

Self-calibration from turn-table sequences in presence of zoom and focus

Xiaochun Cao ^{a,*}, Jiangjian Xiao ^b, Hassan Foroosh ^a, Mubarak Shah ^c

^a *Computational Imaging Lab, University of Central Florida, Orlando, FL 32816, USA*

^b *Sarnoff Corporation, Princeton, NJ 08540, USA*

^c *Computer Vision Lab, University of Central Florida, Orlando, FL 32816, USA*

Received 1 April 2005; accepted 9 January 2006

Available online 31 March 2006

Abstract

This paper proposes a novel method, using constant inter-frame motion, for self-calibration from an image sequence of an object rotating around a single axis with varying camera internal parameters. Our approach makes use of the facts that in many commercial systems rotation angles are often controlled by an electromechanical system, and that the inter-frame essential matrices are invariant if the rotation angles are constant but not necessary known. Therefore, recovering camera internal parameters is possible by making use of the equivalence of essential matrices which relate the unknown calibration matrices to the fundamental matrices computed from the point correspondences. We also describe a linear method that works under restrictive conditions on camera internal parameters, the solution of which can be used as the starting point of the iterative non-linear method with looser constraints. The results are refined by enforcing the global constraints that the projected trajectory of any 3D point should be a conic after compensating for the focusing and zooming effects. Finally, using the bundle adjustment method tailored to the special case, i.e., static camera and constant object rotation, the 3D structure of the object is recovered and the camera parameters are further refined simultaneously. To determine the accuracy and the robustness of the proposed algorithm, we present the results on both synthetic and real sequences.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Constant inter-frame motion; Self-calibration; Turn-table; Conic

1. Introduction

Acquiring 3D models from circular motion sequences, particularly turn-table sequences, has been widely used by computer vision and graphics researchers, e.g., [36,31,4,35], since these methods are simple and robust. Generally, the whole reconstruction procedure includes: first, the determination of camera positions at different viewpoints or, equivalently, the different positions of the rotating device; second, the detection of object boundaries or silhouettes; third, the extraction of a visual hull as the surface model from a volume representation [21]. Fitzgibbon et al. [9] extended the analysis of the circular motion to recover unknown rotation angles from uncalibrated

image sequences based on a projective geometry approach and multi-view geometric constraints. [28,29] recovered the circular motion by using surface profiles. Wong et al. [42] also presented a method for camera calibration using surfaces of revolution, which is related to circular motion since an object placed on a turn-table spans a surface of revolution. Recently, Jiang et al. [19,18] developed new methods to compute single axis motion by either fitting the conic to the locus of the tracked points in at least five images or computing a plane homography from a minimal of two points in four images. Colombo et al. [6] improved the approach [42] in which the calibration of a natural camera, a pinhole camera with zero skew and unit aspect ratio [22], requires the presence of two different surfaces of revolutions in the same view. In addition, the method [6] relaxes the conditions claimed by Bougnoux [19], that three ellipses are needed to compute the imaged circular points.

* Corresponding author.

E-mail address: xc CAO@cs.ucf.edu (X. Cao).

However, most of these methods deal with the case in which a static camera with fixed internal parameters views an object rotating on a turn-table (Fig. 1), and utilize the fixed image entities of the circular motion. These fixed image entities (Fig. 3) include two lines: one is the image of the rotation axis \mathbf{l}_s , a line of fixed points, while the other one, called the horizon line \mathbf{l}_∞ , is the image of the vanishing line of the horizontal planes, e.g., π_1 and π_2 (π_A and π_B in Fig. 1). Unlike the image of the rotation axis, the horizon line is a fixed line, but not a line of fixed points. Under the assumption of the fixed camera internal parameters, the image of the absolute conic is fixed for a rigid motion. Therefore, there are two points, \mathbf{i} and \mathbf{j} , located at the intersection of the absolute conic with the horizon line, and remaining fixed in all images. Actually, these two fixed points are the images of the two circular points on the horizontal planes, and can be found by the intersections of conic loci of corresponding points since the trajectories of space points are circles in 3D space and intersect in the circular points on the plane at infinity. However, these entities are fixed only when the camera has fixed internal parameters. For example, the projected trajectory of a 3D point is not a conic any more when the camera's internal parameters are varying (Fig. 2B).

In this paper, we concentrate on the situation where the stationary camera is free to zoom and focus, and assume that in many commercial systems, rotation angles are often controlled by an electromechanical system [36,31,35,26], i.e., they are constant and even known. We show that the inter-frame essential matrices are invariant if the rotation angle is constant but not necessary known and, therefore, recovering camera internal parameters is possible by making use of the equivalence of essential matrices. We also

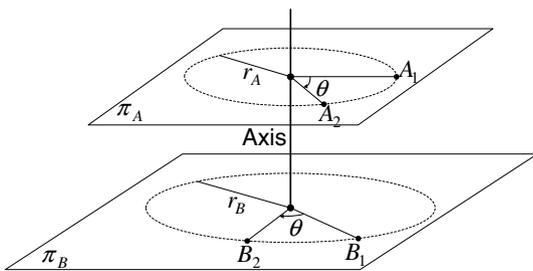


Fig. 1. The geometric configuration in 3D space. The space points A_i and B_i are circularly moving around the fixed rotation axis on two different planes π_A and π_B . To aid in visualization, we assume that the rotation axis is vertical, so that the 3D points rotate in horizontal planes. In our case, the relative angle between views i and $i + 1$ are constant and denoted by θ .

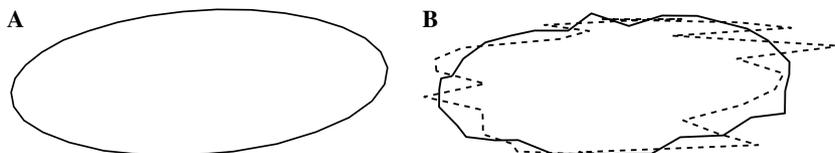


Fig. 2. Both (A and B) are the projected trajectories across 90 views of a typical scene point under the configuration described in Section 5.1. Different from (A) which uses a fixed camera, the solid curve in (B) has focal lengths chosen with a mean value of 1000 pixels and a standard deviation $\sigma_f = 10$ pixels, while the dashed curve in (B) has $\sigma_f = 100$ pixels.

introduce a linear method that works under restrictive conditions on camera internal parameters, such as known camera skew, aspect ratio and principal point, the solution of which can be used as the starting point of the iterative non-linear methods with looser constraints. The results are optimized by enforcing the global constraints that the projected trajectories of 3D points should be conics after compensating the focusing and zooming effects. Finally, using the bundle adjustment method tailored to the special case, i.e., static camera and constant rotation angle, the 3D structure of the object is recovered and the camera parameters are further refined simultaneously.

The rest of the paper is organized as follows. We start with the description of previous work on self-calibration in the next section, and then present the preliminaries of the pinhole camera model and epipolar geometry in Section 3. In Section 4, a practical calibration method, making use of constant inter-frame motion, is developed. A simple linear solution is also given which can be used as an initialization. The method is then validated through the experiments on both computer simulation and real data in Section 5. Finally, Section 6 concludes the paper with perspective of this work.

2. Related work

Self- (or auto-) calibration is the process of determining internal camera parameters directly from a set of uncalibrated images. This differs from conventional calibration, where the camera internal parameters are determined from the image of a known calibration grid, e.g., [41,46,14], or properties of the scene, such as vanishing points of orthogonal directions [5,22]. In self-calibration the metric properties of the cameras are determined directly from constraints on the internal and/or external parameters.

The first self-calibration method, originally introduced into computer vision by Faugeras et al. [8], involves the use of the Kruppa equations. Kruppa equations embrace the constraints that camera intrinsic parameters should be consistent with the underlying projective geometry of a sequence of images. Luong and Faugeras [25] have shown that the Kruppa equations are equivalent to the Trivedi constraints [40] and the Huang and Faugeras constraints [17,11], although the equivalence does not mean that they will produce the same results when used in self-calibration algorithms. Algorithms for computing the focal lengths of two cameras given the corresponding fundamental matrix

and knowledge of the remaining intrinsic parameters are provided by Hartley [11] and Bougnoux [3]. Mendonça [27] generalizes the results [11,3] for an arbitrary number of cameras and introduces a built-in method for the detection of critical motions for each pair of images in the sequence. Thorough analyses of critical motions which would result in ambiguous solutions by Kruppa-based methods are described in [34] and [20].

An alternative direct method for self-calibration estimates the absolute dual quadric, introduced by Triggs [37], over many views. The basic idea is to transfer a constraint on the dual image of absolute conic to a constraint on the absolute dual quadric and, as a result, to determine the matrix representing the absolute dual quadric, from which a rectifying 3D homography can be decomposed that transforms from a projective to metric reconstruction. Heyden and Astrom [15] showed that metric reconstruction was possible knowing only skew and aspect ratio, and [32,16] showed that zero skew alone was sufficient. In addition, Pollefeys et al. [32] developed a practical method for self-calibration of multiple cameras with varying intrinsic parameters, and showed results for real sequences.

Special motions can also be used for self-calibration. Refs. [1,33] solved the self-calibration of a rotating and zooming camera using the infinite homography constraint. Before their work, [12] solved the special case where the camera's internal parameters remain constant throughout the sequence. Frahm and Koch [10] showed it was also possible to solve the problem of generally moving camera with varying intrinsics but known rotation information. Triggs [38] provided a solution for self-calibration from scene planes when the internal parameters are constant. Zisserman et al. [47] presented a method for self-calibration of a stereo rig. For planar motion of a monocular camera, the original method was published by Armstrong et al. [2].

Different from the existing self-calibration methods, our algorithm makes use of constant inter-frame camera motion, i.e., a 3D rigid displacement described by the relative orientation and translation of two cameras. Inter-frame essential matrices are invariant in this case, since the essential matrix depends only on relative camera motion. This paper develops a novel linear algorithm for estimating the relative focal lengths of multiple cameras. The input of the algorithm is only a set of fundamental matrices, and therefore there is no need for projective bundle adjustment before self-calibration.

3. Preliminaries

3.1. Pinhole camera model

A real world camera can be modeled by a pinhole or perspective camera model. A pinhole camera, based on the principle of collinearity, projects a region of \mathbb{R}^3 lying in front of the camera into a region of the image plane \mathbb{R}^2 . As is well known, a 3D point $\mathbf{X} = [X Y Z 1]^T$ and its

corresponding projection $\mathbf{x} = [x y 1]^T$ in the image plane are related via a 3×4 matrix \mathbf{P} as

$$\mathbf{x} \sim \underbrace{\mathbf{K}[\mathbf{R}|\mathbf{t}]}_{\mathbf{P}} \mathbf{X}, \quad \mathbf{K} = \begin{bmatrix} f & \gamma f & u_0 \\ 0 & \lambda f & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (1)$$

where \sim indicates equality up to multiplication by a non-zero scale factor, \mathbf{R} is a 3×3 orthonormal rotation matrix, \mathbf{t} is a translation vector, and \mathbf{K} is a non-singular 3×3 upper triangular matrix containing the five camera intrinsic parameters: the focal length f , the aspect ratio λ , the principal point $\mathbf{u}_0 = [u_0 v_0]^T$ and the skew factor γ accounting for non-rectangular pixels. The intrinsic parameters in \mathbf{K} define the internal imaging geometry of the camera, while the extrinsic parameters (\mathbf{R} and \mathbf{t}) relate the world coordinate frame to that of the camera.

3.2. Epipolar geometry

Among the geometric properties of a set of two cameras, the widely known property in computer vision is the epipolar geometry. It is algebraically represented by the fundamental matrix \mathbf{F} ,

$$\mathbf{x}'^T \mathbf{F} \mathbf{x} = 0, \quad (2)$$

where \mathbf{x} and \mathbf{x}' are a pair of corresponding points. \mathbf{F} is also known as the uncalibrated version of the essential matrix, \mathbf{E} , because

$$\mathbf{F} = \mathbf{K}'^{-T} \mathbf{E} \mathbf{K}^{-1}, \quad (3)$$

where \mathbf{K}' and \mathbf{K} are matrices representing the internal calibration parameters of the stereo cameras. In general, both matrices \mathbf{F} and \mathbf{E} have a rank two. For an arbitrary stereo pair, the rank two constraint is the only constraint on the \mathbf{F} , and thus \mathbf{F} generally has seven degrees of freedom. The essential matrix, $\mathbf{E} = [\mathbf{t}]_{\times} \mathbf{R}$, where the rotation matrix \mathbf{R} and the translation vector \mathbf{t} represent the motion between the two positions of the cameras, has only five degrees of freedom: both \mathbf{R} and \mathbf{t} have three degrees of freedom, but there is an overall scale ambiguity. Note that the essential matrix, \mathbf{E} , must have a zero singular value and two equal non-zero singular values, which is also known as the Huang and Faugeras constraint.

4. Solving for self-calibration

It is well known that when the projective image measurements alone are used it is only possible to recover the scene up to an unknown projective transformation [7,11]. Additional scene, motion or calibration constraints are required for a metric or Euclidean reconstruction. We also use the constraints on camera internal parameters similar to previous self-calibration methods. However, the main difference is that constant inter-frame motion is exploited in this paper.

4.1. Self-calibration using constant inter-frame motion

In this section, we first elaborate on the equivalence between the scenario where the camera is static and the object is rotating around an unknown axis, and the case where the object is fixed while the camera is both rotating and translating.

The i th camera projection matrix can be factorized as $\mathbf{P}_i = \mathbf{K}_i[\mathbf{R}|\mathbf{t}]$, since our camera is static and thus has the same \mathbf{R} and \mathbf{t} through all views. We are interested in the case where the relative rotation angle between views i and $i+1$ are constant (Fig. 1). Let \mathbf{R}_θ denotes the 3×3 orthonormal rotation matrix of the object, which has only one degree of freedom from θ . Therefore, after applying the rotation, the projective transformation of the i th frame becomes $\mathbf{K}_i[\mathbf{R}\mathbf{R}_\theta^i|\mathbf{t}]$. This means that the new camera center is located at $-(\mathbf{R}\mathbf{R}_\theta^i)^T\mathbf{t}$, with new rotation matrix $\mathbf{R}\mathbf{R}_\theta^i$. Note that the equality is also true for non-constant rotations.

Then let us rewrite the i th camera matrix such that the world origin coincides with the i th camera center,

$$\mathbf{x}_i \sim \mathbf{K}_i[\mathbf{R}\mathbf{R}_\theta^i|\mathbf{t}]\mathbf{X} = \mathbf{K}_i[\mathbf{I}|\mathbf{0}]\begin{bmatrix} \mathbf{R}\mathbf{R}_\theta^i & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}\mathbf{X}.$$

For the $(i+1)$ th view, we can derive that

$$\mathbf{x}_{i+1} \sim \mathbf{K}_{i+1}[\mathbf{R}\mathbf{R}_\theta\mathbf{R}^T|(\mathbf{I} - \mathbf{R}\mathbf{R}_\theta\mathbf{R}^T)\mathbf{t}]\begin{bmatrix} \mathbf{R}\mathbf{R}_\theta^i & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}\mathbf{X}.$$

Therefore, we obtain the essential matrix as

$$\mathbf{E}_{i,i+1} = [(\mathbf{I} - \mathbf{R}\mathbf{R}_\theta\mathbf{R}^T)\mathbf{t}]_{\times} \mathbf{R}\mathbf{R}_\theta\mathbf{R}^T, \quad (4)$$

where $[\cdot]_{\times}$ is the notation for the skew symmetric matrix characterizing the cross product. Since \mathbf{R} , \mathbf{t} and \mathbf{R}_θ are all constants, the inter-frame essential matrices are invariant. It is possible to use the invariance property of the inter-frame essential matrices to solve for the camera matrices, \mathbf{K}_i , given the set of fundamental matrices that encapsulate the intrinsic projective geometry between two views. The equality of essential matrices can be expressed as

$$\mathbf{K}_{i+2}^T \mathbf{F}_{i+1,i+2} \mathbf{K}_{i+1} \sim \mathbf{K}_{i+1}^T \mathbf{F}_{i,i+1} \mathbf{K}_i, \quad (5)$$

where $\mathbf{F}_{i,i+1}$ is the fundamental matrix between the i th and $(i+1)$ th views. A solution may be obtained using a non-linear least squares algorithm. The parameters to be computed are the unknown intrinsic parameters of each calibration matrix \mathbf{K}_i and the following criterion should be minimized:

$$\min \sum_{i=1}^{n-2} \|\mathbf{K}_{i+2}^T \mathbf{F}_{i+1,i+2} \mathbf{K}_{i+1} - \mathbf{K}_{i+1}^T \mathbf{F}_{i,i+1} \mathbf{K}_i\|_F^2, \quad (6)$$

where the subscript F indicates the use of the Frobenius norm, and $\mathbf{K}_{i+2}^T \mathbf{F}_{i+1,i+2} \mathbf{K}_{i+1}$ and $\mathbf{K}_{i+1}^T \mathbf{F}_{i,i+1} \mathbf{K}_i$ are both normalized to have unit Frobenius norm. It is also important to enforce that two of the essential matrices' singular values are equal and the third one is zero. In our implementation,

we found that the final results are sensitive to errors in the computed fundamental matrices. Therefore, we recommend the methods [24,45] that minimize the reprojection errors to compute the fundamental matrices between pairs of images.

4.2. Linear approach

To obtain an initial starting point, we propose a linear approach to compute an approximate solution for the calibration. This linear solution can be obtained by assuming zero skew, known aspect ratio and the principal point. For instance, we set the principal point \mathbf{u}_0 to $(0,0)$, and the aspect ratio to one. These assumptions result as follows:

$$\mathbf{K}_{i+1}^T \mathbf{F}_{i,i+1} \mathbf{K}_i \sim \begin{bmatrix} f_{i+1}F_i^1f_i & f_{i+1}F_i^2f_i & f_{i+1}F_i^3 \\ f_{i+1}F_i^4f_i & f_{i+1}F_i^5f_i & f_{i+1}F_i^6 \\ f_iF_i^7 & f_iF_i^8 & F_i^9 \end{bmatrix}, \quad (7)$$

where f_i and f_{i+1} are focal lengths of i th and $(i+1)$ th cameras respectively, and F_i^k denotes, in a row-major order vector, the components of $\mathbf{F}_{i,i+1}$. From the Eq. (7), and the equivalence property of the essential matrix, one obtains

$$\begin{aligned} f_iF_{i-1}^j f_{i-1} &= \lambda_i f_{i+1} F_i^j f_i, & j &= 1, 2, 4, 5, \\ f_iF_{i-1}^j &= \lambda_i f_{i+1} F_i^j, & j &= 3, 6, \\ f_{i-1}F_{i-1}^j &= \lambda_i f_i F_i^j, & j &= 7, 8, \\ \lambda_{i-1} F_{i-1}^9 &= \lambda_{i,i+1} F_i^9, \end{aligned} \quad (8)$$

where $\lambda_i \in \mathbb{R}$. In the cases where the elements F_i^9 of the fundamental matrices are not zero, the focal lengths, f_{i-1} , f_i , and f_{i+1} , can be obtained from equations in (8) by the left null space of the following matrix:

$$\begin{bmatrix} F_i^9 F_{i-1}^1 & F_i^9 F_{i-1}^2 & 0 & F_i^9 F_{i-1}^4 & F_i^9 F_{i-1}^5 & 0 & F_i^9 F_{i-1}^7 & F_i^9 F_{i-1}^8 \\ 0 & 0 & F_i^9 F_{i-1}^3 & 0 & 0 & F_i^9 F_{i-1}^6 & -F_{i-1}^9 F_i^7 & -F_{i-1}^9 F_i^8 \\ -F_{i-1}^9 F_i^1 & -F_{i-1}^9 F_i^2 & -F_{i-1}^9 F_i^3 & -F_{i-1}^9 F_i^4 & -F_{i-1}^9 F_i^5 & -F_{i-1}^9 F_i^6 & 0 & 0 \end{bmatrix}. \quad (9)$$

When more images are available, the linear estimation of the focal lengths $(f_i)_{i=1}^n$ can be given by the null space of the $\mathcal{A}_{8(n-2) \times n}$, where

$$\begin{aligned} \mathcal{A}_{8i-7} &= [\mathbf{0}_{(i-1) \times 1}^T, F_{i+1}^9 F_i^1, 0, -F_i^9 F_{i+1}^1, \mathbf{0}_{(n-2-i) \times 1}^T], \\ \mathcal{A}_{8i-6} &= [\mathbf{0}_{(i-1) \times 1}^T, F_{i+1}^9 F_i^2, 0, -F_i^9 F_{i+1}^2, \mathbf{0}_{(n-2-i) \times 1}^T], \\ \mathcal{A}_{8i-5} &= [\mathbf{0}_{i \times 1}^T, F_{i+1}^9 F_i^3, -F_i^9 F_{i+1}^3, \mathbf{0}_{(n-2-i) \times 1}^T], \\ \mathcal{A}_{8i-4} &= [\mathbf{0}_{(i-1) \times 1}^T, F_{i+1}^9 F_i^4, 0, -F_i^9 F_{i+1}^4, \mathbf{0}_{(n-2-i) \times 1}^T], \\ \mathcal{A}_{8i-3} &= [\mathbf{0}_{(i-1) \times 1}^T, F_{i+1}^9 F_i^5, 0, -F_i^9 F_{i+1}^5, \mathbf{0}_{(n-2-i) \times 1}^T], \\ \mathcal{A}_{8i-2} &= [\mathbf{0}_{i \times 1}^T, F_{i+1}^9 F_i^6, -F_i^9 F_{i+1}^6, \mathbf{0}_{(n-2-i) \times 1}^T], \\ \mathcal{A}_{8i-1} &= [\mathbf{0}_{(i-1) \times 1}^T, F_{i+1}^9 F_i^7, -F_i^9 F_{i+1}^7, \mathbf{0}_{(n-1-i) \times 1}^T], \\ \mathcal{A}_{8i} &= [\mathbf{0}_{(i-1) \times 1}^T, F_{i+1}^9 F_i^8, -F_i^9 F_{i+1}^8, \mathbf{0}_{(n-1-i) \times 1}^T], \end{aligned}$$

here \mathcal{A}_j denotes the j th row of the matrix $\mathcal{A}_{8(n-2) \times n}$.

From the null space of $\mathcal{A}_{8(n-2) \times n}$, we have a solution for the estimation of the focal lengths up to a global scale κ . There are several options to compute κ . One possibility is to pick κ that best enforces the Huang–Faugeras constraint

of equality of singular values of the Essential matrices [3,27] for non-critical motion sequences. In our implementation, we compute the ratios ρ_i of f_i over f_1 , and thus compensate the effects of varying focal lengths for each image point, \mathbf{x}_{ij} of the i th image, as $\hat{\mathbf{x}}_{ij} = \mathbf{x}_{ij}/\rho_i$. We then use the existing method [19] to obtain an initial solution of the focal length, f_1 , of the first image. Other focal lengths, f_i , can be simply computed as $f_i = \rho_i * f_1$.

However, the above method using Eq. (10) will fail when the element F_i^9 of the fundamental matrices are zeros. Note that this special case is easy to be detected since all the essential matrices are equal up to an unknown scale. In other words, we only need to check the element, F_i^9 , of one inter-frame fundamental matrix \mathbf{F} . It is shown by Mendonça [27] that whenever the optical axes of the i th and $(i+1)$ th cameras intersect, F_i^9 is equal to zero, since, in this case, the principal points must satisfy the epipolar constraint, i.e., $\mathbf{u}_0^{i+1T} \mathbf{F}_{i,i+1} \mathbf{u}_0^i = 0$, where $\mathbf{u}_0^{i+1} = \mathbf{u}_0^i = [001]^T$. While the case where the optical axes of the two cameras intersect is a critical motion for Kruppa-based methods [20], the constant inter-frame motion is still able to provide enough constraints for the computation of the relative camera motion. For example, one can check the following equations based on the equivalence of the inter-frame essential matrices:

$$F_i^n F_{i+1}^m f_{i+1} - F_i^m F_{i+1}^n f_i = 0, \quad m = 1, 2, 4, 5; \quad n = 3, 6, \quad (10)$$

$$F_i^n F_{i+1}^m f_{i+2} - F_i^m F_{i+1}^n f_{i+1} = 0, \quad m = 1, 2, 4, 5; \quad n = 7, 8. \quad (11)$$

Similarly to $\mathcal{A}_{8(n-2) \times n}$, we can build another matrix $\mathcal{A}'_{16(n-2) \times n}$, whose null space provides the solution of the focal lengths up to a scale κ . In this case, we still can use existing method [19] to obtain κ , although the Huang–Faugeras constraint will fail. Note that, as in the case of F_i^9 , if any one of the components $F_{i+1}^m, m = 1, \dots, 8$ is zero, it can be easily identified from the fundamental matrix of any image pair and excluded in our computations.

4.3. Two-stage optimization

4.3.1. Conic enforcement after compensation

We first improve the results by enforcing the global constraint, originally proposed in Ref. [19], that the projected trajectories of 3D points should be conics after compensating for the focusing and zooming effects. The main advantage of enforcing conic constraint is that it is intrinsically a multiple view approach as all geometric information from the whole sequence is nicely summarized in the conics as argued by Jiang et al. [19]. Practically, conic enforcement efficiently improves the results as shown in Section 5.1.

The image points $x_{i,j}$ can be compensated as

$$\hat{\mathbf{x}}_{ij} = \mathbf{K}_1 \mathbf{K}_i^{-1} \mathbf{x}_{ij}, \quad (12)$$

where \mathbf{K}_1 is the camera calibration matrix of a reference view, which is without loss of generality assumed to be the first view. After the compensation, the conic property of the correspondence tracks are fully restored, where the

entities related to the conic and plane motion become fixed again, such as the rotation axis \mathbf{l}_s , horizon line \mathbf{l}_∞ , and circular points, \mathbf{i} and \mathbf{j} , shown in Fig. 3. Given a conic \mathbf{C}_j , there is a homography \mathbf{H}_j to map \mathbf{C}_j into an unit circle, \mathbf{O} , such that

$$\mathbf{O} = \mathbf{H}_j^{-T} \mathbf{C}_j \mathbf{H}_j^{-1},$$

where \mathbf{H}_j can be parameterized as [23,19]

$$\mathbf{H}_j = \begin{bmatrix} s_j & 0 & -\mu_j \\ 0 & s_j & -v \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\beta} & -\frac{\alpha}{\beta} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_1 & l_2 & l_3 \end{bmatrix}, \quad (13)$$

where s_j is the scale that enforces the radius of the circle \mathbf{O} to be one, $(\mu_j, v, 1)$ is the pole, o_j (Fig. 3), of \mathbf{l}_∞ with respect to the conic \mathbf{C}_j , $(\alpha + i\beta, 1, 0)^T$ are the circular points \mathbf{i} and \mathbf{j} , and $(l_1, l_2, l_3)^T$ is the vanishing line \mathbf{l}_∞ . Basically, the parameters l_1, l_2, l_3, α and β are fixed, since circular points and vanishing line are fixed entities. In addition, v can be assumed to be constant in that the pole is constrained by the fixed rotation axis \mathbf{l}_s . Given m 3D points, therefore, a total of $6 + 2m$ parameters needs to be estimated by minimizing the following MLE function:

$$\arg \min_{\Theta_1} \sum_{i=1}^n \sum_{j=1}^m d^2(\hat{\mathbf{x}}_{ij}, \mathbf{C}_j), \quad (14)$$

where $\Theta_1 = \{l_1, l_2, l_3, \alpha, \beta, v, s_j, \mu_j\}$, $d^2(\hat{\mathbf{x}}_{ij}, \mathbf{C}_j)$ are distance function from point, $\hat{\mathbf{x}}_{ij}$, to conic \mathbf{C}_j , defined as

$$d^2(\hat{\mathbf{x}}_{ij}, \mathbf{C}_j) = \begin{cases} \frac{(\hat{\mathbf{x}}_{ij} \mathbf{C}_j \hat{\mathbf{x}}_{ij})^2}{4((\mathbf{C}_j \hat{\mathbf{x}}_{ij})_1^2 + (\mathbf{C}_j \hat{\mathbf{x}}_{ij})_2^2)}, & \text{if } (\hat{\mathbf{x}}_{ij} \in \mathbf{C}_j), \\ 0, & \text{otherwise,} \end{cases} \quad (15)$$

where $(\mathbf{C}_j \hat{\mathbf{x}}_{ij})_i$ is the i th component of $\mathbf{C}_j \hat{\mathbf{x}}_{ij}$. After substituting $\hat{\mathbf{x}}_{ij}$ with Eq. (12), we obtain

$$\arg \min_{\Theta_1, \mathbf{K}_i} \sum_{i=1}^n \sum_{j=1}^m d^2(\mathbf{K}_1 \mathbf{K}_i^{-1} \mathbf{x}_{ij}, \mathbf{C}_j). \quad (16)$$

This cost function is minimized using the standard Levenberg–Marquardt algorithm [30].

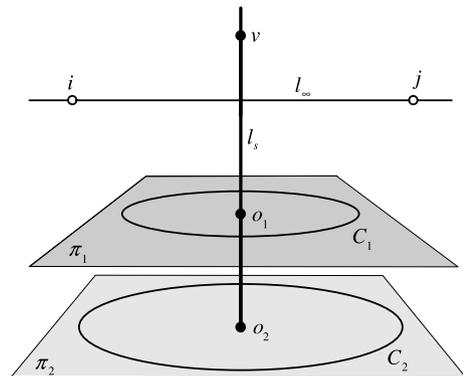


Fig. 3. The entities related to the geometry of a single axis motion observed by a fixed camera. The fixed entities include the rotation axis, \mathbf{l}_s , the horizon line \mathbf{l}_∞ , and the circular points, \mathbf{i} and \mathbf{j} , the vanishing point, v of the rotation axis. The projection, o_j , of the center of one circle \mathbf{C}_i as $o_i = \mathbf{C}_i^{-1} \mathbf{l}_\infty$.

The $6 + 2m$ parameters are initialized as follows. First, the focusing and zooming effects are compensated by using the Eq. (12). Second, each conic is fitted to corresponding points from at least five views. Third, the pole of each conic with respect to the vanishing line is calculated as shown in Fig. 3, and the point on the rotation axis l_s which is nearest to the pole, o_i , is used to estimate the initial value of μ_i . Fourth, the radius of each 3D circle, transformed from each imaged conic, determines the initial value of s_j . Finally, each conic is mapped to a unit circle with center at the origin and the points on the conic is mapped to the points near the unit circle for the optimal procedure.

4.3.2. Reconstruction using bundle adjustment

After the refined camera matrices are obtained, the 3D points or structure can be determined by triangulation from two or more views [13]. To minimize the overall reconstruction errors and to further refine the estimated camera parameters, here we use a bundle adjustment approach [39] explicitly enforcing another available constraint: static camera and constant rotation angle. Given n images and m corresponding image points, the maximum likelihood estimate (MLE) can be obtained

$$\arg \min_{\Theta_2} \sum_{i=1}^n \sum_{j=1}^m d^2(\mathbf{x}_{ij}, \mathbf{K}_i[\mathbf{R}\mathbf{R}_\theta^i|\mathbf{t}]\mathbf{X}_j), \quad (17)$$

where $\Theta_2 = \{\mathbf{K}_i, \mathbf{R}, \mathbf{R}_\theta, \mathbf{t}, \mathbf{X}_j\}$, and $d^2(\cdot, \cdot)$ is the distance function between the image measurement \mathbf{x}_{ij} and the projection of the estimated 3D point \mathbf{X}_j . Nevertheless, as shown in [9,13], the circular motion has the fundamental ambiguity on the vertical apex, v (Fig. 3), which causes unknown ratios between the horizontal and vertical direction for the 3D reconstruction. In order to remove this ambiguity, we assume a unit aspect ratio and zero skew for all cameras and specify a reasonable choice of the aspect ratio of the object.

Similar to most other self-calibration methods, such as [32,1], we also have difficulty to precisely estimate the principal points because the principal point \mathbf{u}_0 is known to be a poorly constrained parameter which tends to fit to noise. In practice, we notice that \mathbf{u}_0 is mostly located close to the image center. Using this prior information, we model the expectation of the principal point as a Gaussian distribution, which has its mean at the image center $\bar{\mathbf{u}}_0$, with the uncertainties $\Sigma_{\mathbf{u}_0} = \text{diag}(\sigma_u^2, \sigma_v^2)$. Therefore, we apply the prior information of principal point on Eq. (17). Consequently, the bundle adjustment is rewritten as

$$\arg \min_{\Theta_2} \sum_{i=1}^n \left(\sum_{j=1}^m d^2(\mathbf{x}_{ij}, \mathbf{K}_i[\mathbf{R}\mathbf{R}_\theta^i|\mathbf{t}]\mathbf{X}_j) + (\mathbf{u}_0^i - \bar{\mathbf{u}}_0)^T \Sigma_{\mathbf{u}_0}^{-1} (\mathbf{u}_0^i - \bar{\mathbf{u}}_0) \right), \quad (18)$$

where \mathbf{u}_0^i is the estimate of the principal point for each view. Without further mention, we use 1/10 image width and height as σ_u and σ_v , and 1/2 image width and height as \bar{u}_0 and \bar{v}_0 , respectively.

Note that our optimization process differs from the general reconstruction in that it explicitly encodes the specific non-general motion, i.e., constant \mathbf{R} , \mathbf{t} , and \mathbf{R}_θ . Consequently, a total of $3m + 4$ parameters must be estimated for m views, where three is the number of degrees of freedom of \mathbf{K} (note that we enforce zero skew, unit aspect ratio), and four includes three rotation angles in \mathbf{R} and one constant but unknown angle θ (Fig. 1) in \mathbf{R}_θ . This is a considerable saving over the $9m$ that would be required for a projective reconstruction of a general motion sequence if we make the same assumptions on the camera internal parameters, which reduce the number of degrees of freedom of a projection matrix \mathbf{P} of a pinhole camera from eleven to nine.

5. Experimental results

The proposed approach has been tested on both simulated and real image sequences. First, a synthetic image sequence is used to assess the quality of the algorithm under simulated circumstances. Both the amount of noise on the projected image points and on the rotation angles of the objects are varied. Then results are given for real image sequences to demonstrate the usability of this proposed solution.

5.1. Computer simulation

The simulations are carried out on a sequence of views of a synthetic scene, which consists of 100 points uniformly distributed on a sphere with a radius of 200 units and centered at the origin. Our synthetic camera is located in front of the scene at a distance of 500 units with three rotation angles (20° , 20° , and 15°) between the world coordinate system and the camera coordinate system. In addition to an unity aspect ratio and zero skew, the camera's other internal parameters are chosen as follows. The focal lengths are different for each view, randomly chosen with an expected value of 1000 (in pixels) and a standard deviation of 250. To avoid the case that the chosen focal lengths fall outside the reasonable range, e.g., below zero, we limit them to vary between 750 and 1250. The principal point, \mathbf{u}_0 , had an expected value of $[0 \ 0]^T$ with a standard deviation of $20\sqrt{2}$. An example view of the equivalent scene, where the camera is moving and the object is stationary, is shown in Fig. 4A.

5.1.1. Performance versus pixel error

To assess the performance versus noises on the projected image points, nine views are generated to compute the camera matrices. Gaussian noise with zero mean and a standard deviation of $\sigma \leq 5.0$ pixels was added to the projected image points. The estimated camera parameters were then compared with the ground truth. As argued by Triggs [38] and Zhang [44], the relative difference with respect to the focal length rather than the absolute error is a geometrically meaningful error measure. Therefore,

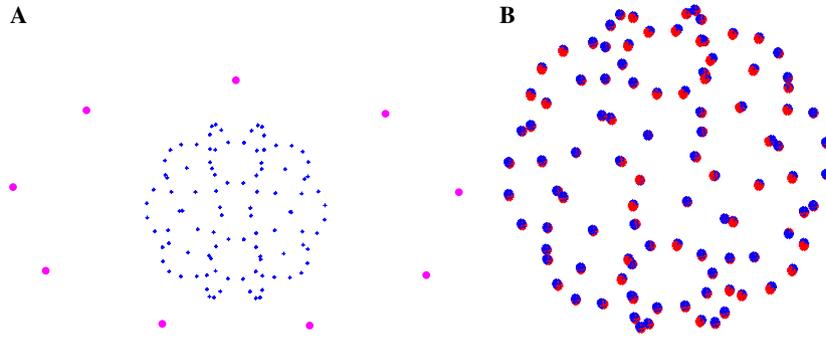


Fig. 4. (A) A view of geometrically equivalent sequences used for simulation, where magenta points denote the positions of cameras. (B) Reconstructed 3D points, where blue cubes denote the ground truth while red cubes are reconstructed ones at the noise level $\sigma = 2.5$ pixels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

we measured the relative error of focal length, f , and the principal point, \mathbf{u}_0 , while varying the noise level from 0.5 pixels to 5.0 pixels. At each noise level, we perform 100 independent trials, and the averaged results of the proposed self-calibration algorithm are shown in Fig. 5. Errors increase almost linearly with respect to the noise level for both focal lengths and principal points. Using our two-stage non-linear optimization, the results are refined from a coarse starting point to a fine level for both f and \mathbf{u}_0 . After the first stage conic enforcement, it reduces on average around 18.5% (range from 7.4% to 36.7%) errors of the estimated focal lengths. Then these errors are further reduced by another average 11% at the second stage bundle adjustment after enforcing the constant rotation angle. In our experiment, we even increase the σ up to 5 pixels. For $\sigma = 2.5$, a typical large noise in the practical calibration, the relative error of focal length f is 1.0%. Fig. 4B shows the 3D reconstructed scene in one trial. The maximum relative error of f (resp. \mathbf{u}_0) is around 2.55% (resp. 1.39%) when $\sigma \leq 5.0$.

5.1.2. Performance versus rotation angle error

Another experiment (nine views) is carried out to evaluate how sensitive the algorithm is to noise in the rotation

angles. Gaussian noise with zero mean and a standard deviation of $\sigma \leq 2.0$ degrees was added to the rotation angles. Considering the fact that extracted feature points will in practice be affected by noise, we also add a typical noise level of $\sigma = 1.0$ pixels to all projected image points. The final results after optimal estimation are shown in Fig. 6. The influence of the orientation noise is larger than that of pixel noise (see Fig. 5), which of course depends on the absolute rotation angle between the views. Note that this coincides with the observation by Frahm and Koch [10], in which case this is more evident since they have smaller rotation angles. The errors in both focal lengths and 3D reconstruction increase almost linearly with respect to the rotation angle noises. Notice also that the errors do not go to zero as noise goes towards zero due to the added noise in image projections.

5.2. Real data

The first real sequence is the Tylenol sequence from Columbia Object Image Library (COIL-20). The COIL sequences have previously resisted structure from motion extraction, due to their low feature counts and variable focal length, which this paper provides the machinery to

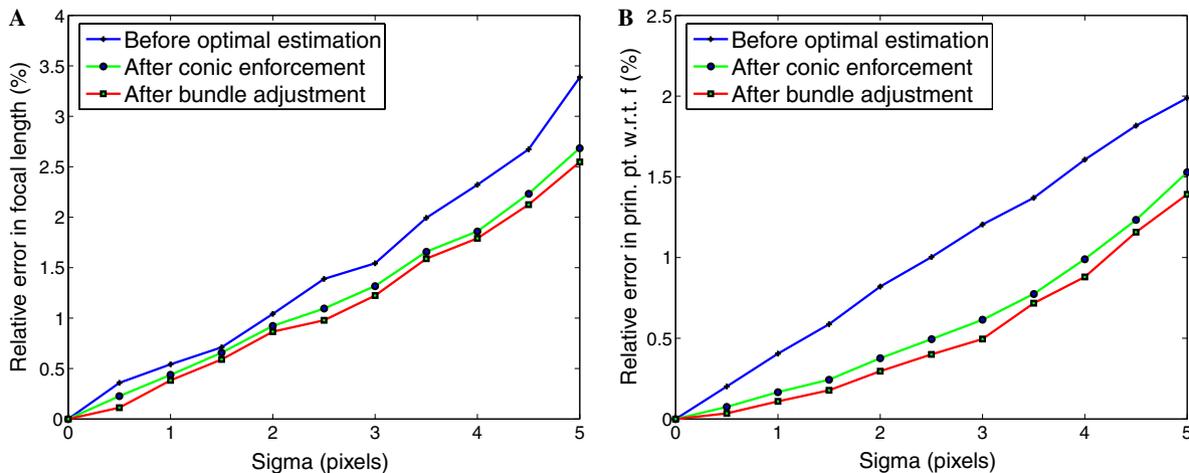


Fig. 5. Performance of the focal length, f , and the principal point, \mathbf{u}_0 , in a function of noise levels: (A) relative error of f , and (B) relative distance of the principal point \mathbf{u}_0 with respect to the true focal lengths.

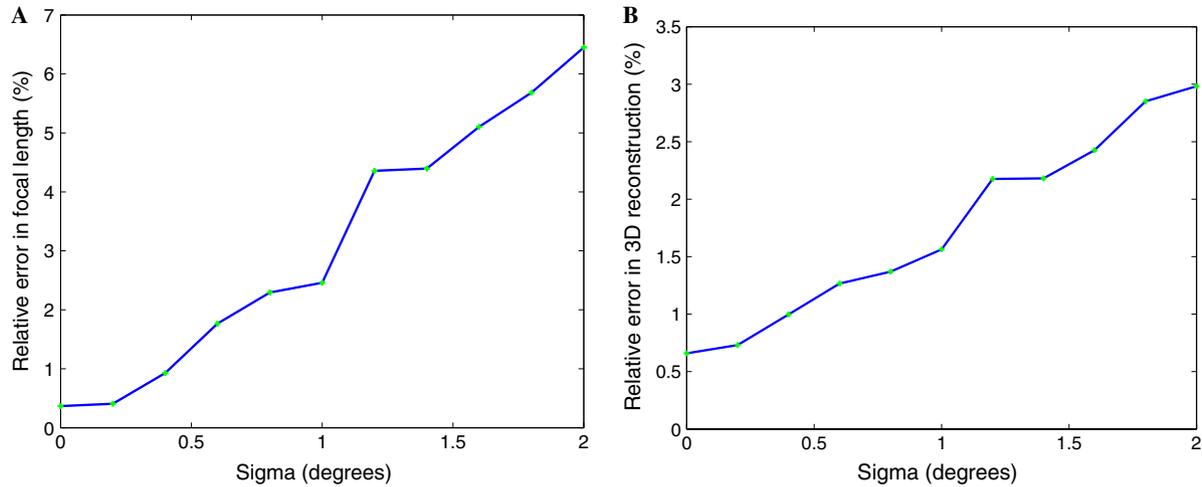


Fig. 6. Performance of focal length and 3D metric reconstruction in a function of rotation angle errors: (A) relative error of focal length and (B) relative 3D metric error. All results shown here are averaged over 100 independent trials.

overcome. We use 18 frames out of the original 72 views of the box as shown in Fig. 7. The tracks of the corresponding points estimated using our previous work [43] are shown in Fig. 9A, and the final determined focal lengths for these images are shown in Fig. 8A. The estimated focal lengths are consistent with the real sequence. For example, the camera zoomed in a lot to capture the 5th and 15th frames while it is zoomed out when shooting the 10th frame. To evaluate the proposed method, we first compensate the frames according to the final estimated calibration matrices by using the 7th frame as the reference. The fitted conics and estimated rotation axis are shown in Figs. 9B–D for three compensated frames (frames 7, 5, and 11). We also show the conics, rotation axis and horizontal line of the compensated frames in Fig. 8B. Finally, piecewise planar model with mapped texture is shown in Fig. 10.

We also tested our approach on the popular dinosaur sequence from the University of Hannover. The sequence contains 36 views of a dinosaur located on a turn-table which is rotating with a constant angular motion of 10 degrees per frame. One frame with tracked points is shown

in Fig. 11A. The computed focal length for the image sequence is shown in Fig. 12A. The results are consistent to the known truth that the focal lengths are fixed.

In another dinosaur sequence, the focal lengths of the camera is set to change in a zigzag fashion (0.8–1.0–1.2), by rescaling the original images. Three consecutive frames are shown in Fig. 13. When the static camera is free to zoom and focus, the 3D circular trajectory is not projected to a conic anymore (Fig. 11B). The computed focal lengths for the dinosaur sequence is shown in Fig. 12B, which is close to the changing pattern in a zigzag fashion. To estimate the correctness of our proposed method, the visual hull of the dinosaur can be computed [31] as shown in Fig. 14. The processing of the volume is performed using a resolution in space of 200^3 unit cubes for the bounding box of the dinosaur. We have no hope to expect the same estimated visual hull by using a calibrated sequence or a sequence captured by a camera with fixed internal parameters. Nevertheless, we extend the power of reconstruction from circular motion even in presence of changing internal parameters.

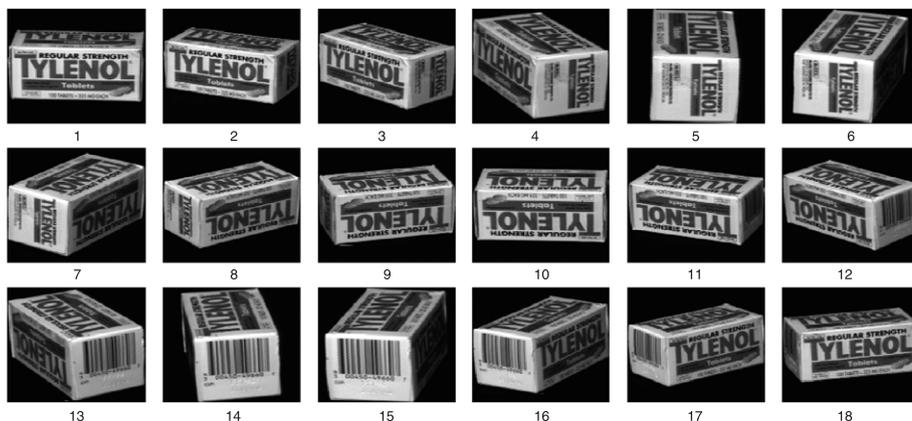


Fig. 7. Eighteen views of the Tylenol sequence.

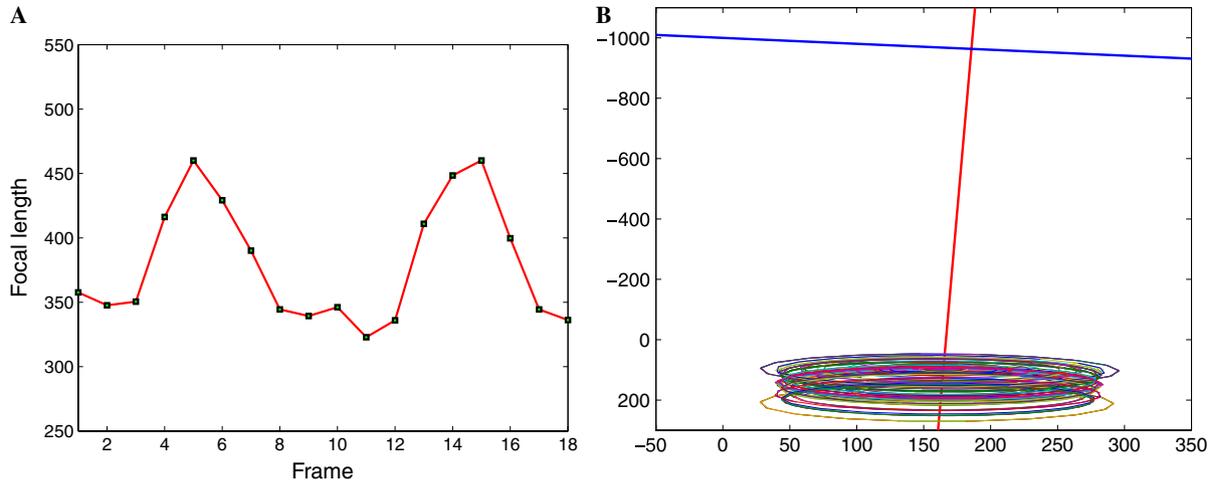


Fig. 8. (A) Computed focal lengths of the Tylenol sequence. (B) The conics, rotation axis, and horizontal line of the compensated frames. Note: we scale the vertical direction to show the all entities.

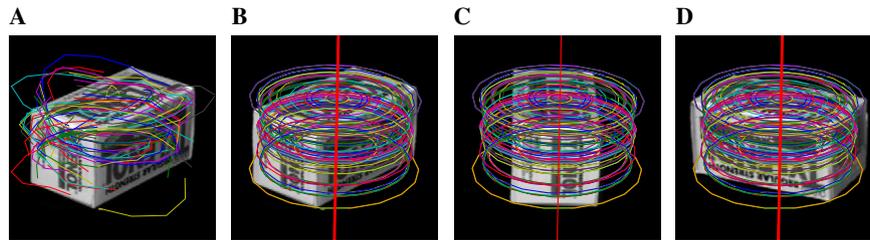


Fig. 9. (A) One original frame (frame 7) of the Tylenol sequence and a subset of the point tracks—only 94 tracks which survived for longer than 4 successive views out of the original 18 views are shown. (B–D) Show the estimated conics on compensated frames 7, 5, and 11. The red vertical line is the rotation axis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)



Fig. 10. Three snapshots of the piecewise planar models with mapped texture of the Tylenol box.

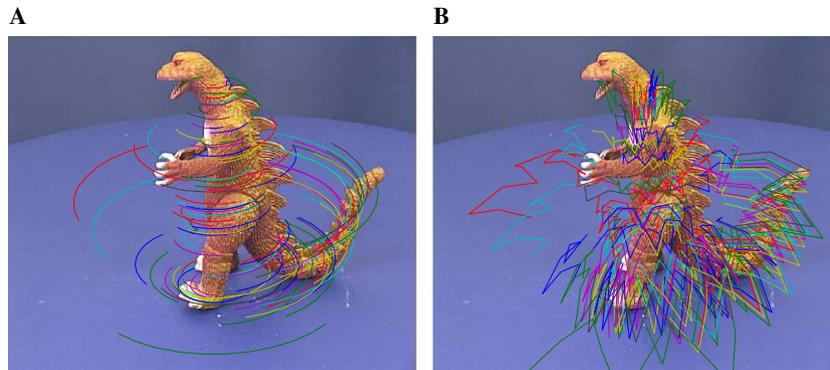


Fig. 11. One frame of the dinosaur sequence and a subset of the point tracks—only 107 tracks which survived for longer than eight successive views are shown. (A) In the case of a static camera with fixed internal parameters, the point tracks are ellipses which are images of circles. (B) When the static camera is free to zoom and focus, the 3D circular trajectories are not projected to conics anymore.

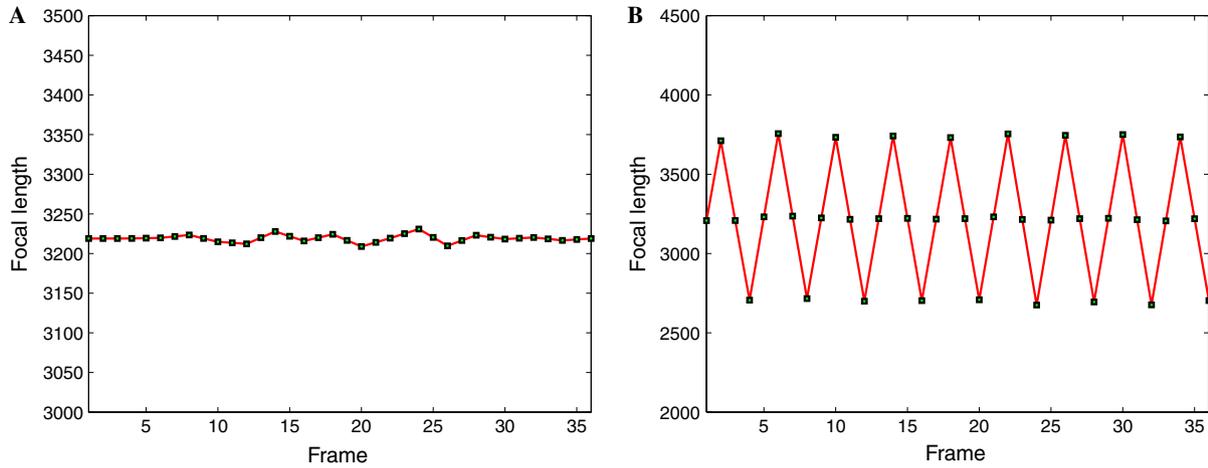


Fig. 12. (A) Computed focal lengths of the dinosaur image sequence with fixed focal length, and (B) computed focal lengths of the dinosaur image sequence with focal length changing in a zigzag fashion.



Fig. 13. Three consecutive sample frames of the zigzag dinosaur sequence. These frames are rescaled from the original images by scales (A) 0.8, (B) 1.0, and (C) 1.2, respectively.

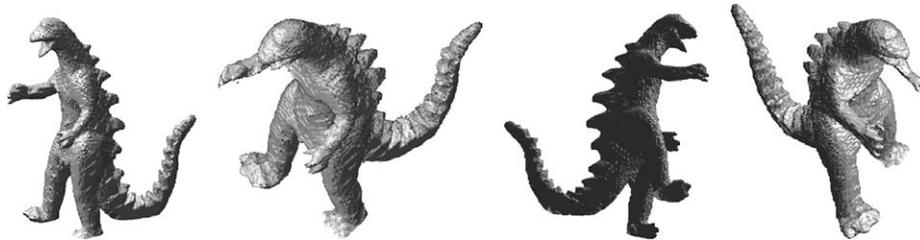


Fig. 14. Four views of the 3D reconstruction of dinosaur from silhouettes.

6. Conclusion

This paper focuses on the problem of self-calibration from an image sequence of an object rotating around a single axis in presence of varying camera internal parameters. Using the invariance property of the inter-frame essential matrices when the rotation angle is constant, we present a new and simple algorithm for camera calibration. Compared to the existing methods, we effectively utilize the prior information, such as constant rotation angle and circular motion, and design a two-stage optimization approach to gradually refine the camera parameters from coarse to fine. The experimental results demonstrate the usability of this proposed solution.

References

- [1] L.D. Agapito, E. Hayman, I. Reid, Self-calibration of rotating and zooming cameras, *Int. J. Comput. Vision* 45 (2) (2001) 107–127.
- [2] M. Armstrong, A. Zisserman, R. Hartley, Self-calibration from image triplets, in: *Proc. ECCV, 1996*, pp. 3–16.
- [3] S. Bounoux. From projective to euclidean space under any practical situation, a criticism of self-calibration, in: *Proc. IEEE ICCV, 1998*, pp. 790–796.
- [4] E. Boyer, Object models from contour sequences, in: *Proc. ECCV, 1996*, pp. 109–118.
- [5] B. Caprile, V. Torre, Using vanishing points for camera calibration, *Int. J. Comput. Vision* 4 (2) (1990) 127–140.
- [6] C. Colombo, A. Bimbo, F. Pernici, Metric 3D reconstruction and texture acquisition of surfaces of revolution from a single uncalibrated view, *IEEE Trans. Pattern Anal. Mach. Intell.* 27 (1) (2005) 99–114.

- [7] O. Faugeras, What can be seen in three dimensions with an uncalibrated stereo rig?, in: Proc. ECCV, 1992, pp. 563–578.
- [8] O. Faugeras, T. Luong, S. Maybank, Camera self-calibration: theory and experiments, in: Proc. of ECCV, 1992, pp. 321–334.
- [9] A.W. Fitzgibbon, G. Cross, A. Zisserman, Automatic 3D model construction for turn-table sequences, in: SMILE Wkshp., 1998, pp. 155–170.
- [10] J. Frahm, R. Koch, Camera calibration with known rotation, in: Proc. IEEE ICCV, 2003, pp. 1418–1425.
- [11] R.I. Hartley, Estimation of relative camera positions for uncalibrated cameras, in: Proc. ECCV, 1992, pp. 579–587.
- [12] R.I. Hartley, Self-calibration of stationary cameras, *Int. J. Comput. Vision* 22 (1) (1997) 5–23.
- [13] R.I. Hartley, A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, MA, 2004.
- [14] J. Heikkila, Geometric camera calibration using circular control points, *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (10) (2000) 1066–1077.
- [15] A. Heyden, K. Astrom, Euclidean reconstruction from image sequences with varying and unknown focal length and principal point, in: Proc. IEEE CVPR, 1997, pp. 438–443.
- [16] A. Heyden, K. Astrom, Flexible calibration: minimal cases for auto-calibration, in: Proc. IEEE ICCV, 1999, pp. 350–355.
- [17] T.S. Huang, O. Faugeras, Some properties of the e matrix in twoview motion estimation, *IEEE Trans. Pattern Anal. Mach. Intell.* 11 (12) (1989) 1310–1312.
- [18] G. Jiang, L. Quan, H.T. Tsui, Circular motion geometry using minimal data, *IEEE Trans. Pattern Anal. Mach. Intell.* 26 (6) (2004) 721–731.
- [19] G. Jiang, H.T. Tsui, L. Quan, A. Zisserman, Single axis geometry by fitting conics, *IEEE Trans. Pattern Anal. Mach. Intell.* 25 (10) (2003) 1343–1348.
- [20] F. Kahl, B. Triggs, K. Åström, Critical motions for auto-calibration when some intrinsic parameters can vary, *J. Math. Imaging Vis.* 13 (2) (2000) 131–146.
- [21] K. Kutulakos, S. Seitz, A theory of shape by space carving, in: Proc. IEEE CVPR, 1999, pp. 307–314.
- [22] D. Liebowitz, A. Zisserman, Combining scene and auto-calibration constraints, in: Proc. IEEE ICCV, 1999, pp. 293–300.
- [23] D. Liebowitz, A. Zisserman, Metric rectification for perspective images of planes, in: Proc. IEEE CVPR, 1998, pp. 482–488.
- [24] Q. Luong, O. Faugeras, The fundamental matrix: theory, algorithms, and stability analysis, *Int. J. Comput. Vision* 17 (1) (1996) 43–75.
- [25] Q. Luong, O. Faugeras, Self-calibration of a moving camera from point correspondences and fundamental matrices, *Int. J. Comput. Vision* 22 (3) (1997) 261–289.
- [26] W. Matusik, H. Pfister, A. Ngan, P. Beardsley, R. Ziegler, L. McMillan, Image-based 3D photography using opacity hulls, in: Proc. ACM SIGGRAPH, 2002, pp. 427–437.
- [27] P.R.S. Mendonça, *Multiview Geometry: Profiles and Self-Calibration*, PhD thesis, University of Cambridge, Cambridge, UK, May 2001.
- [28] P.R.S. Mendonça, K. K. Wong, R. Cipolla, Camera pose estimation and reconstruction from image profiles under circular motion, in: Proc. ECCV, 2000, pp. 864–877.
- [29] P.R.S. Mendonça, K.K. Wong, R. Cipolla, Epipolar geometry from profiles under circular motion, *IEEE Trans. Pattern Anal. Mach. Intell.* 23 (6) (2001) 604–616.
- [30] J. More, *The Levenberg–Marquardt Algorithm, Implementation, and Theory*, numerical analysis ed., Springer, Berlin, 1977.
- [31] W. Niem, Robust and fast modelling of 3d natural objects from multiple views, in: Proc. SPIE, 1994, pp. 388–397.
- [32] M. Pollefeys, R. Koch, L.V. Gool, Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters, *Int. J. Comput. Vision* 32 (1) (1999) 7–25.
- [33] Y. Seo, K. Hong, About the self-calibration of a rotating and zooming camera: theory and practice, in: Proc. IEEE ICCV, 1999, pp. 183–189.
- [34] P. Sturm, Critical motion sequences for monocular self-calibration and uncalibrated euclidean reconstruction, in: Proc. IEEE CVPR, 1997, pp. 1100–1105.
- [35] S. Sullivan, J. Ponce, Automatic model construction, pose estimation, and object recognition from photographs using triangular splines, *IEEE Trans. Pattern Anal. Mach. Intell.* 20 (10) (1998) 1091–1097.
- [36] R. Szeliski, Shape from rotation, in: Proc. IEEE CVPR, 1991, pp. 625–630.
- [37] B. Triggs, Autocalibration and the absolute quadric, in: Proc. IEEE CVPR, 1997, pp. 609–614.
- [38] B. Triggs, Autocalibration from planar scenes, in: Proc. ECCV, 1998, pp. 89–105.
- [39] B. Triggs, P. McLauchlan, R.I. Hartley, A. Fitzgibbon, Bundle adjustment—a modern synthesis, in: *Vision Algorithms: Theory and Practice*, 1999, pp. 298–373.
- [40] H. Trivedi, Can multiple views make up for lack of camera registration? *Image Vision Comput.* 6 (1) (1988) 29–32.
- [41] R. Tsai, A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf tv cameras and lenses, *IEEE J. of Robotics and Automation* 3 (4) (1987) 323–344.
- [42] K. Wong, R. Mendonça, R. Cipolla, Camera calibration from surfaces of revolution, *IEEE Trans. Pattern Anal. Mach. Intell.* 25 (2) (2003) 147–161.
- [43] J. Xiao, M. Shah, Two-frame wide baseline matching, in: Proc. IEEE ICCV, 2003, pp. 603–609.
- [44] Z. Zhang, Camera calibration with one-dimensional objects, *IEEE Trans. Pattern Anal. Mach. Intell.* 26 (7) (2004) 892–899.
- [45] Z. Zhang, Determining the epipolar geometry and its uncertainty: a review, *Int. J. Comput. Vision* 27 (2) (1998) 161–195.
- [46] Z. Zhang, A flexible new technique for camera calibration, *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (11) (2000) 1330–1334.
- [47] A. Zisserman, P. Beardsley, I. Reid, Metric calibration of a stereo rig, in: *IEEE Workshop on Representation of Visual Scenes*, 1995, pp. 93–100.