therefore a function of f and ∇f that we denote by K. It satisfies the condition $K(m_1) = m_2$. The inner product (4) must be modified as follows:

$$\langle I_1, I_2\rangle(f, \nabla f, x, y) = \int^* (I_1(m_1 + m) - \overline{I_1}(m_1))(I_2(K(m_1 + m)) - \overline{I_2}(m_2))dm, \tag{10}$$

Note that, the definition of $\langle I_1, I_2 \rangle$ is no longer symmetric, because of K. In order to make it symmetric, we should define it as:

$$\langle I_1, I_2 \rangle(f, \nabla f, x, y) = \int^* (I_1(m_1 + m) - \overline{I_1}(m_1))(I_2(K(m_1 + m)) - \overline{I_2}(m_2))dm + \int^* (I_1(K^{-1}(m_2 + m')) - \overline{I_1}(m_1))(I_2(m_2 + m') - \overline{I_2}(m_2))dm' \quad (11)$$

The definition (5) of $\overline{I_1}$ (resp. of $\overline{I_2}$) is not modified in the first (resp. second) integral of the right hand side, that of $\overline{I_2}$ (resp. of $\overline{I_1}$), on the other hand, must be modified as follows:

$$\overline{I_2}(m_2) = \int^* I_2(K(m_1 + p))dp \quad \overline{I_1}(m_1) = \int^* I_1(K^{-1}(m_2 + p'))dp' \quad (12)$$

Since this new definition does not modify the fundamental ideas exposed in this paper but makes the computations significantly more complex, we will assume the definition (10) in what follows, acknowledging the fact that in practice (11) should be used.

We now want to minimize the following error measure:

$$C_3(f, \nabla f) = -\int\int \frac{\langle I_1, I_2 \rangle}{\mid I_1 \mid \cdot \mid I_2 \mid}(f, \nabla f, x, y)dxdy = \int\int {}_3\Phi(f, \nabla f, x, y)dxdy \quad (13)$$

Since the functional ${}_3\Phi$ now depends on both f and ∇f, its Euler-Lagrange equations have the form ${}_3\Phi_f - div({}_3\Phi_{\nabla f}) = 0$. We must therefore recompute ${}_3\Phi_f$ to take into account the new dependency of K upon f and compute ${}_3\Phi_{\nabla f}$.

We will simplify the computations by assuming that the collineation K can be well approximated by an affine transformation. Because of the condition $K(m_1) = m_2$, this transformation can be written:

$$K(m_1 + m) \approx m_2 + \mathbf{A}\mathbf{m}$$

where $\mathbf{A}$ is a 2×2 matrix depending upon f and ∇f.

In practice this approximation is often sufficient and we will assume that it is valid in what follows.

Under these assumptions, $\mid I_1 \mid_f$ is not modified but $\langle I_1, I_2 \rangle_f$, and $\mid I_2 \mid_f$ are, because of the new dependency in $\mathbf{A}$. The results are summarized in the following lemma which is proved in Appendix A:

Lemma 2 *The partial derivatives* $\langle I_1, I_2\rangle_f$ *and* $\mid I_2 \mid_f$ *are given by the following formulas:*

$$\langle I_1, I_2\rangle_f = \frac{\partial \mathbf{m}_1}{\partial f}\cdot\langle \nabla I_1,\, I_2\rangle + \frac{\partial \mathbf{m}_2}{\partial f}\cdot\langle I_1,\, \nabla I_2\rangle + \int^* (I_1(m_1+m) - \overline{I_1}(m_1))\nabla I_2(m_2+\mathbf{A}m)^T\mathbf{A}_f\mathbf{m}dm \quad (14)$$

$$\mid I_2 \mid \cdot \mid I_2 \mid_f = \frac{\partial \mathbf{m}_2}{\partial f}\cdot\langle \nabla I_2, I_2\rangle + \int^* (I_2(m_2+\mathbf{A}m) - \overline{I_2}(m_2))\nabla I_2(m_2+\mathbf{A}m)^T\mathbf{A}_f\mathbf{m}dm \quad (15)$$

where the values of the quantities $\langle \nabla I_1, I_2\rangle$, $\langle I_1, \nabla I_2\rangle$ *and* $\langle \nabla I_2, I_2\rangle$ *can be found in Appendix A.*

Note that, not surprisingly, equations (14) and (15) are similar to (7) and (8) with correction terms which are caused by the added complexity of our matching functional. This allows us to compute the first part of the Euler-Lagrange equation, Φ_f. The expression for $\mathbf{A}_f$ is found in the next section.

Concerning the second part, we have $\mid I_1 \mid^2_{\nabla f} = 0$ since I_1 is not a function of ∇f. The following lemma yields the needed results:

Lemma 3 *The partial derivatives* $\langle I_1, I_2\rangle_{\nabla f}$ *and* $\mid I_2 \mid_{\nabla f}$ *are given by the following formulas:*

$$\langle I_1, I_2\rangle_{\nabla f} = \int^* (I_1(m_1+m) - \overline{I_1}(m_1))\nabla I_2(m_2+\mathbf{A}\mathbf{m})^T(\mathbf{A}\mathbf{m})_{\nabla f}dm,$$

and,

$$\mid I_2 \mid \cdot \mid I_2 \mid_{\nabla f} = \int^* (I_2(m_2+\mathbf{A}\mathbf{m}) - \overline{I_2}(m_2))\nabla I_2(m_2+\mathbf{A}\mathbf{m})^T(\mathbf{A}\mathbf{m})_{\nabla f}dm,$$

expressions in which the quantity $(\mathbf{A}\mathbf{m})_{\nabla f}$ *denotes the* 2×2 *matrix defined as* $[\mathbf{A}_{(\nabla f)_x}\mathbf{m}, \mathbf{A}_{(\nabla f)_y}\mathbf{m}]$.

The proof is similar to those given in Appendix A.

These expressions allow us to compute the second part $div(\ {}_3\Phi_{\nabla f})$ of the Euler-Lagrange equations as soon as we will have made explicit the relationship between $\mathbf{A}$ and f and ∇f which is the object of the next section.

3.2.3 The matrix A and its derivatives

We have to compute the derivatives of the matrix $\mathbf{A}$ with respect to f and ∇f. We will use a mixed projective-affine approach and first compute the matrix $\mathbf{K}$ of the collineation K. Let (x_1, y_1) be the affine (image) coordinates of m_1, (x, y) those of m. We denote by $m_1 + m$ the point of affine coordinates $(x_1 + x, y_1 + y)$ or of projective coordinates $(x_1 + x, y_1 + y, 1)$. Let $\mathbf{K}$ be the 3×3 matrix of the collineation K with row vectors $\mathbf{k}_i, i = 1, 2, 3$. The affine coordinates of the point $K(m_1 + m)$ are thus equal to

$$X_2 = \frac{\mathbf{k}_1(\mathbf{m}_1 + \mathbf{m})}{\mathbf{k}_3(\mathbf{m}_1 + \mathbf{m})} \qquad Y_2 = \frac{\mathbf{k}_2(\mathbf{m}_1 + \mathbf{m})}{\mathbf{k}_3(\mathbf{m}_1 + \mathbf{m})} \tag{16}$$

where $\mathbf{m}_1$ is the vector of projective coordinates $(x_1, y_1, 1)$ and $\mathbf{m}$ the vector of coordinates $(x, y, 0)$.

According to the section 3.2.1, we know that the matrix $\mathbf{K}$ can be written as:

$$\mathbf{K} = d\mathbf{R}^T - \mathbf{T}\mathbf{N}^T$$

where

$$\begin{array}{l} d = xf_x + yf_y + f \\ \mathbf{T} = \mathbf{R}^T\mathbf{t} \\ \mathbf{N}^T = (f_x, f_y, 1) \end{array} \tag{17}$$

We have used the fact that $\mathbf{K}$ is defined up to a scale factor to get rid of the term $\frac{1}{\sqrt{1+|\nabla f|^2}}$. We have therefore established the dependency between the collineation K and f and ∇f. We deduce easily that:

$$\begin{array}{l} \frac{\partial \mathbf{K}}{\partial f} = \mathbf{R}^T \\ \frac{\partial \mathbf{K}}{\partial f_x} = x\mathbf{R}^T - [\mathbf{T}\,\mathbf{0}\,\mathbf{0}] \\ \frac{\partial \mathbf{K}}{\partial f_x} = y\mathbf{R}^T - [\mathbf{0}\,\mathbf{T}\,\mathbf{0}] \end{array} \tag{18}$$

Let us now look at the affine approximation. We can assume in general that the point m_2, image of m_1 by K, is not at infinity (can you see why?), hence one divides the numerators and the denominator of X_2 and Y_2 in equation (16) by $\mathbf{k}_3\mathbf{m}_1$ and we introduce the affine coordinates x_2 and y_2 of m_2:

$$K(m_1 + m) = \begin{cases} \frac{x_2 + \frac{\mathbf{k}_1\mathbf{m}}{\mathbf{k}_3\mathbf{m}_1}}{1 + \frac{\mathbf{k}_3\mathbf{m}}{\mathbf{k}_3\mathbf{m}_1}} \\ \frac{y_2 + \frac{\mathbf{k}_2\mathbf{m}}{\mathbf{k}_3\mathbf{m}_1}}{1 + \frac{\mathbf{k}_3\mathbf{m}}{\mathbf{k}_3\mathbf{m}_1}} \end{cases}$$

The affine approximation is obtained by assuming that $\frac{\mathbf{k}_3\mathbf{m}}{\mathbf{k}_3\mathbf{m}_1} \ll 1$, $\frac{\mathbf{k}_1\mathbf{m}}{\mathbf{k}_3\mathbf{m}_1} \ll x_2$ and $\frac{\mathbf{k}_2\mathbf{m}}{\mathbf{k}_3\mathbf{m}_1} \ll y_2$. Expanding X_2 et Y_2 up to the first order with respect to these quantities, we obtain the following expression for the matrix $\mathbf{A}$:

$$\mathbf{A} = \frac{1}{\mathbf{k}_3\mathbf{m}_1} \begin{bmatrix} \mathbf{k}_1 - \mathbf{k}_3 \\ \mathbf{k}_2 - \mathbf{k}_3 \end{bmatrix} \tag{19}$$

This expression, combined with (18) allow us to compute the partial derivatives of $\mathbf{A}$ with respect to f and ∇f and, eventually, $div({}_3\Phi_{\nabla f})$. But we will not pursue this computation since we present in section 4 a more elaborate model that encompasses this one and for which we will perform the corresponding computation.

4 An even more refined model

In this section we consider the case when the objects in the scene are not defined as the graph of a function of x and y as in the previous sections, but as the zero level set of a function $\hat{u} : \mathbf{R}^3 \to \mathbf{R}$ which we assume to be smooth, i.e. C^2. The coordinates (x, y, z) of the points in the scene which are on the surface of the objects present are thus defined by the equation $\hat{u}(x, y, z) = 0$. This approach has at least two advantages. First, by relaxing the graph assumption, it potentially allows us to use an arbitrary number of cameras to analyze the scene and second, it leads very naturally to an implementation of a surface evolution scheme through the level set method as follows.

Let us consider a family of smooth surfaces $S : (v, w, t) \to \mathbf{S}(v, w, t)$ where (v, w) parametrize the surface and t is the time. It is in general not possible to find a single mapping S from $\mathbf{R}^2$ to $\mathbf{R}^3$ that describes the entire surface of the objects (think of the sphere for example where we need at least two) but we do not have to worry about this since our results will in fact be independent of the parametrization we choose. The objects in the scene correspond to a surface $\hat{\mathbf{S}}(v, w)$ and our goal is, starting from an initial surface $\mathbf{S}_0(v, w)$, to derive a partial differential equation

$$\mathbf{S}_t = \beta\mathbf{N}, \tag{20}$$

where $\mathbf{N}$ is the unit normal to the surface, which, when solved with initial conditions $\mathbf{S}(v, w, 0) = \mathbf{S}_0(v, w)$, will yield a solution that closely approximates $\hat{\mathbf{S}}(v, w)$. The function β is determined by the matching functional that we minimize in order to solve the stereo problem. We define such a functional in the next paragraph. An interesting point is that the evolution equation (20) can be solved using the level

set method which has the advantage of coping automatically with several objects in the scene. In detail, the surfaces $\mathbf{S}$ are at each time instant the zero level sets of a function $u : \mathbf{R}^4 \rightarrow \mathbf{R}$:

$$u(\mathbf{S}, t) = 0$$

Taking derivatives with respect to u, v, t, noticing that $\mathbf{N}$ can be chosen such that $\mathbf{N} = -\frac{\nabla u}{|\nabla u|}$, where ∇ is the gradient operator for the first three coordinates of u, one finds easily that the evolution equation for u is:

$$u_t = \beta \mid \nabla u \mid \tag{21}$$

Using the same ideas as in the section 3.2, we can define the following error measure:

$$C_4(\mathbf{S}, \mathbf{N}) = -\sum_{i,j=1, i \neq j}^{n} \int\int \frac{1}{\mid I_i \mid \cdot \mid I_j \mid} \langle I_i, I_j \rangle d\sigma = \int\int {}_4\Phi(\mathbf{S}, \mathbf{N}, v, w) d\sigma \tag{22}$$

In this equation, the indexes i and j range from 1 to n, the number of views. In practice it is often not necessary to consider all possible pairs but it does not change our analysis of the problem. In equation (22), the integration is carried over with respect to the area element $d\sigma$ on the surface S. With the previous notations, we have

$$d\sigma = \mid \mathbf{S}_v \times \mathbf{S}_w \mid dvdw$$

$d\sigma$ plays the role of $dx\,dy$ in our previous analysis, $\mathbf{S}$ that of f, and $\mathbf{N} = \frac{\mathbf{S}_v \times \mathbf{S}_w}{|\mathbf{S}_v \times \mathbf{S}_w|}$, the unit normal vector to the surface S, that of ∇f.

Note that this is a significant departure from what we had before because we are multiplying our previous normalized cross-correlation score with the term $\mid \mathbf{S}_v \times \mathbf{S}_w \mid$. This has two dramatic consequences

1. It automatically regularizes the variational problem like in the geodesic snakes approach [1], and

2. it makes the problem intrinsic, i.e. independent of the parametrization of the objects in the scene.

Note also that each integral that appears in (22) is only computed for those points of the surface S which are visible in the two concerned images. Thus, *visibility and occlusion* are modelled in this approach (see section 5 for more details).

The rest of the derivation is extremely similar, although technically more complicated, to the derivations in the previous section, namely we write the Euler-Lagrange

equations of the variational problem (22), consider their component β along the normal to the surface, set up a surface evolution equation (20) and implement it by a level-set method. This is all pretty straightforward except for the announced result that the resulting value of β is *intrinsic* and does not depend upon the parametrization of the surface S.

We will in fact prove a more general result. Let $\Phi : \mathbf{R}^3 \times \mathbf{R}^3 \longrightarrow \mathbf{R}$ be a smooth function of class at least C^2 defined on the surface S and depending upon the point $\mathbf{S}(v,w)$ and the unit normal $\mathbf{N}(v,w)$ at this point, which we denote by $\Phi(\mathbf{X}, \mathbf{Y})$. Let us now consider the following error measure:

$$C(\mathbf{S}, \mathbf{S}_v, \mathbf{S}_w) = \int\int \Phi(\mathbf{S}(v,w),\ \mathbf{N}(v,w))h(v,w)dv\ dw \tag{23}$$

where the integral is taken over the surface S and $h(v,w) = \mid \mathbf{S}_v \times \mathbf{S}_w \mid$.

We prove in Appendix B the following theorem:

Theorem 1 *Under the assumptions of smoothness that have been made for the function Φ and the surface S, the component of the Euler-Lagrange equations for criterion (23) along the normal to the surface is the product of h with an intrinsic factor, i.e. which does not depend upon the parametrization (v,w). Furthermore, this component is equal to*

$$h(-\Phi_{\mathbf{X}}\mathbf{N} + 2H(\Phi - \Phi_{\mathbf{Y}}\mathbf{N}) - Trace((\Phi_{\mathbf{XY}})_{T_S} + d\mathbf{N} \circ (\Phi_{\mathbf{YY}})_{T_S})) \tag{24}$$

where all quantities are evaluated at the point $\mathbf{S}$ of normal $\mathbf{N}$ of the surface, T_S is the tangent plane to the surface at the point $\mathbf{S}$. $d\mathbf{N}$ is the differential of the Gauss map of the surface, H is its mean curvature, $\Phi_{\mathbf{XY}}$ and $\Phi_{\mathbf{YY}}$ are the second order derivatives of Φ, $(\Phi_{\mathbf{XY}})_{T_S}$ and $(\Phi_{\mathbf{YY}})_{T_S}$ their restrictions to the tangent plane T_S of the surface at the point S.

Note that the error criterion (22) is of the form (23) if we define Φ to be

$$-\sum_{i,j=1,i\neq j}^{n} \frac{1}{\mid I_i \mid \cdot \mid I_j \mid}\langle I_i, I_j\rangle$$

According to the theorem 1, in order to compute the velocity β along the normal in the evolution equations (20) or (21) we only need to compute $\Phi_{\mathbf{S}}$, $\Phi_{\mathbf{N}}$, $\Phi_{\mathbf{SN}}$ and $\Phi_{\mathbf{NN}}$ as well as the second order intrinsic differential properties of the surface S. Using the fact that the function Φ is a sum of functions $\Phi_{ij} = -\frac{1}{|I_i|\cdot|I_j|}\langle I_i,\ I_j\rangle$, the problem is broken down into the problem of computing the corresponding derivatives

of the Φ_{ij}'s, which, for the first order derivatives is extremely similar to what we have done in the section 3.2. The computations are carried out in Appendix C.

In terms of the level set implementation, we ought to make a few remarks. The first is to explain how we compute β in equation (21) at each point (x, y, z) rather than on the surface S. It should be clear that we do not have any problem for computing $\mathbf{N} = -\frac{\nabla u}{|\nabla u|}$ and $2H = div(\frac{\nabla u}{|\nabla u|})$ and $d\mathbf{N}$ which is the differential of the Gauss map of the level set surface going through the point (x, y, z). The vectors $\Phi_{\mathbf{X}}$, $\Phi_{\mathbf{Y}}$, the matrices $\Phi_{\mathbf{XY}}$, $\Phi_{\mathbf{YY}}$ are computed as explained in the Appendix C.

The second remark is that we can now write equation (21) as follows:

$$u_t = \mid \nabla u \mid div(\Phi \frac{\nabla u}{\mid \nabla u \mid}) - div(\frac{\nabla u}{\mid \nabla u \mid}) \nabla_1 \Phi \cdot \nabla u - Trace((\Phi_{XY})_{T_S} + d\mathbf{N} \circ (\Phi_{\mathbf{YY}})_{T_S}) \mid \nabla u \mid \quad (25)$$

where at each point (x, y, z) the tangent plane T_S is that of the level set surface $u = constant$ going through that point. The operator ∇ indicates derivatives with respect to the space variables (x, y, z) whereas ∇_1 indicates derivatives with respect to the coordinates of the normal.

The first term $\mid \nabla u \mid div(\Phi \frac{\nabla u}{|\nabla u|})$ is identical to the one in the work of Caselles, Kimmel, Sapiro and Sbert [2] on the use of minimal surfaces or geodesic snakes to segment volumetric images. Our other terms come from the particular process that we are modelling, i.e. stereo. Because these terms involve only second order differential quantities, we would like to make the following conjecture which we hope can be proved in the near future.

Conjecture 1 *Under some reasonable smoothness assumptions on the image intensity functions, the equation (25) admits a unique stable viscosity solution in* $C(\mathbf{R}^3 \times [0, \infty[) \cap L^\infty(0, T; W^{1,\infty}(\mathbf{R}^3))$ *for all* $T < \infty$. $W^{1,\infty}$ *is the space of bounded Lipschitz functions in* $\mathbf{R}^3$.

5 Two-dimensional implementation

We now present an implementation in two dimensions of the theory that has been described in the previous section. The situation is described in figure 5 which depicts a planar object S with two connected components. This object is observed by the five cameras numbered from 1 to 5 and the level set corresponding to some value k has been drawn to illustrate how the visibility condition is taken into account when

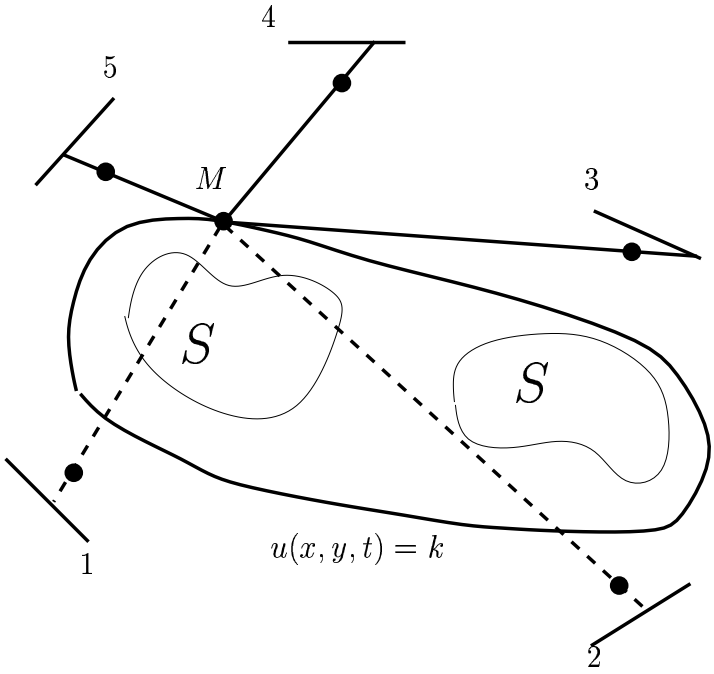


Figure 4: The 2D implementation of the algorithm described in the previous section.

solving equation (25): the point M on the level set is seen only by the cameras 3 to 5 and not by 1 and 2.

Planar objects are viewed by linear cameras. Pixels and images are functions of only one variable. The inner product (10) is:

$$\langle I_i, I_j \rangle(\mathbf{S}, \mathbf{N}) = \frac{1}{2p} \int_{-p}^{p} (I_i(m_i + m) - \overline{I_i}(m_i))(I_j(K_{ij}(m_i + m)) - \overline{I_j}(m_j))dm$$

We use the affine approximation: $K_{ij}(m_i + m) = m_j + \alpha_{ij} m$. The surface $\mathbf{S}(v, w)$ is a plane curve $\mathbf{S}(v)$ and the error measure:

$$\int \Phi(\mathbf{S}(v), \mathbf{N}(v)) h(v) dv$$

where $h(v) = |\mathbf{S}_v|$ and Φ is still $-\sum_{i,j=1, i \neq j}^{n} \frac{1}{|I_i| \cdot |I_j|} \langle I_i, I_j \rangle$. Let σ be the arc-length parameter of $\mathbf{S}$ ($d\sigma = h(v)dv$), $\mathbf{T}$ the unit tangent, and κ its curvature, the normal component β of the Euler-Lagrange equation simplifies to the intrinsic quantity:

$$\beta = \kappa \Phi + [\Phi_{\mathbf{S}} + \kappa(\mathbf{T}\mathbf{T}^T - \mathbf{N}\mathbf{N}^T)\Phi_{\mathbf{N}}^T] \cdot \mathbf{N} + \mathbf{T}^T(\Phi_{\mathbf{SN}} + \kappa \Phi_{\mathbf{NN}})\mathbf{T}$$

In order to implement the evolution equation (21) of $u(x, y, t)$, the following steps are required to get β at point $\mathbf{M} = (x, y)$ at time t:

- Considering the level set $\mathbf{S}(v)$ of φ passing through point $\mathbf{M}$, determine from which cameras is $\mathbf{M}$ visible, let us call them the S-cameras. They are determined by assuming that the level curve going through the point is opaque.
- Compute the normal $\mathbf{N}$ and the curvature κ of $\mathbf{S}$ at $\mathbf{M}$.
- For each S-camera, compute m_i, $\frac{\partial m_i}{\partial \mathbf{S}}$ and $\frac{d}{d\sigma} m_i$.
- For each pair of S-cameras, compute $\alpha_{ij}, \alpha_{ij_{\mathbf{S}}}, \alpha_{ij_{\mathbf{N}}}, \frac{d}{d\sigma}\alpha_{ij}$, and $\frac{d}{d\sigma}\alpha_{ij_{\mathbf{N}}}$.
- Compute $\langle I_i, I_j \rangle, \langle I_i, I_j \rangle_{\mathbf{S}}, \langle I_i, I_j \rangle_{\mathbf{N}}, \frac{d}{d\sigma}\langle I_i, I_j \rangle$ and $\frac{d}{d\sigma}\langle I_i, I_j \rangle_{\mathbf{N}}$
- Compute $\Phi, \Phi_{\mathbf{S}}, \Phi_{\mathbf{N}}$ and $\frac{d}{d\sigma}\Phi_{\mathbf{N}}$ hence β.

This scheme has been tested with synthetic noisy images of 2D objects viewed from a number of cameras located around them and on a number of real images. Some implementation details are discussed in the Appendix E.

We first present our results on three synthetic objects, each one being meant to demonstrate one feature of the algorithm. All those results have been obtained with

18 cameras observing the scene and are presented in an homogeneous manner: on the left hand side of the figure we show the views from some of the cameras, on the right hand side we show the convergence of the zero level set.

We start with figure 5 which demonstrates that the algorithm works for non convex objects. Notice that the background is dark.

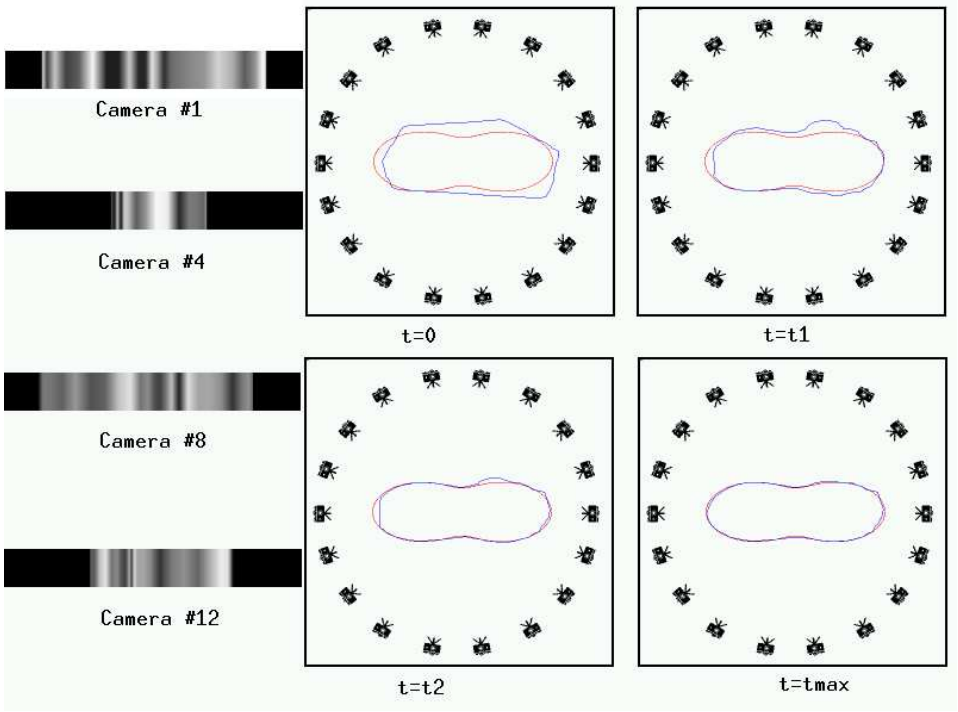


Figure 5: 2D case – Detection of a non convex object.

Figure 6 shows the stereovision process for two circles viewed on a random background. This example shows that our algorithm can deal with multiple objects (note the change in topology at time $t2$) and can cope with a textured background without being fooled.

Figure 7 shows the results for two squares viewed on a black background. This example shows that our modelling can cope somewhat with non smooth objects.

Figure 8 shows a real example in the case of two cameras. The stereo pair of a human face is shown on the left hand side of the figure, the trace of the epipolar plane (vertical in this case) being shown in red. On the right hand side of the figure

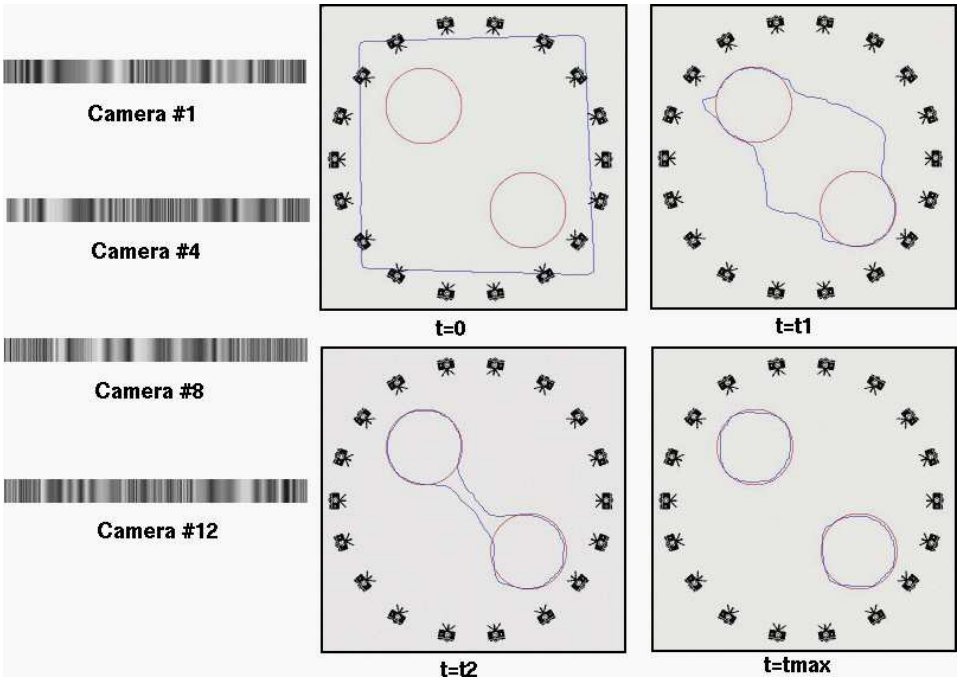


Figure 6: 2D case – Detection of two circular objects.

we show the evolution of the zero level set at four time instants. The red curve is the result of the correlation algorithm described in [5] while the blue curve is the level set. It can be seen that the convergence is satisfactory.

In terms of computing time, it takes between five and ten minutes on a fast workstation to achieve convergence in the examples shown. We are clearly still far from real time video but reaching this goal was clearly not the purpose of this paper.

6 Conclusion

We have presented a novel geometric approach for solving the stereo problem from an arbitrary number of views which degrades gracefully with the number of views. It is based upon writing a variational principle that must be satisfied by the surfaces of the objects to be detected. The design of the variational principle allows us to clearly and cleanly incorporate the hypotheses we make about the objects in the scene and

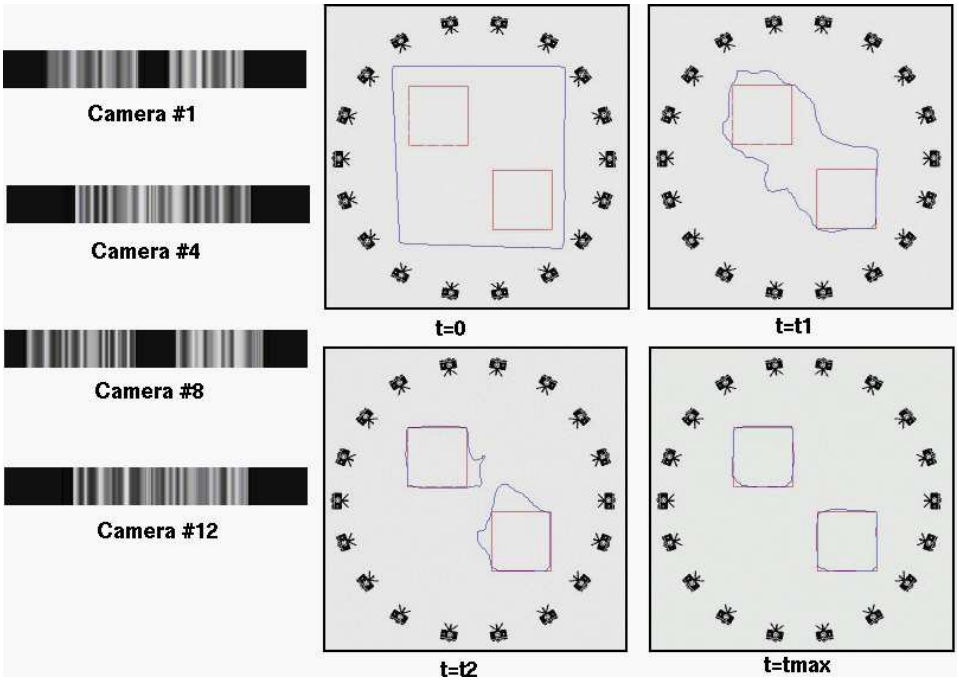


Figure 7: 2D case – Detection of two squares.

how we obtain correspondences between image points. The Euler-Lagrange equations which are deduced from the variational principle provide a set of PDE's which are used to deform an initial set of surfaces which then move towards the objects to be detected. The level set implementation of these PDE's potentially provides an efficient and robust way of achieving the surface evolution and to deal automatically with changes in the surface topology during the deformation.

Our implementation is so far only two-dimensional, i.e. it works only in epipolar planes but the results we have obtained on synthetic noisy planar images look promising enough to lead us into thinking that the approach will also be successful in 3D.

As far as the theory goes, we conjecture that the PDE that we have obtained in the most general case that we have considered has, under some technical smoothness assumptions, a unique solution in the viscosity sense. Issues of stability and correctness of the solution should also be explored.

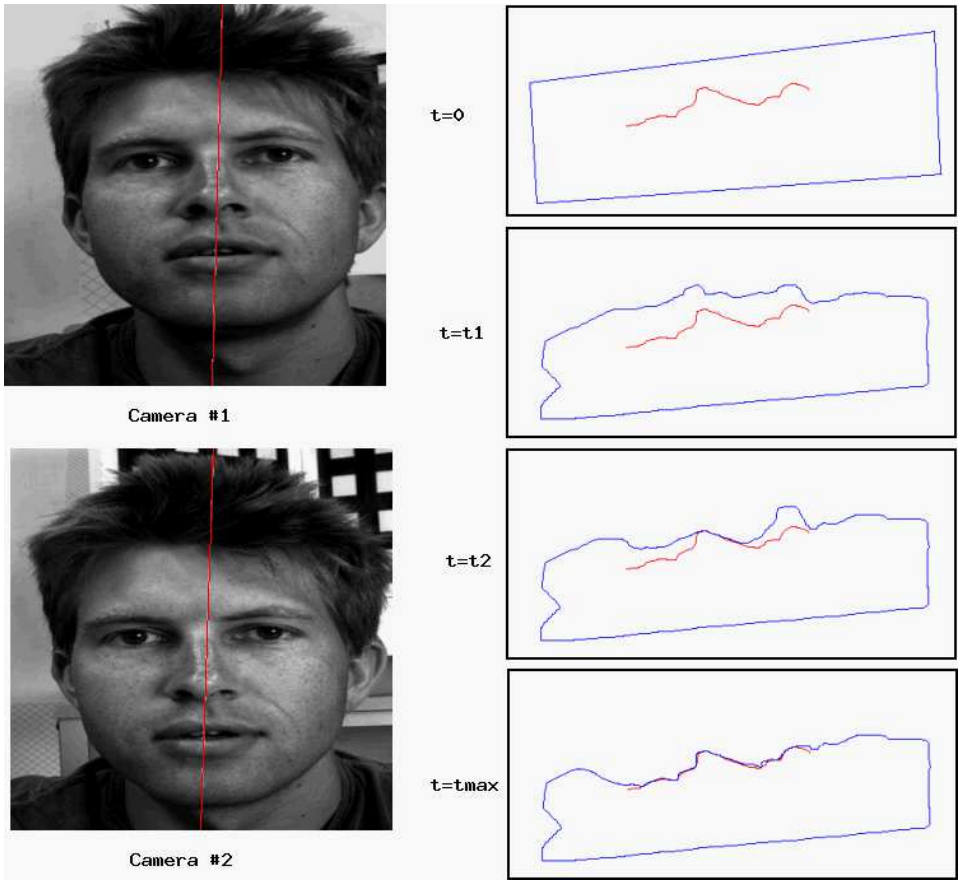


Figure 8: 2D case – Detection of a human face in an epipolar plane (in red).

To finish on a methodological note, we believe that it is by working in such well defined conceptual or mathematical frameworks, as the calculus of variations or the theory of PDE's, where the tools exist to prove (or disprove) the correctness of algorithms, that we will be able to bring computer vision up to the level of predictability where it can be used reliably in real applications and interfaced to other components to build complex systems.

A Proofs of lemmas 1 and 2

We first prove lemma 1. Using the definitions it can be seen that the derivative of $\langle I_1, I_2 \rangle$ is given by:

$$\langle I_1, I_2 \rangle_f = \int^* (\nabla I_1(m_1+m) \cdot \frac{\partial \mathbf{m}_1}{\partial f} - \frac{\partial \overline{I_1}}{\partial f}(m_1))(I_2(m_2+m) - \overline{I_2}(m_2))dm + \int^* (I_1(m_1+m) - \overline{I_1}(m_1))(\nabla I_2(m_2+m) \cdot \frac{\partial \mathbf{m}_2}{\partial f} - \frac{\partial \overline{I_2}}{\partial f}(m_2))dm, \quad (26)$$

while the derivative of $\mid I_k \mid \; k = 1, 2$ can be easily obtained from that of $\mid I_k \mid^2$:

$$\mid I_k \mid \cdot \mid I_k \mid_f = \int^* (\nabla I_k(m_k+m) \cdot \frac{\partial \mathbf{m}_k}{\partial f} - \frac{\partial \overline{I_k}}{\partial f}(m_k))(I_k(m_k+m) - \overline{I_k}(m_k))dm \quad (27)$$

The derivatives $\frac{\partial \overline{I_k}}{\partial f}$ $k = 1, 2$ do not need to be computed since they cancel out from the integrals, for example:

$$\int^* \frac{\partial \overline{I_k}}{\partial f}(m_k)(I_k(m_k+m) - \overline{I_k}(m_k))dm = \frac{\partial \overline{I_k}}{\partial f}(m_k) \int^* (I_k(m_k+m) - \overline{I_k}(m_k))dm = 0$$

This last remark allows us to rewrite (26) and (27) in a simpler manner:

$$\langle I_1, I_2 \rangle_f = \frac{\partial \mathbf{m}_1}{\partial f} \cdot \int^* \nabla I_1(m_1+m)(I_2(m_2+m) - \overline{I_2}(m_2))dm + \frac{\partial \mathbf{m}_2}{\partial f} \cdot \int^* (I_1(m_1+m) - \overline{I_1}(m_1))\nabla I_2(m_2+m)dm \quad (28)$$

$$\mid I_k \mid \cdot \mid I_k \mid_f = \frac{\partial \mathbf{m}_k}{\partial f} \cdot \int^* \nabla I_k(m_k+m)(I_k(m_k+m) - \overline{I_k}(m_k))dm \; k = 1, 2 \quad (29)$$

These are the expressions of lemma 1.

We now prove lemma 2. Similarly to lemma 1, the derivatives of $\overline{I_k}$, $k = 1, 2$ with respect to f are not needed.

A computation alike the one that led to (28) yields:

$$\frac{\partial \langle I_1, I_2 \rangle}{\partial f} = \frac{\partial \mathbf{m}_1}{\partial f} \cdot \int^* \nabla I_1(m_1+m)(I_2(m_2+\mathbf{A}\mathbf{m}) - \overline{I_2}(m_2))dm + \int^* (I_1(m_1+m) - \overline{I_1}(m_1))\nabla I_2(m_2+\mathbf{A}\mathbf{m})^T (\frac{\partial \mathbf{m}_2}{\partial f} + \frac{\partial \mathbf{A}}{\partial f}\mathbf{m})dm$$

which immediatly yields (14). We also need:

$$\mid I_2 \mid \cdot \mid I_2 \mid_f = \int^* (I_2(m_2 + \mathbf{Am}) - \overline{I_2}(m_2))\nabla I_2(m_2 + \mathbf{Am})^T (\frac{\partial \mathbf{m}_2}{\partial f} + \frac{\partial \mathbf{A}}{\partial f}\mathbf{m}) dm$$

which yields (15. where we have written:

$$\langle \nabla I_2, I_2 \rangle = \int^* \nabla I_2(m_2 + \mathbf{Am})(I_2(m_2 + \mathbf{Am}) - \overline{I_2}(m_2)) dm$$

B Proof of theorem 1

We assume for simplicity but without loss of generality that the parametrization of S is orthogonal, i.e. that $\mathbf{S}_v \cdot \mathbf{S}_w = 0$.

Let us write $\psi(\mathbf{S}, \mathbf{S}_v, \mathbf{S}_w)$ for $h(\mathbf{S}_v, \mathbf{S}_w)\Phi(\mathbf{S}, \mathbf{N})$ where

$$h(\mathbf{S}_v, \mathbf{S}_w) = \mid \mathbf{S}_v \times \mathbf{S}_w \mid$$
$$\mathbf{N} = \frac{\mathbf{S}_v \times \mathbf{S}_w}{h}$$

The Euler-Lagrange equations of the error criterion (23) can be written

$$\psi_{\mathbf{S}} - \frac{d}{dv}\psi_{\mathbf{S}_v} - \frac{d}{dw}\psi_{\mathbf{S}_w}$$

Since h does not depend upon $\mathbf{S}$, we have $\psi_{\mathbf{S}} = h\Phi_{\mathbf{S}}$ and

$$\psi_{\mathbf{S}_v} = \Phi h_{\mathbf{S}_v} + h\Phi_{\mathbf{Y}}\mathbf{N}_{\mathbf{S}_v}$$

There is a similar expression for $\psi_{\mathbf{S}_w}$. We first prove the following lemma:

Lemma 4 *The derivatives of h with respect to $\mathbf{S}_v$ and $\mathbf{S}_w$ are given by*

$$h_{\mathbf{S}_v} = \frac{\mid \mathbf{S}_w \mid^2}{h}\mathbf{S}_v^T \quad h_{\mathbf{S}_w} = \frac{\mid \mathbf{S}_v \mid^2}{h}\mathbf{S}_w^T \tag{30}$$

Proof : We use the fact that $h^2 = \mid \mathbf{S}_v \times \mathbf{S}_w \mid^2$. From this it follows that

$$hh_{\mathbf{S}_v} = (\mathbf{S}_v \times \mathbf{S}_w)^T (\mathbf{S}_v \times \mathbf{S}_w)_{\mathbf{S}_v}$$

It is easy to verify that $(\mathbf{x} \times \mathbf{y})_{\mathbf{x}} = -[\mathbf{y}]_\times$ for all vectors $\mathbf{x}$ and $\mathbf{y}$ of $\mathbf{R}^3$. From this follows the result that

$$hh_{\mathbf{S}_v} = -(\mathbf{S}_v \times \mathbf{S}_w)^T [\mathbf{S}_w]_\times$$

But the matrix $[\mathbf{S}_w]_\times$ is antisymmetric, hence $-[\mathbf{S}_w]_\times = [\mathbf{S}_w]_\times^T$. From this it follows that:

$$hh_{\mathbf{S}_v} = [[\mathbf{S}_w]_\times(\mathbf{S}_v \times \mathbf{S}_w)]^T$$

The vector in bracket is:

$$\mathbf{S}_w \times (\mathbf{S}_v \times \mathbf{S}_w) = \mid \mathbf{S}_w \mid^2 \mathbf{S}_v - (\mathbf{S}_v \cdot \mathbf{S}_w)\mathbf{S}_w$$

The result follows from the assumption that the parametrization of S is orthogonal. □

Let us now attack the computation of $\mathbf{N}_{\mathbf{S}_v}$ and $\mathbf{N}_{\mathbf{S}_w}$. We have the following lemma:

Lemma 5 *The derivatives of* $\mathbf{N}$ *with respect to* $\mathbf{S}_v$ *and* $\mathbf{S}_w$ *are given by:*

$$\mathbf{N}_{\mathbf{S}_v} = -\frac{1}{h} h_{\mathbf{S}_v}^T \mathbf{N}^N \quad \mathbf{N}_{\mathbf{S}_w} = -\frac{1}{h} h_{\mathbf{S}_w}^T \mathbf{N}^T \tag{31}$$

Proof : By definition we have

$$\mathbf{N}_{\mathbf{S}_v} = \frac{1}{h}(\mathbf{S}_v \times \mathbf{S}_w)_{\mathbf{S}_v} - \frac{1}{h^2}(\mathbf{S}_v \times \mathbf{S}_w) h_{\mathbf{S}_v}$$

Replacing $(\mathbf{S}_v \times \mathbf{S}_w)_{\mathbf{S}_v}$ and $h_{\mathbf{S}_v}$ by their values in this expression, we obtain

$$\mathbf{N}_{\mathbf{S}_v} = -\frac{1}{h}[\mathbf{S}_w]_\times - \frac{\mid \mathbf{S}_w \mid^2}{h^3}(\mathbf{S}_v \times \mathbf{S}_w)\mathbf{S}_v^T$$

and similarly for $\mathbf{N}_{\mathbf{S}_w}$:

$$\mathbf{N}_{\mathbf{S}_w} = \frac{1}{h}[\mathbf{S}_v]_\times - \frac{\mid \mathbf{S}_v \mid^2}{h^3}(\mathbf{S}_v \times \mathbf{S}_w)\mathbf{S}_w^T$$

Now do the following rewriting:

$$(\mathbf{S}_v \times \mathbf{S}_w)\mathbf{S}_v^T = -([\mathbf{S}_w]_\times \mathbf{S}_v)\mathbf{S}_v^T = -[\mathbf{S}_w]_\times(\mathbf{S}_v \mathbf{S}_v^T)$$

and similarly:

$$(\mathbf{S}_v \times \mathbf{S}_w)\mathbf{S}_w^T = [\mathbf{S}_v]_\times(\mathbf{S}_w \mathbf{S}_w^T)$$

Replacing now in $\mathbf{N}_{\mathbf{S}_v}$ and $\mathbf{N}_{\mathbf{S}_w}$ we obtain

$$\mathbf{N}_{\mathbf{S}_v} = -\frac{1}{h}[\mathbf{S}_w]_\times(\mathbf{I} - \frac{\mid \mathbf{S}_w \mid^2}{h^2}\mathbf{S}_v \mathbf{S}_v^T)$$
$$\mathbf{N}_{\mathbf{S}_w} = \frac{1}{h}[\mathbf{S}_v]_\times(\mathbf{I} - \frac{\mid \mathbf{S}_v \mid^2}{h^2}\mathbf{S}_w \mathbf{S}_w^T)$$

Let us now use the fact that $h^2 = \mid \mathbf{S}_v \mid^2 \mid \mathbf{S}_w \mid^2$. The term $\mathbf{I} - \frac{|\mathbf{S}_v|^2}{h^2}\mathbf{S}_w\mathbf{S}_w^T$ can be rewritten as

$$-\frac{1}{\mid \mathbf{S}_v \mid^2}(\mathbf{S}_v\mathbf{S}_v^T - \mid \mathbf{S}_v \mid^2 \mathbf{I})$$

But $\mathbf{S}_v\mathbf{S}_v^T - \mid \mathbf{S}_v \mid^2 \mathbf{I}$ is equal to $[\mathbf{S}_v]_\times^2$. We obtain a new expression for $\mathbf{N}_{\mathbf{S}_v}$:

$$\mathbf{N}_{\mathbf{S}_v} = \frac{1}{h \mid \mathbf{S}_v \mid^2}[\mathbf{S}_w]_\times[\mathbf{S}_v]_\times^2$$

Because the parametrization is orthogonal, we can find a simpler form for $[\mathbf{S}_w]_\times[\mathbf{S}_v]_\times^2$. Indeed we have

$$\mathbf{S}_w \times (\mathbf{S}_v \times (\mathbf{S}_v \times \mathbf{x})) = -(\mathbf{x}, \mathbf{S}_v, \mathbf{S}_w)\mathbf{S}_v$$

for all vectors $\mathbf{x}$ of $\mathbf{R}^3$. Thus

$$[\mathbf{S}_w]_\times[\mathbf{S}_v]_\times^2 = -\mathbf{S}_v(\mathbf{S}_v \times \mathbf{S}_w)^T$$

Combining this with equations (30), the results follows for $\mathbf{N}_{\mathbf{S}_v}$ and a similar computation yields the result for $\mathbf{N}_{\mathbf{S}_w}$. □

Returning to our original goal, we have now discovered that

$$\psi_{\mathbf{S}_v} = \Phi h_{\mathbf{S}_v} - \Phi_{\mathbf{Y}} h_{\mathbf{S}_v}^T \mathbf{N}^T$$

and

$$\psi_{\mathbf{S}_w} = \Phi h_{\mathbf{S}_w} - \Phi_{\mathbf{Y}} h_{\mathbf{S}_w}^T \mathbf{N}^T$$

We encourage the reader to make sure that he agrees with the dimensions of the various matrices appearing in these two expressions.

We now have to compute $\frac{d}{dv}\psi_{\mathbf{S}_v} + \frac{d}{dw}\psi_{\mathbf{S}_w}$. Applying the chain rule for derivation, we obtain the following terms:

$$\begin{aligned} \Phi(\frac{d}{dv}h_{\mathbf{S}_v} + \frac{d}{dw}h_{\mathbf{S}_w}) + \Phi_X(\mathbf{S}_v h_{\mathbf{S}_v} + \mathbf{S}_w h_{\mathbf{S}_w}) + \\ \Phi_{\mathbf{Y}}[\mathbf{N}_v h_{\mathbf{S}_v} + \mathbf{N}_w h_{\mathbf{S}_w} - h_{\mathbf{S}_v}^T \mathbf{N}_v^T - h_{\mathbf{S}_w}^T \mathbf{N}_w^T - \\ (\frac{d}{dv}h_{\mathbf{S}_v} + \frac{d}{dw}h_{\mathbf{S}_w})^T \mathbf{N}^T] - \\ (\mathbf{S}_v^T \Phi_{\mathbf{XY}} h_{\mathbf{S}_v}^T + \mathbf{S}_w^T \Phi_{\mathbf{XY}} h_{\mathbf{S}_w}^T + \mathbf{N}_v^T \Phi_{\mathbf{YY}} h_{\mathbf{S}_v}^T + \mathbf{N}_w^T \Phi_{\mathbf{YY}} h_{\mathbf{S}_w}^T)\mathbf{N}^T \end{aligned} \tag{32}$$

Remember that we are only interested in the terms along the normal $\mathbf{N}$. This allows to eliminate a number of terms, according to the following lemma:

Lemma 6 *In equation (32), the following term belongs to the tangent plane and can therefore be discarded:*

$$\Phi_{\mathbf{X}}(\mathbf{S}_v h_{\mathbf{S}_v} + \mathbf{S}_w h_{\mathbf{S}_w}) + \Phi_{\mathbf{Y}}[\mathbf{N}_v h_{\mathbf{S}_v} + \mathbf{N}_w h_{\mathbf{S}_w} - h^T_{\mathbf{S}_v}\mathbf{N}^T_v - h^T_{\mathbf{S}_w}\mathbf{N}^T_w]$$

Proof : Indeed, $\Phi_{\mathbf{Y}}\mathbf{N}_v$ and $\Phi_{\mathbf{X}}\mathbf{S}_v$ (resp. $\Phi_{\mathbf{Y}}\mathbf{N}_w$ and $\Phi_{\mathbf{X}}\mathbf{S}_w$) are scalars and $h_{\mathbf{S}_v}$ (resp. $h_{\mathbf{S}_w}$) is the tangent plane to S according to (30). Similarly $\Phi_{\mathbf{Y}}h^T_{\mathbf{S}_v}$ (resp. $\Phi_{\mathbf{Y}}h^T_{\mathbf{S}_w}$) is a scalar and $\mathbf{N}_v$ (resp. $\mathbf{N}_w$) is the derivative of the unit vector $\mathbf{N}$ and is therefore orthogonal to $\mathbf{N}$, hence is in the tangent plane. □

We now study the vector $(\frac{d}{dv}h_{\mathbf{S}_v} + \frac{d}{dw}h_{\mathbf{S}_w})^T$ and show that it is directed along the normal $\mathbf{N}$ to the surface.

Lemma 7 *The vector $(\frac{d}{dv}h_{\mathbf{S}_v} + \frac{d}{dw}h_{\mathbf{S}_w})^T$ is in the direction of the normal $\mathbf{N}$ to the surface. Its component in that direction is equal to $2hH$, where H is the mean curvature of the surface.*

Proof : Let us first compute the normal component. According to equations (30), we have

$$\frac{d}{dv}h^T_{\mathbf{S}_v} = \frac{|\,\mathbf{S}_w\,|^2}{h}\mathbf{S}_{vv} + \frac{d}{dv}(\frac{|\,\mathbf{S}_w\,|^2}{h})\mathbf{S}_v \quad \frac{d}{dw}h^T_{\mathbf{S}_w} = \frac{|\,\mathbf{S}_v\,|^2}{h}\mathbf{S}_{ww} + \frac{d}{dw}(\frac{|\,\mathbf{S}_v\,|^2}{h})\mathbf{S}_w$$

Taking the inner product with the normal $\mathbf{N}$, and using the fact that $h^2 = |\,\mathbf{S}_v\,|^2|\,\mathbf{S}_w\,|^2$ we obtain:

$$(\frac{d}{dv}h_{\mathbf{S}_v} + \frac{d}{dw}h_{\mathbf{S}_w})^T \cdot \mathbf{N} = h(\frac{\mathbf{S}_{vv}\cdot\mathbf{N}}{|\,\mathbf{S}_v\,|^2} + \frac{\mathbf{S}_{ww}\cdot\mathbf{N}}{|\,\mathbf{S}_w\,|^2}) = 2hH$$

as shown in Appendix D.

Let us now consider the component of our vector $(\frac{d}{dv}h_{\mathbf{S}_v} + \frac{d}{dw}h_{\mathbf{S}_w})^T$ in the tangent plane to the surface. This is the sum of four contributions. The terms $\frac{d}{dv}(\frac{|\mathbf{S}_w|^2}{h})\mathbf{S}_v$, $\frac{d}{dw}(\frac{|\mathbf{S}_v|^2}{h})\mathbf{S}_w$ and the projections of $\frac{|\mathbf{S}_w|^2}{h}\mathbf{S}_{vv}$ and $\frac{|\mathbf{S}_v|^2}{h}\mathbf{S}_{ww}$ onto the tangent plane which are given by, given the fact that $\mathbf{S}_v$ and $\mathbf{S}_w$ are orthogonal:

$$\frac{|\,\mathbf{S}_w\,|^2}{h}\frac{\mathbf{S}_{vv}\cdot\mathbf{S}_v}{|\,\mathbf{S}_v\,|^2}\mathbf{S}_v + \frac{|\,\mathbf{S}_w\,|^2}{h}\frac{\mathbf{S}_{vv}\cdot\mathbf{S}_w}{|\,\mathbf{S}_w\,|^2}\mathbf{S}_w,$$

and

$$\frac{|\,\mathbf{S}_v\,|^2}{h}\frac{\mathbf{S}_{ww}\cdot\mathbf{S}_v}{|\,\mathbf{S}_v\,|^2}\mathbf{S}_v + \frac{|\,\mathbf{S}_v\,|^2}{h}\frac{\mathbf{S}_{ww}\cdot\mathbf{S}_w}{|\,\mathbf{S}_w\,|^2}\mathbf{S}_w$$

Let us use the notations of Appendix D) and let $E = \mid \mathbf{S}_v \mid^2$ and $G = \mid \mathbf{S}_w \mid^2$. The component along $\mathbf{S}_v$ is equal to:

$$((\frac{G}{E})^{\frac{1}{2}})_v + \frac{G}{E\sqrt{EG}}(\mathbf{S}_{vv} \cdot \mathbf{S}_v) + \frac{1}{\sqrt{EG}}(\mathbf{S}_{ww} \cdot Sv)$$

But we have $\mathbf{S}_{vv} \cdot \mathbf{S}_v = \frac{1}{2}E_v$ and, because $\mathbf{S}_v \cdot \mathbf{S}_w = 0$, $\mathbf{S}_{ww} \cdot \mathbf{S}_v = -\mathbf{S}_{vw} \cdot \mathbf{S}_w = -\frac{1}{2}G_v$. Substituting in the previous expression, we find:

$$\frac{1}{2}\frac{G_v E - G E_v}{E^2}(\frac{G}{E})^{-\frac{1}{2}} + \frac{1}{2}\frac{G E_v}{E\sqrt{EG}} - \frac{1}{2}\frac{G_v}{\sqrt{EG}} = 0$$

It can be verified that this is also true for the component along $\mathbf{S}_w$. □

Let us now consider the term containing $\Phi_{\mathbf{XY}}$ in equation (32). We prove the following lemma:

Lemma 8 *The term $\mathbf{S}_v^T \Phi_{\mathbf{XY}} h_{\mathbf{S}_v}^T + \mathbf{S}_w^T \Phi_{\mathbf{XY}} h_{\mathbf{S}_w}^T$ is equal to the product of h by the trace of the restriction to T_S of the linear map $\mathbf{R}^3 \longrightarrow \mathbf{R}^3$ defined by $\Phi_{\mathbf{XY}}$.*

Proof : Just replace $h_{\mathbf{S}_v}$ and $h_{\mathbf{S}_w}$ by their values (30). We obtain:

$$\mathbf{S}_v^T \Phi_{\mathbf{XY}} h_{\mathbf{S}_v}^T + \mathbf{S}_w^T \Phi_{\mathbf{XY}} h_{\mathbf{S}_w}^T = h(\mathbf{t}_1 \Phi_{\mathbf{XY}} \mathbf{t}_1 + \mathbf{t}_2 \Phi_{\mathbf{XY}} \mathbf{t}_2)$$

Introducing the two unit vectors $\mathbf{t}_1 = \frac{\mathbf{S}_v}{|\mathbf{S}_v|}$ and $\mathbf{t}_2 = \frac{\mathbf{S}_w}{|\mathbf{S}_w|}$ which form an orthonormal basis of T_S, we obtain the result. □

Let us now consider the term containing $\Phi_{\mathbf{YY}}$ in equation (32). We prove the following proposition:

Proposition 1 *The expression $\mathbf{N}_v^T \Phi_{\mathbf{YY}} h_{\mathbf{S}_v}^T + \mathbf{N}_w^T \Phi_{\mathbf{YY}} h_{\mathbf{S}_w}^T$ is equal to the product of h by the trace of the product of the composition of the differential $d\mathbf{N}$ of the Gauss map of the surface S with the restriction to the tangent plane T_S of the surface of the endomorphism defined by the Hessian $\Phi_{\mathbf{YY}}$ of Φ, respectively.*

Proof : Using again the notations of Appendix D, we express $\mathbf{N}_v$ and $\mathbf{N}_w$, which are in the tangent plane to S, as linear combinations of $\mathbf{S}_v$ and $\mathbf{S}_w$:

$$-\mathbf{N}_v = \frac{e}{\mid \mathbf{S}_v \mid^2}\mathbf{S}_v + \frac{f}{\mid \mathbf{S}_w \mid^2}\mathbf{S}_w$$
$$-\mathbf{N}_w = \frac{f}{\mid \mathbf{S}_v \mid^2}\mathbf{S}_v + \frac{g}{\mid \mathbf{S}_w \mid^2}\mathbf{S}_w$$

Replacing $\mathbf{N}_v$ and $\mathbf{N}_w$ by those values and $h_{\mathbf{S}_v}$ and $h_{\mathbf{S}_w}$ by their values in equations (30), and taking once again into account the fact that $h^2 = \mid \mathbf{S}_v \mid^2 \mid \mathbf{S}_w \mid^2$, we obtain the following expression:

$$-h[\frac{e}{\mid \mathbf{S}_v \mid^4}\mathbf{S}_v^T \Phi_{\mathbf{YY}} \mathbf{S}_v + \frac{2f}{\mid \mathbf{S}_v \mid^2 \mid \mathbf{S}_w \mid^2}\mathbf{S}_w^T \Phi_{\mathbf{YY}} \mathbf{S}_v + \frac{g}{\mid \mathbf{S}_w \mid^4}\mathbf{S}_w^T \Phi_{\mathbf{YY}} \mathbf{S}_w]$$

Using the vectors $\mathbf{t}_1$ and $\mathbf{t}_2$ defined in the previous lemma, we obtain for the term within the bracket:

$$\frac{e}{E}\mathbf{t}_1^T \Phi_{\mathbf{YY}} \mathbf{t}_1 + \frac{2f}{h}\mathbf{t}_2^T \Phi_{\mathbf{YY}} \mathbf{t}_1 + \frac{g}{G}\mathbf{t}_2^T \Phi_{\mathbf{YY}} \mathbf{t}_2$$

Considering the two symmetric matrices

$$\mathbf{B} = \begin{bmatrix} \frac{e}{E} & \frac{f}{h} \\ \frac{f}{h} & \frac{g}{G} \end{bmatrix} \qquad \mathbf{C} = \begin{bmatrix} \mathbf{t}_1^T \Phi_{\mathbf{YY}} \mathbf{t}_1 & \mathbf{t}_1^T \Phi_{\mathbf{YY}} \mathbf{t}_2 \\ \mathbf{t}_2^T \Phi_{\mathbf{YY}} \mathbf{t}_1 & \mathbf{t}_2^T \Phi_{\mathbf{YY}} \mathbf{t}_2 \end{bmatrix},$$

we verify that our expression is the trace of the product $\mathbf{BC}$. The matrix $\mathbf{C}$ represents the restriction of the endomorphism defined by the Hessian $\Phi_{\mathbf{YY}}$ of Φ to the tangent plane of the surface expressed in the orthonormal basis $\mathbf{t}_1$, $\mathbf{t}_2$ of that plane. The matrix $\mathbf{B}$ represents the differential of the Gauss map of the surface S expressed in the same basis. □

If we put together the results of lemmas 6, 7 and proposition 1, we obtain a proof of theorem 1.

C Computation of the derivatives $\Phi_{\mathbf{S}}$, $\Phi_{\mathbf{N}}$, $\Phi_{\mathbf{SN}}$ and $\Phi_{\mathbf{NN}}$

We compute in this Appendix the necessary ingredients in order to apply theorem 1 to our case. As shown in section 4 it is sufficient to focus on one of the Φ_{ij}, say Φ_{12} which we call Φ. This is a function of the point $\mathbf{S} = (S^1, S^2, S^3)$ on the surface through the corresponding pixels m_1 and m_2 and the matrix $\mathbf{A}$ induced by the tangent plane at the point S, and of the unit normal $\mathbf{N}$ to the surface at that point through the matrix $\mathbf{A}$. In effect, the equation (9) still holds and the equations (17) defining $\mathbf{K}$ can be written in terms of $\mathbf{S}$ and $\mathbf{N}$:

$$\begin{aligned} d &= -\mathbf{S} \cdot \mathbf{N} \\ \mathbf{T} &= \mathbf{R}^T \mathbf{t} \\ \mathbf{N} &= (N_1, N_2, N_3) \end{aligned} \tag{33}$$

The derivatives of $\mathbf{K}$ with respect to $\mathbf{S}$ and $\mathbf{N}$ are therefore quite simple:

$$\begin{aligned} \mathbf{K}_{S^i} &= d_{S^i}\mathbf{R}^T = -N_i\mathbf{R}^T \\ \mathbf{K}_{N_i} &= d_{N_i}\mathbf{R}^T - \mathbf{T}\mathbf{N}_{N_i}^T = S^i\mathbf{R}^T - [\mathbf{0}\ \mathbf{T}\ \mathbf{0}] \end{aligned} \tag{34}$$

Using the expression (19) giving the affine approximation $\mathbf{A}$ of the collineation K, we can now compute $\mathbf{A_S}$ and $\mathbf{A_N}$ and $\mathbf{A_{NN}}$ which are needed in order to compute relation (24). These entities are tensors which we will actually not need for reasons that will become clear in the sequel.

We arrange the results as a series of lemmas.

Lemma 9 *The row vector $\Phi_{\mathbf{S}}$ is computed from the following quantities:*

$$\langle I_1,\, I_2\rangle_{\mathbf{S}} = \langle \nabla I_1,\, I_2\rangle^T \frac{\partial \mathbf{m}_1}{\partial \mathbf{S}} + \langle I_1,\, \nabla I_2\rangle^T \frac{\partial \mathbf{m}_2}{\partial \mathbf{S}} + \int^* (I_1(m_1+m) - \overline{I_1}(m_1))\nabla I_2(m_2+\mathbf{A}\mathbf{m})^T(\mathbf{A}\mathbf{m})_{\mathbf{S}}\, dm$$

$$\mid I_2 \mid \cdot \mid I_2 \mid_{\mathbf{S}} = \langle \nabla I_2, I_2\rangle \frac{\partial \mathbf{m}_2}{\partial \mathbf{S}}^T + \int^* (I_2(m_2+\mathbf{A}\mathbf{m}) - \overline{I_2}(m_2))\nabla I_2(m_2+\mathbf{A}\mathbf{m})^T(\mathbf{A}\mathbf{m})_{\mathbf{S}}\, dm$$

Proof : The proof is just a matter of applying the chain rule to the definitions of $\langle I_1,\, I_2\rangle$ and $\mid I_k \mid$, $k = 1, 2$. Note that $\frac{\partial \mathbf{m}_k}{\partial \mathbf{S}}$, $k = 1, 2$ is a 2×3 matrix and that $(\mathbf{A}\mathbf{m})_{\mathbf{S}}$ is the 2×3 matrix $[\mathbf{A}_{S^1}\mathbf{m},\, \mathbf{A}_{S^2}\mathbf{m},\, \mathbf{A}_{S^3}\mathbf{m}]$. □

A similar result holds for $\Phi_{\mathbf{N}}$, as stated in the following lemma:

Lemma 10 *The row vector $\Phi_{\mathbf{N}}$ is computed from the following quantities:*

$$\langle I_1,\, I_2\rangle_{\mathbf{N}} = \int^* (I_1(m_1 + m) - \overline{I_1}(m_1))\nabla I_2(m_2 + \mathbf{A}\mathbf{m})^T(\mathbf{A}\mathbf{m})_{\mathbf{N}}\, dm$$

Proof : Again, this is just a matter of applying the chain rule. Note that $(\mathbf{A}\mathbf{m})_{\mathbf{N}}$ is the 2×3 vector $[\mathbf{A}_{N_1}\mathbf{m},\, \mathbf{A}_{N_2}\mathbf{m},\, \mathbf{A}_{N_3}\mathbf{m}]$. □

For the matrix $\Phi_{\mathbf{SN}}$ or $\Phi_{\mathbf{XY}}$, we have the following lemma:

Lemma 11 *The matrix $\Phi_{\mathbf{SN}}$ can be computed from the following quantities:*

$$\langle I_1,\, I_2\rangle_{\mathbf{SN}} = \int^* \frac{\partial \mathbf{m}_1}{\partial \mathbf{S}}^T \nabla I_1(m_1+m)\nabla I_2^T(m_2+\mathbf{Am})(\mathbf{Am})_{\mathbf{N}} dm - \frac{\partial \overline{I_1}}{\partial \mathbf{S}}^T (m_1) \int^* \nabla I_2^T(m_2+\mathbf{Am})(\mathbf{Am})_{\mathbf{N}} dm + \int^* (I_1(m_1+m)-\overline{I_1}(m_1))[\mathbf{Am})_{\mathbf{S}}^T \mathbf{H}_{I_2}(\mathbf{Am})_{\mathbf{N}} + \nabla I_2^T \mathbf{A}_{\mathbf{SN}}\mathbf{m}] dm$$

$$\frac{\partial \overline{I_1}}{\partial \mathbf{S}} = \int^* \nabla I_1^T(m_1+m')dm' \frac{\partial \mathbf{m}_1}{\partial \mathbf{S}}$$

Proof : Again, it is just a matter of applying the chain rule. The quantity $(\mathbf{Am})_{\mathbf{S}}^T \mathbf{H}_{I_2}(\mathbf{Am})_{\mathbf{N}}$ is the 3×3 matrix $[\mathbf{m}^T \mathbf{A}_{S_i}^T \mathbf{H}_{I_2}\mathbf{A}_{N_j}\mathbf{m}]$ and the quantity $\nabla I_2^T \mathbf{A}_{\mathbf{SN}}\mathbf{m}$ is the 3×3 matrix $[\nabla I_2^T \mathbf{A}_{S_iN_j}\mathbf{m}]$. Note that ∇I_2 and $\mathbf{H}_{I_2}$ are evaluated at the image point $m_2+\mathbf{Am}$. □

Finally, for the Hessian $\Phi_{\mathbf{NN}}$ or $\Phi_{\mathbf{YY}}$, we have the following result:

Lemma 12 *The 3×3 symmmetric matrix $\Phi_{\mathbf{YY}}$ can be computed from the following expressions:*

$$\langle I_1,\, I_2\rangle_{\mathbf{NN}} = \int^* (I_1(m_1+m)-\overline{I_1}(m_1))[(\mathbf{Am})_{\mathbf{N}}^T \mathbf{H}_{I_2}(\mathbf{Am})_{\mathbf{N}} + \nabla I_2^T \mathbf{A}_{\mathbf{NN}}\mathbf{m}]\; dm$$

Proof : Again, it is just a matter of applying the chain rule. The quantity $(\mathbf{A_N m})^T \mathbf{H}_{I_2}(\mathbf{A_N m})$ is the 3×3 symmetric matrix $[\mathbf{m}^T \mathbf{A}_{N_i}^T \mathbf{H}_{I_2}\mathbf{A}_{N_j}\mathbf{m}]$. The matrix is symmetric because if we transpose its ijth element we obtain $\mathbf{m}^T \mathbf{A}_{N_j}^T \mathbf{H}_{I_2}^T \mathbf{A}_{N_i}\mathbf{m}$ which, because the Hessian $\mathbf{H}_{I_2}$ is symmetric, is also equal to the jith term. Similarly, $\nabla I_2^T \mathbf{A}_{\mathbf{NN}}\mathbf{m}$ is the 3×3 symmetric matrix whose ijth element is $\nabla I_2^T \mathbf{A}_{N_iN_j}\mathbf{m}$. It is symmetric because of the Schwarz equality ($\mathbf{A}_{N_iN_j} = \mathbf{A}_{N_jN_i}$). Note that ∇I_2 and $\mathbf{H}_{I_2}$ are evaluated at the image point $m_2+\mathbf{Am}$. □

D Results on the differential geometry of surfaces

We use the notations of [6]. Let U be an open set of $\mathbf{R}^2$ and $\mathbf{S} : \longrightarrow S$ be a parametrization of the surface S. The vectors $\mathbf{S}_v$ and $\mathbf{S}_w$ span the tangent plane

$T_M(S)$ at the point M of S. The *first fundamental form* of S at point M is the quadratic form I_p defined in $T_M(S)$ by the 2×2 symmetric matrix:

$$\begin{bmatrix} E & F \\ F & G \end{bmatrix}$$

The unit normal at point M is $\mathbf{N} = \frac{\mathbf{S}_v \times \mathbf{S}_w}{|\mathbf{S}_v \times \mathbf{S}_w|}$ and defines a mapping from S into the unit sphere, called the Gauss map. The differential of this mapping is a linear mapping from $T_M(S)$ into the tangent plane to the unit sphere at the point $\mathbf{N}$. Since this plane is parallel to $T_M(S)$ by construction, we think of this linear mapping as one from $T_M(S)$ into $T_M(S)$. The second order intrinsic properties of the surface at point M, such as the principal directions and the principal curvatures are obtained from this linear mapping expressed in the basis $\mathbf{S}_v$, $\mathbf{S}_w$ of $T_M(S)$. The 2×2 matrix $\mathbf{B} = [b_{ij}]$, $i, j = 1, 2$ is given by the equations of Weingarten:

$$\begin{array}{ll} b_{11} = \frac{fF - eG}{EG - F^2} & b_{12} = \frac{gF - fG}{EG - F^2} \\ b_{21} = \frac{eF - fE}{EG - F^2} & b_{22} = \frac{fF - gE}{EG - F^2} \end{array}$$

where

$$\begin{aligned} e &= -\mathbf{N}_v \cdot \mathbf{S}_v \\ f &= -\mathbf{N}_w \cdot \mathbf{S}_v = -\mathbf{N}_v \cdot \mathbf{S}_w \\ g &= -\mathbf{N}_w \cdot \mathbf{S}_w \end{aligned}$$

The principal curvatures k_1 and k_2 are the eigenvalues of the matrix $\mathbf{B}$ and the mean curvature $H = \frac{1}{2}(k_1 + k_2)$ is easily shown to be equal to:

$$H = \frac{1}{2} \frac{eG - 2fF + gE}{EG - F^2}$$

which reduces to $\frac{1}{2}(\frac{e}{E} + \frac{g}{G})$ when the parametrization is orthogonal, i.e. $F = 0$.

E Bells and whistles for the 2D implementation

In practice, β has been corrected with two terms: we used $\beta - \epsilon_1 \kappa - \epsilon_2 \mathcal{G}_{\Phi_{max}}(\Phi)$ where ϵ_i are positive constants and $\mathcal{G}_{\Phi_{max}}$ is an increasing function, null for values less than Φ_{max} ($-1 < \Phi_{max}$).

- The term $-\epsilon_1 \kappa$ makes the solution smoother, allowing to take greater time steps thus accelerating the convergence.

- Assuming Φ has been normalized with respect to the number of cameras used for each particular point, so that Φ is equal to -1 where the correlation is maximum, the term $-\epsilon_2 \mathcal{G}_{\Phi_{max}}(\Phi)$ leads to an inward evolution where the correlation is not good enough. ($\Phi_{max} < \Phi$). This also accelerates the convergence and avoids local minima. The initial curve can be chosen far from the objects (but must not intersect them).

References

[1] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. In *Proceedings of the 5th International Conference on Computer Vision*, pages 694–699, Boston, MA, June 1995. IEEE Computer Society Press.

[2] V. Caselles, R. Kimmel, G. Sapiro, and C. Sbert. 3d active contours. In M-O. Berger, R. Deriche, I. Herlin, J. Jaffre, and J-M. Morel, editors, *Images, Wavelets and PDEs*, volume 219 of *Lecture Notes in Control and Information Sciences*, pages 43–49. Springer, June 1996.

[3] David L. Chopp. Computing minimal surfaces via level set curvature flow. *Journal of Computational Physics*, 106:77–91, 1993.

[4] Rachid Deriche, Stéphane Bouvin, and Olivier Faugeras. A level-set approach for stereo. In *Fisrt Annual Symposium on Enabling Technologies for Law Enforcement and Security - SPIE Conference 2942 : Investigative Image Processing.*, Boston, Massachusetts USA, November 1996.

[5] Frédéric Devernay and Olivier Faugeras. Computing differential properties of 3-D shapes from stereoscopic images without 3-D models. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, pages 208–213, Seattle, WA, June 1994. IEEE.

[6] M. P. DoCarmo. *Differential Geometry of Curves and Surfaces*. Prentice-Hall, 1976.

[7] Olivier Faugeras. *Three-Dimensional Computer Vision: a Geometric Viewpoint*. The MIT Press, 1993.

[8] M. Gage and R.S. Hamilton. The heat equation shrinking convex plane curves. *J. of Differential Geometry*, 23:69–96, 1986.

[9] M. Grayson. The heat equation shrinks embedded plane curves to round points. *J. of Differential Geometry*, 26:285–314, 1987.

[10] W.E.L. Grimson. *From Images to Surfaces.* MIT Press : Cambridge, 1981.

[11] H. L. F. von Helmholtz. *Treatise on Physiological Optics.* New York: Dover, 1925. Translated by J.P. Southall.

[12] Berthold Klaus Paul Horn. *Robot Vision.* MIT Press, 1986.

[13] Bela Julesz. *Foundations of Cyclopean perception.* The University of Chicago Press, Chicago and London, 1971.

[14] M. Kass, A. Witkin, and D. Terzopoulos. SNAKES: Active Contour Models. *The International Journal of Computer Vision*, 1:321–332, January 1988.

[15] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi. Gradient flows and geometric active contour models. In *Proc. Fifth International Conference on Computer Vision*, Boston, MA, June 1995. IEEE Computer Society Press.

[16] S. Osher and J. Sethian. Fronts propagating with curvature dependent speed : algorithms based on the Hamilton-Jacobi formulation. *Journal of Computational Physics*, 79:12–49, 1988.

[17] L. Robert and R. Deriche. Dense depth map reconstruction: A minimization and regularization approach which preserves discontinuities. In Bernard Buxton, editor, *Proceedings of the 4th European Conference on Computer Vision*, Cambridge, UK, April 1996.

[18] L. Robert, R. Deriche, and O.D. Faugeras. Dense depth recovery from stereo images. In *Proceedings of the European Conference on Artificial Intelligence*, pages 821–823, Vienna, Austria, August 1992.

[19] Guillermo Sapiro and Allen Tannenbaum. Affine Invariant Scale Space. *The International Journal of Computer Vision*, 11(1):25–44, August 1993.

[20] J. A. Sethian. *Level Set Methods.* Cambridge University Press, 1996.

[21] J.A. Sethian. Numerical algorithms for propagating interfaces: Hamilton-jacobi equations and conservation laws. *Journal of Differential Geometry*, 31:131–161, 1990.

[22] J.A. Sethian. Theory, algorithms, and applications of level set methods for propagating interfaces. Technical Report PAM-651, Center for Pure and Applied Mathematics, University of California, Berkeley, August 1995. To appear Acta Numerica.



Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399