

RECONSTRUCTION OF SCULPTURE FROM UNCALIBRATED IMAGE PROFILES

Roberto Cipolla and Kwan-Yee K. Wong

Department of Engineering, University of Cambridge,
Trumpington Street, Cambridge, CB2 1PZ, UK
[cipolla|kykw2]@eng.cam.ac.uk
<http://svr-www.eng.cam.ac.uk/~cipolla>

ABSTRACT

Profiles of a sculpture provide rich information about its geometry, and can be used for model reconstruction under known camera motion. By exploiting correspondences induced by epipolar tangents on the profiles, a successful solution to motion estimation has been developed for the case of circular motion. Arbitrary general views can then be incorporated to refine the model built from circular motion.

1. INTRODUCTION

Profiles (also known as outlines, or silhouettes) are often a dominant feature in images. They can be extracted relatively easily and reliably from the images, and provide rich information about both the shape and motion of an object. Classical techniques [1] for model reconstruction and motion estimation depend on point and/or line correspondences, and hence cannot be applied directly to profiles, which are viewpoint dependent. This calls for the development of a completely different set of algorithms specific to profiles. This paper will give a brief review of some of the state-of-art algorithms for model building and motion estimation from profiles.

2. PROFILES OF SURFACES

Profiles are projections of contour generators [2], which depend on both the surface geometry and camera positions. In general, 2 contour generators on a surface, associated with 2 different camera positions, will be 2 distinct space curves, and thus the corresponding profiles on the images do not readily provide point correspondences. A *frontier point* [2] is the intersection of 2 contour generators and lies on an epipolar plane tangent to the surface (see fig. 1). It follows that a frontier point will project to a point on the profile which is also on an epipolar tangent [3]. Epipolar tangencies thus provide point correspondences on profiles, and can be exploited for motion estimation.

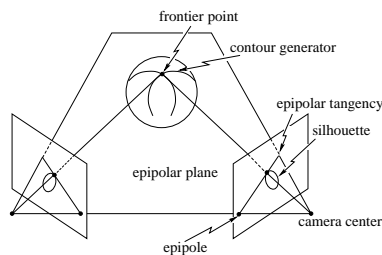


Fig. 1. A frontier point is the intersection of two contour generators and lies on an epipolar plane which is tangent to the surface. It follows that a frontier point will project to a point on the profile which is also on an epipolar tangent.

3. MODEL RECONSTRUCTION

The image profiles of an object provide rich information about its shape. Under known camera motion, it is possible to reconstruct a model of the object from its profiles. For continuous camera motion and simple smooth surfaces, a surface representation can be obtained from the profiles using the epipolar parameterization [4]. Cipolla and Blake [4] developed a simple numerical method for estimating depth from a minimum of 3 discrete views by determining the osculating circle on each epipolar plane. Vaillant and Faugeras [5] developed a similar algorithm which uses the radial plane instead of the epipolar plane. Boyer and Berger [6] derived a depth formulation from a local approximation of the surface up to order 2, which allows the local shape to be estimated from 3 consecutive views by solving a pair of simultaneous equations. In [7], Wong *et al* proposed to use simple triangulation for reconstruction from profiles, and showed that results from simple triangulation are comparable to those from Boyer and Berger's method when the camera motion is small.

Alternatively, for discrete motion and objects with more complex geometry, a volumetric model can be obtained by

an octree carving algorithm [8]. This technique is chosen in Section 4 for illustrating the results of reconstruction using the motion estimated from profiles, and thus is described in more details here. Initially, the octree consists of a single cube in space which encloses the model to be reconstructed. The cube is projected onto each images and classified as either (a) completely outside 1 or more profiles, (b) completely inside all the profiles, or (c) ambiguous. If the cube is classified as type (c), it is subdivided into 8 sub-cubes (hence the name octree) each of which is again projected onto the images and classified. This process is repeated until a preset maximum level (resolution) is reached. Cubes classified as type (a) are thrown away, leaving type (b) and (c) cubes which constitute the volumetric model of the object. Surface triangles, if needed, can be extracted from type (c) cubes using a marching cubes algorithm [9]. The octree carving technique is summarized in algorithm 1, and an octree carving software can be downloaded free at: <http://svr-www.eng.cam.ac.uk/research/vision>.

Algorithm 1 Octree Carving from Profiles

```

initialize a cube that enclose the model;
while max level not reached do
  for each cube in the current level do
    project the cube onto each image;
    classify the cube as either:
      (a) completely outside 1 or more profiles,
      (b) completely inside all the profiles, or
      (c) ambiguous;
    if the cube is classified as type (c) then
      subdivide the cube into 8 sub-cubes;
      add the sub-cubes to the next level;
    end if
  end for
  increase the level count;
end while

```

It is worth noting that both the surface and volumetric models, estimated only from the profiles of the object, correspond to the visual hull of the object with respect to the set of camera positions. Concavities cannot be recovered as they never appear as part of the profiles. In order to "carve" away the concavities, methods like space carving [10, 11] should be used instead.

4. MOTION ESTIMATION

A practical algorithm for motion estimation from profiles, in the case of complete circular motion, was introduced in [12]. In [13], the profiles from (incomplete) circular motion were exploited for the registration of any arbitrary general view, and a complete system for model acquisition from uncalibrated profiles under both circular and general motion

was developed. A summary of the techniques reported in [12, 13] is given below.

The 3 main image features in circular motion, namely the image of the rotation axis \mathbf{l}_s , the horizon \mathbf{l}_h and a special vanishing point \mathbf{v}_x (see [12] for details), are fixed throughout the sequence and satisfy

$$\mathbf{v}_x \cdot \mathbf{l}_h = 0, \text{ and} \quad (1)$$

$$\mathbf{v}_x = \mathbf{K}\mathbf{K}^T\mathbf{l}_s, \quad (2)$$

where \mathbf{K} is the 3×3 camera calibration matrix. The fundamental matrix can be parameterized explicitly in terms of these features [14, 12], and is given by

$$\mathbf{F} = [\mathbf{v}_x]_{\times} + \kappa \tan \frac{\theta}{2} (\mathbf{l}_s \mathbf{l}_h^T + \mathbf{l}_h \mathbf{l}_s^T), \quad (3)$$

where θ is the angle of rotation, and κ is a constant which can be determined from the camera intrinsic parameters. A sequence of N images taken under circular motion, with known camera intrinsic parameters, can hence be described by $N + 2$ motion parameters. By using the 2 outer epipolar tangents [13], the N images will provide $2N$ (or 2 when $N = 2$) independent constraints on these parameters, and a solution will be possible when $N \geq 3$.

The circular motion will generate a *web of contour generators* around the object, which can be exploited for registering any new arbitrary general view. Given an arbitrary general view, the associated contour generator will intersect with this web and form frontier points. If the camera intrinsic parameters are known, the 6 motion parameters of the new view can be fixed if there are 6 or more frontier points on the associated contour generator. This corresponds to having a minimum of 3 views under circular motion, each providing 2 outer epipolar tangents to the profile in the new general view (see fig. 2).

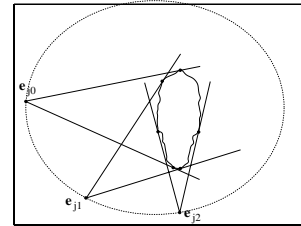


Fig. 2. Three views from circular motion provide 6 outer epipolar tangents to the profile in the new general view for estimating its pose.

The motion estimation proceeds as an optimization which minimizes the reprojection errors of epipolar tangents. For view i and view j , a fundamental matrix \mathbf{F}_{ij} is formed from the current estimate of the motion parameters, and the epipoles \mathbf{e}_{ij} and \mathbf{e}_{ji} are obtained from the right

and left nullspaces of \mathbf{F}_{ij} . The outer epipolar tangent points \mathbf{t}_{ij0} , \mathbf{t}_{ij1} and \mathbf{t}_{ji0} , \mathbf{t}_{ji1} are located in view i and j respectively (see fig. 3). The reprojection errors are then given by the geometric distances between the epipolar tangent points and their epipolar lines,

$$d_{ijk} = \frac{\mathbf{t}_{jik}^T \mathbf{F}_{ij} \mathbf{t}_{ijk}}{\sqrt{(\mathbf{F}_{ij}^T \mathbf{t}_{jik})_1^2 + (\mathbf{F}_{ij}^T \mathbf{t}_{jik})_2^2}}, \quad (4)$$

$$d_{jik} = \frac{\mathbf{t}_{jik}^T \mathbf{F}_{ij} \mathbf{t}_{ijk}}{\sqrt{(\mathbf{F}_{ij}^T \mathbf{t}_{ijk})_1^2 + (\mathbf{F}_{ij}^T \mathbf{t}_{ijk})_2^2}}. \quad (5)$$

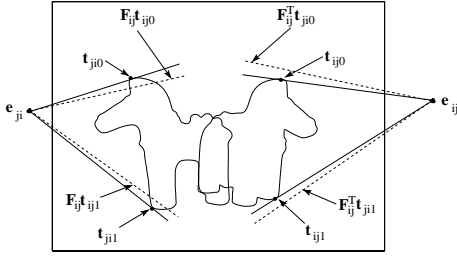


Fig. 3. The motion parameters can be estimated by minimizing the reprojection errors of epipolar tangents, which are given by the geometric distances between the epipolar tangent points and their epipolar lines.

For a sequence of N images taken under circular motion, the image of the rotation axis and the horizon are initialized approximately, and the angles are arbitrarily initialized. The cost function is given by

$$C_{cm}(\mathbf{x}) = \frac{1}{4} \sum_{i=1}^N \sum_{j=i+1}^{\min(i+3, N)} \sum_{k=1}^2 d_{ijk}(\mathbf{x})^2 + d_{jik}(\mathbf{x})^2, \quad (6)$$

where \mathbf{x} consists of the $N + 2$ motion parameters.

For arbitrary general motion, the 6 motion parameters can be initialized by roughly aligning the projection of the 3D model built from the estimated circular motion with the image. The cost function of general motion for view j is given by

$$C_j(\mathbf{x}') = \sqrt{\frac{\sum_{i=1}^N f_{ij} \sum_{k=1}^2 d_{ijk}(\mathbf{x}')^2 + d_{jik}(\mathbf{x}')^2}{4 \sum_{i=1}^N f_{ij}}}, \quad (7)$$

where \mathbf{x}' consists of the 6 motion parameters. f_{ij} is 0 if the baseline formed with view i passes through the object, otherwise it is 1.

The motion estimation procedure is summarized in algorithm 2. Fig. 5 and fig. 6 show some examples of reconstruction from the motion estimated using profiles (see fig. 4 for the image sequences used).

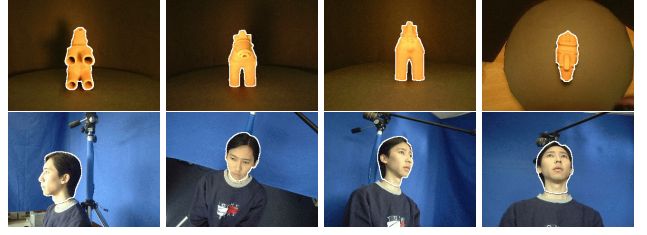


Fig. 4. Top: 4 images from an uncalibrated sequence of a haniwa. Bottom: 4 images from an uncalibrated sequence of a human head.



Fig. 5. Different views of the model reconstructed from the haniwa sequence using the motion estimated from the profiles.



Fig. 6. Different views of the mode reconstructed from the head sequence using the motion estimated from the profiles.

Algorithm 2 Motion Estimation from Profiles

extract the profiles using cubic B-spline snakes [4] ;
initialize I_s , I_h and the $N - 1$ angles for the circular motion;
while not converged **do**
 compute the cost for circular motion using (6);
 update the $N + 2$ motion parameters to minimize the cost;
end while
Form the essential matrices from the fundamental matrices using the calibration matrix;
Decompose the essential matrices to obtain the projection matrices;
Build a partial model using an octree carving algorithm;
for each arbitrary general view **do**
 initialize the 6 motion parameters using the partial model from circular motion;
 while not converged **do**
 compute the cost for the general motion using (7);
 update the 6 motion parameters to minimize the cost;
 end while
end for
Refine the model using the now calibrated images from general motion.

5. CONCLUSIONS

The incorporation of arbitrary general views reveals information which is concealed under circular motion, and greatly improves both the shape and textures of the 3D models. Since only profiles have been used in both the motion estimation and octree carving, no corner detection nor matching is necessary. This means that the algorithm is capable of reconstructing any kind of objects, including *smooth* and *textureless* surfaces. Experiments on various objects have produced convincing 3D models, demonstrating the practicality of the algorithm.

6. REFERENCES

- [1] O. D. Faugeras, *Three-Dimensional Computer Vision: a Geometric Viewpoint*. Cambridge, MA: MIT Press, 1993.
- [2] R. Cipolla and P. J. Giblin, *Visual Motion of Curves and Surfaces*. Cambridge, UK: Cambridge University Press, 1999.
- [3] J. Porrill and S. B. Pollard, "Curve matching and stereo calibration," *Image and Vision Computing*, vol. 9, no. 1, pp. 45–50, February 1991.
- [4] R. Cipolla and A. Blake, "Surface shape from the deformation of apparent contours," *Int. Journal of Computer Vision*, vol. 9, no. 2, pp. 83–112, November 1992.
- [5] R. Vaillant and O. D. Faugeras, "Using extremal boundaries for 3D object modeling," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 157–173, February 1992.
- [6] E. Boyer and M. O. Berger, "3d surface reconstruction using occluding contours," *Int. Journal of Computer Vision*, vol. 22, no. 3, pp. 219–233, March 1997.
- [7] K.-Y. K. Wong, P. R. S. Mendonça, and R. Cipolla, "Reconstruction and motion estimation from apparent contours under circular motion," in *Proc. British Machine Vision Conference*, T. Pridmore and D. Elliman, Eds., vol. 1, Nottingham, UK, September 1999, pp. 83–92.
- [8] R. Szeliski, "Rapid octree construction from image sequences," *Computer Vision, Graphics and Image Processing*, vol. 58, no. 1, pp. 23–32, July 1993.
- [9] W. E. Lorensen and H. E. Cline, "Marching cubes: a high resolution 3D surface construction algorithm," *ACM Computer Graphics*, vol. 21, no. 4, pp. 163–169, July 1987.
- [10] K. N. Kutulakos and S. M. Seitz, "A theory of shape by space carving," *Int. Journal of Computer Vision*, vol. 38, no. 3, pp. 197–216, July 2000.
- [11] A. Broadhurst, T. W. Drummond, and R. Cipolla, "A probabilistic framework for space carving," in *Proc. 8th Int. Conf. on Computer Vision*, vol. I, Vancouver, BC, Canada, July 2001, pp. 388–393.
- [12] P. R. S. Mendonça, K.-Y. K. Wong, and R. Cipolla, "Epipolar geometry from profiles under circular motion," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 604–616, June 2001.
- [13] K.-Y. K. Wong and R. Cipolla, "Structure and motion from silhouettes," in *Proc. 8th Int. Conf. on Computer Vision*, vol. II, Vancouver, BC, Canada, July 2001, pp. 217–222.
- [14] T. Vieville and D. Lingrand, "Using singular displacements for uncalibrated monocular visual systems," in *Proc. 4th European Conf. on Computer Vision*, ser. Lecture Notes in Computer Science, B. Buxton and R. Cipolla, Eds., vol. 1065. Cambridge, UK: Springer-Verlag, April 1996, pp. 207–216.