

1D Camera Geometry and Its Application to the Self-Calibration of Circular Motion Sequences

Kwan-Yee K. Wong, *Member, IEEE*, Guoqiang Zhang, *Student-Member, IEEE*,
Chen Liang, *Student-Member, IEEE*, and Hui Zhang *Member, IEEE*

Kwan-Yee K. Wong, Guoqiang Zhang and Chen Liang are with the Department of Computer Science, The University of Hong Kong, Pokfulam Road, Hong Kong. Email: [kykwong,gqzhang,cliang]@cs.hku.hk

Hui Zhang is with the Division of Science and Technology, United International College, 28 Jinfeng Road, Tangjiawan, Zhuhai, Guangdong, P.R. China, 519 085. Email: amyzhang@uic.edu.hk

Abstract

This paper proposes a novel method for robustly recovering the camera geometry of an uncalibrated image sequence taken under circular motion. Under circular motion, all the camera centers lie on a circle and the mapping from the plane containing this circle to the horizon line observed in the image can be modelled as a 1D projection. A 2×2 homography is introduced in this paper to relate the projections of the camera centers in two 1D views. It is shown that the two imaged circular points of the motion plane and the rotation angle between the two views can be derived directly from such a homography. This way of recovering the imaged circular points and rotation angles is intrinsically a multiple view approach, as all the sequence geometry embedded in the epipoles is exploited in the estimation of the homography for each view pair. This results in a more robust method compared to those computing the rotation angles using adjacent views only. The proposed method has been applied to self-calibrate turntable sequences using either point features or silhouettes, and highly accurate results have been achieved.

Index Terms

1D projective camera, homography, self-calibration, circular motion.

I. INTRODUCTION

Turntable is commonly used for showcasing a 360° -view of an object in museums, exhibitions and retail stores. It also provides a convenient way for acquiring images around an object. In both computer vision and graphics, such a sequence is often exploited to reconstruct a 3D model of the object. The most critical step to a successful reconstruction is camera calibration, which estimates the relative pose and intrinsic parameters of the camera, as well as the rotation angles of the turntable. Note that turntable motion (a.k.a. circular motion and single axis motion) is a particular case of critical motions [1], which often cannot be solved by a more general algorithm. Nonetheless, by explicitly taking the special properties of the motion into account, it is possible to estimate the motion with relative ease.

Many studies have been conducted in the recovery of circular motion. Early works ([2], [3]) obtained the sequence geometry using special calibration patterns and/or mechanical measurements. In [4], Fitzgibbon et al. studied the projective geometry of turntable motion and developed a method for the case of an uncalibrated camera with unknown rotation angles. Their method depends on the fundamental matrix and trifocal tensor between adjacent views for recovering the image invariants and rotation angles respectively. In [5], Jiang et al. showed that the image invariants can be recovered from the conic trajectories of the tracked points.

Their method then applies the Laguerre's formula to recover the rotation angle from the tracked points in two views. In order to fit a conic to the tracked points, their method requires tracking at least two points in five images. In [6], Jiang et al. showed that the image invariants can be recovered from the homography relating two points tracked in four images. Although their new method reduces the number of images in which a point must be tracked from five to four, this requirement still cannot be easily satisfied especially when the rotation angles are large (e.g., for a rotation step of 20° - 30° , it means a point must stay visible after a rotation of 60° - 90°). In [7] and [8], the authors considered the problem of recovering circular motion from silhouettes by exploiting outer epipolar tangents. Their methods involve nonlinear optimizations to locate the epipolar tangents and require the prior information of the camera intrinsics to recover the rotation angles. In [9], Zhang et al. extended the work of [7] by performing self-calibration. Note that all the aforementioned methods estimate the image invariants using multiple views, but the recovery of the rotation angle depends only on the information provided by adjacent views. Hence the solution is often not optimal. In [10], Hernández et al. proposed a silhouette coherence constraint which uses the joint effort of all the silhouettes to recover the camera motion as well as the focal length. In [11] and [12], the authors considered the problem of general motion recovery from silhouettes of a dynamic object. A RANSAC approach [11] is proposed for initializing the nonlinear optimization in locating the epipolar tangents. Nonetheless, such methods are not suitable for static objects.

This paper proposes using a 1D geometry to solve the circular motion problem from either point correspondences or silhouettes. Under circular motion, all the camera centers lie on a circle, which is the trajectory of the rotating camera center. The projection of the plane containing the camera centers onto the image plane gives a line known as the *horizon*. It follows that the images of the camera centers (i.e., the epipoles) must all lie on this line. This results in a 1D projective geometry. A 2×2 homography is introduced to relate the projections of the camera centers in two 1D views. It will be shown that both the imaged circular points of the motion plane and the rotation angle between the two views can be obtained directly from this homography. Unlike previous methods ([4], [5], [6], [7], [8], [9]) which only depend on adjacent views for recovering the rotation angle, the proposed method exploits information provided by all views to obtain an optimal solution for the homography and hence the rotation angle. Like [4], the proposed method only require point correspondences across three views, which is a more relaxed requirement than that of [5] and [6]. In [10], 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. This avoids potential problems of local minima often encountered in high-dimensional optimizations. Preliminary results of this work have been published in [13].

II. 1D PROJECTIVE CAMERA

This section gives a brief review on the 1D camera model (see [14] for details). Similar to a 2D projective camera which projects a point in \mathcal{P}^3 to a point in \mathcal{P}^2 , a 1D camera maps a point in \mathcal{P}^2 to a point in \mathcal{P}^1 . A 1D camera can be modelled by a 2×3 matrix [14] in the following form

$$\mathbf{P} = \mathbf{KR}(\rho)(\mathbf{I} \quad -\mathbf{t}), \quad (1)$$

where $\mathbf{K} = \begin{pmatrix} f & u_0 \\ 0 & 1 \end{pmatrix}$ is the calibration matrix with focal length f and principal point u_0 , \mathbf{t} is the position vector of the camera center, and $\mathbf{R}(\rho) = \begin{pmatrix} \cos \rho & \sin \rho \\ -\sin \rho & \cos \rho \end{pmatrix}$ is a rotation matrix describing the orientation of the camera with respect to the world coordinate system. The scene space for a 1D camera is a projective plane, and the two circular points \mathbf{I} and \mathbf{J} on the line at infinity \mathbf{L}_∞ are invariant under any similarity transformation. The imaged circular points \mathbf{i} and \mathbf{j} can be expressed in terms of the intrinsic parameters of a 1D camera as

$$\mathbf{i}, \mathbf{j} \sim \mathbf{P}(\pm j \quad 1 \quad 0)^T \sim e^{-j\rho} \begin{pmatrix} u_0 \pm jf \\ 1 \end{pmatrix}, \quad (2)$$

where $j^2 = -1$.

III. 1D HOMOGRAPHY IN CIRCULAR MOTION

Without loss of generality, consider a camera located along the negative Z -axis of the world coordinate system, with its image plane arbitrarily tilted with respect to the Y -axis (see fig. 1(a)). By considering the image of the X - Z plane, the camera can be simplified to a 1D camera with a projection matrix $\mathbf{P} = \mathbf{KR}(\rho)(\mathbf{I} \quad -\mathbf{t})$, where $\mathbf{t} = (0 \quad -t)^T$. Note that t determines the scale of the scene. Without loss of generality, let $t = 1$. If the camera is now being rotated about the Y -axis, its projection matrix can be written as $\mathbf{P}_i = \mathbf{KR}(\rho)(\mathbf{R}(\theta_i) \quad -\mathbf{t})$ where θ_i denotes the rotation angle. The trajectory of its center will be a unit circle lying on the X - Z plane and centered at the origin O .

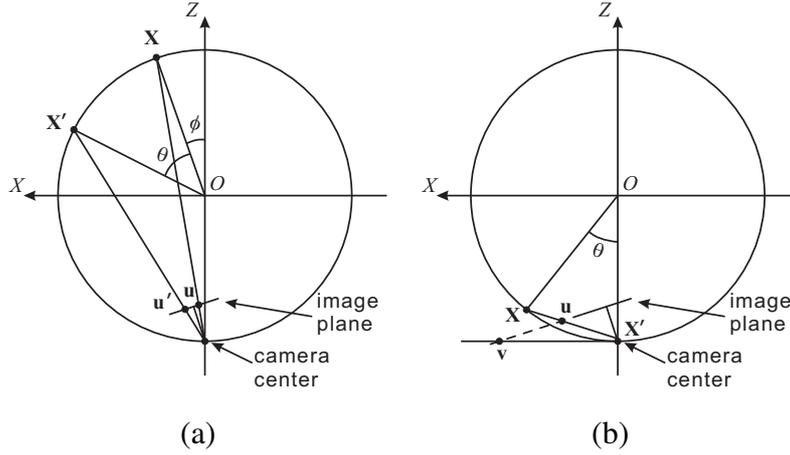


Fig. 1. (a) The relative positions of a 1D camera and the points \mathbf{X} and \mathbf{X}' in the coordinate system. (b) Special case where \mathbf{X}' overlaps with the camera center.

Proposition 1: The 2D points \mathbf{X} , \mathbf{X}' and the camera center are on a circle as shown in fig. 1(a). The rotation angle between the two points with respect to the center of the circle is θ . Their projections \mathbf{u} and \mathbf{u}' on the image are related by a 1D homography \mathbf{H} as $\mathbf{u}' \sim \mathbf{H}\mathbf{u}$ where $\mathbf{H} = \mathbf{K}\mathbf{R}(\theta/2)\mathbf{K}^{-1}$. If either \mathbf{X} or \mathbf{X}' overlaps with the camera center as illustrated in fig. 1(b), the corresponding projection will become the vanishing point \mathbf{v} of the tangent direction at the camera center.

Proof: Point \mathbf{X} has coordinates $(\sin \phi \ \cos \phi \ 1)^T$ and its image \mathbf{u} is given by

$$\begin{aligned} \mathbf{u} &\sim \mathbf{K}\mathbf{R}(\rho)(\mathbf{I} \ -\mathbf{t})(\sin \phi \ \cos \phi \ 1)^T \\ &\sim \mathbf{K}\mathbf{R}(\rho)(\sin \phi \ \cos \phi + 1)^T \\ &\sim \mathbf{K}\mathbf{R}(\rho)(\tan \frac{\phi}{2} \ 1)^T. \end{aligned} \quad (3)$$

Similarly, point \mathbf{X}' has coordinates $(\sin(\theta + \phi) \ \cos(\theta + \phi) \ 1)^T$ and its image \mathbf{u}' can be expressed as

$$\begin{aligned} \mathbf{u}' &\sim \mathbf{K}\mathbf{R}(\rho)(\tan(\frac{\theta+\phi}{2}) \ 1)^T \\ &\sim \mathbf{K}\mathbf{R}(\rho) \begin{pmatrix} \tan \frac{\theta}{2} + \tan \frac{\phi}{2} \\ 1 - \tan \frac{\theta}{2} \tan \frac{\phi}{2} \end{pmatrix} \\ &\sim \mathbf{K}\mathbf{R}(\rho) \begin{pmatrix} 1 & \tan \frac{\theta}{2} \\ -\tan \frac{\theta}{2} & 1 \end{pmatrix} \begin{pmatrix} \tan \frac{\phi}{2} \\ 1 \end{pmatrix} \\ &\sim \mathbf{K}\mathbf{R}(\rho) \begin{pmatrix} \cos \frac{\theta}{2} & \sin \frac{\theta}{2} \\ -\sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{pmatrix} \begin{pmatrix} \tan \frac{\phi}{2} \\ 1 \end{pmatrix} \end{aligned}$$

$$\sim \mathbf{KR}(\rho)\mathbf{R}(\theta/2)\left(\tan\frac{\phi}{2} \ 1\right)^T. \quad (4)$$

Substituting (3) into (4) gives

$$\mathbf{u}' \sim \mathbf{KR}(\rho)\mathbf{R}(\theta/2)\mathbf{R}^{-1}(\rho)\mathbf{K}^{-1}\mathbf{u} \sim \mathbf{Hu}, \quad (5)$$

where $\mathbf{H} = \mathbf{KR}(\theta/2)\mathbf{K}^{-1}$. The special case where \mathbf{X} or \mathbf{X}' overlaps with the camera center can be proved in a similar manner. ■

Note that \mathbf{H} has only three degrees of freedom (dof). Its eigenvalues are $\{e^{\pm j\theta/2}\}$, which are functions of the rotation angle θ , and its eigenvectors are the two imaged circular points \mathbf{i} and \mathbf{j} .

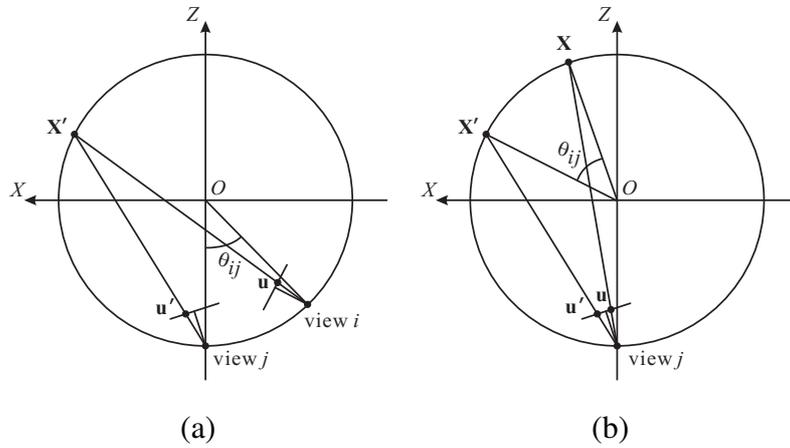


Fig. 2. (a) Two views and one point configuration. View j is obtained by rotating view i by an angle of θ_{ij} . (b) One view and two points configuration. The setup of \mathbf{X}' and the camera is the same as in (a).

Now consider the projection of a camera center \mathbf{X}' onto two views i and j as shown in fig. 2(a). The rotation angle between the two views is θ_{ij} . It is easy to see that the projection of \mathbf{X}' in view i in fig. 2(a) is equivalent to the projection of \mathbf{X} in view j in fig. 2(b) where the rotation angle between \mathbf{X} and \mathbf{X}' is θ_{ij} . From **Proposition 1**, \mathbf{u} and \mathbf{u}' satisfies $\mathbf{u}' \sim \mathbf{H}_{ij}\mathbf{u}$, where $\mathbf{H}_{ij} = \mathbf{KR}(\theta_{ij}/2)\mathbf{K}^{-1}$. It indicates that one such pair of \mathbf{u} and \mathbf{u}' induced by a camera center \mathbf{X}' provides one constraint on the homography \mathbf{H}_{ij} relating the two views. Given N images taken under circular motion, there will be N constraints for estimating \mathbf{H}_{ij} , and the rotation angle θ_{ij} and the imaged circular points $\{\mathbf{i}, \mathbf{j}\}$ can be obtained from the eigenvalues and eigenvectors of \mathbf{H}_{ij} respectively. Note that \mathbf{H}_{ij} is different for different view pair as it depends on the rotation angle θ_{ij} . **Algorithm 1** summarizes the procedure for recovering camera intrinsics as well as rotation angle between two views using the proposed 1D geometry.

Algorithm 1 Recovering camera intrinsics and rotation angle between view i and view j
for each view k in the circular motion sequence **do**

 Compute the projections $\{\mathbf{e}_{ki}, \mathbf{e}_{kj}\}$ of its camera center onto views i and j ;

end for

 Compute the homography \mathbf{H}_{ij} relating views i and j using the constraints $\mathbf{e}_{kj} \sim \mathbf{H}_{ij}\mathbf{e}_{ki}$;

 Compute the eigenvalues $\{e^{\pm j\theta_{ij}/2}\}$ and eigenvectors $\{\mathbf{i}, \mathbf{j}\}$ of \mathbf{H}_{ij} ;

 Compute the camera intrinsics from $\{\mathbf{i}, \mathbf{j}\}$;

 Compute the rotation angle θ_{ij} from $\{e^{\pm j\theta_{ij}/2}\}$;

Given two views, if the vanishing point \mathbf{v} as illustrated in **Proposition 1** is known, the epipoles $\{\mathbf{e}_{ij}, \mathbf{e}_{ji}\}$ of the two views and \mathbf{v} together can provide two independent constraints on \mathbf{H}_{ij} . Under this situation, one more constraint provided by another view is enough to determine \mathbf{H}_{ij} . The computation of \mathbf{H}_{ij} therefore requires a minimum of three views, which is the same as in [15]. Practically, \mathbf{H}_{ij} relating a view pair can be computed more accurately by incorporating constraints induced by as many views as possible.

IV. CIRCULAR MOTION ESTIMATION

Previous section describes how the proposed 1D homography can be used to solve the circular motion problem. The key step is the computation of epipoles from all view pairs, which is equivalent to computing the associated fundamental matrices. In this section, we will summarize the methods for computing the fundamental matrices from either point correspondences or silhouettes.

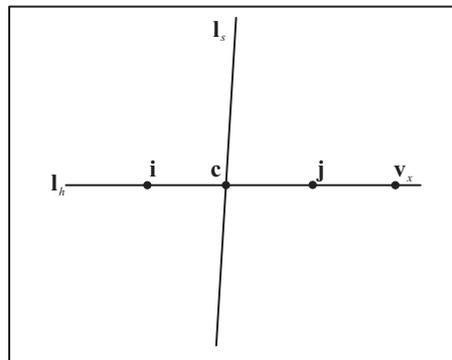


Fig. 3. Image invariants under circular motion.

Fig. 3 shows the image invariants under circular motion: \mathbf{l}_s is the image of the rotation axis; \mathbf{l}_h is the vanishing line (horizon) for the plane Π_h containing the camera centers; \mathbf{c} (given by the intersection of \mathbf{l}_s and \mathbf{l}_h) is the image of the intersection point between the rotation axis and Π_h ; \mathbf{v}_x is the vanishing point for the normal direction of the plane defined by the camera center and the rotation axis; \mathbf{i} and \mathbf{j} are the two imaged circular points of Π_h . The fundamental matrix \mathbf{F}_{ij} associated with views i and j can be parameterized in terms of these image invariants as [4], [7]

$$\mathbf{F}_{ij} = [\mathbf{v}_x]_{\times} + \mu_{ij}(\mathbf{l}_s \mathbf{l}_h^T + \mathbf{l}_h \mathbf{l}_s^T), \quad (6)$$

where $\mu_{ij} = \kappa \tan \frac{\theta_{ij}}{2}$, with θ_{ij} being the rotation angle between the two views and κ an unknown fixed scalar accounting for the different scales used in the homogeneous representation of these features. Hence for different view pairs, the fundamental matrices only differ by the scalar μ_{ij} which depends on the rotation angle.

Given point correspondences in a view pair, a fundamental matrix \mathbf{F} can be estimated robustly using a RANSAC method. The image invariants can be obtained from \mathbf{F} as follows [4]: \mathbf{v}_x can then be recovered from the skew-symmetric part of \mathbf{F} ; \mathbf{l}_h can be obtained as a line passing through the two epipoles; and \mathbf{l}_s can be extracted from the symmetric part of \mathbf{F} . Note that due to noise and wrong correspondences, the image invariants computed from different view pairs will in general be different. To assess the quality of the image invariants recovered from a particular view pair, the fundamental matrices for all other view pairs are estimated by fixing these invariants, and the mean symmetric transfer error [16] of the fundamental matrices is computed. For each view pair, a set of the image invariants is computed and the one that produces the smallest mean symmetric transfer error is then chosen.

Alternatively given a set of silhouettes, \mathbf{v}_x and \mathbf{l}_s can be obtained from the harmonic homography \mathbf{W} associated with the envelope of the silhouettes [7]. This harmonic homography also allows the location of outer epipolar tangents to silhouettes in view pairs. \mathbf{l}_h can then be obtained as a line robustly fitted to a set of epipoles obtained as the intersections of the epipolar tangents (see [7] for details).

The invariants \mathbf{v}_x , \mathbf{l}_s and \mathbf{l}_h thus obtained are used to initialize a nonlinear optimization which minimizes the symmetric transfer errors of the fundamental matrices for all view pairs. The parameters of the optimization are \mathbf{v}_x (2 dof), \mathbf{l}_s (2 dof), \mathbf{l}_h (1 dof, as $\mathbf{v}_x^T \mathbf{l}_h = 0$) and the scalars μ_{ij} . For point correspondences, the transfer error is given by the distance between a point and the epipolar line defined by its correspondent point. For silhouettes, the transfer

error is given by the distance between an epipolar tangent point and the epipolar line defined by its correspondent epipolar tangent point. After the optimization converges, the fundamental matrices for all view pairs can be obtained and the 1D homographies relating adjacent views can be computed.

V. METRIC RECONSTRUCTION

The imaged circular points \mathbf{i} and \mathbf{j} obtained from the eigenvectors of the 1D homography are 1D points lying on the horizon \mathbf{l}_h . Their 2D image coordinates can be recovered from their 1D coordinates using the parametrization of \mathbf{l}_h . They completely encode the metric information of the horizontal plane on which they lie, and provide two independent constraints on the imaged absolute conic ω in the form of [16]

$$\mathbf{i}^T \omega \mathbf{i} = \mathbf{j}^T \omega \mathbf{j} = 0. \quad (7)$$

The vanishing point \mathbf{v}_x and the imaged axis \mathbf{l}_s have a pole-polar relationship w.r.t. ω , and they provide another two constraints on ω in the form of

$$\mathbf{l}_s = \omega \mathbf{v}_x. \quad (8)$$

Note that, however, these two constraints are not completely independent of those induced by the imaged circular points. This is because \mathbf{v}_x and \mathbf{l}_s also encode metric information on the horizontal plane in the form of

$$\mathbf{c}^T \omega \mathbf{v}_x = 0, \quad (9)$$

where \mathbf{c} is the intersection of \mathbf{l}_s and \mathbf{l}_h . Hence \mathbf{v}_x , \mathbf{l}_s , \mathbf{i} and \mathbf{j} altogether provide only three independent constraints on ω . Nonetheless, by assuming the camera has unit aspect ratio and zero skew, ω will only have three dof and can thus be estimated. The 3×3 camera calibration matrix \mathbf{K} can then be recovered from ω using Cholesky decomposition.

With the estimated \mathbf{K} , \mathbf{v}_x , \mathbf{l}_h , \mathbf{l}_s and the set of rotation angles $\{\theta_{i,j}\}$ between adjacent views, the set of projection matrices compatible with the image sequence can be written as

$$\mathbf{P}_i = \mathbf{K} \begin{pmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 \end{pmatrix} \begin{pmatrix} \cos \theta_i & 0 & \sin \theta_i & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \theta_i & 0 & \cos \theta_i & t \end{pmatrix}, \quad (10)$$

where $\theta_1 = 0$ and $\theta_i = \sum_{n=1}^{i-1} \theta_{n,n+1}$ for $i = 2, 3, \dots, N$. The scalar t only affects the scale of the reconstruction and can therefore be chosen arbitrarily. The rotation matrix $\begin{pmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 \end{pmatrix}$

describes the orientation of the camera and its columns can be recovered as

$$\mathbf{r}_1 = \frac{\mathbf{K}^{-1}\mathbf{v}_x}{|\mathbf{K}^{-1}\mathbf{v}_x|}, \mathbf{r}_3 = \frac{\mathbf{K}^{-1}\mathbf{c}}{|\mathbf{K}^{-1}\mathbf{c}|}, \mathbf{r}_2 = \mathbf{r}_3 \times \mathbf{r}_1, \quad (11)$$

where $\mathbf{c} = \mathbf{I}_h \times \mathbf{I}_s$.

VI. EXPERIMENTAL RESULTS



(a) Dinosaur sequence



(b) Head sequence

Fig. 4. Turntable sequences used in the experiments. (a) Dinosaur sequence (image courtesy of the University of Hannover). (b) Head sequence.

The first experimental sequence (see fig. 4(a)) is the popular dinosaur sequence from the University of Hannover (obtained from the University of Oxford along with a file of point correspondences). It contains 36 images taken under circular motion with a constant rotation angle of 10° . The angular accuracy is about 0.05° [2]. Using the method described in Sec. IV, the image invariants obtained from the fundamental matrix relating views 22 and 23 were chosen automatically to initialize the nonlinear optimization for recovering the fundamental matrices of all view pairs. The homographies relating adjacent view pairs were then estimated using **Algorithm 1**. For each homography, the imaged circular points $\{\mathbf{i}, \mathbf{j}\}$ were obtained from its eigenvectors. Histograms for the real and imaginary components of the imaged circular points were constructed from the set of $\{\mathbf{i}, \mathbf{j}\}$ obtained from all the homographies (see fig. 5). The mean values of those in the bins with the highest frequency were then used to form the best estimate for $\{\mathbf{i}, \mathbf{j}\}$. Note that the distributions of the two histograms are

close to a Gaussian distribution, and one observes that the best solutions (denoted by the dashed lines) are very distinct. Fig. 6(a) shows the recovered rotation angles, and the rms error is just 0.073° .

Note that in [4], Fitzgibbon et al. achieved a rms error of 0.040° in the rotation angles for the same sequence. Their method first recovered the image invariants and rotation angles from the fundamental matrices and trifocal tensors of adjacent views respectively, and used these values to initialize a bundle adjustment that minimized the reprojection errors. The rms error achieved here is of the same order as in [4], which approaches the mechanical accuracy of the turntable. This demonstrates that by exploiting multiple view information in every step of the estimation as in the proposed method, a highly accurate solution can be obtained without carrying out any bundle adjustment. A visual hull model was constructed from the silhouettes using octree carving (see [8] for details), and a triangulated mesh was extracted from the octree using the standard marching cubes algorithm. Fig. 7 shows three views of the reconstructed model.

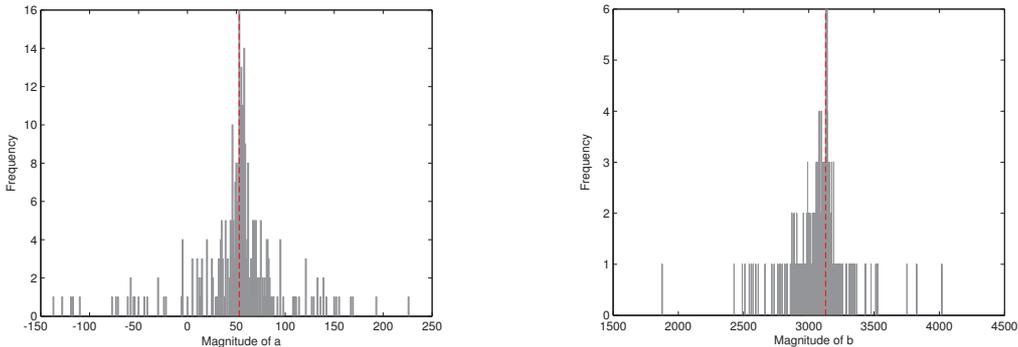


Fig. 5. Histograms for the real and imaginary components of the 1D imaged circular points $(a \pm jb, 1)^T$ estimated from the homographies of the dinosaur sequence. The dashed lines indicate the best solutions computed.

TABLE I

THE CAMERA INTRINSIC PARAMETERS ESTIMATED FROM THE HEAD SEQUENCE.

-	f	u_0	v_0
Ground-truth	2091.53	368.35	279.13
Estimated value	2078.89	343.56	376.29
Percentage error	0.60%	1.19%	4.65%

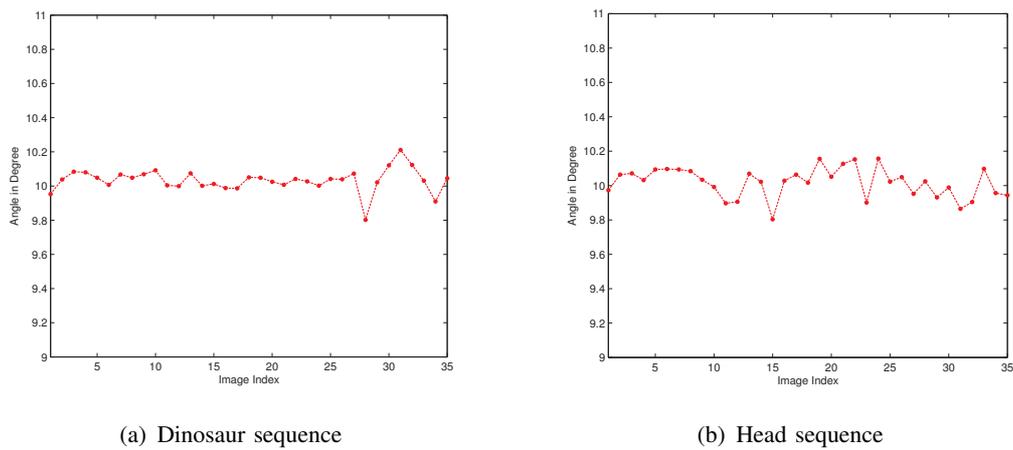


Fig. 6. Estimated rotation angles. (a) The rms error for the dinosaur sequence is 0.073° . (b) The rms error for the head sequence is 0.087° .



Fig. 7. Reconstructed dinosaur model using the camera geometry estimated from point correspondences.



Fig. 8. Reconstructed head model using the camera geometry estimated from the silhouettes.

The second experimental sequence (see fig. 4(b)) consists of 36 images of a head model rotating on an electronic turntable with a constant rotation angle of 10° . The angular accuracy is about 0.1° . The silhouettes of the head model were extracted using Cubic B-spline snake [17]. The fundamental matrices of all view pairs were estimated from the silhouettes using the method described in Sec. IV. The remaining steps of the estimation procedure were identical to those in the first experiment. Fig. 6(b) shows the recovered rotation angles, and the rms error is 0.087° which approaches the mechanical accuracy of the turntable. Note that in [9], Zhang et al. achieved a much larger rms error of 0.233° in the rotation angles for the same sequence. This is expected as they only considered adjacent views for the recovery of rotation angle. In order to improve the solution, Zhang et al. resorted to carrying out a bundle adjustment like approach to optimize all the image invariants and rotation angles simultaneously by minimizing the transfer errors of the epipolar tangents, and achieved a slightly improved rms error of 0.171° . This again demonstrates that the proposed multiple view approach can produce a highly accurate solution using silhouettes. Table I shows the camera intrinsic parameters estimated from the head sequence and the ground-truth values obtained by classical camera calibration technique [16] 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}_s will result in the same error in the location of the principal point. Fig. 8 shows different views of the reconstructed model.

VII. CONCLUSION

In this paper, we have investigated the circular motion estimation problem from a new viewpoint. By considering the horizon line which is the projection of the plane containing the camera centers, a 2D projective camera can be simplified to a 1D projective camera. We have shown that there exists a 1D homography relating the images of any two 1D projective cameras in the sequence. The eigenvectors of such a homography are shown to be the imaged circular points, and the eigenvalues are functions of the rotation angle between the two cameras.

We have demonstrated that this 1D geometry can be nicely applied to circular motion estimation. To compute the rotation angle between two particular views in the sequence, the

images of all other camera centers (i.e., the epipoles) are computed first. They are exploited to compute a homography relating the two views robustly. The eigenvectors and eigenvalues of the homography then give the imaged circular points and the rotation angle respectively. Experiments using both point correspondences and silhouettes, respectively, demonstrate that the recovered rotation angles achieve a very high precision. This is because the algorithm efficiently and effectively exploits the underlying multiple view information. The high quality of the reconstructed models demonstrate the practicality of the proposed algorithm.

REFERENCES

- [1] P. Sturm, "Critical motion sequences for monocular self-calibration and uncalibrated euclidean reconstruction," in *Proc. Conf. Computer Vision and Pattern Recognition*, 1997, pp. 1100–1105.
- [2] W. Niem, "Robust and fast modelling of 3d natural objects from multiple views," in *SPIE Proceedings - Image and Video Processing II*, vol. 2182, 1994, pp. 388–397.
- [3] R. Szeliski, "Rapid octree construction from image sequences," *Computer Vision, Graphics and Image Processing*, vol. 58, no. 1, pp. 23–32, July 1993.
- [4] 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, European Workshop SMILE'98*, ser. Lecture Notes in Computer Science, R. Koch and L. Van Gool, Eds., vol. 1506. Freiburg, Germany: Springer-Verlag, June 1998, pp. 155–170.
- [5] G. Jiang, H. T. Tsui, L. Quan, and A. Zisserman, "Geometry of single axis motions using conic fitting," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, pp. 1343–1348.
- [6] G. Jiang, L. Quan, and H. T. Tsui, "Circular motion geometry using minimal data," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 721–731, June 2004.
- [7] 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.
- [8] K.-Y. K. Wong and R. Cipolla, "Structure and motion from silhouettes," in *Proc. 8th Int. Conf. on Computer Vision*, vol. II, Vancouver, BC, Canada, July 2001, pp. 217–222.
- [9] H. Zhang, G. Zhang, and K.-Y. K. Wong, "Auto-calibration and motion recovery from silhouettes for turntable sequences," in *Proc. British Machine Vision Conference*, vol. I, Oxford, UK, September 2005, pp. 79–88.
- [10] C. Hernández, F. Schmitt, and R. Cipolla, "Silhouette coherence for camera calibration under circular motion," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 29, no. 2, pp. 343–349.
- [11] S. N. Sinha, M. Pollefeys, and L. McMillan, "Camera network calibration from dynamic silhouettes," in *Proc. Conf. Computer Vision and Pattern Recognition*, vol. I, 2004, pp. 195–202.
- [12] E. Boyer, "On using silhouettes for camera calibration," in *Proc. Asian Conf. on Computer Vision*, vol. I, 2006, pp. 1–10.
- [13] 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.
- [14] O. Faugeras, L. Quan, and P. Sturm, "Self-calibration of a 1D projective camera and its application to the self-calibration of a 2D projective camera," in *Proc. 5th European Conf. on Computer Vision*, June 1998, pp. 36–52.

- [15] M. Armstrong, A. Zisserman, and R. Hartley, "Self-calibration from image triplets," in *Proc. 4th European Conf. on Computer Vision*, ser. Lecture Notes in Computer Science, B. Buxton and R. Cipolla, Eds., vol. 1064. Cambridge, UK: Springer-Verlag, April 1996, pp. 3–16.
- [16] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, UK: Cambridge University Press, 2000.
- [17] R. Cipolla and A. Blake, "Surface shape from the deformation of apparent contours," *Int. Journal of Computer Vision*, vol. 9, no. 2, pp. 83–112, November 1992.