Dense Depth Map Reconstruction : A Minimization and Regularization Approach which Preserves Discontinuities

Luc ROBERT and Rachid DERICHE

INRIA. B.P. 93. 06902 Sophia-Antipolis. FRANCE.
Phone : (+33) 93 65 78 32 - Fax: (+33) 93 65 78 45
E-mail: {lucr,der}@sophia.inria.fr

Abstract. We present a variational approach to dense stereo reconstruction which combines powerful tools such as regularization and multi-scale processing to estimate directly depth from a number of stereo images, while preserving depth discontinuities. The problem is set as a regularization and minimization of a non-quadratic functional. The *Tikhonov* quadratic regularization term usually used to recover smooth solution is replaced by a function of the gradient depth specifically derived to allow depth discontinuities formation in the solution. Conditions to be fulfilled by this specific regularizing term to preserve discontinuities are also presented. To solve this problem in the discrete case, a PDE-based explicit scheme for moving iteratively towards the solution has been developed. This approach presents the additional advantages of not introducing any intermediate representation such as disparity or rectified images: depth is computed directly from the grey-level images and we can also deal with any number (greater than two) of cameras. Promising experimental results illustrate the capabilities of this approach.

1 Introduction

Over the years numerous algorithms for passive stereo have been proposed, which use different strategies:

Feature-based: Those algorithms establish correspondences between features extracted from the images, like edge pixels [20, 23, 18], line segments [17, 4] or curves [8, 25] for instance. Their main advantage is to yield accurate information and to manipulate reasonably small amounts of data, thus gaining in time and space complexity. Their main drawback is the sparseness of the recovered depth information.

Area-based: In these approaches, dense depth maps are provided by correlating the grey levels of image patches in the views being considered, assuming that they present some similarity [19, 13]. These methods are well adapted for relatively textured areas; however, they generally assume that the observed scene is locally fronto-parallel, which causes problems for slanted surfaces and in particular near the occluding contours of the objects. Lastly, the matching process does not take into account the edge information.

Energy-based: A third kind of approach which does not suffer any of the inconvenients presented above, consists of expressing the correspondence problem as a minimization and regularization one [6, 30]. An iterative solution of the discrete version of the associated Euler-Lagrange equation is then used in order to estimate depth.

The method which we present in this paper follows the third strategy with the important following issues:

- It computes depth directly from the grey-level images intensities. No intermediate process such as rectification [5] or disparity estimation is used. The system of cameras is supposed to be calibrated, and the depth information is directly issued as a depth function $m \rightarrow Z(m)$ of the image point.

- The method addresses the problem of accurately determining depth near discontinuities. It is well known that using the classical *Tikhonov* regularization approach [28] by considering a quadratic regularizing term in the energy function, leads to smoothing the depth image across the discontinuities, yielding a destruction of these important characteristics in the resulting depth image. We address this important problem by replacing the energy quadratic regularizing term by a function specially designed in order to allow the minimization process to preserve the original discontinuities in the depth map. It is shown that in this case, the minimization process involves an isotropic smoothing step in the homogeneous regions (i.e with small depth gradient), and an anisotropic smoothing step in the inhomogeneous regions (i.e high depth gradient).

- In order to speed up convergence and avoid possible local minima,a multi-scale approach is also used.

2 Formalism of the matching process

2.1 Notations for one camera

We assume that the imaging system follows the pinhole model. The projection matrix $\tilde{\mathbf{P}}$ of a camera with respect to a reference frame $\mathcal{R}_w$ is computed during a calibration phase. It allows finding $m = [u, v]^T$, projection onto the retina of the point $M = [X, Y, Z]^T$ expressed in world coordinates (s is a scale factor):

$$[su, sv, s]^T = \tilde{\mathbf{P}}[X, Y, Z, 1]^T$$

The 3×4 matrix $\tilde{\mathbf{P}}$ can be decomposed as follows:

$$\tilde{\mathbf{P}} = [\mathbf{J}\ \mathbf{0_3}] \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}_3^T & 1 \end{bmatrix} = [\mathbf{J}\ \mathbf{0_3}]\,\mathbf{D}$$

The 4×4 matrix $\mathbf{D}$ involving $\mathbf{R}$ and $\mathbf{t}$ changes world coordinates into camera coordinates (see figure 1). It represents the *extrinsic* information about the camera. 0_3 is the

null 3-vector. The other matrix contains the *intrinsic* information [29]: the 3×3 matrix **J** changes camera coordinates into pixel units. It is well known that in the generic case,

$$\mathbf{J} = \begin{bmatrix} \alpha_u & -\alpha_u \cot\theta & u_0 \\ 0 & \alpha_v / \sin\theta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

where θ is the angle between pixel rows and columns.

Thus, the camera coordinates of a point M in the 3D space can be derived from its projection m and its depth:

$$[X, Y, Z]^T = Z\mathbf{J}^{-1}[u, v, 1]^T$$

All these matrices can be computed with good accuracy with the method described in [24].

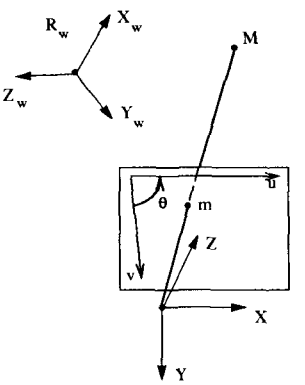

Fig. 1. The pinhole model and the different frames

2.2 Correspondence

In the remainder of this article, we consider a calibrated stereo rig. All the quantities introduced in the previous paragraph and related to one camera are assigned the index of the considered camera.

If m_i, m_j are the two-dimensional projections of a three-dimensional point M on

retinae i, j, we can easily express m_j as a function $f_{i \to j}$ of (m_i, Z_i). Indeed,

$$\begin{bmatrix} su_j \\ sv_j \\ s \end{bmatrix} = Z_i \, [\mathbf{J_j} \; \mathbf{0_3}] \, \mathbf{D_j} \mathbf{D_i^{-1}} \begin{bmatrix} \mathbf{J_i^{-1}} \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} \\ 1 \end{bmatrix}$$

Thus, finding a correspondent for a point of image i is completely equivalent to finding depth Z_i at that point.

3 The energy function

Let us assume for a while that the world is made of Lambertian objects, i.e. of objects that look equally bright from all viewing directions. Finding a correspondent in image 2 for each point of image 1 turns into finding a Z function that minimizes:

$$M_{12}(Z) = \iint \|\mathbf{I}_1(m_1) - \mathbf{I}_2(f_1(m_1, Z(m_1)))\|^2 \, dm_1$$

($\mathbf{I}_i$ is the intensity in image i). This is of course not sufficient, since it may first lead to chaotic solutions, in which each point of image 1 finds its correspondent in image 2 independently from its neighbors. We need to add a constraint on the shape of the depth function. As a consequence, the matching problem is expressed as a constrained minimization one, where the minimized functional is the following:

$$E(Z) = M_{12}(Z) + \lambda S(Z)$$

The first term holds for similarity of the image intensities at corresponding points. In practice, the use of additional attributes like intensity gradient or rgb grey-level images helps establishing the right correspondences by reducing ambiguity. More generally, we have:

$$M_{12}(Z) = \iint \sum_k \left\|\mathbf{F}_1^k(m_1) - \mathbf{F}_2^k(f_1(m_1, Z(m_1)))\right\|^2 dm_1$$

where $\mathbf{F}_i^k$ is a field of attributes, scalar - e.g. intensity - or vectorial - e.g. intensity gradient - extracted from image i.

If we have more than two views, we can very easily take into account the information of all the images, by defining the similarity term of the functional as

$$M(Z) = \sum_j M_{1j}(Z)$$

where j varies from 2 to the number of images.

Please note that the simplicity of this expression emerges from the fact that we are searching for a depth function Z defined with respect to one reference image, which is matched with any other image. With classical techniques based on a disparity representation, things would not be as simple. Indeed, to define disparity, one needs two images.

In the last section of the article, we will show the importance of using more than two views in image-based surface reconstruction.

The second term stands for the constraint to be applied on the depth map. A classical constraint is the smoothness assumption on the resulting depth map. That is the case for the well-known *Tikhonov* regularization term [28]:

$$S(Z) = \int\int \mid \nabla_{m_1} Z \mid^2 dm_1$$

This term leads to a solution where discontinuities of the depth function are smoothed. This is not really desirable if one wants to preserve these discontinuities and recover the original scene as accurately as possible. The next section is devoted to presenting an original approach which tackles in an efficient way this important problem.

4 Regularizing the solution and preserving discontinuities

In order to preserve the discontinuities while regularizing the solution, a natural way to proceed is to forbid regularizing and smoothing across such discontinuities. One way of taking into account these technical remarks is by looking for a function $\Phi(.)$ such that the following regularization term:

$$S(Z) = \int\int \Phi(\mid \nabla_{m_1} Z \mid) dm_1$$

preserves those discontinuities (For instance, a quadratic function as the one used in the *Tikhonov* case [28], does clearly not correspond to such type of functions).

In this section, we summarize our variational approach to 3D recovery from stereo images. This approach is inspired from the approaches developed for image restoration purpose in [26, 21, 31, 7, 9, 10, 3]. A detailed review of all these approaches can be found in [12].

The key idea to deal with such a problem is first to consider the functional to be minimized, written as follows:

$$E(Z) = \int\int F(u, v, Z, \Phi(\mid \nabla_m Z \mid)) du dv$$

A necessary condition for it to be extremal is the derived Euler-Lagrange differential equation:

$$F_Z - \frac{\partial}{\partial u} F_{Z_u} - \frac{\partial}{\partial v} F_{Z_v} = 0$$

This yields the following equation:

$$\sum_k \left(\mathbf{F}_1^k(m) - \mathbf{F}_2^k(f_1(m, Z(m)))\right) . \frac{\partial \mathbf{F}_2}{\partial Z} - \frac{\lambda}{2} \left(div\left(\frac{\phi'(\mid \nabla Z \mid)}{\mid \nabla Z \mid} \nabla Z\right)\right) = 0 \quad (1)$$

Where the term div denotes the divergence operator. Developping and simplifying, the term on the left can be rewritten as

$$\sum_k \left(\mathbf{F}_1^k(m) - \mathbf{F}_2^k(f_1(m, Z(m)))\right) . \frac{\partial \mathbf{F}_2}{\partial Z} - \frac{\lambda}{2} \left(\frac{\Phi'(\mid \nabla_m Z \mid)}{\mid \nabla_m Z \mid} Z_{\xi\xi} + \Phi''(\mid \nabla_m Z \mid) Z_{\eta\eta} \right) \quad (2)$$

where Φ' and Φ'' represent respectively the first and second derivatives of $\Phi(s)$ with respect to the parameter s. $Z_{\eta\eta}$ represents the second order directional derivatives of $Z(m)$ in the direction of the gradient $\eta = \frac{\nabla Z}{|\nabla Z|}$, and $Z_{\xi\xi}$ is the second order directional derivatives of $Z(m)$ in the direction ξ orthogonal to the gradient. Boundary conditions have also to be considered in order to solve this equation.

In order to regularize the solution and preserve discontinuities, one would like to smooth isotropically the solution inside homogeneous regions and preserve the discontinuities in the inhomogeneous regions. Assuming that the function $\Phi''(.)$ exists, the condition on smoothing in an isotropic way inside homogeneous regions can be achieved by imposing the following conditions on the $\Phi(.)$ function:

$$\lim_{|\nabla_m Z|\to 0} \frac{\Phi'(|\nabla_m Z|)}{|\nabla_m Z|} = \lim_{|\nabla_m Z|\to 0} \Phi''(|\nabla_m Z|) = \Phi''(0) > 0 \tag{3}$$

Therefore, at the points where the depth gradient is small, Z is solution of:

$$\sum_k \left(\mathbf{F}_1^k(m) - \mathbf{F}_2^k(f_1(m, Z(m)))\right) . \frac{\partial \mathbf{F}_2}{\partial Z} - \frac{\lambda}{2}\Phi''(0)(Z_{\xi\xi} + Z_{\eta\eta}) = 0$$

This process corresponds to the case where the function $\Phi(s)$ is quadratic [28]. Note that the coefficients are required to be positive, otherwise the regularization part will act as an inverse heat equation notably known as an instable process.

In order to preserve the discontinuities near inhomogeneous regions presenting a strong depth gradient, one would like to smooth along the isophote and not across them. This leads to stopping the diffusion in the gradient direction, i.e setting the weight $\Phi''(|\nabla_m Z|)$ to zero, while keeping a stable diffusion along the tangential direction to the isophote, i.e setting the weight $\frac{\Phi'(|\nabla_m Z|)}{|\nabla_m Z|}$ to some positive constant:

$$\lim_{|\nabla_m Z|\to\infty} \Phi''(|\nabla_m Z|) = 0 \;\; ; \;\; \lim_{|\nabla_m Z|\to\infty} \frac{\Phi'(|\nabla_m Z|)}{|\nabla_m Z|} = \beta > 0 \tag{4}$$

Therefore, at the points where the depth gradient is strong, Z will be the solution of the following equation:

$$\sum_k \left(\mathbf{F}_1^k(m) - \mathbf{F}_2^k(f_1(m, Z(m)))\right) . \frac{\partial \mathbf{F}_2}{\partial Z} - \frac{\lambda}{2}\beta Z_{\xi\xi} = 0$$

which yields Z as a regularized solution in the ξ direction. Note that the positiveness of the β coefficient is also required to generate a stable smoothing process in the ξ direction.

Unfortunately, the two conditions of (4) cannot be satisfied simultaneously by a function $\Phi(|\nabla_m Z|)$. However, the following conditions can be imposed in order to decrease the effects of the diffusion along the gradient more rapidly than those associated with the diffusion along the isophotes:

$$\lim_{|\nabla_m Z|\to\infty} \Phi''(|\nabla_m Z|) = \lim_{|\nabla_m Z|\to\infty} \frac{\Phi'(|\nabla_m Z|)}{|\nabla_m Z|} = \lim_{|\nabla_m Z|\to\infty} \frac{\Phi''(|\nabla_m Z|)}{\frac{\Phi'(|\nabla_m Z|)}{|\nabla_m Z|}} = 0 \tag{5}$$

The conditions given by Equations (3) and (5) are those one would like to impose in order to deal with a regularization process while preserving the existing discontinuities. As it has been shown very recently in [3], these conditions are also sufficient to prove that the model is well-posed mathematically, and the existence and uniqueness of a solution is also guaranteed by these conditions.

A certain number of functions have already been proposed in the literature in order to address the problem of discontinuities. Table 4 illustrates the most commonly used functions. One can see easily that only the last three functions fulfill all the conditions mentioned above. The *Tikhonov* function and the *Aubert* function will be used in our experimental section.

Author	$\Phi(s)$	$\Phi'(s)/s$	$\Phi''(s)$
Perona-Malik [22]	$\frac{-k^2}{2}(e^{-(s/k)^2}-1)$	$e^{-(s/k)^2}$	$(1-2(\frac{s}{k})^2)e^{-(\frac{s}{k})^2}$
Perona-Malik [22]	$\frac{k^2}{2}log(1+(s/k)^2)$	$\frac{1}{1+(\frac{s}{k})^2}$	$\frac{k^2(k^2-s^2)}{(k^2+s^2)^2}$
Geman et Reynolds [14]	$\frac{(s/k)^2}{1+(s/k)^2}$	$\frac{2k^2}{(k^2+s^2)^2}$	$-\frac{2k^2(-k^2+3s^2)}{(k^2+s^2)^3}$
Alvarez [1]	...	$g(s)$	$(1-h(s))g(s)$
Tikhonov [28]	$s^2/2$	1	1
Green [15]	$logcosh(s/k)$	$\frac{tanh(s/k)}{ks}$	$k^{-2}\left(\cosh(\frac{s}{k})\right)^{-2}$
Rudin [26]	s	$\frac{1}{s}$	0
Aubert [10]	$\sqrt{1+(s/k)^2}-1$	$\frac{1}{\sqrt{\frac{k^2+s^2}{k^2}}}k^{-2}$	$\frac{\lvert k\rvert}{(k^2+s^2)^{3/2}}$

5 Minimization of the Energy Function

This section presents the numerical scheme developed to solve the Euler-Lagrange equation (1,2) associated to the energy function. A time-dependent approach has been developed to solve this non-linear PDE. We consider the associated evolution equation, or equivalently the gradient descent method. This leads us to consider the following equation:

$$Z_t = \sum_k \left(\mathbf{F}_1^k(m) - \mathbf{F}_2^k(f_1(m, Z(m)))\right).\frac{\partial \mathbf{F}_2}{\partial Z} - \frac{\lambda}{2}\left(\frac{\Phi'(|\nabla_m Z|)}{|\nabla_m Z|}Z_{\xi\xi} + \Phi''(|\nabla_m Z|)Z_{\eta\eta}\right)$$

The corresponding explicit numerical scheme is then implemented:

$$\begin{cases} Z_{i,j}^{n+1} = Z_{i,j}^n + \Delta t \left(\sum_k (\mathbf{F}_1^k(m) - \mathbf{F}_2^k(f_1(m, Z(m)))).\frac{\partial \mathbf{F}_2}{\partial Z} - \frac{\lambda}{2}(\frac{\Phi'(|\nabla_m Z|)}{|\nabla_m Z|}Z_{\xi\xi} + \Phi''(|\nabla_m Z|)Z_{\eta\eta})\right)_{i,j}^n \\ \oplus \text{ Boundary conditions on the depth} \end{cases} \quad (6)$$

Finally, we apply a Gauss-Seidel relaxation method for moving iteratively towards the solution of this problem [27].

5.1 Discretization Scheme

In the following we present the consistent way to discretize the divergence term that appears in the equation (1).

Denoting by θ the angle that the unit gradient $\nabla Z/(\mid \nabla Z \mid)$ makes with the x axis, we have the well known expressions:

$$\begin{aligned} Z_{\xi\xi} &= sin(\theta)^2 Z_{xx} - 2sin(\theta)cos(\theta)Z_{xy} + cos(\theta)^2 Z_{yy} \\ Z_{\eta\eta} &= cos(\theta)^2 Z_{xx} + 2sin(\theta)cos(\theta)Z_{xy} + sin(\theta)^2 Z_{yy} \end{aligned} \tag{7}$$

In order to deal with a consistent discrete approximation of these second directional derivatives, we proceed as follows:

We look for some constant $(\lambda_i)_{i=0..4}$ and $(\lambda'_i)_{i=0..4}$ such that:

$$Z_{\xi\xi_{i,j}} \simeq \begin{bmatrix} \lambda_4 & \lambda_2 & \lambda_3 \\ \lambda_1 & -4\lambda_0 & \lambda_1 \\ \lambda_3 & \lambda_2 & \lambda_4 \end{bmatrix} * Z \text{ and } Z_{\eta\eta_{i,j}} \simeq \begin{bmatrix} \lambda'_4 & \lambda'_2 & \lambda'_3 \\ \lambda'_1 & -4\lambda'_0 & \lambda'_1 \\ \lambda'_3 & \lambda'_2 & \lambda'_4 \end{bmatrix} * Z \tag{8}$$

where the sign $*$ denotes the convolution operation. Using The Taylor expansion of (8) up to the second order and identifying the coefficients with (7) yields the values of the following λ_i and λ'_i parameters we are looking for:

$$\begin{cases} \lambda_1 = 2\lambda_0 - \frac{\delta^x Z}{|\nabla Z|}^2 \\ \lambda_2 = 2\lambda_0 - \frac{\delta^y Z}{|\nabla Z|}^2 \\ \lambda_3 = \frac{1}{2}(-\frac{\delta^y Z}{|\nabla Z|}\frac{\delta^x Z}{|\nabla Z|} + 1 - 2\lambda_0) \\ \lambda_4 = \frac{1}{2}(1 - 2\lambda_0 + \frac{\delta^y Z}{|\nabla Z|}\frac{\delta^x Z}{|\nabla Z|}) \end{cases} \quad \begin{cases} \lambda'_1 = 2\lambda'_0 - \frac{\delta^y Z}{|\nabla Z|}^2 \\ \lambda'_2 = 2\lambda'_0 - \frac{\delta^x Z}{|\nabla Z|}^2 \\ \lambda'_3 = \frac{1}{2}(\frac{\delta^y Z}{|\nabla Z|}\frac{\delta^x Z}{|\nabla Z|} + 1 - 2\lambda'_0) \\ \lambda'_4 = \frac{1}{2}(1 - 2\lambda'_0 - \frac{\delta^y Z}{|\nabla Z|}\frac{\delta^x Z}{|\nabla Z|}) \end{cases}$$

Then denoting:

$$c_\eta = \Phi''(\mid \nabla Z \mid) \quad c_\xi = \frac{\Phi'(\mid \nabla Z \mid)}{\mid \nabla Z \mid} \tag{9}$$

We obtain the following coefficients $(\Sigma_{Zi})_{i=0..4}$ for the divergence term:

$$\left(div(\frac{\phi'(\mid \nabla Z \mid)}{\mid \nabla Z \mid}\nabla Z)\right) = (c_\eta Z_{\eta\eta} + c_\xi Z_{\xi\xi})_{i,j} \simeq \begin{bmatrix} \Sigma_{Z4} & \Sigma_{Z2} & \Sigma_{Z3} \\ \Sigma_{Z1} & \Sigma_{Z0} & \Sigma_{Z1} \\ \Sigma_{Z3} & \Sigma_{Z2} & \Sigma_{Z4} \end{bmatrix}$$

Where:

$$\begin{cases} \Sigma_{Z0} = -4(c_\xi \lambda_0 + c_\eta \lambda'_0) \\ \Sigma_{Z1} = 2(c_\xi \lambda_0 + c_\eta \lambda'_0) - (c_\xi \frac{\delta^x Z}{|\nabla Z|}^2 + c_\eta \frac{\delta^y Z}{|\nabla Z|}^2) \\ \Sigma_{Z2} = 2(c_\xi \lambda_0 + c_\eta \lambda'_0) - (c_\xi \frac{\delta^y Z}{|\nabla Z|}^2 + c_\eta \frac{\delta^x Z}{|\nabla Z|}^2) \\ \Sigma_{Z3} = -(c_\xi \lambda_0 + c_\eta \lambda'_0) + \frac{c_\xi}{2}(1 - \frac{\delta^y Z}{|\nabla Z|}\frac{\delta^x Z}{|\nabla Z|}) + \frac{c_\eta}{2}(1 + \frac{\delta^y Z}{|\nabla Z|}\frac{\delta^x Z}{|\nabla Z|}) \\ \Sigma_{Z4} = -(c_\xi \lambda_0 + c_\eta \lambda'_0) + \frac{c_\xi}{2}(1 + \frac{\delta^y Z}{|\nabla Z|}\frac{\delta^x Z}{|\nabla Z|}) + \frac{c_\eta}{2}(1 - \frac{\delta^y Z}{|\nabla Z|}\frac{\delta^x Z}{|\nabla Z|}) \end{cases} \quad (10)$$

Therefore, the divergence term $c_\eta Z_{\eta\eta} + c_\xi Z_{\xi\xi}$ can be implemented as a convolution of the image Z by a 3*3 mask. It is important to note that the coefficients of this convolution mask are not constant. The values of the Σ_{Zi} depend on the gradient $|\nabla Z|$ through the function Φ. Hence, in the regions where the depth is homogeneous, $|\nabla Z|$ is small and the two coefficients c_η and c_ξ are equal to a constant c (this is due to the constraints imposed on the function $\Phi(.)$) and if one note $\lambda = \lambda_0 + \lambda'_0$, the coefficients for the divergence term simplify to:

$$\begin{cases} \Sigma_{w0} = -4c\lambda \quad \Sigma_{w1} = c(2\lambda - 1) \quad \Sigma_{w2} = c(2\lambda - 1) \\ \Sigma_{w3} = c(1 - \lambda) \quad \Sigma_{w4} = c(1 - \lambda) \end{cases}$$

As expected, these coefficients depend on the λ parameter. Here are some configurations for the divergence filter: These filters look like the classical Laplacian operator.

	$\lambda = \frac{1}{4}$	$\lambda = \frac{1}{2}$	$\lambda = \frac{3}{4}$	$\lambda = 1$
filter	$\begin{matrix} \frac{3c}{4} & -\frac{c}{2} & \frac{3c}{4} \\ -\frac{c}{2} & -c & -\frac{c}{2} \\ \frac{3c}{4} & -\frac{c}{2} & \frac{3c}{4} \end{matrix}$	$\begin{matrix} \frac{c}{2} & 0 & \frac{c}{2} \\ 0 & -2c & 0 \\ \frac{c}{2} & 0 & \frac{c}{2} \end{matrix}$	$\begin{matrix} \frac{c}{4} & \frac{c}{2} & \frac{c}{4} \\ \frac{c}{2} & -3c & \frac{c}{2} \\ \frac{c}{4} & \frac{c}{2} & \frac{c}{4} \end{matrix}$	$\begin{matrix} 0 & c & 0 \\ c & -4c & c \\ 0 & c & 0 \end{matrix}$

Table 1. Filters obtained for different values of λ

This is what we expected in the design of the regularization scheme: In homogeneous depth regions, the regularization must allow an isotropic smoothing. In inhomogeneous regions the configuration of the mask will not look like a Laplacian operator but as a directional second derivative in the direction normal to the gradient. This will perform an anisotropic diffusion i.e a diffusion only along the ISO-depth and not across the depth discontinuities.

Another directional numerical scheme has also been developed. It uses only 6 neighbors and not all the 8 neighbors as the one presented here and inspired from the discretization scheme proposed by L.Alvarez in [2] for multi-scale image analysis purpose. For more details about these two numerical schemes, one can refer to [11] and [16] where the two numerical schemes are presented in details.

6 Further Implementation Details

The following two additional points have to be mentioned regarding the implementation part:

First, it appears that the only use of the intensity field is not sufficient for dealing with real images, where the Lambertian assumption is almost always violated. The use of intensity gradient information generally helps a great deal in getting close to the correct solution. We have studied the evolution of the minimization process using different additional fields of attributes such as: intensity gradient along the epipolar line, intensity gradient vector, Laplacian of the intensity. It turns out that the intensity gradient vector field gives the best results.

Secondly, a multi scale approach speeds up convergence and helps avoiding local minima. The initial depth function at the coarsest scale is assigned a constant value. Then, the solution at each scale is used as initial value for the next, finer scale. In practice, we start with 32×32 images and double the image size at each level of the pyramid until we reach the final resolution.

7 Experimental Results

This section illustrates the experimental results that we have obtained on triplet of synthetic views of a pyramid. The baseline of cameras 1 and 2 (resp. 1 and 3) is aligned with the X (resp. Y) axis of the pyramid, so the epipolar lines have the directions represented in Figure 2 (top). Image intensity is very different from one stage of the pyramid to the next, and at each stage it is almost constant. The small variations due to lighting simulation cannot be perceived when looking at the images.

The three bottom lines of Figure 2 represent depth maps ((u,v,Z(u,v)) (left column) and dense 3D reconstructions (center and right columns) obtained, from top to bottom,

1. with cameras 1 and 2, and smoothing using the *Tikhonov* regurarization term (see how strong the smoothing is across the discontinuities)

2. with cameras 1, 2 and 3, and smoothing using the *Tikhonov* regurarization term (see how the use of a third camera improves the results but we still have the smoothing across the discontinuities)

3. with cameras 1, 2 and 3, and smoothing using the *Aubert* regurarization term (see how the use of this function improves the preservation of the discontinuities.)

The top two lines show the importance of using more than two images. Indeed, in the binocular case, there is almost no matching information along the epipolar lines, except close to the vertical edges. Thus, the shape is recovered only at these places. Elsewhere, since the image information is very weak, the depth-recovery process almost completely relies on the regularization term. In the trinocular case, depth is recovered everywhere on the pyramid. However, the smoothing term is too important for the small intensity variations within each pyramid level to be of any use to the algorithm. In the third case, we see how the algorithm can recover the shape of the pyramid when it also takes into account depth discontinuities.

8 Conclusion

A variational approach to dense stereo reconstruction has been presented. It allows to estimate directly the depth from a number of stereo images, while preserving depth discontinuities. The problem has been set as a regularization and minimization of a non quadratic functional. The *Tikhonov* regularization term usually used to recover smooth solution has been replaced by a function of the gradient depth specifically derived to allow depth discontinuities formation in the solution. The conditions that must be be fulfilled by this specific regularizing term to preserve discontinuities have been presented and a PDE based explicit scheme for moving iteratively towards the solution has been developed. Promising experimental results have been obtained. These results clearly illustrated the capabilities of this promising approach: direct depth recovery, depth discontinuities preservation and possibility to use more than 2 cameras.

References

1. L. Alvarez, P.L. Lions, and J.M.. Morel. Image selective smoothing and edge detection by nonlinear diffusion (ii). *SIAM J. Numer Anal*, 29(3):845–866, 1992.
2. L. Alvarez and F. Morales. Affine multiscale analysis of corners. *Prepublicacion, Ref: 9402, Universidad de Las Palmas de Gran Canaria*, 1994.
3. G. Aubert and L. Lazaroaia. A variational method in image recovery. TR 423, URA au CNRS no 168,laboratoire J-A Dieudonne, Université de Nice Sophia-Antipolis-CNRS, June 1995.
4. N. Ayache and B. Faverjon. Efficient Registration Of Stereo Images by Matching Graph Descriptions of Edge Segments. *The Int'l Journal of Comp. Vision*, 1(2):107–131, April 1987.
5. N. Ayache and C. Hansen. Rectification of Images for Binocular and Trinocular Stereovision. In *Proc. Int'l Conf. on Pattern Recognition*, pages –, October 1988. 9th, Rome, Italy.
6. S.T. Barnard. Stochastic Stereo Matching Over Scale. *The Int'l Journal of Comp. Vision*, 3(1):17–32, May 1989.
7. L. Blanc-Ferraud, P. Charbonnier, G. Aubert, and M. Barlaud. Nonlinear image processing : Modeling and fast algorithm for regularization with edge detection. TR 1, Laboratoire I3S, URA 1376 du CNRS, Université de Nice Sophia-Antipolis-CNRS, Janvier 1995.
8. A.T. Brint and M. Brady. Stereo Matching of Curves. In *IARP – 1st Int'l Workshop on Multi-sensor Fusion and Environment Modelling*, October 1989.
9. P. Charbonnier. *Reconstruction d'image : Régularisation avec prise en compte des discontinuités*. Thèse, Université of Nice Sophia-Antipolis, September 1994.
10. P. Charbonnier, G. Aubert, M. Blanc-Ferraud, and M. Barlaud. Two-Deterministic Half-Quadratic Regularizatiob Algorithms for Computed Imaging. In *Proc. IEEE Int. Conf. on Image Proc. ICIP*, volume II of III, pages 168–172, Austin(Texas), November 13-16 1994.
11. P. Delacourt. Applications des EDP en Analyse Multi-Echelle et en Traitement d'Images. Rapport de stage de DEA. proposé et encadré par R. Deriche, INRIA, Université Nice-Sophia Antipolis, France, Juin 1995.
12. R. Deriche and O. Faugeras. Les EDP en traitement des images et vision par ordinateur. Technical Report 2697, INRIA, November 1995.
13. P. Fua. Combining Stereo and Monocular Information to Compute Dense Depth Maps that Preserve Depth Discontinuities. In *Proc. Int'l Joint Conf. on Artificial Intelligence*, pages –, Sydney, Australia, August 1991.

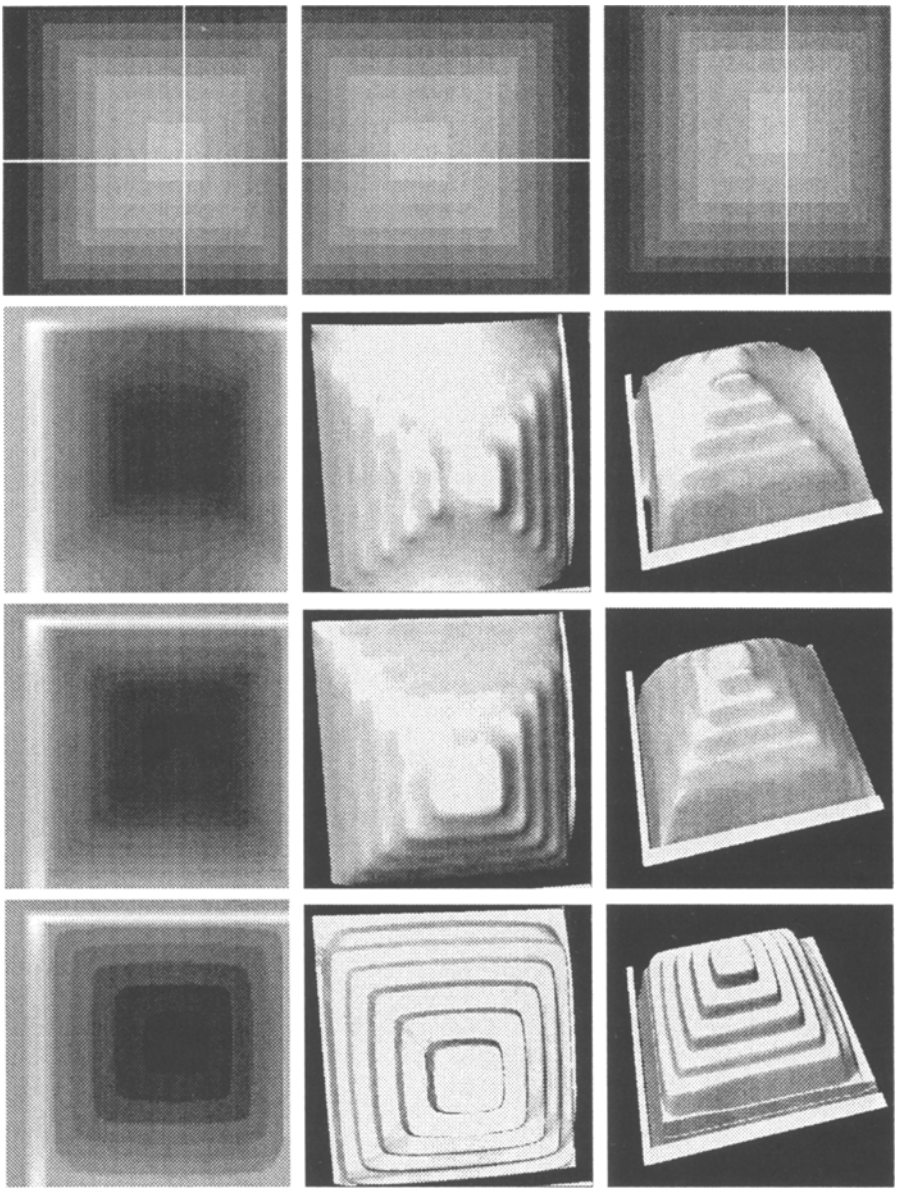

Fig. 2. Triple of images, depth maps and reconstructed surfaces (see text).

14. S. Geman and G. Reynolds. Constrained Restauration and the Recovery of discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14, 1992.
15. P.J. Green. Bayesian Reconstruction From Emission Tomography Data Using A Modified EM Algorithm. *IEEE Trans. Med. Imaging*, MI-9, 1990.
16. P Kornprobst. Calcul de Flot Optique avec Préservation des Discontinuités. Rapport de stage de DEA. Proposé et encadré par R. Deriche et G. Aubert, INRIA, Université Nice-Sophia Antipolis, France, Juin 1995.
17. F. Lustman. *Vision Stéréoscopique Et Perception Du Mouvement En Vision Artificielle*. PhD thesis, Université de Paris-Sud Centre d'Orsay, December 1988.
18. A. Meygret, M. Thonnat, and M. Berthod. A Pyramidal Stereovision Algorithm Based on Contour Chain Points. In *Proc. European Conf. on Comp. Vision*, Antibes, France, April 1990. Springer Verlag.
19. H.K. Nishihara and T. Poggio. Stereo Vision for Robotics. In *Int'l Symposium on Robotics Research*, Bretton Woods, New Hampshire, 1983.
20. Y. Ohta and T. Kanade. Stereo by Intra- and Inter-Scanline Search. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7, No 2:139–154, 1985.
21. S. Osher and L. Rudin. Shocks and Other Nonlinear Filtering Applied to Image Processing. In *Proc SPIE Vol 1567 : Applications of Digital Image Processing XIV*, pages 414–431, San-Diego, 1991.
22. P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.
23. S.B. Pollard, J.E.W. Mayhew, and J.P. Frisby. PMF: a Stereo Correspondance Algorithm using a Disparity Gradient Limit. *Perception*, 14:449–470, 1985.
24. L Robert. Camera calibration without feature extraction. *Computer Vision, Graphics, and Image Processing*, 1995. to appear, also INRIA Technical Report 2204.
25. L. Robert and O.D. Faugeras. Curve-Based Stereo: Figural Continuity And Curvature. In *Proc. Int'l Conf. on Comp. Vision and Pattern Recognition*, pages 57–62, Maui, Hawai, June 1991. IEEE.
26. L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
27. D. Terzopoulos. Image Analysis Using Multigrid Relaxation Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(2):129–139, March 1986.
28. A.N. Tikhonov and V.Y. Arsenin. In *Solutions of Ill-Posed Problems*. Wiley, New-York, 1977.
29. G. Toscani. *Systèmes de Calibration et Perception du Mouvement en Vision Artificielle*. PhD thesis, Université de Paris-Sud Centre d'Orsay, December 1987.
30. N. Yokoya. Stereo Surface Reconstruction by Multiscale-Multistage Regularization. Technical Report TR-90-45, ETL Electrotechnical Laboratory, November 1990.
31. Y.L. You, M. Kaveh, W.Y. Xu, and A. Tannenbaum. Analysis and Design of Anisotropic Diffusion for Image Processing. In *Proc. IEEE Int. Conf. on Image Proc. ICIP-94*, volume II of III, pages 497–501, Austin(Texas), November 13-16 1994.