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

Kwan-Yee Kenneth WONG† and Roberto CIPOLLA††, Nonmembers

SUMMARY 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 proposed technique does not require the camera motion to be perfectly circular (e.g., turntable sequence). It employs an image rectification step before the motion estimation step 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.

key words: silhouette, approximate circular motion, motion estimation, model reconstruction

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 [18]. 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 [3], [5]. A frontier point is given by the intersection of two contour generators from two distinct viewpoints, and is thus visible in both images (see Fig. 1). 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 [1], [11]. In previous work [9], 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 motion estimation step to provide 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 [14], 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 a 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

\*The author is with the Department of Computer Sc. & Information Systems, University of Hong Kong, HK.
\††The author is with the Department of Engineering, University of Cambridge, UK.
\*This paper was presented at IAPR Workshop on Machine Vision Applications (MVA2002) [15].
scenture. Finally, conclusions are given in Section 5.

2. Circular Motion

Consider a pinhole camera rotating about a fixed axis. Let \( v_x \) be the vanishing point corresponding to the normal direction \( N_x \) of the plane \( \Pi_h \) that contains the axis of rotation and the camera center, and \( l_h \) be the horizon which is the image of the plane \( \Pi_h \) that contains the trajectory of the camera center (see Fig. 2). By definition, the epipoles are the projections of the camera center and must therefore lie on \( l_h \). Besides, since \( N_x \) is parallel to the plane \( \Pi_h \), it follows that \( v_x \) also lies on \( l_h \), i.e.,

\[
 v_x \cdot l_h = 0. \tag{1}
\]

The plane \( \Pi_s \) will be projected onto the image plane as a line \( l_s \), which is also the image of the rotation axis. In [17], it has been shown that there exists a pole-polar relationship, with respect to the absolute conic, between the vanishing point \( v_x \) and the vanishing line \( l_s \) of \( \Pi_s \), i.e.,

\[
 \omega \cdot v_x = l_s, \tag{2}
\]

where \( \omega = K^{-T}K^{-1} \) is the image of the absolute conic and \( K \) is the \( 3 \times 3 \) camera calibration matrix.

The algorithm for estimating circular motion from silhouettes is summarized in Algorithm 1. Note that the above parameterization also leads to a trivial initialization as all the parameters bear physical meaning. Using such an algorithm, we have successfully implemented an user-friendly software for building 3D models from uncalibrated turntable sequences using silhouettes alone [9]. 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
Algorithm 1 Circular motion estimation.

extract the silhouettes using cubic B-spline snakes;
initialize \( l_s \), \( l_h \) and the \( N-1 \) rotation angles;
while not converged do
  for each view in the sequence do
    form the fundamental matrices with the next 2 views;
    locate the outer epipolar tangents in each view pair;
    determine the reprojection errors of the epipolar tangents;
  end for
  compute the cost as \( \sum (\text{reprojection errors})^2 \);
  update the \( N+2 \) motion parameters to minimize the cost
  using conjugate gradient method;
end while
form fundamental matrices from the estimated parameters;
convert fundamental matrices to essential matrices using the
camera calibration matrix;
decompose essential matrices and form projection matrices;

(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. 4).

![Fig. 4](image.png)

For an outdoor sculpture, an approximate circular mo-
tion of the camera can be achieved by using a string, a peg and
and a tripod.

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 esti-
mate for the image of the rotation axis \( l_s \) (see Fig. 5).
Despite the fact that the camera center roughly follows
a circular path, the orientation of the camera is, how-
ever, unconstrained and hence the image of the rotation
axis \( l_s \) and the horizon \( l_h \) will not be fixed throughout
the image sequence (see Fig. 6).

In order to apply the above circular motion al-
gorithm for estimating an approximate circular mo-
tion, each image in the sequence must be rectified in-
dependently to make the image of the rotation axis \( l_s \), the horizon \( l_h \) and the special vanishing point \( v_x \)
fixed throughout the resulting sequence. This can be
achieved by transforming each image by a planar ho-
mography induced by two rotations. The first rotation
rotates the camera about its optical center until the ro-
tation axis of the circular motion lies on the \( y-z \) plane
of the camera coordinate system\(^1\). This leaves only

\(^1\)Here we assume a right-hand coordinate system, where
the optical center is at the origin, the \( x \)-axis and the \( y \)-axis
point right and down, respectively, and the \( z \)-axis is the
viewing direction.
1 degree of freedom in the orientation of the camera which is a rotation about its $x$-axis. The second rotation then fixes this degree of freedom by rotating the camera about its $x$-axis such that the image of the fixed end of the string becomes a fixed point throughout the rectified sequence.

To compute a planar homography for rectification, an image is first normalized by $K^{-1}$ to remove the effects of the intrinsic parameters of the camera, and the image of the rotation axis becomes $I_n^0 = KR_0 l_0$. The normalized image is then transformed by a rotation matrix $R$ that brings $x_0$, a point (in homogeneous coordinates) along $I_n^0$ which is closest to the image origin $x_0 = [0 0 1]^T$, to $x_0$. The axis and angle of the rotation $R$ are given by $n_l = \frac{x_b - x_0}{|x_b - x_0|}$ and $\phi_l = \arccos(\frac{x_b^T s_0}{|x_b||s_0|})$, respectively. This transformation corresponds to rotating the camera until it points directly towards the rotation axis of the circular motion [16]. The image of the rotation axis now becomes $I_n^1 = RL_n^0 = [\cos \theta \quad \sin \theta \quad 0]$, which is a line passing through the image origin (i.e., the principal point$^1$ in the normalized image). The resulting image is then rotated about the image origin until the image of the rotation axis aligns with the $y$-axis, and the transformation is given by a rotation matrix $R_x$, which is a rotation about the $z$-axis by an angle $-\theta$. This corresponds to rotating the camera about its $z$-axis until the rotation axis of the circular motion lies on its $y$-$z$ plane. Finally, the image is transformed by a rotation matrix $R_z$ that brings $x_f$, the image of the fixed point of the string, to $x_h$, which is the image of the fixed point in the first rectified image. The rotation matrix $R_z$ is a rotation about the $x$-axis by an angle $\phi_x = \arccos(\frac{x_h^T s_0}{|x_h||s_0|})$. To complete the transformation, the camera calibration matrix is re-applied and the overall transformation (i.e., the planar homography) is therefore given by $H = KR_x R_z R K^{-1}$.

After the rectification, the resulting image sequence will then resemble a circular motion sequence, in which the image of the rotation axis $l_0$ (plotted as a solid line), the horizon $l_h$, and the special vanishing point $v_z$ are fixed throughout the sequence (see Fig. 7).

The algorithm for circular motion estimation [9] 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 $l_1$ which is the image of the string in the rectified sequence, and the vanishing point $v_z$ is then computed using Eq. (2). A point $x_0$ along the line $l_1$ at roughly the same height as the camera in 3D is picked, and an initial estimate for the horizon $l_h$ is then given by $v_z \times x_0$. Finally, the rotation angles are arbitrarily initialized. After the optimization, which minimizes the reprojection errors of the epipolar tangents, a rough estimate of the camera poses can be obtained (see Algorithm 1).

In [14], 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 (see Algorithm 2). 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 (3 degrees of freedom) and translation (3 degrees of freedom). 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 [12], and a triangulated mesh of the reconstructed model can be extracted by the marching cubes.

$^1$The principal point is the point at which the optical axis pierces the image plane, and hence it indicates the pointing direction of the camera.
Algorithm 2 Iterative refinement of camera poses.

\begin{algorithm}
\begin{algorithmic}
\While {improved}
\For {each view in the sequence}
\While {not converged}
\State form fundamental matrices with all other views;
\State locate the outer epipolar tangents in each view pair;
\State determine the reprojection errors of the epipolar tangents;
\State compute the cost as $\sum (\text{reprojection errors})^2$;
\State update the 6 motion parameters (3 for rotation and 3 for translation) to minimize the cost using conjugate gradient method;
\EndWhile
\EndFor
\EndWhile
\end{algorithmic}
\end{algorithm}

4. Experimental Results

The experimental sequence consists of 14 images of a 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 Sect. 3, and Fig. 5 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. 7). 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. 8. Using the estimated motion, an octree representation of the horse was built using a space carving algorithm (see Fig. 10) and a surface representation was then extracted using the marching cubes algorithm. Figure 11 shows different novel views of the reconstructed sculpture model with texture-mapping, and Fig. 9 shows the underlying triangulated mesh, which demonstrates the quality of both the motion estimated and the model reconstructed.

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 from the rectified sequence using a circular motion estimation algorithm [9]. An iterative refinement process [14] 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 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.

Acknowledgement

The work described in this paper was partially sup-
Fig. 10  A volumetric model of the horse was built by an octree carving algorithm using
the silhouettes and the estimated camera poses.

Fig. 11  Different views of the reconstructed sculpture model with texture-mapping.

ported by a grant from the Research Grants Council of
the Hong Kong Special Administrative Region, China
(Project No. HKU 7155/03E).

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 deforma-
tion of apparent contours. Int. Journal of Computer Vision,

Surfaces. Cambridge University Press, Cambridge, UK,
1999.

matic 3D model construction for turn-table sequences. In
R. Koch and L. Van Gool, editors, 3D Structure from Mul-
tiple Images of Large-Scale Environments, European Work-
shop SMILE’98, volume 1506 of Lecture Notes in Computer
Springer–Verlag.

an unknown axis of rotation from the profiles of a rotating


resolution 3D surface construction algorithm. ACM Com-

[8] Q. T. Luong and O. D. Faugeras. The fundamental matrix:


