A Variational Approach to 3D Geometry Reconstruction from Two or Multiple Views

L. Alvarez, C. Cuenca, A. Salgado, J. Sanchez
Dpto. Informatica y Sistemas
Universidad de Las Palmas de Gran Canaria
{lalvarez,ccuenca,a2652,jsanchez}@dis.ulpgc.es

Abstract

In the last years we have developed some methods for 3D reconstruction. First we began with the problem of reconstructing a 3D scene from a stereoscopic pair of images. We developed some methods based on energy functionals which produce dense disparity maps by preserving discontinuities from image boundaries. Then we passed to the problem of reconstructing a 3D scene from multiple views (more than 2). The method for multiple view reconstruction relies on the method for stereoscopic reconstruction. For every pair of consecutive images we estimate a disparity map and then we apply a robust method that searches for good correspondences through the sequence of images. Recently we have proposed several methods for 3D surface regularization. This is a postprocessing step necessary for smoothing the final surface, which could be afected by noise or mismatch correspondences. These regularization methods are interesting because they use the information from the reconstructing process and not only from the 3D surface. We have tackled all these problems from an energy minimization approach. We investigate the associated Euler–Lagrange equation of the energy functional, and we approach the solution of the underlying partial differential equation (PDE) using a gradient descent method.

1 Introduction

This paper is about three different main topics: The first topic is the 3D reconstruction from two views. The second is the problem of reconstructing a 3D scene from multiple views – more than two – that, in our case, make use of the previous method. The third one deals with the problem of regularizing a 3D surface. This method takes into account the information from the two previous topics.

For the first problem we present a variational approach to recover a dense disparity map from a set of two *weakly calibrated* stereoscopic images. To solve this problem, we first make full use of the knowledge of the so-called *fundamental matrix* to derive the equations that relate corresponding pixels in the two views, and then combine regularization and scale-space tools to estimate iteratively and hierarchically the disparity map. The solution obtained at a coarse spatial scale is used to restrict searching at finer scales. We minimize an energy term that takes into account the epipolar line constraint as well as the edge information constraint through an appropriate regularization term. In order to reduce the risk to be trapped within some irrelevant local minima during the iterations, we use a focusing strategy based on a linear scale-space. This method is explained in paper [1]. We have also implemented a symmetric method for computing the disparity map in both senses (see [4]). In papers [2] and [3] we have proposed methods for computing the optical flow between two images which are very similar to the disparity map estimation method except that we do not use any geometric constraint as the epipolar geometry.

For the second problem – reconstruction from multiple views – we have developed a robust method to recover a 3D model. After computing the disparity maps for every two consecutive frames we search for the best sequences of corresponding points through the set of frames. We estimate sequences of corresponding points across the multiple view image sequence. Basically, we try to connect points between images following the disparity map estimation. We select sequences of correspondent points for which the forward and backward disparity estimations are coherent. That is, if we take the initial point and go through the sequence using the forward disparity estimations and then go back using the backward disparity estimations we have to arrive to the same point (modulus a threshold parameter). From each selected corresponding point sequence we recover a 3D point by intersecting the projection lines of the points in the sequence. By collecting the 3D points obtained from each sequence we recover an unstructured set of 3D points. Recently, a new accurate technique based on a variational approach has been proposed in [17, 18]. Using a level set approach, this technique optimizes a 3D surface by minimizing an energy that takes into account the surface regularity as well as the projection of the surface on different images.In this paper we propose a different approach which is also based on a variational formulation but only using a disparity estimation between images and without defining explicitly any 3D surface.

For the third problem we present a method for the regularization of a set of unstructured 3D points obtained from a sequence of stereo images. Typically, the recovered set of 3D points is noisy, because of errors in the camera calibration process, errors in the disparity estimations, errors in the corresponding point sequences computations, etc., so some kind of regularization is needed. The regularization model we propose is a variational approach. We propose a model based on an energy. This method takes into account the information supplied by the disparity maps computed between pair of images to constraint the regularization of the set of 3D points. As in the first problem there is a regularization term that relies on an operator that is very similar to the Nagel–Enkelmann operator which allows for the regularization of the set of 3D points by preserving discontinuities presented on the disparity maps. One interesting advantage of this approach is that we regularize the set of 3D points by only using the 2D image projection information and, in particular, we do not need to define any 3D triangulation on the set of 3D points. In paper [5] we proposed a general method for regularizing a set of 3D points according to the information of the disparity maps, and in paper [6] there is an explanation of a 3D regularization method for cylindrical surfaces. In this paper we present the first approach for general surfaces.

The paper is organized as follows: In Section 2.1, we present the model for stereoscopic reconstruction from two views. In Section 2.2 we explain the 3D reconstruction method from multiple views. In Section 3 we present the regularization model and in Section 4 the conclusions.

2 Reconstruction model

2.1 Reconstruction from two views

In order to estimate a dense disparity map between two images we present an energy based approach. This energy also preserves discontinuities resulting from image boundaries. We derive a simplified expression for the disparity that allows us to easily estimate it from a stereo pair of images using an energy minimization approach. We assume that the epipolar geometry is known, and we include this information in the energy model. Discontinuities are preserved by means of a regularization term based on the Nagel–Enkelmann operator. We investigate the associated Euler–Lagrange equation of the energy functional, and we approach the solution of the underlying partial differential equation (PDE) using a gradient descent method. In order to reduce the risk to be trapped within some irrelevant local minima during the iterations, we use a focusing strategy based on a linear scale-space.

In order to estimate the disparity ($\lambda(x, y)$), one can proceed in a classical way and try to recover this important information using a simple correlation scheme. Unfortunately, this naive solution will not provide a correct and accurate solution, in particular in the regions where the disparity map may present some discontinuities, as is often the case close to image edges. It is well known that the disparity map obtained using this classical method tends to be very smooth across the boundaries of the images. The idea we would like to formalize and develop here is to estimate a $\lambda(x, y)$ function which is smooth only along the image boundaries and not across them. This leads us to consider the minimization of the following energy functional:

$$\begin{aligned} E(\lambda) &= \int_\Omega \left(I_l(x,y) - I_r(x+u(\lambda), y+v(\lambda))\right)^2 dxdy \\ &+ \; C \int_\Omega \Phi\left(\nabla I_l, \nabla\lambda\right)\, dx\, dy \end{aligned} \tag{1}$$

where Ω is the image domain, C is a positive constant, and $\Phi(\nabla I_l, \nabla\lambda)$ determines the regularization term. This function includes a diffusion tensor first proposed by Nagel and Enkelmann that guides the diffusion along the contours at image boundaries and in all directions at homogeneus regions. The associated Euler-Lagrange equations give us a diffusion PDE, which is then embedded into a gradient descend process to reach the solution:

$$\begin{aligned} \frac{\partial\lambda}{\partial t} &= C\,\mathrm{div}\left(D\left(\nabla I_l\right)\, \nabla\lambda\right) + \left(I_l(x,y) - I_r^\lambda(x,y)\right) \\ & \quad \frac{a\left(\frac{\partial I_r}{\partial y}\right)^\lambda (x,y) - b\left(\frac{\partial I_r}{\partial x}\right)^\lambda (x,y)}{\sqrt{a^2+b^2}}. \end{aligned} \tag{2}$$

In Figure 1 we show an example of disparity map computation from a stereoscopic pair of a human face. We also show the result given by a common correlation based technique.

In Figure 2 we show four views of the 3D reconstruction of the previous stereo pair.

Figure 1: Top: the original stereo pair. Bottom left: the computed disparity map using a correlation window of size 13×13. Bottom right: the result from our method, with the correlation result as initialization.

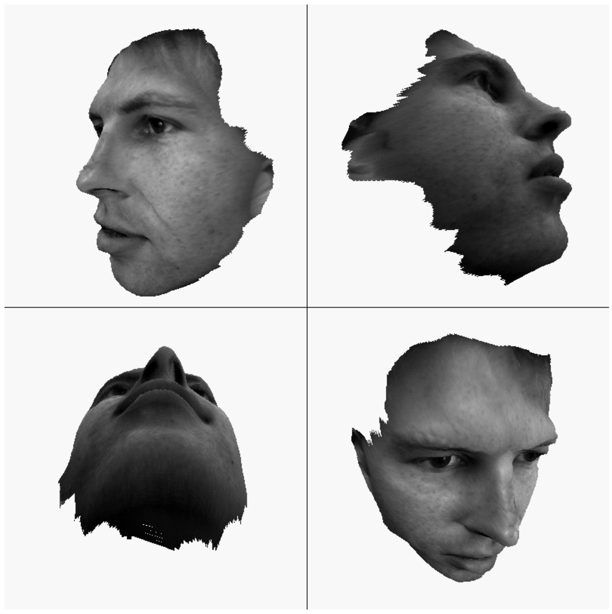

Figure 2: Four views of the 3-D reconstruction of the stereo pair in Figure 1, using the disparity map from our method.

2.2 Reconstruction from several views

We have developed a robust method to recover a 3D model from several views. All the views have been taken at the same time and all of them pointing to a common 3D scene. We make use of the previous model for stereoscopic images. This method follows these steps:

- For each pair of consecutive images, we estimate a dense disparity map using the accurate technique developed in [1]. We estimate such disparity map forward and backward, that is, from one image to the next one and in the opposite direction.

- We estimate sequences of corresponding points across the multiple view image sequence. Basically, we try to connect points between images following the disparity map estimation. We select sequences of correspondent points for which the forward and backward disparity estimations are coherent. That is, if we take the initial point and go through the sequence using the forward disparity estimations and then go back using the backward disparity estimations we have to arrive to the same point (modulus a threshold parameter). We keep trace of the pixels belonging to a sequence in order to avoid that the same pixel is included in different groups.

- From each selected corresponding points sequence we recover a 3D point by intersecting the projection lines of the points in the sequence. By collecting the 3D points obtained from each sequence we recover an unstructured set of 3D points.

- Typically, the recovered set of 3D points is noisy, because of errors in the camera calibration process, errors in the disparity estimations, errors in the corresponding point sequences computations, etc., so some kind of regularization is needed. In this paper, we propose a new variational model to smooth the unstructured set of 3D points. This regularization model is based on the 2D image information and does not require to define any kind of geometric relation between the 3D points.

In Figure 4, we show the front and profile views of the reconstruction of the Bust sequence.

3 Regularization model

In this section we present a method for the regularization of a set of unstructured 3D points obtained from a sequence of stereo images. This is a postprocessing step to the reconstruction from multiple views explained in the previous section. This method takes into account the information supplied by the disparity maps computed between pair of images to constraint the regu-

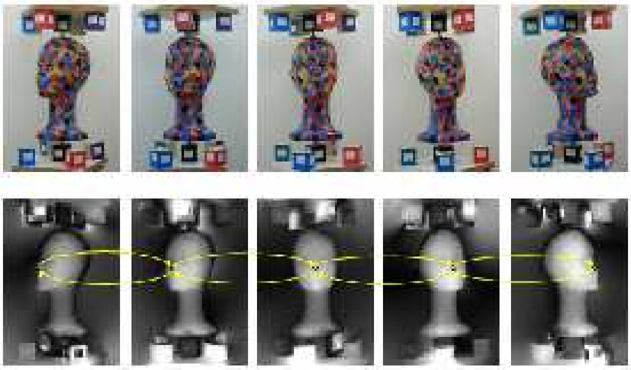

Figure 3: Some images of the Bust sequence and the corresponding disparity maps associated to them. We search for sequences of corresponding points which have a small error going forward and backward, therefore we have to compute disparity maps for every pair of images in both senses.

Figure 4: Front and profile views of the 3D reconstruction of the Bust sequence.

larization of the set of 3D points. We propose a model based on an energy which is composed of an attachment term that minimizes the distance from 3D points to the projective lines of camera points, and a second term which relies on an operator that is very similar to the Nagel–Enkelmann operator. This second term allows for the regularization of the set of 3D points by preserving discontinuities presented on the disparity maps. One interesting advantage of this approach is that we regularize the set of 3D points by only using the 2D image projection information and, in particular, we do not need to define any 3D triangulation on the set of 3D points. We embed this energy in a 2D finite element method to take advantage of the underlying float precision mesh. Delauny triangulations are managed to give support to this finite element method.

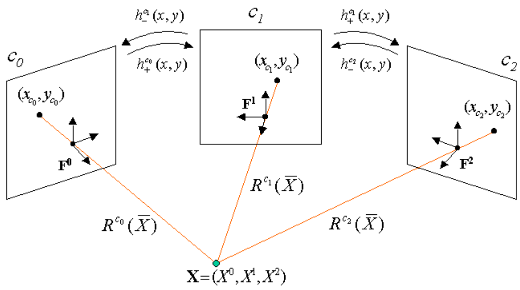

Figure 5: Notation

After minimizing, this 2D finite element method results in a large system of equations that can be optimized for fast computations over no-null values. We derive an efficient implicit numerical scheme which reduces the number of calculations and memory allocations.

The energy to be minimized for the regularization of the set of 3D point is:

$$
\begin{aligned}
E\left(\bar{X}^0., \bar{X}^{N_c-1}\right) &= \sum_{c=0}^{N_c-1}\left(\int_\Omega dist(\bar{X}^c, \bar{R}^c)^2 + \right. \\
&\quad \left. \alpha \sum_{i=0}^{2}\int_\Omega \nabla^T X^c D(\bar{h}^c)\nabla X^c\right) \qquad (3)
\end{aligned}
$$

where α is a parameter that states the balance between the two terms and $dist(\bar{X}^c, \bar{R}^c)$ denotes the distance from point $\bar{X}^c$ to the straight line $\bar{R}^c$ and is given by formula

$$
\begin{aligned}
dist(\bar{X}^c, \bar{R}^c)^2 &= \sum_{i=0}^{2}\left(X^{c,i} - F^{c,i}\right)^2 \\
&\quad - \left(\sum_{i=0}^{2} l^{c,i}\left(X^{c,i} - F^{c,i}\right)\right)^2 . \qquad (4)
\end{aligned}
$$

In Figure 5 is the notation that we have used for equations (3) and (4) and in Figure 6 we show the result of applying this method to the 3D reconstruction on Figure 4.

4 Conclusions

Our method for 3D reconstruction from a pair of stereoscopic images combines some techniques developed in the context of optic flow estimation [2, 22] with some other techniques developed in the context of dense disparity map estimation which take into account the geometric constraints associated to a stereo pair. We think

Figure 6: Front and profile views of a regularization of the 3-D reconstruction on Figure 4

that the combination of these ideas is fruitful in that it produces new tools to estimate dense disparity fields which benefit from the research efforts in stereo vision as well as in optic flow estimation.

The method for 3D reconstruction from multiple views is very robust in the sense that it intensively looks for the best matching sequences of points. This yields a set of 3D points that are probably the most accurate ones and discards those points that have not good matches or do not appear in enough views – maybe due to occlusions.

In this paper we have presented a novel method for the regularization of a set of 3D points. We have established an energy in a traditional attachment–regularizing couple of terms. In the regularizing term we have made used of an operator similar to the Nagel–Enkelmann operator for 3D regularizations.

This energy model has been embedded into a 2D finite element approach to take advantage of the underlying precision of data. Then we have managed to derive this energy and propose a very efficient and optimal numerical scheme that allows us to speed up the process and reduce the memory needs.

One of the main advantages of the method is that it regularizes sets of unstructured 3D points without using any geometric relation in 3D. We only use the information of the projection of the points in the cameras. In particular, this method could be used as a preprocessing step before the construction of 3D surfaces fitting the 3D points. We notice that most of the techniques for such surface reconstruction are very sensitive to the noise in the 3D points representation, and they require the set of 3D points to be regular enough to work properly.

5 Acknowledgments

This work has been partially supported by the research project PI2002/193 founded by the Consejería de Educación, Cultura y Deportes of the Canary Islands Government.

References

[1] L. Alvarez, R. Deriche, J. Sánchez and J. Weickert, *Dense disparity map estimation respecting image discontinuities: a PDE and scale-space based approach*, Journal of Visual Communication and Image Representation, 13:3–21, January 2002. Also published as Inria Research Report n^o 3874

[2] L. Alvarez, J. Weickert, J. Sánchez, *A scale-space approach to nonlocal optical flow calculations*, M. Nielsen, P. Johansen, O.F. Olsen, J. Weickert (Eds.), Scale-space theories in computer vision, Lecture Notes in Computer Science, Springer, Berlin, Vol. 1682, 235–246, 1999.

[3] L. Alvarez, J. Weickert, J. Sánchez, *Reliable Estimation of Optical Flow for Large Displacements*, International Journal of Computer Vision, 39(1):41–56, 2000. An extended version maybe be found at Technical Report No. 2, Cuadernos del Instituto Universitario de Ciencias y Tecnologías Cibernéticas, Universidad de Las Palmas de Gran Canaria, Spain, 1999

[4] L. Alvarez, R. Deriche, T. Papadopoulo and J. Sánchez, *Symmetrical dense optical flow estimation with oclussions detection*, ECCV 2002 Congress (Lectures Notes in Computer Science), Cophenaghe, Denmark, May 2002

[5] L. Alvarez, C. Cuenca, J. Sánchez, *Regularizing a Set of Unstructured 3D Points from a Sequence of Stereo Images*, Scale-Space 2003 Congress, Isle of Skye, England, June 2003

[6] L. Alvarez, C. Cuenca, J. Sánchez, *Regularization of 3D Cylindrical Surfaces*, IBPRIA 2003 Congress, Palma de Mallorca, Spain, June 2003

[7] P. Charbonnier, L. Blanc–Féraud, G. Aubert, M. Barlaud, *Two deterministic half-quadratic regularization algorithms for computed imaging*, Proc. IEEE Int. Conf. Image Processing (ICIP–94, Austin, Nov. 13–16, 1994), Vol. 2, IEEE Computer Society Press, Los Alamitos, 168–172, 1994.

[8] Frédéric Devernay and Olivier D. Faugeras. Computing differential properties of 3-D shapes from stereoscopic images without 3-D models. Proc. IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, June 21–23, 1994, pages 208–213.

[9] W. Enkelmann, *Investigation of multigrid algorithms for the estimation of optical flow fields in image sequences*, Computer Vision, Graphics and Image Processing, Vol. 43, 150–177, 1988.

[10] Olivier Faugeras, Bernard Hotz, Hervé Mathieu, Thierry Viéville, Zhengyou Zhang, Pascal Fua, Eric Théron, Laurent Moll, Gérard Berry, Jean Vuillemin, Patrice Bertin, and Catherine Proy. Real time correlation based stereo: algorithm implementations and applications. *International Journal of Computer Vision*. To appear.

[11] D. Geman, G. Reynolds, *Constrained restoration and the recovery of discontinuities*, IEEE Trans. Pattern Anal. Mach. Intell., 14, 367–383, 1992

[12] D.G. Jones, J. Malik, *A computational framework for determining stereo correspondence from a set of linear spatial filters*, G. Sandini (Ed.), Computer vision – ECCV '92, Lecture Notes in Computer Science, Vol. 588, Springer, Berlin, 395–410, 1992.

[13] N. Amenta and M. Bern. Surface Reconstruction by Voronoi Filtering. *14th ACM Symposium on computational Geometry*, pages 248–253, June 1998

[14] F. Bernardini, J. Mittleman and H. Rushmeier and C. Silva. The Ball Pivoting Algorithm for Surfarce Reconstruction. *IEEE Transactions on Visualization and Computer Graphics*, 5(4):349–359, Oct.–Dec. 1999

[15] H. Edelsbrunner and E. P. Mucke. Three-dimensional Alpha Shapes. *ACM Transactions on Graphics*, 13(1):43–72, January 1994

[16] O. Faugeras. Three-Dimensional Computer Vision: A Geometric Viewpoint. MIT Press, 1993

[17] O. Faugeras and R. Keriven. Complete Dense Stereovision Using Level Set Methods. *Proceedings of Fifth European Conference on Computer Vision*, 1998

[18] O. Faugeras and R. Keriven. Variational principles, surface evolution, PDEs, level set methods and the stereo problem. *IEEE Transactions on Image Processing. Special issue on Geometry driven diffusion and PDEs in image processing*, 7(3):336–344, 1998

[19] O. Faugeras, and Q. Luong and T. Papadopoulo. The Geometry of Multiple Images. Mit Press, 2001

[20] P. George and H. Borouchaki. Triangulation de Delaunay et Maillage. Hermes, 1997

[21] R. Hartley and A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge University Press, 2000

[22] H.H. Nagel and W. Enkelmann. An Investigation of Smoothness Constraints for the Estimation of Displacement Vector Fields from Images Sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(5):565–593, 1986