Self-Calibration of Turntable Sequences from Silhouettes

Hui Zhang, *Member, IEEE,* and Kwan-Yee K. Wong, *Member, IEEE*

The authors are with the Department of Computer Science, The University of Hong Kong, Pokfulam Road, Hong Kong. Email: {hzhang,kykwong}@cs.hku.hk

Abstract

This paper addresses the problem of recovering both the intrinsic and extrinsic parameters of a camera from the silhouettes of an object in a turntable sequence. Previous silhouette-based approaches have exploited correspondences induced by epipolar tangents to estimate the image invariants under turntable motion and achieved a weak calibration of the cameras. In order to recover the rotation angles and obtain a Euclidean reconstruction, these approaches require the prior knowledge of the camera intrinsics. In this paper, we propose a novel approach for recovering the rotation angles precisely in the absence of the camera intrinsics. It is known that the fundamental matrix relating any two views in a turntable sequence can be expressed explicitly in terms of the image invariants, the rotation angle, and a fixed scalar. It will be shown that the imaged circular points for the turntable plane can also be formulated in terms of the same image invariants and fixed scalar. This allows the imaged circular points to be recovered directly from the estimated image invariants. The imaged circular points and the image invariants provide constraints for the estimation of the imaged absolute conic, and the camera calibration matrix can thus be recovered. A robust method for estimating the fixed scalar from image triplets is introduced, and a method for recovering the rotation angles using the estimated imaged circular points and epipoles is presented. Using the estimated camera intrinsics and extrinsics, a Euclidean reconstruction can be obtained. Experimental results on real data sequences are presented, which demonstrate the high precision achieved by the proposed method.

Index Terms

Self-calibration, circular points, silhouettes, turntable motion.

I. INTRODUCTION

3D object reconstruction from image sequences has always been a hot research topic in both computer vision and computer graphics. Due to its ease of operation, turntable has been commonly used with a fixed camera for acquiring images around an object in the modelling process. The most important and difficult step in object reconstruction from a turntable sequence is the recovery of the relative pose of the camera, the camera intrinsic parameters, and the rotation angles. Traditional approaches accomplish this task by attaching some special calibration patterns to the turntable for calibrating the camera [1], [2]. In [3], Fitzgibbon et al. introduced a point-based method to handle the case with an uncalibrated camera and unknown rotation angles. Their method is based on the projective geometry of single axis motion, and it involves the computation of both fundamental matrices and trifocal tensors from point correspondences. Jiang et al. [4], [5] further extended this approach by

making use of the conic trajectories of the rotating point features, and developed an algorithm that requires neither the computation of fundamental matrices nor trifocal tensors. Note that all the above approaches require the matching of point correspondences across the sequence, and would therefore be not applicable to smooth objects with sparse surface textures.

For smooth objects like sculptures, silhouettes are the predominant and stable image feature. Silhouettes are the projections of contour generators [6] which are viewpoint dependent. For two distinct viewpoints, the contour generators of an object are, in general, two distinct space curves. Hence they provide no point correspondences between the corresponding silhouettes, save for the frontier points [7] which are the points of intersection between the two contour generators. Silhouette-based methods for motion estimation generally exploit epipolar tangents [8], [9], [10], [11] to locate the images of the frontier points for deriving point correspondences between images. However, such an approach requires the presence of at least seven pairs of epipolar tangents, which may not be always available for objects with simple shape. Besides, the location of epipolar tangents in a pair of uncalibrated images involves a nonlinear optimization with a non-trivial initialization. In [12], Sinha et al. proposed to avoid the problem of insufficient frontier points by using silhouettes of a dynamic object, and they also proposed a RANSAC approach for initializing the nonlinear optimization in locating the epipolar tangents. Nonetheless, such a method is not suitable for a static object.

In [13], Mendonça et al. developed a multi-step algorithm for recovering the turntable motion. The number of required epipolar tangents per image pair is only two, which can be easily satisfied in almost all situations. By exploiting the symmetry properties exhibited in the image of a surface of revolution (SoR) generated by the rotating object, a simple method for locating epipolar tangents based on a one-dimensional search is introduced. Note that this method requires a *dense* turntable sequence in order to approximate the image of the SoR by the envelope of the silhouettes. In [14], Wong et al. further extended the work in [13] to handle the case of sparse sequences, and their method also allows the incorporation of general views. The tradeoff is the involvement of a high-dimensional optimization. Note that in both [13] and [14], the recovery of the rotation angles and furthermore the Euclidean reconstruction still require the knowledge of the camera intrinsics. In [15], Hernández et al. considered the problem of recovering both the turntable motion and the camera focal length from silhouettes. Rather than using the epipolar tangency constraint as in [13] and [14], they proposed a novel *silhouette coherence* constraint and developed a method which optimizes both the camera poses and focal length simultaneously by maximizing the coherence between

the silhouettes and the projections of the visual hull built from the silhouettes and the current estimate of the camera parameters. Like [14], their method involves a high-dimensional optimization. It surpasses [14] by exploiting all the information contained in the contours of the silhouettes (rather than just the epipolar tangent points), and can handle incomplete silhouettes. Experimental results reported improvement over those obtained using epipolar tangents. In [16], Furukawa et al. proposed a RANSAC-based method for recovering general motion from silhouettes. They simplified the problem of locating potential epipolar tangents by restricting the viewing geometry to *orthographic projection* under which all epipolar lines are parallel, and used *signature representation* of the dual image outlines to facilitate the matching of epipolar tangents. Although they also outlined how to extend their method to handle the weak-perspective and affine case, only results for the orthographic case were reported.

In this paper, we propose a novel algorithm for recovering both the intrinsic and extrinsic parameters of a camera from the silhouettes of an object in a turntable sequence. The proposed method follows the same approach as in [13], which first estimates the image invariants using epipolar tangents and recovers the fundamental matrices in terms of these invariants. Based on the fact that the epipoles in one view are the images of the other camera centers, a simple method is proposed for precise estimation of the image invariants and rotation angles in the absence of the camera intrinsics. Such a method can also handle the degenerate case as described in [17] and [18], where the baseline passes through the scene object. Besides, it will be shown that the imaged circular points for the turntable plane can be obtained directly in terms of the image invariants and a fixed scalar used in the formula of the fundamental matrix. This allows the imaged circular points to be recovered directly from the estimated image invariants. The imaged circular points, together with the image invariants, provide constraints for the estimation of the imaged absolute conic, and the camera calibration matrix can thus be recovered. This can be used to upgrade the weakly calibrated cameras to fully calibrated ones, and a Euclidean reconstruction follows. Preliminary results of this work have been published in [19].

The work presented here is most closely related to the work of Hernández et al. [15]. In [15], the authors considered the recovery of the focal length, which was estimated simultaneously with other motion parameters via an optimization in a high-dimensional space. In this work, we consider recovering three camera intrinsic parameters, namely the focal length and the coordinates of the principal point. We also decouple the estimation of the

camera intrinsic parameters from the extrinsic parameters, and demonstrate that the intrinsic parameters can be recovered from the image invariants derived directly from the estimated motion parameters. This avoids potential problems of local minima often encountered in high-dimensional optimizations.

The remainder of this paper is organized as follows. Section II rehashes the image invariants under turntable motion and a special parametrization of the fundamental matrix in terms of the image invariants. Section III briefly reviews two existing algorithms for recovering the turntable motion from silhouettes upon which the algorithm proposed in this paper is based. Section IV derives a novel formula for the imaged circular points expressed in terms of the image invariants and a fixed scalar, and presents an algorithm for recovering the imaged circular points from the estimated fundamental matrices. Besides, a robust method for computing the rotation angles in the absence of the camera intrinsics is also introduced. Section V shows the experimental results, followed by conclusions in Section VI.

II. Theoretical Background

The geometry of a stationary camera viewing an object on a rotating turntable is equivalent to that of a camera rotating about the same axis and viewing a stationary object. For the sake of clear notations, we will refer to the geometry of the latter in the rest of this paper.

A. Image Invariants under Turntable Motion

Consider a reference camera $\mathbf{C}_1$ lying on the negative Z-axis of a world coordinate system, and rotating about the Y-axis (see fig. 1(a)). The relative positions of the camera center describe a circle on a plane $\mathbf{\Pi}_\mathrm{h}$ orthogonal to the rotation axis. The image of $\mathbf{\Pi}_\mathrm{h}$ is the line $\mathbf{l}_\mathrm{h}$, which is the vanishing line (i.e., the *horizon*) for the turntable plane (see fig. 1(b)). Since all the camera center positions lie on $\mathbf{\Pi}_\mathrm{h}$, all the epipoles $\mathbf{e}_{ij}$ which are images of the rotating camera center must lie on $\mathbf{l}_\mathrm{h}$, i.e.,

$$\mathbf{l}_\mathrm{h}^\mathrm{T}\mathbf{e}_{ij} = 0 \ \forall i, j. \tag{1}$$

Let the plane defined by the camera center and the rotation axis be $\mathbf{\Pi}_\mathrm{s}$, and consider three orthogonal directions $\mathbf{N}_x$, $\mathbf{N}_y$ and $\mathbf{N}_z$, given by the normal direction of $\mathbf{\Pi}_\mathrm{s}$, the Y-axis, and $\mathbf{N}_x \times \mathbf{N}_y$, respectively. These three directions will have vanishing points $\mathbf{v}_x$, $\mathbf{v}_y$ and $\mathbf{v}_z$, respectively. The image of $\mathbf{\Pi}_\mathrm{s}$ is $\mathbf{l}_\mathrm{s}$, which is also the image of the rotation axis. Since $\mathbf{N}_x$

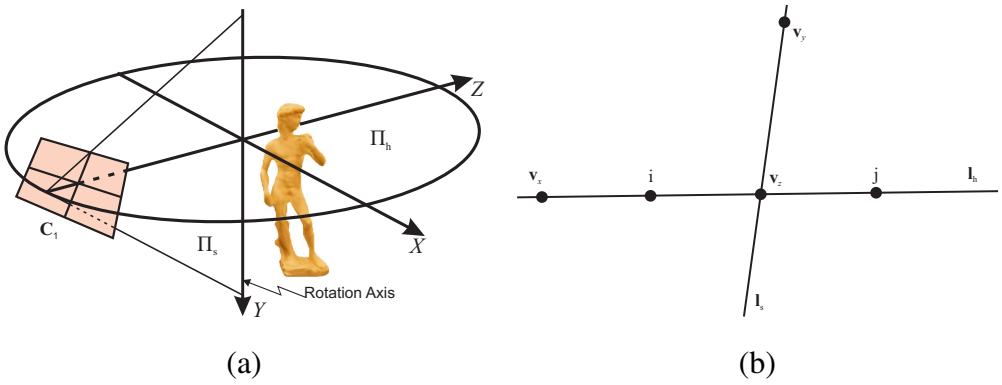


Fig. 1. Image invariants under turntable motion. (a) The reference camera $\mathbf{C}_1$ is initially lying on the negative Z-axis of the world coordinate system, and it rotates about the Y-axis. The plane $\mathbf{\Pi}_\mathrm{h}$ is defined by the positions of the rotating camera center, and the plane $\mathbf{\Pi}_\mathrm{s}$ is defined by the Y-axis and the camera center. (b) The lines $\mathbf{l}_\mathrm{h}$ and $\mathbf{l}_\mathrm{s}$ are the images of $\mathbf{\Pi}_\mathrm{h}$ and $\mathbf{\Pi}_\mathrm{s}$ respectively. The points $\mathbf{v}_x$, $\mathbf{v}_y$ and $\mathbf{v}_z$ are the vanishing points for the normal direction of $\mathbf{\Pi}_\mathrm{s}$, the normal direction of $\mathbf{\Pi}_\mathrm{h}$, and a direction orthogonal to both the former two directions respectively. The points $\mathbf{i}$ and $\mathbf{j}$ are the imaged circular points for $\mathbf{\Pi}_\mathrm{h}$.

is orthogonal to $\mathbf{\Pi}_\mathrm{s}$, it follows that $\mathbf{v}_x$ and $\mathbf{l}_\mathrm{s}$ form a pole-polar relationship with respect to the image of the absolute conic $\boldsymbol{\omega}$ [20], i.e.,

$$\mathbf{l}_\mathrm{s} \sim \boldsymbol{\omega}\mathbf{v}_x. \tag{2}$$

Similarly, $\mathbf{N}_y$ is orthogonal to $\mathbf{\Pi}_\mathrm{h}$, and $\mathbf{v}_y$ and $\mathbf{l}_\mathrm{h}$ also form a pole-polar relationship with respect to the image of the absolute conic $\boldsymbol{\omega}$, i.e.,

$$\mathbf{l}_\mathrm{h} \sim \boldsymbol{\omega}\mathbf{v}_y. \tag{3}$$

By construction, $\mathbf{N}_x$ is parallel to $\mathbf{\Pi}_\mathrm{h}$, $\mathbf{N}_y$ is parallel to $\mathbf{\Pi}_\mathrm{s}$, and $\mathbf{N}_z$ is parallel to both $\mathbf{\Pi}_\mathrm{h}$ and $\mathbf{\Pi}_\mathrm{s}$. Hence, $\mathbf{v}_x$ and $\mathbf{v}_y$ must lie on $\mathbf{l}_\mathrm{h}$ and $\mathbf{l}_\mathrm{s}$, respectively, and $\mathbf{v}_z$ must lie on both $\mathbf{l}_\mathrm{h}$ and $\mathbf{l}_\mathrm{s}$, i.e.,

$$\mathbf{l}_\mathrm{h}^\mathrm{T}\mathbf{v}_x = 0, \tag{4}$$

$$\mathbf{l}_\mathrm{s}^\mathrm{T}\mathbf{v}_y = 0, \text{and} \tag{5}$$

$$\mathbf{v}_z \sim \mathbf{l}_\mathrm{h} \times \mathbf{l}_\mathrm{s}. \tag{6}$$

Every circle on a plane will intersect the line at infinity at the circular points with canonical coordinates $\mathbf{I} = [1, j, 0]^\mathrm{T}$ and $\mathbf{J} = [1, -j, 0]^\mathrm{T}$ (where $j^2 = -1$) [21]. The coordinates of these two conjugate complex points are invariant (as a pair) under any similarity transformation.

The pair of imaged circular points $\mathbf{i}$ and $\mathbf{j}$ for the plane $\mathbf{\Pi}_{\rm h}$ must lie on its vanishing line $\mathbf{l}_{\rm h}$, i.e.,

$$\mathbf{l}_{\rm h} \sim \mathbf{i} \times \mathbf{j}. \tag{7}$$

Since a plane will intersect the absolute conic at the two circular points, it follows that the imaged circular points will lie on the image of the absolute conic, i.e.,

$$\mathbf{i}^{\rm T}\boldsymbol{\omega}\mathbf{i} = \mathbf{j}^{\rm T}\boldsymbol{\omega}\mathbf{j} = 0. \tag{8}$$

If the intrinsic parameters of the camera are kept constant, due to symmetry in the configuration, the aforementioned vanishing points ($\mathbf{v}_x$, $\mathbf{v}_y$ and $\mathbf{v}_z$), vanishing lines ($\mathbf{l}_{\rm h}$ and $\mathbf{l}_{\rm s}$) and the imaged circular points ($\mathbf{i}$ and $\mathbf{j}$) will remain unchanged throughout the sequence.

B. Fundamental Matrix under Turntable Motion

The fundamental matrix relating a view pair under turntable motion can be expressed explicitly in terms of the image invariants, and is given by [22], [3], [17]

$$\mathbf{F}_{ij} = \frac{1}{\det(\mathbf{K})}[\mathbf{v}_x]_\times + \tan\frac{\theta_{ij}}{2}(\mathbf{l}_{\rm s}\mathbf{l}_{\rm h}^{\rm T} + \mathbf{l}_{\rm h}\mathbf{l}_{\rm s}^{\rm T}), \tag{9}$$

where $\mathbf{K}$ is the 3×3 calibration matrix of the camera, and θ_{ij} is the rotation angle between the two views. Since $\mathbf{F}_{ij}$ is expressed as a sum of a skew-symmetric part and a symmetric part, the scales used in the homogeneous representations of the image invariants are important and cannot be ignored. In (9), $\mathbf{v}_x$, $\mathbf{l}_{\rm s}$ and $\mathbf{l}_{\rm h}$ must have the following forms:

$$\mathbf{v}_x = \mathbf{KR}[\ 1 \quad 0 \quad 0\ ]^{\rm T}, \tag{10}$$

$$\mathbf{l}_{\rm s} = (\mathbf{KR})^{-{\rm T}}[\ 1 \quad 0 \quad 0\ ]^{\rm T}, \text{and} \tag{11}$$

$$\mathbf{l}_{\rm h} = (\mathbf{KR})^{-{\rm T}}[\ 0 \quad 1 \quad 0\ ]^{\rm T}, \tag{12}$$

where $\mathbf{R}$ is a 3×3 rotation matrix transforming vectors from the world coordinate system to the reference camera coordinate system. Given only the image invariants in an uncalibrated sequence (i.e., without the knowledge of $\mathbf{K}$ and $\mathbf{R}$), a scale factor κ has to be introduced into the above formula to account for the different scales used in the homogeneous representations of the image invariants, and the expression becomes

$$\mathbf{F}_{ij} \sim [\mathbf{v}_x]_\times + \kappa\tan\frac{\theta_{ij}}{2}(\mathbf{l}_{\rm s}\mathbf{l}_{\rm h}^{\rm T} + \mathbf{l}_{\rm h}\mathbf{l}_{\rm s}^{\rm T}) \tag{13}$$

$$\sim [\mathbf{v}_x]_\times + \lambda(\mathbf{l}_{\rm s}\mathbf{l}_{\rm h}^{\rm T} + \mathbf{l}_{\rm h}\mathbf{l}_{\rm s}^{\rm T}), \tag{14}$$

where $\lambda = \kappa \tan \frac{\theta_{ij}}{2}$. If the scales used in the homogeneous representations of the image invariants are kept constant, the scale factor κ will remain unchanged and will be independent of the particular view pair being considered. From (14), it is easy to see that a fundamental matrix relating a pair of views under turntable motion has only 6 degrees of freedom (dof): 2 dof for $\mathbf{l}_\mathrm{s}$, 3 dof for $\mathbf{l}_\mathrm{h}$ and $\mathbf{v}_x$ (due to the constraint imposed by (4)) and 1 dof for λ. If the image invariants have been recovered, only 1 dof is needed to fix λ and hence $\mathbf{F}_{ij}$.

III. Weak Calibration from Silhouettes

In this section, two recent approaches for estimating turntable motion from silhouettes will be reviewed. Both methods estimate the fundamental matrices associated with the views under turntable motion by minimizing the symmetric transfer errors of the epipolar tangents to the silhouettes, and successful solutions have been developed by exploiting the special parametrization of the fundamental matrix. Note that in order to recover the rotation angles between different views and achieve a Euclidean reconstruction, the camera calibration matrix is needed in both methods to upgrade the weak calibration to a full calibration. In the next section, a novel method will be introduced for recovering the camera calibration matrix from the estimated fundamental matrices. A robust method for recovering the rotation angles in the absence of the camera intrinsics will also be presented. These enhance the current algorithms by removing the requirement of known camera intrinsics, and allow a Euclidean reconstruction from silhouettes in an uncalibrated turntable sequence.

A. Dense Sequence

Given a dense sequence, Mendonça et al. [17] introduced a multi-step algorithm for estimating the turntable motion. Their method first recovers the imaged rotation axis $\mathbf{l}_\mathrm{s}$ and the vanishing point $\mathbf{v}_x$ from the image profile ρ of a SoR generated by the rotating object. They approximate such an image by the envelope of the deforming silhouettes. By exploiting the symmetry exhibited in ρ, $\mathbf{l}_\mathrm{s}$ and $\mathbf{v}_x$ can be easily obtained by locating bitangents to ρ (see fig. 2). Since ρ is invariant to the transformation induced by a harmonic homology $\mathbf{W}$ with axis $\mathbf{l}_\mathrm{s}$ and vertex $\mathbf{v}_x$, defined as

$$\mathbf{W} = \mathbf{I} - 2\frac{\mathbf{v}_x \mathbf{l}_\mathrm{s}^\mathrm{T}}{\mathbf{v}_x^\mathrm{T} \mathbf{l}_\mathrm{s}}, \tag{15}$$

$\mathbf{l}_\mathrm{s}$ and $\mathbf{v}_x$ can be further optimized by minimizing the transformation error of ρ brought about by $\mathbf{W}$. In the second step, the fact that corresponding epioplar lines are being mapped by

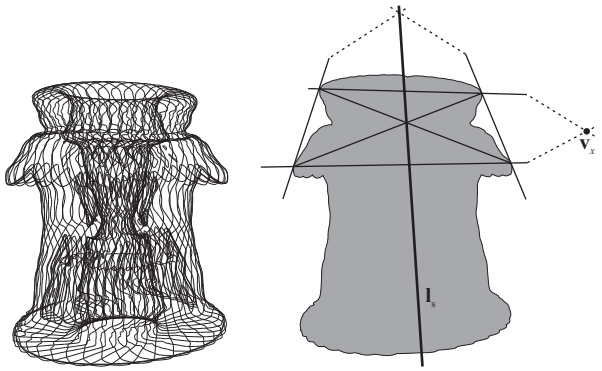


Fig. 2. (a) The image profile of the surface of revolution generated by the rotating object can be approximated by the envelope of the deforming silhouettes of the object. (b) The image invariants $\mathbf{v}_x$ and $\mathbf{l}_s$ can be obtained from the intersections of the bitangent lines and lines formed from the bitangent points.

$\mathbf{W}^{-\mathrm{T}}$ is exploited to locate the outer epipolar tangents to the silhouettes. The objective is to locate a line $\mathbf{l}$ tangent to the silhouette in one view which will be transformed by $\mathbf{W}^{-\mathrm{T}}$ to a line $\mathbf{l}' = \mathbf{W}^{-\mathrm{T}}\mathbf{l}$ tangent to the silhouette in the second view (see fig. 3). This process can be carried out as a one dimensional search in which the single search parameter is the angle that defines the orientation of the line $\mathbf{l}$. The epipoles can then be obtained as the intersection point of the two outer epipolar tangents in each view, and the horizon $\mathbf{l}_\mathrm{h}$ is recovered as a line robustly fitted to the set of epipoles obtained from some randomly sampled image pairs (see fig. 4).

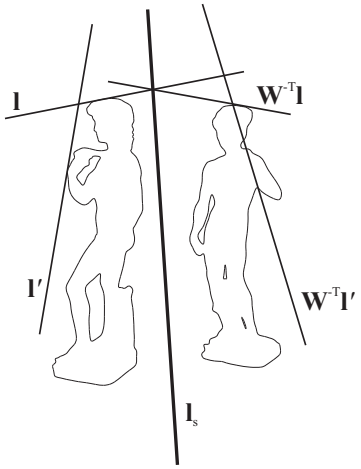


Fig. 3. An epipolar tangent is located by finding a line tangent to the silhouette in the first view which is transformed by the harmonic homology $\mathbf{W}^{-\mathrm{T}}$ to a line tangent to the silhouette in the second view.

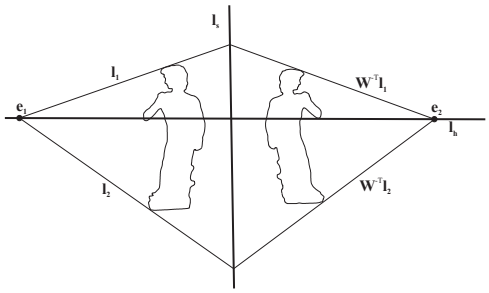


Fig. 4. The epipole is obtained as the intersection point of the two outer epipolar tangents in one view. The horizon $\mathbf{l}_{\mathrm{h}}$ is recovered as a line robustly fitted to the set of epipoles obtained from some randomly sampled image pairs.

After the first two steps, the image invariants $\mathbf{v}_x$, $\mathbf{l}_{\mathrm{s}}$ and $\mathbf{l}_{\mathrm{h}}$ have been recovered. The only missing term in the parametrization of the fundamental matrix $\mathbf{F}_{ij}$ given in (14) is $\lambda = \kappa \tan \frac{\theta_{ij}}{2}$. In the final step, a one dimensional search along the parametric direction λ is performed to optimize $\mathbf{F}_{ij}$ by minimizing the symmetric transfer errors of the epipolar tangents. After this optimization, the set of epipoles obtained from the resulting fundamental matrices will all lie on the horizon $\mathbf{l}_{\mathrm{h}}$.

With known camera intrinsics, the essential matrix can be formed from the fundamental matrix. The rotation angles of the camera can then be recovered by decomposing the essential matrices. A Euclidean reconstruction can be obtained using the silhouettes and the set of projection matrices resulting from the decompositions.

The advantages of the above approach are that it requires the presence of only two epipolar tangents per image pair, and involves only two one-dimensional searches with trivial initializations. Nonetheless, it requires a dense (say with rotation angles less than 10°) and complete (i.e., performing a complete 360° rotation) sequence to approximate the image of a SoR generated by the rotating object, and requires the prior knowledge of the camera calibration matrix to recover the rotation angles and achieve a Euclidean reconstruction. Further, the method will fail if the baseline of an image pair passes through the object. Under this degenerate case, the epipoles will be located inside the silhouettes and the location of the outer epipolar tangents becomes impossible (see fig. 5).

B. Sparse Sequence

The method described in the previous subsection works on a dense and complete turntable sequence. In [18], Wong et al. considered the case of a sparse and incomplete turntable sequence. Their method exploits the special parametrization of the fundamental matrix given

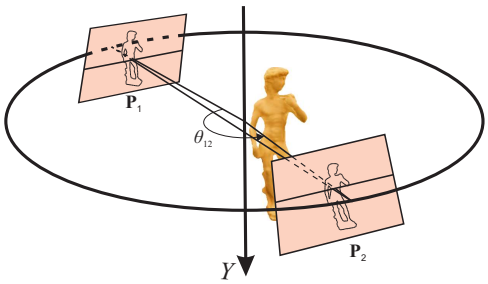


Fig. 5. In the degenerate case, the baseline passes through the object and outer epipolar tangents do not exist.

in (13), and assumes the camera calibration matrix $\mathbf{K}$ is known. Given the camera calibration matrix $\mathbf{K}$, the image of the absolute conic can be recovered as $\omega \sim (\mathbf{K}\mathbf{K}^{\mathrm{T}})^{*}$, and $\mathbf{v}_x$ can be uniquely determined by $\mathbf{l}_{\mathrm{s}}$ using (2). The horizon $\mathbf{l}_{\mathrm{h}}$ can be defined by $\mathbf{v}_x$ and its intersection ($\mathbf{v}_z$) with $\mathbf{l}_{\mathrm{s}}$. Hence $\mathbf{v}_x$, $\mathbf{l}_{\mathrm{s}}$ and $\mathbf{l}_{\mathrm{h}}$ altogether have only 3 dof. Note that with $\mathbf{v}_x$, $\mathbf{l}_{\mathrm{s}}$ and $\mathbf{l}_{\mathrm{h}}$ being properly normalized, the fixed scalar κ can be determined from the determinant of $\mathbf{K}$. Given a sequence of N images, there are $N-1$ rotation angles between adjacent views, and therefore totally $3+(N-1)=N+2$ parameters are needed to describe all the fundamental matrices (i.e., motion parameters) associated with the sequence. The method in [18] begins by manually initializing $\mathbf{l}_{\mathrm{s}}$ (and hence $\mathbf{v}_x$) and selecting a point on $\mathbf{l}_{\mathrm{s}}$ to define $\mathbf{l}_{\mathrm{h}}$. The $N-1$ rotation angles between the viewpoints are arbitrarily initialized. These $N+2$ parameters are then optimized by minimizing the symmetric transfer errors of the epipolar tangents to the silhouettes. After the optimization, the fundamental matrices between adjacent viewpoints are formed, which are then upgraded to the essential matrices using the camera calibration matrix. A set of projection matrices compatible with the turntable motion can then be obtained by decomposing the resulting essential matrices. In [18], Wong et al. also proposed a method for registering new general views with the silhouettes in the turntable sequence which optimizes the 6 extrinsic parameters of the camera by minimizing the symmetric transfer errors of the epipolar tangents. Interested readers are referred to [14] and [18] for details of this algorithm. Under the assumptions of unit aspect ratio and the principal point being at the image center, a minor modification by including also the focal length f in the optimization parameters (i.e., total $N+3$ parameters) will produce an algorithm similar to [15].

Compared with the dense sequence method, the above method can handle sparse and incomplete turntable sequences. However, it involves a nonlinear optimization in a high-dimensional $(N+2)$ space. Nonetheless, the fact that these $N+2$ parameters all have

physical meanings (e.g., image of the rotation axis, the horizon and rotation angles) makes the initialization of the nonlinear optimization an easy task. Like the dense sequence method, this method requires the prior knowledge of the camera calibration matrix to recover the rotation angles and achieve a Euclidean reconstruction, and will fail under the previously described degenerate case.

IV. Self-Calibration of Turntable Motion

In this section, a novel parametrization for the imaged circular points of the turntable plane will be introduced. This allows the imaged circular points to be recovered directly from the estimated image invariants. The imaged circular points thus obtained, together with other image invariants, will be exploited to develop a method for recovering the camera calibration matrix. Finally, a robust method for recovering the rotation angles in the absence of the camera intrinsics is presented. These extend existing methods for turntable motion estimation from silhouettes to handle uncalibrated sequences, by removing the restrictive requirement of the prior knowledge of the camera intrinsics.

A. Parametrization for the Imaged Circular Points

In this subsection, a new formula for the imaged circular points of the turntable plane will be derived. It will be shown that the imaged circular points can be expressed in terms of the image invariants and the fixed scalar used in the special parametrization of the fundamental matrix for turntable sequence. The exact expression for the imaged circular points of the turntable plane in terms of the image invariants is given below in Proposition 1.

Proposition 1: Given a turntable sequence, the imaged circular points $\mathbf{i}$ and $\mathbf{j}$ of the turntable plane can be expressed as

$$\mathbf{i}, \mathbf{j} \sim \mathbf{v}_x \pm j\kappa(\mathbf{l}_\mathrm{s} \times \mathbf{l}_\mathrm{h}), \tag{16}$$

where $j^2 = -1$, and $\mathbf{v}_x$, $\mathbf{l}_\mathrm{s}$, $\mathbf{l}_\mathrm{h}$ and κ are the image invariants and fixed scale factor used in the special parametrization of the fundamental matrix for turntable motion as given in (13).

Proof: Without loss of generality, consider two views in the turntable sequence with projection matrices $\mathbf{P}_1 = \mathbf{K}\mathbf{R}[\ \mathbf{I}\ \ -\mathbf{C}\ ]$ and $\mathbf{P}_2 = \mathbf{K}\mathbf{R}[\ \mathbf{R}_Y(\theta_{12})\ \ -\mathbf{C}\ ]$, respectively, where $\mathbf{R}_Y(\theta_{12})$ is a rotation about the Y-axis by an angle θ_{12} and $\mathbf{C} = [\ 0\ \ 0\ \ -1\ ]^\mathrm{T}$ is the camera center (see fig. 6). The image $\mathbf{x}$ of a point $\mathbf{X} = [\ X\ \ 0\ \ Z\ \ 1\ ]^\mathrm{T}$ on the X-Z

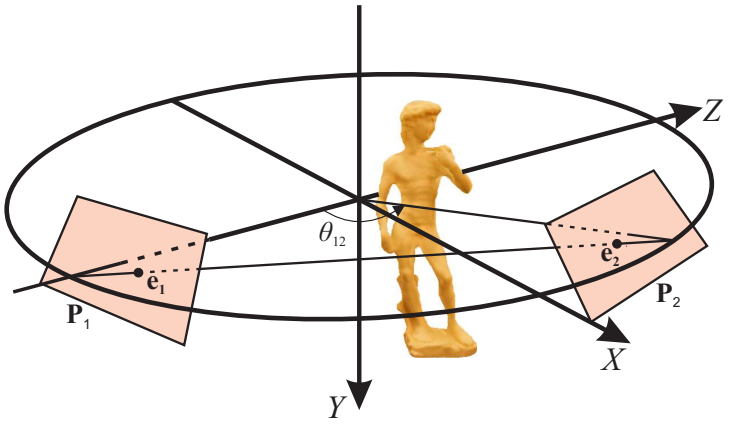


Fig. 6. A pair of cameras $\mathbf{P_1}$ and $\mathbf{P_2}$ related by a rotation with an angle θ_{12} about the Y-axis.

plane under $\mathbf{P}_1$ is given by

$$\begin{aligned}\mathbf{x} &\sim \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3 & \mathbf{p}_4 \end{bmatrix} \begin{bmatrix} X \\ 0 \\ Z \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_3 \end{bmatrix} \begin{bmatrix} X \\ Z+1 \end{bmatrix}, \end{aligned} \tag{17}$$

where $\mathbf{p}_c$ $(c = 1, 2, 3, 4)$ are the columns of $\mathbf{P}_1$ and $\mathbf{p}_3 = \mathbf{p}_4$. Since under $\mathbf{P}_1$, $\mathbf{v}_x$ and $\mathbf{v}_z$ are the vanishing points of the X-axis and Z-axis respectively, it follows that $\mathbf{p}_1 \sim \mathbf{v}_x$ and $\mathbf{p}_3 \sim \mathbf{v}_z$ [3]. The epipole $\mathbf{e}_i$ on the image of $\mathbf{P}_i$ is the projection of the point $\begin{bmatrix} \sin(\pi + (-1)^i\theta_{12}) & 0 & \cos(\pi + (-1)^i\theta_{12}) & 1 \end{bmatrix}^{\mathrm{T}}$, and is given by

$$\begin{aligned}\mathbf{e}_i &\sim \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_3 \end{bmatrix} \begin{bmatrix} \sin(\pi + (-1)^i\theta_{12}) \\ \cos(\pi + (-1)^i\theta_{12}) + 1 \end{bmatrix} \\ &\sim \begin{bmatrix} \mathbf{v}_x & \beta\mathbf{v}_z \end{bmatrix} \begin{bmatrix} \sin(\pi + (-1)^i\theta_{12}) \\ \cos(\pi + (-1)^i\theta_{12}) + 1 \end{bmatrix} \\ &\sim \begin{bmatrix} \mathbf{v}_x & \beta(\mathbf{l}_\mathrm{s} \times \mathbf{l}_\mathrm{h}) \end{bmatrix} \begin{bmatrix} \sin(\pi + (-1)^i\theta_{12}) \\ \cos(\pi + (-1)^i\theta_{12}) + 1 \end{bmatrix} \\ &\sim \mathbf{v}_x - (-1)^i\beta \tan\frac{\theta_{12}}{2}(\mathbf{l}_\mathrm{s} \times \mathbf{l}_\mathrm{h}), \end{aligned} \tag{18}$$

where β is an unknown fixed scalar used to fix the relative scales between $\mathbf{v}_x$ and $\mathbf{v}_z$, and $\mathbf{v}_z \sim \mathbf{l}_\mathrm{s} \times \mathbf{l}_\mathrm{h}$ (see (6)). Note that $\mathbf{e}_1$ and $\mathbf{e}_2$ are the right and left nullspaces of the fundamental

matrix $\mathbf{F}_{12}$ relating the two views. From (13), $\mathbf{F}_{12}$ has the form

$$\mathbf{F}_{12} \sim [\mathbf{v}_x]_\times + \kappa \tan\frac{\theta_{12}}{2}(\mathbf{l}_\mathrm{s}\mathbf{l}_\mathrm{h}^\mathrm{T} + \mathbf{l}_\mathrm{h}\mathbf{l}_\mathrm{s}^\mathrm{T}). \tag{19}$$

Solving $\mathbf{F}_{12}\mathbf{e}_1 = \mathbf{0}$ and $\mathbf{F}_{12}^\mathrm{T}\mathbf{e}_2 = \mathbf{0}$ gives $\beta = \kappa$. Finally, the imaged circular points $\mathbf{i}, \mathbf{j}$ can be obtained by projecting the circular points $\mathbf{I}, \mathbf{J} = [\ 1 \quad 0 \quad \pm j \quad 0\ ]^\mathrm{T}$ under $\mathbf{P}_1$, i.e.,

$$\begin{aligned} \mathbf{i}, \mathbf{j} &\sim \mathbf{p}_1 \pm j\mathbf{p}_3 \\ &\sim \mathbf{v}_x \pm j\kappa(\mathbf{l}_\mathrm{s} \times \mathbf{l}_\mathrm{h}), \end{aligned}$$

and the proof is completed. ■

B. Recovery of the Camera Calibration Matrix

By exploiting the parametrization introduced in the previous subsection, the imaged circular points of the turntable plane can be recovered once the fundamental matrix has been estimated. In this work, the image invariants and the fundamental matrices are first estimated using the dense sequence method [17] as described in Section III-A. The unknown scale factor κ used in the formula of the imaged circular points can be recovered by considering a triplet of views. Let the three views have relative rotation angles θ_{pq} where $p, q = 1, 2, 3$ and $p \neq q$ (see fig. 7), and $\lambda_1 = \kappa \tan\frac{\theta_{12}}{2}$, $\lambda_2 = \kappa \tan\frac{\theta_{23}}{2}$, and $\lambda_3 = \kappa \tan\frac{\theta_{13}}{2}$, respectively. Using the fact that $\theta_{13} = \theta_{12} + \theta_{23}$, the fixed scalar κ can be obtained as

$$\kappa = \sqrt{\frac{\lambda_1\lambda_2\lambda_3}{\lambda_3 - \lambda_1 - \lambda_2}}. \tag{20}$$

Due to the existence of noises, κ estimated from different triplets may have different values. In this work, a set of κ is recovered from some randomly sampled image triplets, and the mode is then chosen as the best estimate.

Since the imaged circular points lie on the image of the absolute conic (IAC) $\boldsymbol{\omega}$, hence

$$\mathbf{i}^\mathrm{T}\boldsymbol{\omega}\mathbf{i} = \mathbf{j}^\mathrm{T}\boldsymbol{\omega}\mathbf{j} = 0. \tag{21}$$

Besides, the imaged rotation axis $\mathbf{l}_\mathrm{s}$ and the vanishing point $\mathbf{v}_x$ define a pole-polar relationship w.r.t. the IAC [21], hence

$$\mathbf{l}_\mathrm{s} \sim \boldsymbol{\omega}\mathbf{v}_x. \tag{22}$$

(21) and (22) together provide three independent constraints on $\boldsymbol{\omega}$, which allow $\boldsymbol{\omega}$ to be estimated under the assumption of a natural camera with zero skew and unit aspect ratio. Finally, the camera calibration matrix $\mathbf{K}$ can be obtained by Cholesky decomposition [23] of $\boldsymbol{\omega}$.

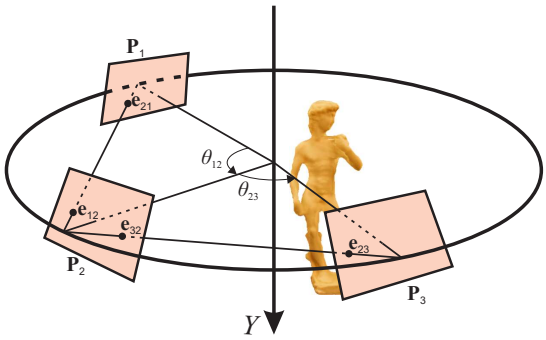


Fig. 7. Three views under turntable motion. Epipole $\mathbf{e}_{pq}$ is the projection of the camera center of view p onto view q. θ_{pq} is the rotation angle between view p and view q ($p, q = 1, 2, 3\ \ p \neq q$).

C. Recovery of the Rotation Angles

Having estimated the fixed scalar κ, a simple way for recovering the rotation angles is to divide the term $\lambda = \kappa \tan(\theta/2)$ in (14) by κ [19]. However, such an approach cannot be applied to the degenerate case shown in fig. 5, since the fundamental matrix (and hence λ) cannot be recovered in such a case. This problem can be avoided by acquiring the images in a way such that the horizon is above or below all the silhouettes. In this paper, this problem is solved by making use of the available epipoles from all image pairs.

Consider the top view of the turntable plane (i.e., the X-Z plane, see fig. 8(a)). The reference camera $\mathbf{C}_1$ (lying on the negative Z-axis) rotates about the origin $\mathbf{O}$ and the relative positions of the camera center $\mathbf{C}_i$ ($i = 1, ..., m$) describe a circle on the X-Z plane. Consider the camera center at two positions $\mathbf{C}_1$ and $\mathbf{C}_2$ with a relative rotation angle θ_{12}, and a camera center at another position $\mathbf{C}_i$ where $i \neq 1, 2$. The epipoles $\mathbf{e}_{i1}$ and $\mathbf{e}_{i2}$ are the projections of the camera center $\mathbf{C}_i$ in the images taken at $\mathbf{C}_1$ and $\mathbf{C}_2$, respectively. By rotating both $\mathbf{C}_2$ and $\mathbf{C}_i$ about the origin by an angle θ_{12} towards $\mathbf{C}_1$, $\mathbf{C}_2$ will become coincide with $\mathbf{C}_1$ (see fig. 8(b)), and let the new position of $\mathbf{C}_i$ be $\mathbf{C}_i'$. It is easy to see that the rotation angle between $\mathbf{C}_i$ and $\mathbf{C}_i'$ is θ_{12}, and $\mathbf{e}_{i2}$ is equivalent to the projection of $\mathbf{C}_i'$ in the image taken at $\mathbf{C}_1$ [24]. Let the viewing rays from $\mathbf{C}_1$ to $\mathbf{C}_i$ and $\mathbf{C}_i'$ be $\mathbf{L}_{i1}$ and $\mathbf{L}_{i2}$, respectively. By simple trigonometry, the angle ϕ_{12} between $\mathbf{L}_{i1}$ and $\mathbf{L}_{i2}$ is equal to half of the rotation angle θ_{12}, i.e., $\phi_{12} = \frac{\theta_{12}}{2}$. The angle ϕ_{12} can be obtained using the Laguerre's formula

$$\phi_{12} = \frac{1}{2j} \log\{\mathbf{L}_{i1}, \mathbf{L}_{i2}; \mathbf{L}_{\mathbf{I}1}, \mathbf{L}_{\mathbf{J}1}\}, \tag{23}$$

where $\{\mathbf{L}_{i1}, \mathbf{L}_{i2}; \mathbf{L}_{\mathbf{I}1}, \mathbf{L}_{\mathbf{J}1}\}$ denotes a cross-ratio, $j^2 = -1$, and $\mathbf{L}_{\mathbf{I}1}$ and $\mathbf{L}_{\mathbf{J}1}$ are the isotropic

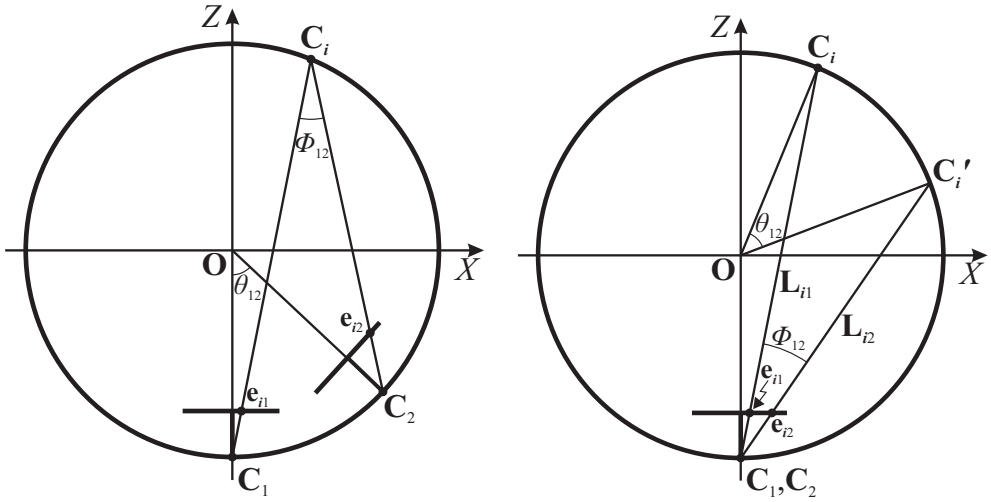


Fig. 8. The projection of camera center(s). (a) Two views and one point configuration: the epipoles $\mathbf{e}_{i1}$ and $\mathbf{e}_{i2}$ are the projections of the camera center $\mathbf{C}_i$ in the images taken at $\mathbf{C}_1$ and $\mathbf{C}_2$, respectively, where $\mathbf{C}_2$ is obtained by rotating $\mathbf{C}_1$ about the origin by an angle θ_{12}. (b) Two points and one view configuration: by rotating both $\mathbf{C}_2$ and $\mathbf{C}_i$ about the origin by an angle θ_{12} towards $\mathbf{C}_1$, $\mathbf{C}_2$ will become coincide with $\mathbf{C}_1$ and let the new position of $\mathbf{C}_i$ be $\mathbf{C}_i'$, $\mathbf{e}_{i2}$ can be regarded as the projection of the camera center $\mathbf{C}_i'$ in the image taken at $\mathbf{C}_1$.

lines through $\mathbf{C}_1$ (i.e., the lines from $\mathbf{C}_1$ to the circular points $\mathbf{I}$ and $\mathbf{J}$ respectively) [25]. Given the epipoles $\{\mathbf{e}_{i1}, \mathbf{e}_{i2}\}$ and the imaged circular points $\{\mathbf{i}, \mathbf{j}\}$, ϕ_{12} can therefore be estimated as [4]

$$\phi_{12} = \frac{1}{2j} \log\{\mathbf{e}_{i1}, \mathbf{e}_{i2}; \mathbf{i}, \mathbf{j}\}, \tag{24}$$

and the rotation angle θ_{12} follows. Note that the accuracy of the recovered angle θ_{12} depends on the accuracy of the epipoles $\mathbf{e}_{i1}$ and $\mathbf{e}_{i2}$ (i.e., the accuracy of the estimated fundamental matrices $\mathbf{F}_{1i}$ and $\mathbf{F}_{2i}$) being used. To handle the degenerate case robustly and obtain an accurate estimate of the rotation angle, all available image pairs are considered to compute a set of estimated angles and the median is then chosen as the robust estimate of θ_{12}.

V. Experimental Results

The first experimental sequence consists of 72 images of a vase (see fig. 9(a)). The images have a resolution of 640×480. Each image was taken by sequentially rotating the object by 5° on a manually operated turntable with a resolution of 0.01°. The silhouettes were extracted using cubic B-spline snakes [26], [27] with manual initializations. The image invariants $\mathbf{l}_\mathrm{s}$, $\mathbf{v}_x$ and $\mathbf{l}_\mathrm{h}$ were first estimated from the whole sequence using the dense sequence method as

described in Section III-A. A subsequence consisting of 18 images with successive rotation angles of 20° was then extracted for the Euclidean reconstruction of the vase. The fixed scalar κ and the rotation angles were estimated from this subsequence using the methods described in Section IV-B and Section IV-C respectively. All the parameters ($\mathbf{v}_x$, $\mathbf{l}_\mathrm{s}$, $\mathbf{l}_\mathrm{h}$, κ and rotation angles) were then optimized in a bundle adjustment which minimized the symmetric transfer errors of the epipolar tangents to the silhouettes using (13). The camera calibration matrix was then recovered from the optimized image invariants using the method described in Section IV-B. Finally, a set of projection matrices compatible with the turntable motion were obtained from the decomposition of the essential matrices, and a Euclidean reconstruction was achieved using a visual hull from silhouettes method as described in [28].

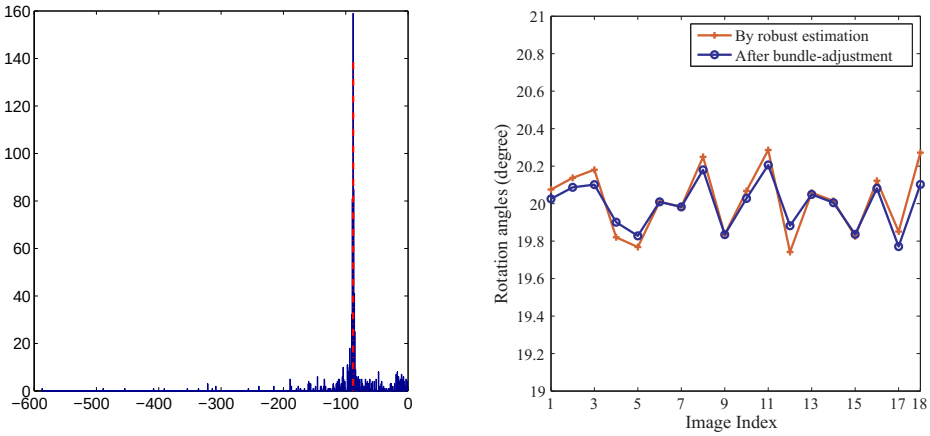
(a) (b)

Fig. 9. (a) An image form the vase sequence. (b) An image from the David sequence.

Fig. 10(a) shows the histogram of the estimated values for κ from some randomly sampled image triplets, and the peak indicates the mode which was taken as the robust estimate of κ. Fig. 10(b) shows both the rotation angles obtained before and after the bundle adjustment. The rms errors of the estimated rotation angles before and after the bundle adjustment are 0.134° and 0.112° respectively, which are better than the rms error of 0.192° reported in [17] for the same vase sequence.

The intrinsics recovered from the optimized image invariants are shown in Table I, which also shows the ground-truth values obtained with a classical calibration method [21] using a L-shape calibration pattern. It can be seen that the focal length f and the u_0 coordinate of the principal point were both precisely estimated, while v_0 was not. This is due to the high uncertainty in the estimated coordinates of $\mathbf{v}_x$, which was far away from the image center for a camera looking at a direction close to the rotation axis. Note that under the assumption of

a natural camera, any error in the location of $\mathbf{v}_x$ in a direction parallel to $\mathbf{l}_\mathrm{s}$ will result in the same error in the location of the principal point. Fig. 11(a) shows the recovered positions and orientations of the rotating camera relative to the model, and fig. 11(b) shows three views of the 3D model reconstructed from the estimated motion.



(a) (b)

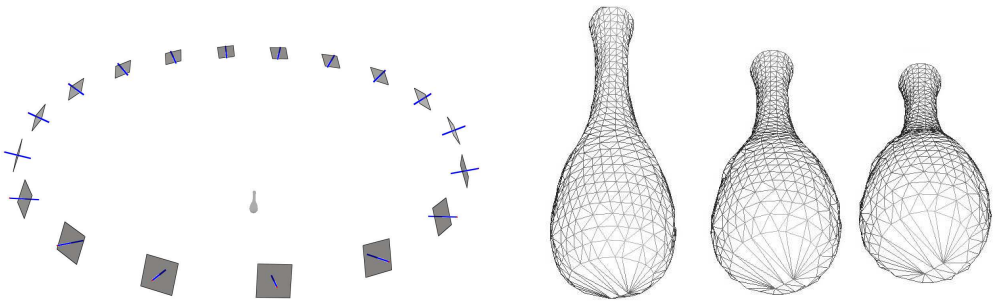
Fig. 10. Results for the vase sequence. (a) Histogram of the estimated values for κ, and the dashed line indicates the "best estimate" κ_0. (b) Angles estimated using the robust method and those obtained after bundle adjustment. The rms errors of the estimated rotation angles before and after the bundle adjustment are 0.134° and 0.112° respectively.

TABLE I

ESTIMATED AND GROUND-TRUTH VALUES FOR THE INTRINSICS OF THE VASE SEQUENCE.

-	f	u_0	v_0
Ground-truth	2389.8	342.83	255.32
Estimated value	2366.5	336.76	393.35
Percentage error	0.97%	0.25%	5.78%

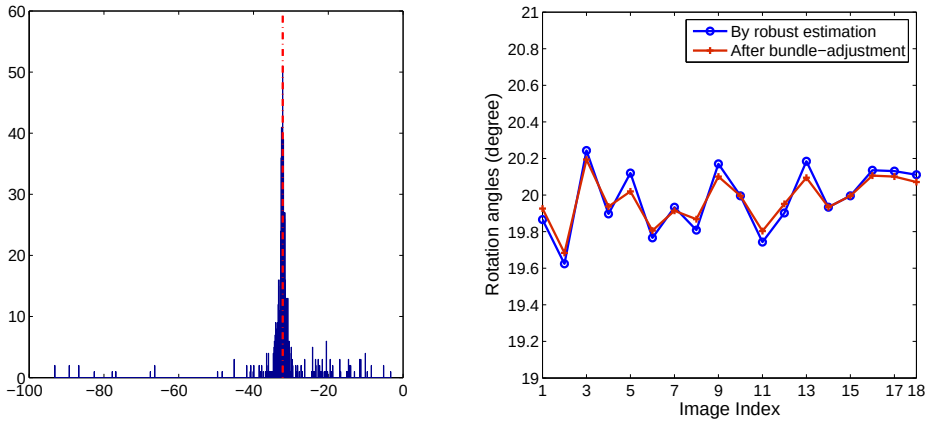
The second sequence consists of 72 images of a David model with successive rotation angles of 5° (see fig. 9(b)). The electronic turntable used has a resolution of 0.2°. Similar to the first experiment, after obtaining $\mathbf{l}_\mathrm{s}$, $\mathbf{v}_x$ and $\mathbf{l}_\mathrm{h}$ from the whole dense sequence, a subsequence consisting of 18 images with successive rotation angles of 20° was extracted, and the same procedures for motion estimation, self-calibration and Euclidean reconstruction were applied to this subsequence.



(a) (b)

Fig. 11. Results for the vase sequence. (a) The recovered camera positions and orientations for the vase sequence. (b) Different views of the 3D model of the vase built from the estimated turntable motion.

Fig. 12(a) shows the histogram of the estimated values for the fixed scalar κ. Fig. 12(b) shows both the rotation angles obtained before and after the bundle adjustment. The rms errors of the estimated rotation angles before and after the bundle adjustment are 0.170° and 0.120° respectively, which are better than the rms error of 0.272° achieved using the algorithm described in [17]. The calibration result is shown in Table II. Fig. 13(a) shows the recovered positions and orientations of the rotating camera relative to the model. The reconstructed model is shown in Fig. 13(b), which reflects good qualities of our estimated parameters.



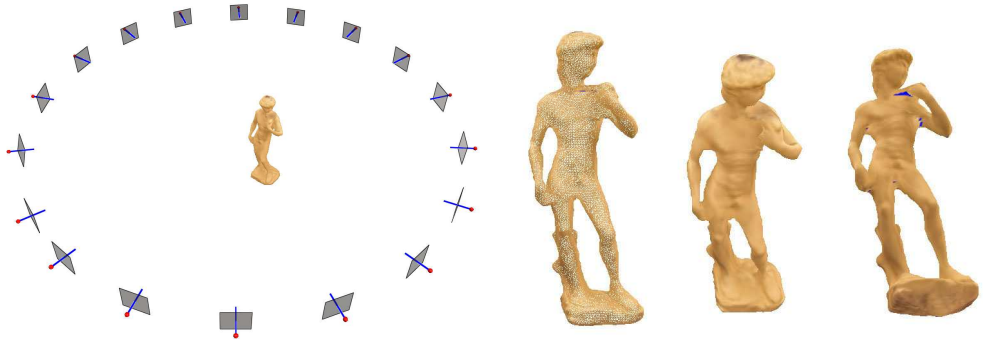
(a) (b)

Fig. 12. Results for the David sequence. (a) Histogram of the estimated values for κ, and the dashed line indicates the "best estimate" κ_0. (b) Angles estimated using the robust method and those obtained after bundle adjustment. The rms errors of the estimated rotation angles before and after the bundle adjustment are 0.170° and 0.120° respectively

TABLE II

ESTIMATED AND GROUND-TRUTH VALUES FOR THE INTRINSICS OF THE DAVID SEQUENCE.

-	f	u_0	v_0
Ground-truth	814.99	414.14	198.70
Estimated value	813.68	412.06	268.47
Percentage error	0.16%	0.26%	8.56%



(a) (b)

Fig. 13. Results for the David sequence. (a) The recovered camera positions and orientations for the David sequence. (b) Different views of the 3D model of David built from the estimated turntable motion.

The third experiment was to test the robustness of the algorithm under different rotation angles. The dense sequence consisting of 72 vase images used in the first experiment was employed here. It provided ten possible subsequences with 72, 36, 24, 18, 12, 9, 8, 6, 4 and 3 images respectively. The corresponding rotation angle θ between adjacent images in these subsequences are $5°$, $10°$, $15°$, $20°$, $30°$, $40°$, $45°$, $60°$, $90°$ and $120°$, respectively.

For each subsequence, the camera intrinsics and the rotation angles were computed in exactly the same way as in the first two experiments, except that the bundle adjustment step was omitted. The results are shown in Table III. It can be seen that the camera calibration matrix can generally be estimated with high accuracy. However, the rotation angles could not be estimated accurately when the rotation angles are large. This is due to the fact that under large rotation angles, the chance of having the degenerate case increases, and at the same time the number of epipoles available for robust estimation of the angles decreases. Hence the accuracy of the estimated rotation angles decreases accordingly.

TABLE III

CALIBRATION RESULTS AND THE ESTIMATED ANGLES UNDER DIFFERENT ROTATION ANGLES. COLUMN 2-4: ESTIMATED INTRINSIC PARAMETERS. THE VALUES IN BRACKETS ARE THE PERCENTAGE ERRORS RELATIVE TO THE GROUND TRUTH FOCAL LENGTH. COLUMN 5: THE ROOT MEAN SQUARE ERRORS OF THE ESTIMATED ROTATION ANGLES. THE VALUES IN BRACKETS ARE THE ROOT MEAN SQUARE ERRORS OF THE ESTIMATED ROTATION ANGLES RELATIVE TO THE GROUND TRUTH ANGLES.

θ	f	u_0	v_0	rms err of angles
5°	2336.78(2.22%)	350.87(0.34%)	487.80(9.73%)	0.08(1.57%)
10°	2339.88(2.09%)	350.82(0.33%)	487.80(9.73%)	0.09(0.95%)
15°	2339.10(2.12%)	350.83(0.33%)	487.80(9.73%)	0.16(1.05%)
20°	2345.88(1.84%)	350.72(0.33%)	487.81(9.73%)	0.13(0.67%)
30°	2331.59(2.44%)	350.95(0.34%)	487.79(9.73%)	0.22(0.74%)
40°	2364.02(1.08%)	350.43(0.32%)	487.84(9.73%)	0.49(1.22%)
45°	2364.24(1.07%)	350.42(0.32%)	487.84(9.73%)	0.41(0.90%)
60°	2402.15(0.52%)	349.80(0.29%)	487.91(9.73%)	1.56(2.59%)
90°	2342.98(1.96%)	350.77(0.33%)	487.81(9.73%)	19.72(21.91%)
120°	2266.31(5.17%)	351.99(0.38%)	487.68(9.72%)	56.10(46.75%)

VI. DISCUSSIONS AND CONCLUSIONS

In this paper, a simple and practical approach is introduced for recovering the camera intrinsics and the relative rotation angles from silhouettes in a turntable sequence. A special parametrization for the imaged circular points of the turntable plane in terms of the image invariants and a fixed scalar is derived. This allows the camera intrinsics to be obtained directly from the estimated image invariants. Besides, a robust method is developed for estimating the rotation angles. This method does not depend on the knowledge of the camera intrinsics, and can handle the degenerate case in which the fundamental matrix is not recoverable. The proposed algorithm takes an uncalibrated turntable sequence as input, and self-calibration is carried out to achieve a Euclidean reconstruction. Experiments on two image sequences showed that both the camera intrinsics and the rotation angles could be estimated with a high precision, and convincing 3D models had been constructed using these estimated parameters.

REFERENCES

[1] R. Szeliki, "Shape from rotation," in *Proc. of IEEE International Conference on Computer Vision and Pattern Recognition*, 1991, pp. 625–630.

[2] W. Niem, "Robust and fast modelling of 3d natural objects from multiple views," in *SPIE Proceedings - Image and Video Processing II*, vol. 2182, San Jose, February 1994, pp. 388–397.

[3] A. W. Fitzgibbon, G. Cross, and A. Zisserman, "Automatic 3d model construction for turn-table sequences," in *3D Structure from Multiple Images of Large-Scale Environments*, ser. LNCS 1506. Springer Verlag, 1998, pp. 155–170.

[4] G. Jiang, H. T. Tsui, L. Quan, and A. Zisserman, "Single axis geometry by fitting conics," in *Proc. of European Conference on Computer Vision*, 2002, pp. 482–488.

[5] G. Jiang, L. Quan, and H. Tsui, "Circular motion geometry by minimal 2 points in 4 images," in *Proc. of International Conference on Computer Vision*, 2003, pp. 221–227.

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

[7] R. Cipolla, K. *Å*ström, and P. Giblin, "Motion from the frontier of curved surfaces," in *Proc. of 5th International Conference on Computer Vision*, Boston, USA, June 1995, pp. 269–275.

[8] J. H. Rieger, "3 dimensional motion from fixed points of a deforming profile curve," *Optics Letters*, vol. 3, no. 11, pp. 123–125, 1986.

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

[10] K. *Å*ström, R. Cipolla, and P. Giblin, "Generalised epipolar constraints," *International Journal of Computer Vision*, vol. 33, no. 1, pp. 51–72, September 1999.

[11] T. Joshi, N. Ahuja, and J. Ponce, "Structure and motion estimation from dynamic silhouettes under perspective projection," *International Journal of Computer Vision*, vol. 31(1), pp. 31–50, February 1999.

[12] S. Sinha, M. Pollefeys, and L. McMillan, "Camera network calibration from dynamic silhouettes," in *Proc. of IEEE International Conference on Computer Vision and Pattern Recognition*, June/July 2004, pp. 195–202.

[13] P. R. S. Mendonça, K.-Y. K. Wong, and R. Cipolla, "Camera pose estimation and reconstruction from image profiles under circular motion," in *Proc. of European Conference on Computer Vision*, vol. II, 2000, pp. 864–877.

[14] K.-Y. K. Wong and R. Cipolla, "Structure and motion from silhouettes," in *Proc. of 8th IEEE International Conference on Computer Vision*, vol. II, Vancouver, Canada, July 2001, pp. 217–222.

[15] C. Hernández, F. Schmitt, and R. Cipolla, "Silhouette coherence for camera calibration under circular motion," vol. 29, no. 2, pp. 343–349.

[16] Y. Furukawa, A. Sethi, J. Ponce, and D. J. Kriegman, "Robust structure and motion from outlines of smooth curved surfaces," pp. 302–315, February 2006.

[17] 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.

[18] K.-Y. K. Wong and R. Cipolla, "Reconstruction of sculpture from its profiles with unknown camera positions," *IEEE Trans. on Image Processing*, vol. 13, no. 3, pp. 381–389, March 2004.

[19] H. Zhang, G. Zhang, and K.-Y. K. Wong, "Auto-calibration and motion recovery from silhouettes for turntable sequences," in *Proc. of British Machine Vision Conference*, vol. I, Oxford, UK, September 2005, pp. 79–88.

[20] 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*, vol. 25(2), pp. 147–161, February 2003.

[21] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. UK: Cambridge University Press, 2004.

[22] 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.

[23] J. Gentle, *Numerical Linear Algebra for Applications in Statistics*. Springer-Verlag, 1998.

[24] G. Zhang, H. Zhang, and K.-Y. K. Wong, "1d camera geometry and its application to circular motion estimation," in *Proc. British Machine Vision Conference*, vol. I, Edinburgh, UK, September 2006, pp. 67–76.

[25] J. G. Semple and G. T. Kneebone, *Algebraic Projective Geometry*. Oxford, UK: Clarendon Press, 1998.

[26] R. Cipolla and A. Blake, "The dynamic analysis of apparent contours," in *Proc. 3rd Int. Conf. on Computer Vision*, Osaka, Japan, December 1990, pp. 616–623.

[27] ——, "Surface shape from the deformation of apparent contours," *International Journal of Computer Vision*, vol. 9(2), pp. 83–112, November 1992.

[28] C. Liang and K.-Y. K. Wong, "Complex 3d shape recovery using a dual-space approach," in *Proc. of IEEE International Conference on Computer Vision and Pattern Recognition*, vol. 2, San Diego, CA, USA, June 2005, pp. 878–884.