Regularization of 3D Cylindrical Surfaces

Luis Alvarez, Carmelo Cuenca, and Javier Sánchez

Departamento de Informática y Sistemas
Universidad de Las Palmas de G.C.
Campus Universitario de Tafira
35017, Las Palmas
{lalvarez,ccuenca,jsanchez}@dis.ulpgc.es
http://serdis.dis.ulpgc.es/~{alvarez,jsanchez}

Abstract. In this paper we present a method for the regularization of 3D cylindrical surfaces. By a cylindrical surface we mean a 3D surface that can be expressed as an application $S(l,\theta) \rightarrow R^3$, where (l,θ) represents a cylindrical parametrization of the 3D surface. We built an initial cylindrical parametrization of the surface. We propose a new method to regularize such cylindrical surface. 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. We propose a model based on an energy which is composed of two terms: an attachment term that minimizes the difference between the image coordinates and the disparity maps and a second term that enables a regularization by means of anisotropic diffusion. One interesting advantage of this approach is that we regularize the 3D surface by using a bi-dimensional minimization problem.

1 Introduction

This paper deals with the problem of 3D geometry reconstruction from multiple 2D views. Recently, a new accurate technique based on a variational approach has been proposed in [4]. Using a level set approach, this technique optimizes a 3D surface by minimizing an energy that takes into account the regularity of the set of points as well as the projection of the set of points 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. We will assume that the 3D surface we want to recover has a cylindrical geometry, that is, it can be expressed as an application $S(l,\theta) \rightarrow R^3$, where (l,θ) represents a cylindrical parametrization of the 3D surface. Of course, this is an important limitation in term of the surface geometry, but it simplifies in a strong way the complexity of the problem and it can be applied in a lot of situations like for instance, human face reconstruction as we will show in the experimental results. We will also assume that the cameras are calibrated (see [3], [5] or [6]). Very accurate techniques to estimate the disparity map in a stereo pair of images have been proposed. To extend these techniques to the case of multiple views is not

a trivial problem. The 3D geometry estimation that we propose can be divided in the following steps:

- For every pair of consecutive images, we estimate a dense disparity map using the accurate technique developed in [1]. We estimate such disparity maps forward and backward. From these disparity maps we obtain a 3D surface for every pair of stereoscopic images.
- Based on the camera configuration we estimate a 3D cylinder and we project in such cylinder the 3D surfaces obtained in the previous step. From these projections we estimate an initial cylindrical parametrization of the surface. This cylindrical parametrization is based on the distance between the 3D point and the cylinder axis. In fact, for each cylinder coordinates (l, θ) we average such distance for all 3D points which are projected in (l, θ) .
- Typically, the recovered set of 3D points is noisy, because of errors in the camera calibration process, errors in the disparity map estimations, etc., so some kind of regularization is needed. In this paper, we propose a new variational model to smooth cylindrical surfaces. This regularization model is based on the disparity estimations.

The regularization model we propose is based on a variational approach. We start with an energy that has two terms, an attachment and a regularizing term. The former minimizes the difference by respect to the disparity map computed for every pair of stereoscopic images. This term is responsible for maintaining the final 3D regularized point close to the information supported by the disparity maps. The latter enables a regularization by preserving discontinuities on the cylindrical function. The regularizing term is similar to the terms used in other fields like stereoscopic reconstruction [1], optical flow estimation [2], etc.

Deriving this energy yields a PDE (Partial Differential Equation) that is then embedded into a gradient descend method to look for the solution. We develop an explicit numerical scheme based on finite differences to implement the method.

In Sect. 2 we introduce the cylindrical coordinate system necessary for the representation of the cylindrical function and the relation with the projective camera model. In Sect. 3 we study the model by proposing an energy deriving it and embedding the resulting PDE into a gradient descend method. In Sect. 3.2 there is an explanation of the explicit numerical scheme. Finally in Sect. 4 we present the experimental results for the bust sequence.

2 The Cylinder Structure

2.1 The Cylindrical Coordinate System and the Projective Camera

Using the notation expressed in Fig. (1) we note by $\bar{N}_1, \bar{N}_2$ and $\bar{N}_3$ the orthogonal axis of the coordinate system and by $\bar{Q}_0$ the origin of the system. $\bar{N}_1$ represents the cylinder axis. The cylindrical coordinates are expressed by means of a list of three candidates (l, θ, r) where l is the displacement on the cylinder axis, θ is an angle (as it is outlined in Fig. (1)) and r is the distance from a 3D point to the

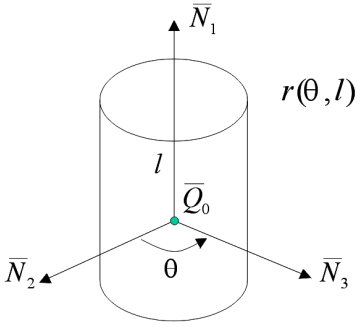

Fig. 1. Cylindrical and cartesian coordinate systems

cylinder axis. A cylindrical surface $S(l,\theta)$ will be given by a cylindrical function $r(\theta,l)$ in the following way :

$$S(l,\theta) = \bar{Q}_0 + l\bar{N}_1 + r(l,\theta)\left(\cos\theta\bar{N}_2 + \sin\theta\bar{N}_3\right) . \tag{1}$$

With this relation we may transform a cylindrical function $r\,(l,\theta)$ into a function in the cartesian coordinate system. So, to provide a cylinder surface is equivalent to provide a cylinder function $r(l,\theta)$. We will see later that our method make use of the disparity maps between pairs of stereoscopic images to constraint the regularization of the cylindrical function. The disparity maps are expressed in image coordinates associated to every camera. We assume the projective model for the cameras. In our problem we have N_c different projective cameras and every camera is represented by a projection matrix $\mathbf{P}_c$ of dimensions $3x4$ that projects 3D points into the projection plane. In projective coordinates these projections can be represented as follows:

$$\tilde{m}_c(l,\theta) = \mathbf{P}_c\,(S(l,\theta),1)^t\ . \tag{2}$$

2.2 Building the Cylindrical Function

We suppose that for every stereoscopic pair we have a 3D surface. Our first problem is to transform the 3D surfaces into a unique cylindrical function. The main steps for computing the cylindrical function are:

1. Compute the coordinate system by estimating $\bar{Q}_0, \bar{N}_1, \bar{N}_2$ and $\bar{N}_3$.
2. Adapt the resolution of the cylindrical image. The cylindrical function will be represented through an image. This is what we call the cylindrical image. The rows and columns of this image are given by the $\bar{N}_1$ axis and the angle, θ, respectively.
3. Create the cylindrical function, $r(\theta,l)$. Once we have carried out the previous steps we have to merge the information of all the 3D surfaces in one function. We compute an average for all coincident 3D points projections in one pixel in the cylinder coordinate system (l,θ).

The first step is to estimate the position, $\bar{Q}_0$ and axis, $\bar{N}_1, \bar{N}_2$ and $\bar{N}_3$, of the cylindrical coordinate system. We have supposed that the camera configuration system is cylindrical in the sense that all the cameras are situated around the scene and looking at the center. We also suppose that the focus of the cameras are situated close to a common plane. $\bar{Q}_0$ is estimated as the average of the 3D points of all surfaces. $\bar{N}_1$ is the cylindrical axis and is computed accordingly to the configuration of the focuses, $\bar{N}_2$ is the unitary vector that points to the focus of the first camera and $\bar{N}_3$ is orthogonal to the others.

In the second step we are concern with the problem of representing the cylindrical function through a bi-dimensional image. We have to compute the dimensions of an image that will allocate the values of the 3D points in cylindrical coordinates. To calculate the number of rows the lowest and highest 3D points in the $\bar{N}_1$ component are computed. The difference between them defines the size of the cylindrical axis. The number of columns are estimated knowing that $2 \cdot \pi \cdot radius$ is the length for the cylinder. We adapt the value of $radius$ in order to obtain an image with regular pixels (same pixel height and width). This value depends on the dimension of the image in the $\bar{N}_1$ axis. This image represents the $r(\theta, l)$ function.

The last step consist of assigning a value to every pixel on the image. This process is carried out by representing the 3D points in cylindrical coordinates and computing an average for coincident points on a pixel. There may be some locations where no 3D point is projected, so a post-processing to fill these pixels is necessary. These are filled from the values of the surrounding pixels.

3 The Regularizing Method

3.1 Energy Minimization

The regularization of the cylindrical function $r(l, \theta)$ is equivalent to regularize the cylindrical surface $S(l, \theta)$. We propose a variational formulation to look for the regularized solution. This solution is the result of a minimization problem. Our model is composed of two terms: the attachment term that uses the disparity maps to constraint the process; and the regularizing term that is used to obtain a smooth solution. This term is designed to regularize the surface by preserving the discontinuities of the cylindrical function which are related to the varying depth of the 3D surface.

The energy model proposed is

$$\begin{aligned} E(r) = \beta \Bigg(& \sum_{c=1}^{N} \int\int \left\| \bar{m}_{c+1}(l,\theta) - \bar{m}_c(l,\theta) - h_+^c(\bar{m}_c) \right\|^2 dl d\theta \\ & + \sum_{c=1}^{N} \int\int \left\| \bar{m}_c(l,\theta) - \bar{m}_{c+1}(l,\theta) - h_-^c(\bar{m}_{c+1}) \right\|^2 dl d\theta \Bigg) \\ & + \alpha \int\int \phi(\|\nabla r\|)\ dl d\theta\,. \end{aligned} \tag{3}$$

$\bar{m}_{c+1}(l,\theta)$ is the image coordinate for camera $c+1$ denoted by (2) and $\bar{m}_c(l,\theta)$ is the correspondent for camera c. Vectors $\bar{h}^c_{+/-}(\bar{m}_c) = \begin{pmatrix} u_{+/-}(\bar{m}_c) \\ v_{+/-}(\bar{m}_c) \end{pmatrix}$ represent the optical flow estimations for pixel $\bar{m}_c$ on camera c. Sign $+$ corresponds to the optical flow from camera c to $c+1$ and sign $-$ to the optical flow from camera c to camera $c-1$.

After minimizing this energy we obtain the associated Euler–Lagrange equation that is given by the following PDE:

$$\begin{aligned}
&\beta \cdot \left(\sum_{c=1}^{N_c} \left(\left(\bar{m}_{c+1} - \bar{m}_c - \bar{h}^c_+(\bar{m}_c)\right)^t \cdot \left(\frac{\partial \bar{m}_{c+1}}{\partial r} - \frac{\partial \bar{m}_c}{\partial r} - \mathcal{J}\bar{h}^c_+ \frac{\partial \bar{m}_c}{\partial r} \right) \right) \right. \\
&\left. + \sum_{c=1}^{N_c} \left(\left(\bar{m}_c - \bar{m}_{c+1} - \bar{h}^{c+1}_-(\bar{m}_{c+1})\right)^t \cdot \left(\frac{\partial \bar{m}_c}{\partial r} - \frac{\partial \bar{m}_{c+1}}{\partial r} - \mathcal{J}\bar{h}^{c+1}_- \frac{\partial \bar{m}_{c+1}}{\partial r} \right) \right) \right) \\
&-\alpha \cdot \operatorname{div}\left(\frac{\phi'\left(\|\nabla r\|\right)}{\|\nabla r\|} \nabla r \right) = 0
\end{aligned} \tag{4}$$

where $\mathcal{J}\bar{h} = \mathcal{J}\begin{pmatrix} u(x,y) \\ v(x,y) \end{pmatrix} = \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix}$.

In order to search for the solution we implement a gradient descend method in the way $\frac{\partial r}{\partial t} = -\frac{\partial E(r)}{\partial r}$. The divergence term is well known and acts like a diffusion scheme. If we expand the divergence expression we obtain

$$\operatorname{div}\left(\frac{\phi'\left(\|\nabla r\|\right)}{\|\nabla r\|} \nabla r \right) = \frac{\phi'\left(\|\nabla r\|\right)}{\|\nabla r\|} r_{\xi\xi} + \phi''\left(\|\nabla r\|\right) r_{\eta\eta} \tag{5}$$

where $\eta = \frac{\nabla r}{\|\nabla r\|}$ and $\xi = \eta^{\perp}$ are the unitary vectors in the directions parallel and perpendicular to the gradient, respectively.

Playing with function $\phi(s)$ it is possible to achieve an anisotropic diffusion at contours. The first in proposing this kind of diffusion equation were Perona and Malik [7] in where they introduced a decreasing function to avoid diffusion at contours.

3.2 Numerical Scheme

In this section we are going to see how to implement an explicit numerical scheme for this method. We derive $\frac{\partial \bar{m}}{\partial r}$ analytically from (2). Regarding (5) the divergence is divided in two terms and the values for both of them are given by the following expressions:

$$r_{\xi\xi} = \frac{r_{xx} r_y^2 - 2 r_x r_y r_{xy} + r_{yy} r_x^2}{r_x^2 + r_y^2}, \qquad r_{\eta\eta} = \frac{r_{yy} r_y^2 + 2 r_x r_y r_{xy} + r_{xx} r_x^2}{r_x^2 + r_y^2}. \tag{6}$$

Fig. 2. Bust configuration: This figure shows the 3D reconstructed bust and the distribution of the projection planes corresponding to the 47 cameras

The first and second derivates on x and y and the derivates of the components of the optical flow, $\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial v}{\partial x}, \frac{\partial v}{\partial y}$ and $\frac{\partial u}{\partial x}$, have been approximated by finite differences.

The final numerical scheme is implemented by means of an explicit scheme in the following way:

$$r_{t+1} = r_t + dt \cdot \left(\alpha \left(r_{\xi\xi} + g\left(\|\nabla r\| \right) r_{\eta\eta} \right) - \beta \left(\sum_{c=1}^{N_c} \left(\left(\bar{m}_c - \bar{m}_{c+1} - \bar{h}_-^{c+1}(\bar{m}_{c+1}) \right)^t \cdot \left(\frac{\partial \bar{m}_c}{\partial r} - \frac{\partial \bar{m}_{c+1}}{\partial r} - \mathcal{J}\bar{h}_-^{c+1} \frac{\partial \bar{m}_{c+1}}{\partial r} \right) \right) + \sum_{c=1}^{N_c} \left(\left(\bar{m}_{c+1} - \bar{m}_c - \bar{h}_+^{c}(\bar{m}_c) \right)^t \cdot \left(\frac{\partial \bar{m}_{c+1}}{\partial r} - \frac{\partial \bar{m}_c}{\partial r} - \mathcal{J}\bar{h}_+^{c} \frac{\partial \bar{m}_c}{\partial r} \right) \right) \right) \right) . \quad (7)$$

Function $g\left(s\right)$ is a decreasing function that disables isotropic diffusion for big values of the gradient.

4 Experimental Results

In this section we show the results of regularizing a bust sequence. In this case the sequence is composed of 47 images taken around a bust. Figure (2) shows the configuration of this sequence with the projection planes of the cameras. This is a close sequence in where the first and last images are correlatives.

In Fig. (4) we may see the original Bust reconstruction and a regularized version for $\alpha = 0,1$ and $s = 0.1$.

From Fig. (5) we may appreciate several regularizations for $\alpha = 3,0$ and different values of s. The β parameter is much smaller and is used to normalize the variation between the two terms. In these experiences $\beta = 10^{-4}$.

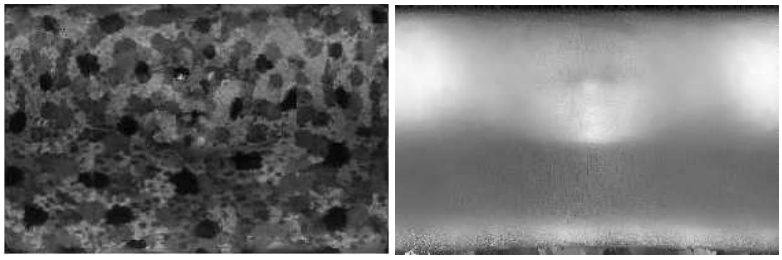

Fig. 3. The left image represents the texture of the Bust sequence projected on a cylindrical image. The right image is the cylindrical function represented in gray levels (the white color is associated to the highest values)

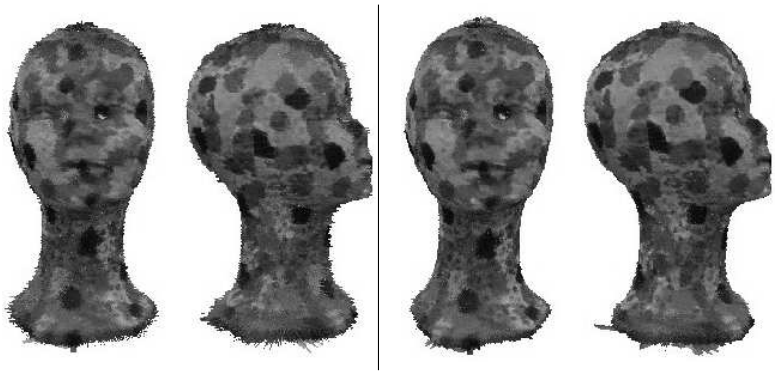

Fig. 4. Left two images: Front and profile views of the Bust reconstruction. Right two images: Front and profile views of a 3D regularization for the Bust sequence using $\alpha = 0,1$ and $s = 0.1$

5 Conclusions

In this paper we have presented a novel and simple method for the representation and regularization of cylindrical surfaces. This method is ideally suited for convex surfaces and also be appropriated for surfaces that have not deep clefts. We have taken advantage of the simplicity of cylindrical coordinates to represent the set of 3D points. Once the cylindrical function is built the problem of regularizing the set of 3D points is reduced to the problem of regularizing a bi-dimensional function.

We have established an energy in a traditional attachment–regularizing couple of terms. From this energy we have derived a diffusion-reaction PDE. We have shown in the experiments that varying the α parameter results in a more regular set of points and varying the λ parameter implies a more regular set of points by preserving the cylindrical function discontinuities as we have expected from the results obtained in other fields. The use of α and λ parameters are

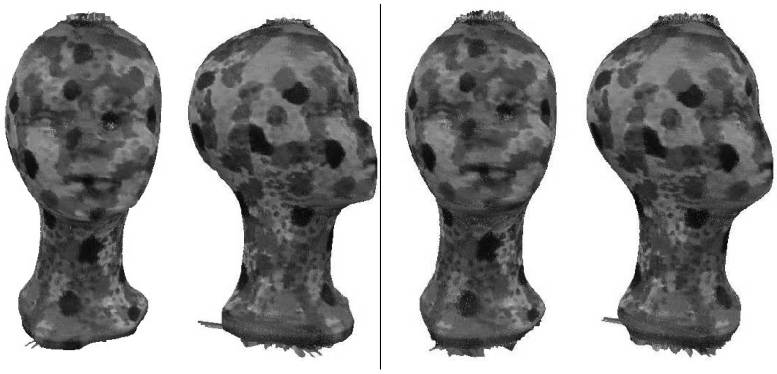

Fig. 5. Several views of different 3D regularizations for the Bust sequence. Left two images: $\alpha = 3,0$ and $s = 0,5$; Right two images: $\alpha = 3,0$ and $s = 1,0$

simple. α refers to the smoothness of the final set of points and λ refers to the way the regularization is carried out at the contours.

6 Acknowledgments

This work has been partially supported by the Spanish research project TIC 2000-0585 founded by the Ministerio de Ciencia y Tecnología and by the research project PI2002/193 founded by the Canary Islands Government.

References

1. Alvarez, L., Deriche, R., Sánchez, J., and Weickert, J.: Dense disparity map estimation respecting image derivatives: a PDE and scale-space based approach. Journal of Visual Communication and Image Representation **13** (2002) 3–21. Also published as Inria Research Report n^o 3874
2. Alvarez, L., Weickert, J., and Sánchez, J.: Reliable Estimation of Dense Optical Flow Fields with Large Displacements. International Journal of Computer Vision, Vol. 39, **1** (2000) 41–56. An extended version maybe be found at Technical Report n^{o}2 del Instituto Universitario de Ciencias y Tecnologías Cibernéticas
3. Faugeras, O.: Three-Dimensional Computer Vision: A Geometric Viewpoint. MIT Press (1993)
4. Faugeras, O. and Keriven, R.: Complete Dense Stereovision Using Level Set Methods. Proceedings of Fifth European Conference on Computer Vision (1998)
5. Faugeras, O., and Luong, Q., and Papadopoulo, T.: The Geometry of Multiple Images. Mit Press (2001)
6. Hartley, R. and Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press (2000)
7. Perona, P. and Malik, J.: Scale-Space and Edge Detection Using Anisotropic Diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence **12** (1990) 429–439