

SMOOTH SURFACE RECONSTRUCTION FROM IMAGE SEQUENCES

Edmond Boyer and Marie Odile Berger

CRIN-CNRS / INRIA Lorraine
Campus scientifique, BP 239
54506 Vandoeuvre les Nancy Cedex, France
e-mail: boyere@loria.fr, berger@loria.fr

ABSTRACT

This paper addresses the problem of 3D surface reconstruction using image sequences. It has been shown that shape recovery from three or more occluding contours of the surface is possible given a known camera motion. Recent algorithms allow the estimation of a 3D depth map of points on extremal contours. We present our approach which is based on local approximations up to order two of the surface. Such reconstruction must then be followed by a global surface description stage. This is done by constructing a triangulation of the reconstructed points and by performing a regularisation step in order to correct perturbations which affect the reconstruction. Experiments on real data were conducted which have proved the reliability of the method.

1. INTRODUCTION

Recovering three-dimensional shapes from visual data is an important intermediate goal of many vision systems involved in robotics, surveillance, guidance, modelling, etc. For such applications, surface reconstruction is often used as an intermediate step in the more important task of 3D representation and recognition.

For non-polyhedral objects, rich and robust information on the shape are provided by the occluding contours. In fact, if some *a priori* knowledge on the object is available, such as parametric descriptions (using quadrics, generalised cylinders, ...), only one view is required to perform reconstruction [1, 2]. Otherwise, a sequence of occluding contours obtained with known camera motion must be considered to recover the shape.

We assume that the object surface is smooth and curved¹ and that the imaging system is based on the pinhole model. Our aim is to reconstruct the 3D sur-

face of an object from the observations of its occluding contours. Such contours are known to be a rich source of geometric information [3]: the corresponding extremal contours on the surface are viewpoint dependent and defined by the fact that the optical rays of their points are tangential to the surface. Moreover, it has been shown that local shape recovery from three or more occluding contours of the surface is theoretically possible given a known camera motion. Several algorithms [4, 5, 6] allow such a local reconstruction under the assumption of a linear camera motion. Zhao [7], in more recent works, attempts to globally recover the 3D surface and proposes to search the B-spline surface which best fits all the occluding contours, however, this method supposes that a complete parametrisation of the surface is *a priori* available which is not, in general, possible.

In this context, we propose a new method which consists of two major phases:

1. We first reconstruct discrete points on extremal contours. Our approach [8] is based on a local approximation of the surface. This allows a linear estimation of depth to be derived, given three consecutive occluding contours. Unlike other local methods, no assumption is made on the camera motion or on the local surface shape. In addition, local methods present advantages compared to a global surface reconstruction. First, it does not require a complete parametrisation of the surface which is usually not available. Secondly, it allows a local control of the reconstruction stability and thus, elimination of highly corrupted points in the case of noisy data.
2. The complete surface is then described as a triangular mesh, which does not require any *a priori* knowledge on the parametrisation. Finally, point positions are optimised in order to correct reconstruction errors. This is done by considering the regularity of the surface through the minimisation of an energy function.

¹We suppose that the surface is at least C^2 and that it is not locally a plane.

The reconstruction process is described in section 2. In section 3, we present the regularisation procedure. Results on real data are shown in section 4.

2. RECONSTRUCTION

In this section, we summarise our method, see [8] for more details.

In order to reconstruct a point P belonging to an extremal contour C , three occluding contours are required: the corresponding occluding contour and its two neighbouring contours in the sequence. The epipolar correspondence [8] define two correspondents of the point P on the neighbouring contours. Given these three points and the camera motion, we can compute an approximation up to order two of the surface at P . This allows linear equations to be derived for both depth and normal curvature in the viewing direction at P .

Our approach is based on a local parametrisation $(x, y, z = h(x, y))$ of the surface S at P , where the x axis is directed by the viewing direction, the y axis is directed by the tangent to the extremal contour at P and the z axis is directed by the normal of S at P . In this coordinate frame, an approximation up to order two of S can be written :

$$\mathcal{Q} = \{(x, y, z), z = \frac{1}{2}(k_t x^2 + k_s y^2)\}.$$

where k_t is normal curvature along the viewing direction and k_s is the normal curvature of the extremal contour at P . This surface is called *the osculating quadric of S at P* .

The epipolar correspondents of P define two epipolar planes. The intersection of such plane with the osculating quadric is, up to order two, a parabola given by [8]:

$$g(x) = \frac{1}{2} \frac{k_t}{\cos \beta} x^2,$$

where β is the angle between the normal N to the surface at P and the projection of N in the epipolar plane.

In each epipolar plane at P , two tangents of the parabola are defined by viewing directions at P and its epipolar correspondent. By writing the tangency equations at these points, we get a linear equation linking the depth λ of P along the viewing direction and the curvature k_t :

$$f(\lambda, \frac{1}{k_t}) = 0$$

Since P has two epipolar correspondents, the tangency condition gives two equations and we finally ob-

tain a system of the form [8]:

$$\begin{cases} f_1(\lambda, \frac{1}{k_t}) = 0, \\ f_2(\lambda, \frac{1}{k_t}) = 0, \end{cases}$$

which is linear in $(\lambda, \frac{1}{k_t})$.

This allows the reconstruction of a 3D depth map. However, it appears that solutions are very noise sensitive in the vicinity of points where epipolar planes are tangent to the surface [8]. In the case of real image sequences, reconstruction is highly biased in the proximity of such points. This points out that the reconstruction stability should be controlled in order to eliminate highly corrupted points.

3. REGULARISATION

The result of the reconstruction is a set of 3D contours points. A parametric description of this set of points is required for a global surface representation. Since no parametrisation of the surface is *a priori* available we choose to triangulate these points. Starting with the reconstructed contour points $\{P_{i,j}\}$, where i denotes the contour and j the position on this contour, we construct a triangular mesh with respect to contour information:

- two 3D points can be connected if and only if they are on two consecutive 3D contours.

The resulting triangular mesh approximates the part of the surface which was covered by the observed extremal contours. However, this reconstructed surface might present perturbations such as wrinkles. This is due to different reasons including :

- the noise which is present in the acquisition system,
- camera calibration errors,
- contours tracking errors,

In order to correct these defaults, positions of mesh vertices are optimised by minimising a functional E :

$$E = E_{dist} + \alpha E_{reg},$$

where E_{dist} controls the fitness to the data and E_{reg} the smoothness of the reconstructed surface.

In the context of a reconstruction from occluding contours, data consists of image point positions. Thus the fidelity to the data can be characterised by the distance between the image point and the projection

of the corresponding mesh vertex $P_{i,j}$ onto the image plane. Hence:

$$E_{dist}(P_{i,j}) = \sum_{i,j} |M_i P_{i,j} - p_{i,j}|^2$$

where $\{M_i\}$ are the calibration matrix (i.e., perspective projection matrix) of the different image planes, and $\{p_{i,j}\}$ the image data points. This expression is consistent with the fact that original data are viewing directions and not 3D reconstructed points. In the optimisation procedure, Surface point displacements are therefore not limited to a closed neighbourhood of the reconstructed point, but to a closed neighbourhood of the corresponding viewing line.

In order to optimise surface point positions and thus, smooth the reconstructed surface we introduce a regularising energy. Classically, such energies are based on curvatures or, equivalently, second derivatives of surface point position function [9]. However, in the case of triangular mesh, second derivatives are not easy to compute. Furthermore, errors such as surface wrinkles may not be corrected by considering discrete curvatures. We therefore introduce a term which is based on the triangles area. Hence, the regularising energy is given by:

$$E_{reg}(T_{k \in 1, N_t}) = \sum_{k=1}^{N_t} S(T_k)^2,$$

where N_t is the number of triangles and $S(T_k)$ the surface of a triangle. This energy, and its derivatives, are easy to compute. Consequently, it can be minimised using a classical optimisation method.

Finally, the total energy can be written :

$$E = \sum_{i,j} |M_i P_{i,j} - p_{i,j}|^2 + \alpha \sum_{k=1}^{N_t} S(T_k)^2.$$

The parameter α controls the trade-off between fidelity to the data and smoothness of the the surface and should be set by the user.

4. EXPERIMENTAL RESULTS

We present here results for two real image sequences (see figure 1). They were taken with unknown camera motion. Therefore, a preliminary calibration step is performed using a calibration pattern which is present in every image. The occluding contours were tracked using *snakes* [10] (see figure 2). The reconstructed points are shown in figure 3. We used a conjugate gradient method to minimise the energy E , results of the regularisation are shown in figure 4.

5. CONCLUSION

We have described a reconstruction procedure that produces a smooth surface from image sequences. This makes it possible to recover complete or partial object surface description and thus, to achieve several tasks such as recognition, build object models, guidance. Results on real data demonstrate the reliability of the method on simple object surfaces. In future works, we plan to reconstruct more complex surfaces.

6. REFERENCES

- [1] R. Glachet, M. Dhome, and J.T. Lapreste. Finding the Pose of an Object of Revolution. In *Proceedings of Second European Conference on Computer Vision, Santa Margherita Ligure (Italy)*, pages 681–686, May 1992.
- [2] M. Zerroug and R. Nevatia. Quasi-Invariant Properties and 3-D Shape Recovery of Non-Straight, Non-Constant Generalized Cylinders. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, New York (USA)*, pages 96–103, 1993.
- [3] J.J. Koenderink. What Does the Occluding Contour Tell us About Solid Shape? *Perception*, 13:321–330, 1984.
- [4] R. Cipolla and A. Blake. The Dynamic Analysis of Apparent Contours. In IEEE, editor, *Proceedings of 3rd International Conference on Computer Vision, Osaka (Japan)*, pages 616–623, December 1990.
- [5] R. Vaillant and O. Faugeras. Using Extremal Boundaries for 3-D Object Modeling. *IEEE Transactions on PAMI*, 14(2):157–173, February 1992.
- [6] R. Szeliski and R. Weiss. Robust Shape Recovery from Occluding Contours Using a Linear Smoother. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, New York (USA)*, 1993.
- [7] C. Zhao and R. Mohr. Relative 3D Regularized B-Spline Surface Reconstruction Through Image Sequences. In *Proceedings of Third European Conference on Computer Vision (Stockholm, Sweden)*, volume 2, pages 417–426, May 1994. Lecture Notes in Computer Science 801.
- [8] E. Boyer and M.-O. Berger. 3D Surface Reconstruction Using Occluding Contours. Rapport interne 95-R-013, Centre de Recherche en Informatique de Nancy, Vandœuvre-lès-Nancy, January 1995.
- [9] T. Poggio, V. Torre, and C. Koch. Computational Vision and Regularization theory. *Nature*, pages 314–319, 1985.
- [10] M.-O. Berger. How to Track Efficiently Piecewise Curved Contours with a View to Reconstructing 3D Objects. In *Proceedings of the 12th International Conference on Pattern Recognition, Jerusalem (Israel)*, volume 1, pages 32–36, 1994.

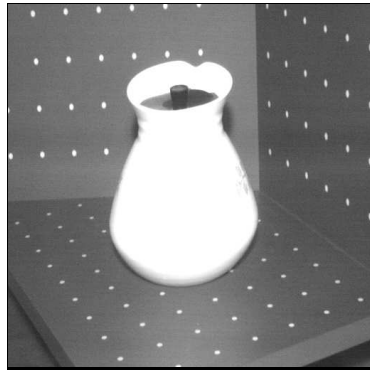


Figure 1: Image of the sequences

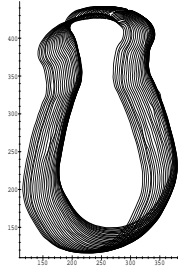
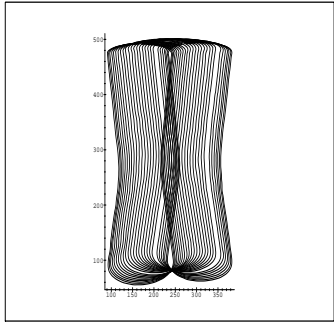


Figure 2: Tracked occluding contours

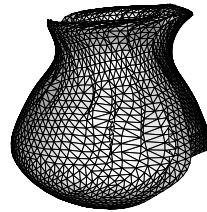
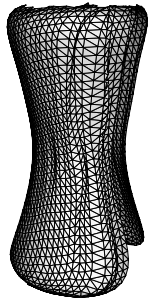


Figure 3: Reconstructed points



Figure 4: Regularized surfaces