Reconstruction of Outdoor Sculptures from Silhouettes under Approximate Circular Motion of an Uncalibrated Hand-Held Camera

Kwan-Yee Kenneth Wong*
Dept. of Comp. Science & Info. Systems
The University of Hong Kong

Roberto Cipolla†
Department of Engineering
University of Cambridge

Abstract

This paper presents a novel technique for reconstructing an outdoor sculpture from an uncalibrated image sequence acquired around it using a hand-held camera. The technique introduced here uses only the silhouettes of the sculpture for both motion estimation and model reconstruction, and no corner detection nor matching is necessary. This is very important as most sculptures are composed of smooth textureless surfaces, and hence their silhouettes are very often the only information available from their images. Besides, as opposed to previous works, the technique here does not require the camera motion to be perfectly circular (e.g., turntable sequence). It employs an image rectification step before the circular motion estimation to obtain a rough estimate of the camera motion which is only *approximately* circular. A refinement process is then applied to obtain the true general motion of the camera. This allows the technique to handle large outdoor sculptures which cannot be rotated on a turntable, making it much more practical and flexible.

1 Introduction

This paper addresses the problem of reconstructing outdoor sculptures from uncalibrated views using their silhouettes [6]. Silhouettes (also known as profiles or outlines) are often a dominant image feature, and can be extracted relatively easily and reliably from the images. They provide rich information about both the shape and motion of an object, and are indeed the only information available in the case of smooth textureless surfaces (e.g., sculptures). Nonetheless, silhouettes are projections of contour generators [2] which are viewpoint dependent, and hence they do not readily provide point correspondences for the computation of the epipolar geometry [16]. As a result, *structure and motion from silhouettes* has always been a challenging problem.

One possible approach to the above problem is to make use of the point correspondences induced by frontier points [5, 3]. A frontier point is given by the intersection of two contour generators from two distinct viewpoints, and is thus visible in both images. It lies on an epipolar plane tangent to the surface, and hence it will be projected onto a point in the silhouette which is also on an epipolar tangent [10, 1]. In previous work [8], we have exploited epipolar tangents to locate point correspondences in the silhouettes, and have developed a practical solution to the problem of structure and motion from silhouettes in the special case of *circular motion*.

In this paper, we will show how an image rectification step can be applied before the circular motion estimation to obtain a rough estimate of the camera motion which is only *approximately* circular. An iterative refinement process, which minimizes the reprojection errors of the epipolar tangents [13], can then be applied to obtain the true general motion of the camera. This allows the technique introduced here to handle large outdoor sculptures which cannot be rotated on a turntable, making it much more practical and flexible.

This paper is organized as follows. Section 2 gives a brief review on circular motion estimation from silhouettes. Section 3 first describes the simple procedure for acquiring and rectifying an approximate circular motion sequence around an outdoor sculpture. It then gives the algorithms for estimating the camera motion from the rectified sequence. Section 4 shows the experimental results of reconstructing an outdoor horse sculpture. Finally, conclusions are given in Section 5.

2 Circular Motion

Consider a pinhole camera rotating about a fixed axis. Let $\mathbf{v}_x$ be the vanishing point corresponding to the normal direction $\mathbf{N}_x$ of the plane Π_s that contains the axis of rotation and the camera center, and $\mathbf{l}_\mathrm{h}$ be the *horizon* which is the image of the plane Π_h that contains the trajectory of the camera center. By definition, the epipoles are the projections of the camera center and must therefore lie on $\mathbf{l}_\mathrm{h}$. Besides, since $\mathbf{N}_x$ is parallel to the plane Π_h, it follows that $\mathbf{v}_x$ also lies on $\mathbf{l}_\mathrm{h}$. The plane Π_s will be projected onto the image plane as a line $\mathbf{l}_\mathrm{s}$, which is also the image of the rotation axis. Note that $\mathbf{l}_\mathrm{s}$ is related to $\mathbf{v}_x$ via the camera calibration matrix $\mathbf{K}$, and is given by [15]

$$\mathbf{v}_x = \mathbf{K}\mathbf{K}^\mathrm{T}\mathbf{l}_\mathrm{s}. \tag{1}$$

If the intrinsic parameters of the camera are assumed to be fixed, due to symmetry in the configuration, $\mathbf{l}_\mathrm{s}$, $\mathbf{l}_\mathrm{h}$ and $\mathbf{v}_x$ will be fixed throughout the image sequence (see fig. 1). The fundamental matrix $\mathbf{F}_{ij}$ associated with any pair of views ij in the circular motion sequence can be parameterized explicitly in terms of these fixed features, and is given by [12, 4]

$$\mathbf{F}_{ij} = [\mathbf{v}_x]_\times + (\det \mathbf{K}) \tan \frac{\theta_{ij}}{2} (\mathbf{l}_\mathrm{s}\mathbf{l}_\mathrm{h}^\mathrm{T} + \mathbf{l}_\mathrm{h}\mathbf{l}_\mathrm{s}^\mathrm{T}), \tag{2}$$

where θ_{ij} is the rotation angle between view i and view j.

*Address: Chow Yei Ching Building, Pokfulam Road, Hong Kong. Email: kykwong@csis.hku.hk

†Address: Trumpington Street, Cambridge CB2 1PZ, United Kingdom. E-mail: cipolla@eng.cam.ac.uk

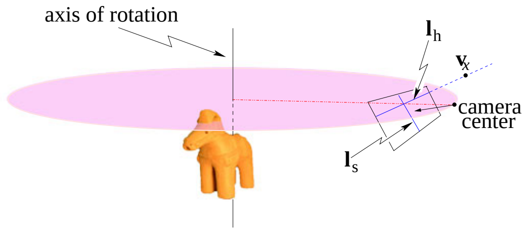

Figure 1: Under circular motion and fixed intrinsic parameters of the camera, the image of the rotation axis $\mathbf{l}_\mathrm{s}$, the horizon $\mathbf{l}_\mathrm{h}$ and the vanishing point $\mathbf{v}_x$ will be fixed throughout the sequence. The fundamental matrix relating any pair of views in the sequence can be parameterized explicitly in terms of these fixed features.

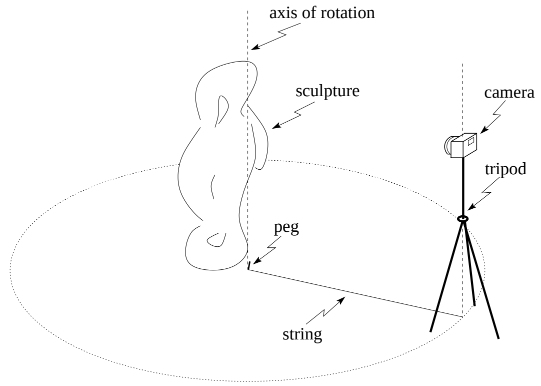

Figure 2: For an outdoor sculpture, an approximate circular motion of the camera can be achieved by using a string, a peg and a tripod.

Such a parameterization greatly reduces the dimension of the search space for the motion estimation problem, which can now be solved by minimizing the reprojection errors of the two outer epipolar tangents [13]. Besides, it also leads to a trivial initialization as all the parameters bear physical meaning. Using such a parameterization, we have successfully implemented an user-friendly software for building 3D models from uncalibrated turntable sequences using silhouettes alone [8]. Such a software is extremely useful for reconstructing 3D models of smooth textureless objects that are small enough to be rotated on a turntable. However for larger objects like outdoor sculptures, it is not always possible to rotate the object on a turntable so as to constrain the camera motion to be *perfectly circular*. As a result, modelling of outdoor sculptures is not as straightforward as the indoor turntable sequence case.

3 Approximate Circular Motion

For an outdoor sculpture, an *approximate circular motion* of the camera can be achieved by using a string, a peg and a tripod. First, one end of the string is fixed to the ground by the peg, and this point will serve as the center of rotation. Next, a circular path on the ground can then be traced out by rotating the free end of the string about its fixed end. With the help of the tripod (optional), images can then be acquired by positioning the camera roughly at a fixed height above the free end of the rotating string and by pointing it towards the sculpture (see fig. 2).

Note that since the camera center, the string and the axis of rotation are roughly coplanar, the image of the string in each image will provide a very good estimate for the image of the rotation axis $\mathbf{l}_\mathrm{s}$ (see fig. 3). Despite the fact that the camera center roughly follows a circular path, the orientation of the camera is, however, unconstrained and hence the image of the rotation axis $\mathbf{l}_\mathrm{s}$ and the horizon $\mathbf{l}_\mathrm{h}$ will not be fixed throughout the image sequence (see fig. 4).

In order to allow the camera motion to be estimated using the circular motion algorithm, each image is first rectified by a planar homography induced by a rotation that brings the image of the string (i.e., the image of the rotation axis) into a vertical line $\mathbf{l}_\mathrm{f}$ passing through the principal point. This corresponds to rotating each camera about its optical center until the rotation axis of the circular motion lies on the y-z plane of the camera coordinate system [14]. For each such rectified image, a transformation induced by a rotation about the x-axis of the camera coordinate system is then computed and applied to bring the image of the fixed end of the string to a fixed point on $\mathbf{l}_\mathrm{f}$ throughout the rectified sequence. The resulting image sequence will then resemble a circular motion sequence, in which the image of the rotation axis $\mathbf{l}_\mathrm{s}$, the horizon $\mathbf{l}_\mathrm{h}$ and the special vanishing point $\mathbf{v}_x$ are fixed throughout the sequence (see fig. 5).

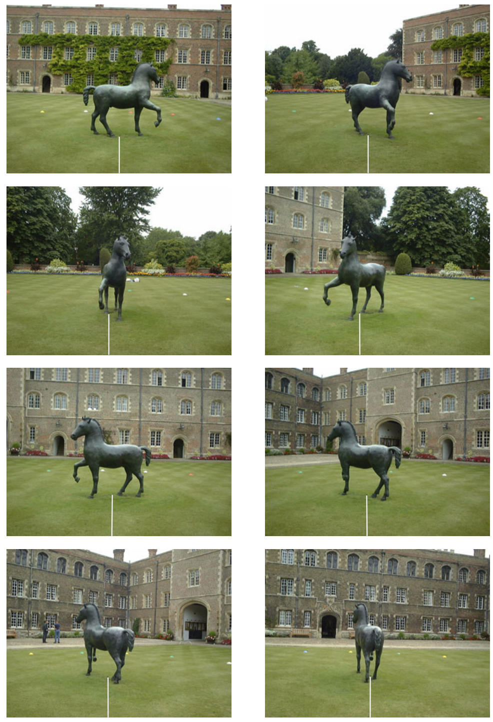

Figure 3: An image sequence of an outdoor sculpture acquired under *approximate* circular motion of a hand-held camera. The image of the string in each image provides a very good estimate for the image of the rotation axis $\mathbf{l}_\mathrm{s}$.

The algorithm for circular motion estimation [8] can then be applied to this rectified sequence, with the motion parameters initialized as follows. The image of the rotation axis is initialized to the line $\mathbf{l}_\mathrm{f}$, and the vanishing point $\mathbf{v}_x$ is then computed using equation (1). A point $\mathbf{x}_\mathrm{h}$ along the line $\mathbf{l}_\mathrm{f}$ at roughly the same height as the camera in 3D is picked, and an initial estimate for the horizon $\mathbf{l}_\mathrm{h}$ is then given by $\mathbf{v}_x \times \mathbf{x}_\mathrm{h}$. Finally, the angles of rotation are arbitrarily initialized. After the optimization, which minimizes the reprojection errors of the epipolar tangents, a rough estimate of

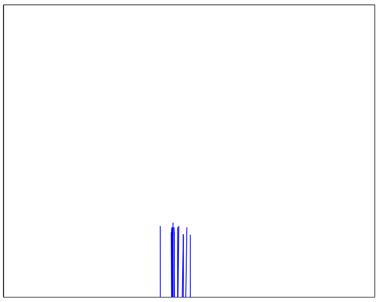

Figure 4: The images of the string in the outdoor sequence do not coincide, and this implies that the image of the rotation axis is not fixed throughout the sequence.

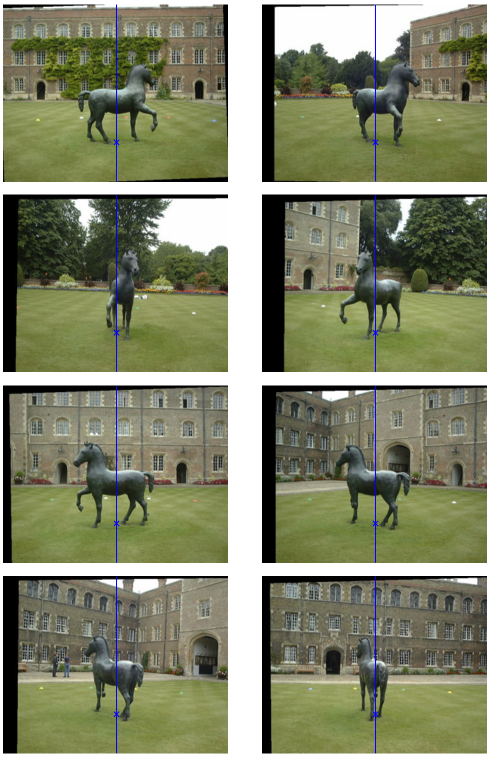

Figure 5: The rectified sequence resembles a circular motion sequence in which the image of the rotation axis $\mathbf{l}_{\mathrm{s}}$ (plotted as a solid line), the horizon $\mathbf{l}_{\mathrm{h}}$ and the special vanishing point $\mathbf{v}_x$ (plotted as a cross) are fixed throughout the sequence.

the camera poses can be obtained.

In [13], we have introduced an algorithm for registering a silhouette under arbitrary general motion with a set of silhouettes under *known* or *estimated* motion. Here we employ the same algorithm for iteratively refining the approximate circular motion. Each camera pose obtained from the circular motion estimation is refined in turn by minimizing the reprojection errors of the two outer epipolar tangents resulting from pairing it with each of the other views in the sequence. Note that now the camera poses are no longer constrained to a circular motion, and the motion parameters to be refined consist of both the independent rotation and translation. This refinement process is repeated until no further improvement can be made. Finally, a volumetric model can be reconstructed from the silhouettes and the estimated motion using an octree carving algorithm [11], and a triangulated mesh of the reconstructed model can be extracted by the marching cubes algorithm [7, 9].

4 Experimental Results

The experimental sequence consists of 14 images of the horse sculpture located at the First Court of Jesus College in Cambridge, UK. The image sequence was acquired using the simple setup as described in Section 3, and fig. 3 shows the 2nd, 4th, 6th, 8th, 10th, 12th, 13th and 14th images in the sequence. The image of the string in each image was picked manually, and the whole sequence was then rectified so that all the images of the string in the rectified sequence became coincident (see fig. 5). The circular motion estimation algorithm was then applied to this rectified sequence, followed by the iterative refinement process. The final camera configuration estimated from the rectified sequence is shown in fig. 6. Using the estimated motion, a volumetric model of the horse was built using the octree carving algorithm. Figure 7 shows the triangulated mesh extracted from the octree representation using the marching cubes algorithm, and fig. 8 shows different novel views of the reconstructed sculpture model with texture-mapping, demonstrating the quality of both the motion estimated and the model reconstructed.

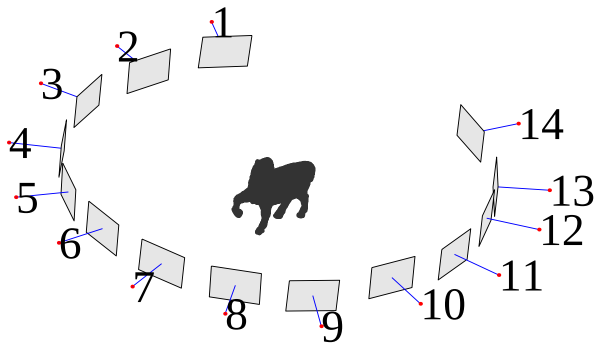

Figure 6: Camera poses estimated from the rectified sequence.

5 Conclusions

In this paper, a novel technique for reconstructing an outdoor sculpture from an uncalibrated image sequence is introduced. For an outdoor sculpture, an *approximate* circular motion of the camera around the sculpture can be achieved by using a simple setup consisting of a string, a peg and a tripod. The image of the string in each image provides a very good estimate for the image of the rotation axis, and can be exploited for rectifying the image sequence into one resembling a circular motion sequence in which the image of the rotation axis, the horizon and the special vanishing point are fixed throughout the sequence. This allows a rough estimate of the camera poses to be obtained form the rectified sequence using a circular motion estimation algorithm [8]. An iterative refinement process [13] can then be applied to obtain the true general motion of the camera. The technique introduced here uses only the silhouettes of the sculptures for both motion estimation and model reconstruction, and does not require the camera motion to be *perfectly* circular. This allows the technique to handle large

Figure 7: Triangulated mesh of the 3D model built from the silhouettes and the estimated motion.

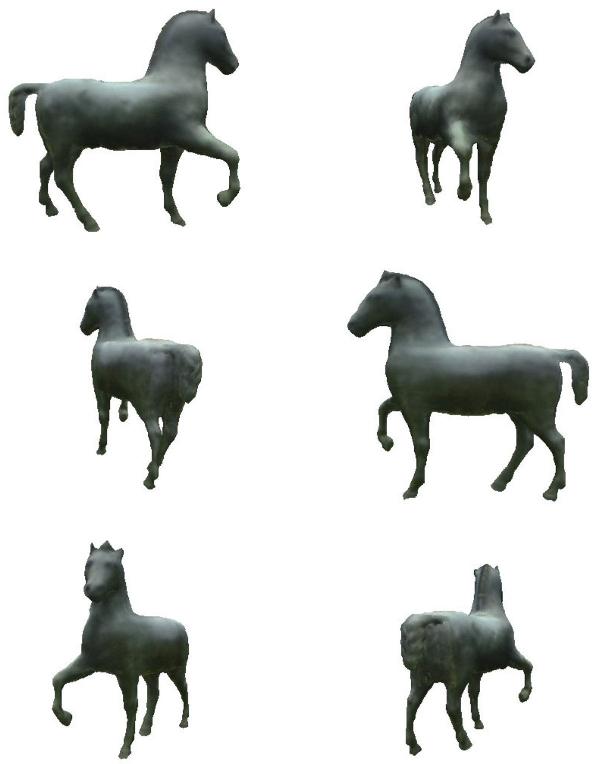

Figure 8: Different views of the reconstructed sculpture model with texture-mapping.

outdoor sculptures which cannot be rotated on a turntable, making it much more practical and flexible. Experimental results on a real outdoor sculpture are presented, which demonstrate the feasibility and practicality of the proposed technique.

References

[1] R. Cipolla, K. E. Åström, and P. J. Giblin. Motion from the frontier of curved surfaces. In *Proc. 5th Int. Conf. on Computer Vision*, pages 269–275, Cambridge, MA, USA, June 1995.

[2] R. Cipolla and A. Blake. Surface shape from the deformation of apparent contours. *Int. Journal of Computer Vision*, 9(2):83–112, November 1992.

[3] R. Cipolla and P. J. Giblin. *Visual Motion of Curves and Surfaces*. Cambridge University Press, Cambridge, UK, 1999.

[4] A. W. Fitzgibbon, G. Cross, and A. Zisserman. Automatic 3D model construction for turn-table sequences. In R. Koch and L. Van Gool, editors, *3D Structure from Multiple Images of Large-Scale Environments, European Workshop SMILE'98*, volume 1506 of *Lecture Notes in Computer Science*, pages 155–170, Freiburg, Germany, June 1998. Springer–Verlag.

[5] P. J. Giblin, F. E. Pollick, and J. E. Rycroft. Recovery of an unknown axis of rotation from the profiles of a rotating surface. *Journal of Optical Soc. of America A*, 11(7):1976–1984, July 1994.

[6] J. J. Koenderink. What does the occluding contour tell us about solid shape? *Perception*, 13:321–330, 1984.

[7] W. E. Lorensen and H. E. Cline. Marching cubes: a high resolution 3D surface construction algorithm. *ACM Computer Graphics*, 21(4):163–169, July 1987.

[8] 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*, 23(6):604–616, June 2001.

[9] C. Montani, R. Scateni, and R. Scopigno. A modified look-up table for implicit disambiguation of marching cubes. *The Visual Computer*, 10(6):353–355, 1994.

[10] J. Porrill and S. B. Pollard. Curve matching and stereo calibration. *Image and Vision Computing*, 9(1):45–50, February 1991.

[11] R. Szeliski. Rapid octree construction from image sequences. *Computer Vision, Graphics and Image Processing*, 58(1):23–32, July 1993.

[12] T. Vieville and D. Lingrand. Using singular displacements for uncalibrated monocular visual systems. In B. Buxton and R. Cipolla, editors, *Proc. 4th European Conf. on Computer Vision*, volume 1065 of *Lecture Notes in Computer Science*, pages 207–216, Cambridge, UK, April 1996. Springer–Verlag.

[13] K.-Y. K. Wong and R. Cipolla. Structure and motion from silhouettes. In *Proc. 8th Int. Conf. on Computer Vision*, volume II, pages 217–222, Vancouver, BC, Canada, July 2001.

[14] K.-Y. K. Wong, P. R. S. Mendonça, and R. Cipolla. Reconstruction of surfaces of revolution from single uncalibrated views. In P. L. Rosin and D. Marshall, editors, *Proc. British Machine Vision Conference 2002*, volume 1, pages 93–102, Cardiff, UK, September 2002. British Machine Vision Association.

[15] K.-Y. K. Wong, P. R. S. Mendonça, and R. Cipolla. Camera calibration from surfaces of revolution. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, to appear.

[16] Z. Zhang. Determining the epipolar geometry and its uncertainty: A review. *Int. Journal of Computer Vision*, 27(2):161–195, March 1998.