

Camera Calibration using One-Dimensional Information and Its Applications in both Controlled and Uncontrolled Environments

En Peng* and Ling Li

Department of Computing, Curtin University of Technology, Perth, WA, Australia

*Corresponding author. Tel: +61(8)92662948; Fax: +61(8)92662819.

E-mail addresses: e.peng@curtin.edu.au (En Peng), l.li@curtin.edu.au (Ling Li).

Abstract

Camera model and its calibration are required in many applications for coordinate conversions between the two-dimensional image and the real three-dimensional world. Self-calibration method is usually chosen for camera calibration in uncontrolled environments because the scene geometry could be unknown. However when no reliable feature correspondences can be established or when the camera is static in relation to the majority of the scene, self-calibration method fails to work. On the other hand, object-based calibration methods are more reliable than self-calibration methods due to the existence of the object with known geometry. However, most object-based calibration methods are unable to work in uncontrolled environments because they require the geometric knowledge on calibration objects. Though in the past few years the simplest geometry required for a calibration object has been reduced to a 1D object with at least three points, it is still not easy to find such an object in an uncontrolled environment, not to mention the additional metric/motion requirement in the existing methods. Meanwhile, it is very easy to find a 1D object with two end points in most scenes. Thus, it would be very worthwhile to investigate an object-based method based on such a simple object so that it would still be possible to calibrate a camera when both self-calibration and existing object-based calibration fail to work. We propose a new camera calibration method which requires only an object with two end points, the simplest geometry that can be extracted from many real-life objects. Through observations of such a 1D object at different positions/orientations on a plane which is fixed in relation to the camera, both intrinsic (focal length) and extrinsic (rotation angles and translations) camera parameters can be calibrated using the proposed method. The proposed method has been tested on simulated data and real data from both controlled and uncontrolled environments, including situations where no

explicit 1D calibration objects are available, e.g. from a human walking sequence. Very accurate camera calibration results have been achieved using the proposed method.

Keywords: Camera calibration; pinhole camera model; one-dimensional calibration object

1. Introduction

Camera calibration is an important component in computer vision applications that aims to extract 3D metric information of the camera from 2D images. Existing camera calibration methods can be classified into two main categories: self-calibration (a.k.a. 0D object-based calibration) and object-based calibration (3D/2D/1D object-based).

Self-calibration [1-6] is normally formed by moving a camera around an unknown static scene. The camera can be calibrated through point/line correspondences between images. The idea here is to find camera parameters that satisfy the pairwise epipolar constraints, e.g. Kruppa equations [2] and Huang-Faugeras' constraints [6], which can be obtained from point/line correspondences. This approach can be applied easily to most pre-existing images. However, one significant limitation of this approach is that a sufficient number of correspondences between views must be established - which is impossible when the scene is featureless, the structure of the scene is dynamic, or the scene and the camera is relatively static. Another type of self-calibration is through the vanishing points [7-9], when there are some known artifacts (with edges that are parallel or perpendicular to each other, e.g. buildings) in the scene being observed. This type of calibrations attempts to find camera parameters to associate the vanishing points with some parallel lines. This approach can also be applied to many pre-existing images, as long as objects with visible parallel/perpendicular edges are present in the scene. While the existing self-calibration methods are useful for calibrating the camera from many pre-recorded images, they cannot be used for camera calibrations in situations where the required features (e.g. point/line correspondences between images from the static scene, parallel/perpendicular lines) are unavailable in the observed scene, or when the camera model is relatively static to the majority of the scene features.

Object-based calibration utilizes a calibration object whose geometric information is pre-known. The camera calibration process can thus become more reliable without involving a large number of uncertain features. The calibration objects may have a dimension of three, two or one. The 3D object-based camera calibration [10-13] is a traditional calibration technique which requires the three-dimensional geometric knowledge of a reference object prior to the calibration process. In this technique the 3D points and their 2D projections are associated through a projection matrix. When a sufficient number of 3D-2D correspondences are obtained, the projection matrix can be solved using Direct Linear Transformation (DLT). Subsequently the camera parameters can be computed. Very accurate results can be produced by this technique. The 2D object-based camera calibration [14-15] makes use of known planar patterns instead of a complex three-dimensional object which makes the calibration process more flexible. The user needs to paste some pre-known patterns on a cardboard and move the cardboard in front of a camera several (at least two) times for the camera to be calibrated. The idea behind this technique is to solve the camera parameters from the homography between the points on the pre-known plane pattern and the corresponding image points. Sufficiently accurate calibration results can be produced by the 2D object-based calibration, despite the simpler preparation process. The 1D object-based calibration [16-22] was proposed a few years ago, which only required a one-dimensional object consisting of at least three feature points. The motion of the 1D object is usually strictly constrained, e.g., moving around a fixed point [16-20] or moving under gravity only [21, 22]. Under such constraints, some approaches [17-22] try to calibrate all intrinsic parameters (focal length, pixel aspect ratio, skew, and an image principle point) with the known metric information on the object; while some other approaches [16] calibrates a simpler camera model (e.g. unit pixel aspect ratio, zero skew, etc.) by eliminating the need for knowing the metrics of the linear points. With the existing 1D object-based calibration techniques, it is possible to use a very simple calibration object (a 1D object with three feature points) to calibrate a reasonably modeled camera without pre-measurement.

Among the existing camera calibration techniques, the self-calibration methods are better suited for certain uncontrolled environments, while the object-based methods are more reliable but are only suitable for controlled environments. To our knowledge, there are no object-based camera calibration

methods that can easily work in uncontrolled environments, due to their restrictions on the geometry of the calibration object. There is certainly a need to investigate an object-based camera calibration method, which could be an alternative choice especially in some uncontrolled situations where the self-calibration methods fail to work. This paper proposes such a method that is able to calibrate both the intrinsic and extrinsic (in relation to the plane) parameters of an ideal pin-hole camera. To make it possible for the proposed method to be employed in uncontrolled environment, the calibration object is relaxed to a 1D line segment consisting of two end points. This object is required to remain on a fixed plane whenever it is observed. The 1D object satisfying such a requirement can be easily identified by an ordinary user as follows. If there is a visible plane in the scene, such as a desk or a floor, it can be easily perceived if there is a 1D object on the plane (e.g. a pen on the desk) or if a 1D object is in a plane parallel to the plane (e.g. the bottom edge of a cabinet door which always remains parallel to the floor). As a matter of fact, the 1D object does not need to remain on a fixed plane all the time - it can move freely as long as at each observation it lies within the same plane that is relatively fixed to the camera. For example, the line from the tip to the toe of a human foot when it is on the floor could be used as a calibration object, although the foot can be lifted up from the floor from time to time.

The proposed method differs from existing camera calibration techniques in the following aspects:

1) Unlike the existing self-calibration technique, the proposed method does not concern whether the structure of the whole scene is artificial/natural or static/dynamic, as long as the two end points of a 1D object can be observed. The proposed method also does not require the camera to move around the scene, and is therefore suitable for applications with a fixed camera, e.g. surveillance camera. In particular, the self-calibration technique using line correspondences normally requires multiple observations of multiple line segments [1], while the proposed method only requires multiple observations of a single line segment.

2) Unlike the existing object-based calibration technique, the proposed method does not require special preparations or measurements of the calibration object. Any object with a line segment will be sufficient. Though the proposed method requires the 1D object to remain on a 2D plane for every observation, it is significantly different from the traditional 2D object-based calibration methods. A 1D

object with two end points cannot be considered a “pattern” on a plane, because observations of such an object do not provide enough 2D point correspondences for deriving the homography matrix required in 2D object-based methods.

The rest of this paper is organized as follows: Section 2 provides the preliminary knowledge for this project; Section 3 introduces the proposed calibration algorithm; Section 4 describes the experimental results on both simulated and real data; and Section 5 concludes this paper.

2. Preliminaries

2.1 World-image projection

An ideal pinhole camera is widely used in computer vision. It relates a 3D point $[x,y,z,1]^T$ and its 2D image point $[u,v,1]^T$ by a 3×4 projection matrix \mathbf{P} :

$$[u,v,1]^T \sim \mathbf{P}[x,y,z,1]^T. \quad (1)$$

The matrix \mathbf{P} is determined by an intrinsic parameter (focal length f) and six extrinsic parameters encoding the Camera Coordinate System (CCS) with respect to the World Coordinate System (WCS). The transformation between the WCS and CCS is defined as follows: CCS is initially aligned with WCS. It is then translated to $[t_x, t_y, t_z, 1]^T$, followed by a rotation about the Y -axis by pan angle α , then a rotation around the X -axis by tilt angle β , and finally, a rotation around the Z -axis by roll angle γ . Accordingly, the projection matrix \mathbf{P} can be represented as:

$$\mathbf{P} = \mathbf{K}[\mathbf{R}_z \mathbf{R}_x \mathbf{R}_y | -\mathbf{R}_z \mathbf{R}_x \mathbf{R}_y \mathbf{T}]$$

$$\mathbf{K} = \begin{bmatrix} f & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{R}_z = \begin{bmatrix} c & \gamma & -s & \gamma & 0 \\ sr & \gamma & ic & \gamma n & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{R}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & \beta & -s & \beta \\ 0 & s & \beta & ic & \beta \end{bmatrix}, \mathbf{R}_y = \begin{bmatrix} c & \alpha & 0 & s & \alpha \\ n & 0 & 1 & 0 \\ -s & s & \alpha & 0 & c & n & \alpha \end{bmatrix}, \mathbf{T} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}. \quad (2)$$

Given a set of correspondences between image points and 3D points, the projection matrix \mathbf{P} can be derived using DLT algorithm, and subsequently, the camera parameters can be calibrated. In this research the 3D coordinates of all observed points are unknown. However, they are assumed to lie on the same plane for all observations. The problem can then be simplified into the plane-image mapping. The specific plane containing the observed points will be referred to as “the reference plane” in the rest of this paper.

2.2 Plane-image homography

Assuming that all 3D points lie on the XZ plane, an arbitrary point $[x,0,z,1]^T$ can be represented as $[x,z,1]^T$ for simplification. Denoting the i^{th} column of $[\mathbf{R}_z\mathbf{R}_x\mathbf{R}_y | -\mathbf{R}_z\mathbf{R}_x\mathbf{R}_y\mathbf{T}]$ as \mathbf{E}_i , a point $[x,z,1]^T$ and its projection $[u,v,1]^T$ are related by a 3×3 homography \mathbf{H} :

$$[u \ v \ 1]^T = \mathbf{H}[x \ z \ 1]^T, \quad \mathbf{H} \sim \mathbf{K}[\mathbf{E}_1 \ \mathbf{E}_3 \ \mathbf{E}_4] \quad (3)$$

If the homography can be obtained, the points on the planes can be calculated directly from $[x,z,1]^T = \mathbf{H}^{-1}[u,v,1]^T$. The problem of obtaining the points on a plane from the images is often called plane rectification and will be discussed next.

2.3 Plane rectification

The image-plane homography, \mathbf{H}^{-1} , can be decomposed uniquely into a concatenation of three matrices: namely, the similarity transformation matrix \mathbf{S}_k , the affine transformation matrix \mathbf{A}_k , and the projective transformation matrix \mathbf{P}_k :

$$\mathbf{H}^{-1} = \mathbf{S}_k \mathbf{A}_k \mathbf{P}_k \quad (4)$$

The projective transformation matrix \mathbf{P}_k is required for the affine rectification of the image. It is defined as:

$$\mathbf{P}_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_1 & l_2 & l_3 \end{bmatrix} \quad (5)$$

The parameters l_1 , l_2 , and l_3 form a homogeneous vector $\mathbf{l}_v = [l_1, l_2, l_3]^T$, which defines the vanishing line of the plane on the image. If the vanishing line of the plane can be identified, the matrix \mathbf{P}_k can be obtained. The vanishing line can be determined by two or more vanishing points, which lie on the vanishing line. A vanishing point can be conventionally obtained as the intersection of projections from two parallel lines in the reference plane. When there are no pre-known parallel lines in the reference plane, it is still possible to obtain the vanishing point in some specific cases. Reference [23] proposes a way to obtain the vanishing point from an object with uniform linear motion: a point \mathbf{p} on the world path and the point \mathbf{p}' on the projection can be related by a 1D homography \mathbf{G} : $\mathbf{p} = \mathbf{G}\mathbf{p}'$. The matrix \mathbf{G} can be estimated by the DLT algorithm when there are sufficient correspondences on the same path. The vanishing point \mathbf{p}_v' can then be computed as $\mathbf{p}_v' = \mathbf{G}^{-1}[1, 0]^T$. With the vanishing points obtained, the matrix \mathbf{P}_k can be computed. The image geometry is affine-rectified by applying the matrix \mathbf{P}_k , i.e., affine properties such as length ratios etc. are restored on parallel segments. In order to restore the length ratios on non-parallel segments, a further rectification – metric rectification is required. The affine transformation matrix \mathbf{A}_k is used to perform metric rectification from affine-rectified geometry and is defined as:

$$\mathbf{A}_k = \begin{bmatrix} 1/\beta' & -\alpha'/\beta' & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6)$$

The parameters α' and β' are used to represent circular points. The circular points are a pair of complex conjugate points on the line at infinity and are invariant to similarity transformation. They are transformed from metric coordinates $[1, \pm i, 0]^T$ to affine coordinates $[\alpha' \mp i\beta', 1, 0]^T$. These two parameters can be estimated from images under three types of constraints as proposed in [24]: (1) a known angle on the reference plane; (2) equal but unknown angles on the reference plane; and (3) known length ratio of two non-parallel segments on the reference planes. The angle constraints are not applicable in this research since a single 1D object does not provide any information on angles. Only the third type of constraints is of interests in this project. Assume that there are two non-parallel

segments in the world plane and their projections on the image are $l_1:(u_{11}, v_{11})-(u_{12}, v_{12})$ and $l_2:(u_{21}, v_{21})-(u_{22}, v_{22})$. Denoting Δu_i for $u_{n1}-u_{n2}$ and Δv_i for $v_{n1}-v_{n2}$, the point (α_i, β_i) lies on a circle with centre $(c_{\alpha'}, c_{\beta'})$ on the axis and radius r :

$$(c_{\alpha'}, c_{\beta'}) = \left(\frac{\Delta u_1 \Delta v_1 - \Delta u_2 \Delta v_2}{\Delta v_1^2 - \Delta v_2^2}, 0 \right) \quad (7)$$

$$r = \left| \frac{\Delta u_2 \Delta v_1 - \Delta u_1 \Delta v_2}{\Delta v_1^2 - \Delta v_2^2} \right| \quad (8)$$

Thus, (α', β') can be solved as the intersection of at least two such circles. With α' and β' computed, the affine-rectified plane can be metric-rectified using the matrix \mathbf{A}_k . The final matrix is a similarity transformation:

$$\mathbf{S}_h = \begin{bmatrix} s\mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}, \quad (9)$$

where s is a scaling parameter, \mathbf{R} is a 2D rotation matrix, and \mathbf{t} is a translation vector.

It can be seen from above that only matrix $\mathbf{A}_h \mathbf{P}_h$ needs to be estimated if the reference plane geometry is to be recovered from the image points. The existing methods cannot recover the vanishing points when no parallel lines on world plane are pre-known or a 1D object with metric information of three or more feature points (the object with uniform linear motion in [23] can in fact be treated as an object with three or more points) is unavailable. Nevertheless, it has been proven that the world plane pattern can be recovered from the image points of line segments if vanishing lines on the plane can be estimated.

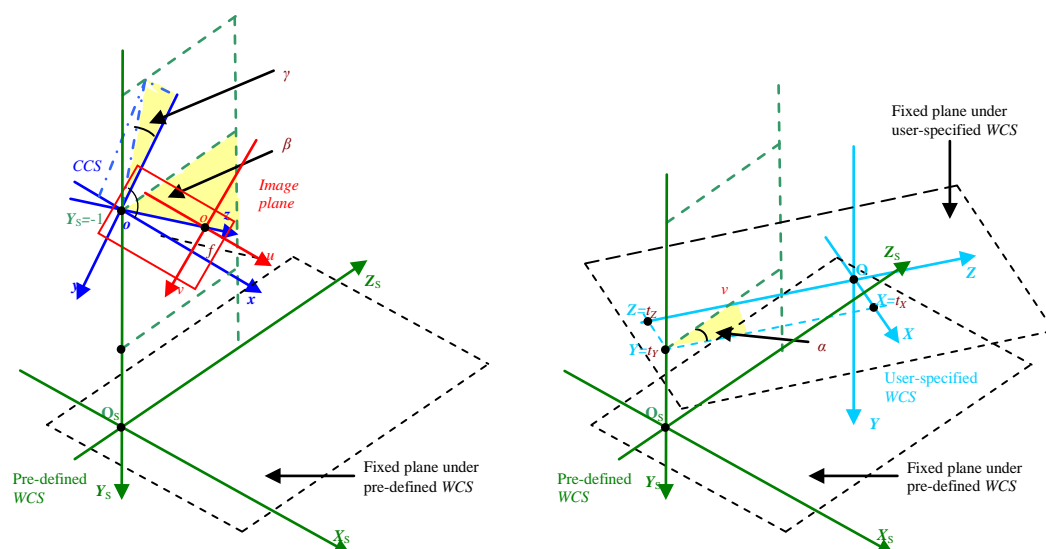
In the next section, the proposed calibration algorithm is introduced, which is able to calibrate the camera and rectify the image using only a 1D object with two end points, without the need to recover the vanishing points.

3. Proposed Calibration Algorithm

As discussed above, if the vanishing line can be determined, a 2D pattern on the reference plane can be recovered. Due to the existence of noises, both the affine and metric rectifications require non-linear optimizations after initiations [23]. The vanishing line has three degrees of freedom: two parameters for a 2D point on the vanishing line and one parameter for the rotation around this point. One way to estimate these parameters is through non-linear optimizations. However, the first two parameters are not bounded. It is difficult to perform non-linear optimizations on them. The idea of the proposed algorithm is to work directly on the camera parameters themselves, rather than deriving them through the vanishing line. The details of the proposed algorithm are introduced in the following sections.

3.1 Camera model

We adopt the pin-hole camera model introduced in Section 2.1. There are seven camera parameters to calibrate: focal length f , tilt angle β , roll angle γ , pan angle α and camera displacement parameters t_x , t_y , and t_z . The origin of WCS is placed on the fixed reference plane and the Y -axis of WCS is set to be perpendicular to the plane pointing downwards. Since some users may prefer to use their own WCS, two WCSs are presented here: the pre-defined WCS and the user-specified WCS. With the pre-defined WCS, pre-define values are given to some camera parameters while preserving the relationship between the image plane and the reference plane up to a scale factor. Not all camera parameters need to be calibrated in this case. With the user-specified WCS, all camera parameters need to be calibrated. The pre-defined WCS and user-specified WCS used in this project are illustrated in Figure 1.



(a)

(b)

Figure 1: Image plane, CCS, pre-defined WCS with reference plane, user-specified WCS with reference plane: (a) CCS and the pre-defined WCS; (b) Pre-defined WCS and user-specified WCS.

We divide these parameters into two groups: critical parameters (f , β , and γ) which can uniquely represent the relationship between the camera and the plane up to a scale factor, and non-critical parameters (α , t_x , t_y , and t_z) which metrically relate the camera model with the plane. Our attempt to calibrate these parameters can be simply summarized as follows. If a pre-defined WCS is assumed by setting the non-critical parameters with pre-defined values ($\alpha=0$ and $\mathbf{T}_S=[0,-1,0]^T$), only the critical parameters are calibrated based on these values. If the user specifies some relevant information about the WCS on the image as detailed later in Section 3.3, all non-critical parameters will be calibrated, in addition to the critical parameters, to satisfy the preferences of the user.

3.2 Computation of Critical Parameters

As introduced in Section 2.2, any arbitrary point on the reference plane can be represented as $[x_S, 0, z_S, 1]^T$ and can be simplified as $[x_S, z_S, 1]^T$ since y coordinate is always zero. A point $[x_S, z_S, 1]^T$ on the reference plane and its projection $[u, v, 1]^T$ are related by a homography as shown in Equation (3). $[x_S, z_S, 1]^T$ can be computed from Equation (3) as:

$$\begin{cases} x_S = \frac{uc - \phi + \beta s \gamma}{-uc - \phi s \gamma + \beta v c - \phi c s \gamma + f s \beta} \\ z_S = \frac{us - \beta \gamma \phi + \gamma v s \beta - \phi \alpha \gamma + f c \phi}{-uc - \phi s \gamma + \beta v c - \phi c s \gamma + f s \beta} \end{cases} \quad (10)$$

We denote the two end points of a 1D object as A and B and let the image points of A and B at the i^{th} observation be $[u_{A_i}, v_{A_i}, 1]^T$ and $[u_{B_i}, v_{B_i}, 1]^T$ respectively. Since the 1D object is always on the reference plane at the time of observations, by denoting the corresponding end point positions on the plane as $[x_{A_i}, z_{A_i}, 1]^T$ and $[x_{B_i}, z_{B_i}, 1]^T$, we can define a function to compute the back-projected length of the line segment as:

$$l([u_A, v_A, 1]^T, [u_B, v_B, 1]^T, f, \beta, \gamma) = \sqrt{(x_B - x_A)^2 + (z_B - z_A)^2}, \quad (11)$$

where x_A, z_A, x_B and z_B can be derived using Equation (6).

Denoting the length of a 1D object as L , it can be obtained as:

$$L = k ([u_A, v_A, 1]^T, [u_B, v_B, 1]^T, f, \beta, \gamma), \quad (12)$$

where k is an arbitrary scaling factor.

If there are N observations of the 1D object at different positions/orientations, and since L is constant for the same 1D object and k is invariant for the same camera settings, a set of $(N-1)$ equations can be derived as:

$$l([u_A, v_A, 1]^T, [u_B, v_B, 1]^T, f, \beta, \gamma) = l([u_{A(i+1)}, v_{A(i+1)}, 1]^T, [u_{B(i+1)}, v_{B(i+1)}, 1]^T, f, \beta, \gamma), \quad i = 1, 2, \dots, N-1. \quad (13)$$

By stacking three such equations, it is theoretically possible to solve for the three camera parameters (f , β , and γ). However, the computation efforts of solving for these unknowns are very huge due to the high complexity of the equations. Fortunately, all these three camera parameters are bounded. The normal range of the angle of view of a lens is between 10° and 100° . The range of focal length (f) for an image with the dimension of $W \times H$ can be computed as between $D \cdot \cot(100^\circ/2)/2$ and $D \cdot \cot(10^\circ/2)/2$ where $D = (W^2 + H^2)^{1/2}$. The normal range of the tilt angle (β) is between -60° and 60° while the normal range of the roll angle (γ) is between -15° and 15° . Hence the parameters can be obtained through optimization.

An object function needs to be derived for the optimization process to evaluate the fitness of the given parameters. The minimum input error tolerance required for the successful reconstruction from at least a specified percentage of all input observations of the calibration object is also reflected in the object function. This percentage, denoted as λ , is normally set at 90% to accommodate occasional extraction error of the calibration object while maintaining the reconstruction accuracy. Defining δ as the input error tolerance for the image points and $\omega_\lambda(f, \beta, \gamma, \delta)$ as the minimum back-projected length discrepancy tolerance for a subset of image point pairs, the object function is constructed as:

$$\mathcal{Q}_\lambda(f, \beta, \gamma) = \arg \min_{\delta} \omega_\lambda(f, \beta, \gamma, \delta). \quad (14)$$

Defining M as the complete set of the input line segments, $q_\lambda(f, \beta, \gamma, \delta)$ as a subset of M such that $|q_\lambda(f, \beta, \gamma, \delta)| \geq \lambda|M|$ and all its element image points are able to be reconstructed within the tolerance δ under the given camera model, and $\Psi(Q, f, \beta, \gamma, \delta)$ as the minimum back-projected length discrepancy tolerance for the set Q , the function $\omega_\lambda(f, \beta, \gamma, \delta)$ is defined as:

$$\omega_\lambda(f, \beta, \gamma, \delta) = \Psi(Q_\lambda, f, \beta, \gamma, \delta), \quad \text{where } Q_\lambda = q_\lambda(f, \beta, \gamma, \delta). \quad (15)$$

It is assumed that there are n observations of the 1D object in the set Q , and the image points of the 1D object in the i^{th} observation are denoted as $\mathbf{m}_{Ai} = [u_{Ai}, v_{Ai}, 1]^T$ and $\mathbf{m}_{Bi} = [u_{Bi}, v_{Bi}, 1]^T$. Function $\varphi(\mathbf{m}_A, \mathbf{m}_B, f, \beta, \gamma, \delta)$ is defined to represent the back-projected length interval of the line segment under the given camera parameters and the given input image noise tolerance, and $\psi(\delta, \varphi_a, \varphi_b)$ is defined to determine whether φ_a and φ_b overlap under the given noise tolerance δ . The function $\Psi(Q, f, \beta, \gamma, \delta)$ can then be defined as:

$$\Psi(Q, f, \beta, \gamma, \delta) = \max \psi(\delta, \varphi(\mathbf{m}_{Ai}, \mathbf{m}_{Bi}, f, \beta, \gamma, \delta), \varphi(\mathbf{m}_{Aj}, \mathbf{m}_{Bj}, f, \beta, \gamma, \delta)) \quad \text{with } 1 \leq i, j \leq n. \quad (16)$$

The function $\psi(\delta, \varphi_a, \varphi_b)$ is defined as:

$$\psi(\delta, \varphi_a, \varphi_b) = \begin{cases} \delta, & \text{w } \varphi_a \cap \varphi_b \neq \emptyset; \\ \infty, & \text{w } \varphi_a \cap \varphi_b = \emptyset. \end{cases} \quad (17)$$

Denoting $\boldsymbol{\varepsilon} = [u_\varepsilon, v_\varepsilon, 0]^T$ as the vector of an input noise, the function $\varphi(\mathbf{m}_A, \mathbf{m}_B, f, \beta, \gamma, \delta)$ is constructed as:

$$\varphi(\mathbf{m}_A, \mathbf{m}_B, f, \beta, \gamma, \delta) = \begin{cases} \emptyset, & l_{\min} = \infty; \\ [l_{\min}, l_{\max}], & l_{\min} \neq \infty. \end{cases} \quad (18)$$

$$\text{w } \begin{cases} l_{\min} = \min_{\mathbf{h}} l(\mathbf{m}_A + \boldsymbol{\varepsilon}_1, \mathbf{m}_B + \boldsymbol{\varepsilon}_2, f, \beta, \gamma) \\ l_{\max} = \min_{\mathbf{h}} l(\mathbf{m}_A + \boldsymbol{\varepsilon}_3, \mathbf{m}_B + \boldsymbol{\varepsilon}_4, f, \beta, \gamma) \end{cases} \quad \|\boldsymbol{\varepsilon}_k\|_\infty \leq \delta \quad (k = 1, 2, 3, 4).$$

When using Equation (14) for optimization, it is desirable to compute $\boldsymbol{\varepsilon}_1$, $\boldsymbol{\varepsilon}_2$, $\boldsymbol{\varepsilon}_3$, and $\boldsymbol{\varepsilon}_4$ directly, or even better, to compute l_{\min} and l_{\max} directly, especially when δ is non-zero. The following method is proposed for computing l_{\min} and l_{\max} directly.

- Computation of l_{\min} :

1) When $\|\mathbf{m}_A - \mathbf{m}_B\|_{\infty} \leq 2\delta$, which indicates that two end points can overlap under the input noise tolerance, $l_{\min} = 0$;

2) When all of $\mathbf{m}_A \pm [\delta, \pm\delta, 0]^T$ or all of $\mathbf{m}_B \pm [\delta, \pm\delta, 0]^T$ are back-projected to invalid positions (infinity or positions that are behind the camera), $l_{\min} = \infty$;

3) In other cases, the four sides of the polygon back-projected from $\mathbf{m}_A \pm [\delta, \pm\delta, 0]^T$ (or $\mathbf{m}_B \pm [\delta, \pm\delta, 0]^T$) are denoted as S_{Ai} (or S_{Bi}) with $i=1, 2, 3,$ and 4 . The shortest distance between each line segment pair of S_{Ai} and S_{Bi} can be easily obtained. This shortest distance among all pairs is chosen as l_{\min} .

- Computation of l_{\max} :

1) When any of $\mathbf{m}_A \pm [\delta, \pm\delta, 0]^T$ or any of $\mathbf{m}_B \pm [\delta, \pm\delta, 0]^T$ is back-projected to an invalid position, $l_{\max} = \infty$;

2) In other cases, the longest distance between each line segment pair of S_{Ai} and S_{Bi} can be obtained. This longest distance among all pairs is chosen as l_{\max} .

With the object function defined and the bounded search spaces for all parameters specified, the dynamic hill climbing (DHC) method [25] is implemented for non-linear optimization. The basic idea of the optimization process using DHC is to randomly generate an initial 3-tuple (f, β, γ) within the 3D searching space for n times in the outer loop of DHC. With every randomly generated starting value, the inner loop of the DHC algorithm will be executed. The inner loop will stop if a solution has been found or the local minimum has been reached (no hill to climb). The outer loop will terminate if a solution has been found or n times of retry have been performed. If no solution can be found, the 3-tuple producing the best local minima will be returned.

3.3 Computation of Non-Critical Parameters

The user can specify his/her preferences for defining the WCS. Such specification should be made as follows: First of all, the user specifies a 2D point $O=[u_o, v_o, 1]^T$ on the image to be the projection of the WCS origin. Next, the user specifies a 2D point $P=[u_p, v_p, 1]^T$ in the image such that OP is the projection of a positive unit vector on the Z -axis. The definition of ‘‘a unit’’ here is up to the user. E.g., if the user specifies the unit vector OP as the projection of a pen, any metric measurement is base on the length of the pen. In that case, if the camera displacement in the X direction is 20 units, it means that this displacement is 20 times of the length of the pen. With the user specified OP , the proposed algorithm can construct the user-specified WCS based on the pre-defined WCS.

Denoting the i^{th} column of $[\mathbf{R}_z \mathbf{R}_x \mathbf{R}_y | -\mathbf{R}_z \mathbf{R}_x \mathbf{R}_y \mathbf{T}]$ as \mathbf{E}_i , a point $[x, z, 1]^T$ on the reference plane and its projection $[u, v, 1]^T$ are related by a 3×3 homography \mathbf{H} :

$$[u, v, 1]^T = \mathbf{H}[x, z, 1]^T, \text{ where } \mathbf{H} \sim \mathbf{K}[\mathbf{E}_1 \ \mathbf{E}_3 \ \mathbf{E}_4]. \quad (19)$$

The homography \mathbf{H}_U relating the coordinates of the same point represented in user-specified WCS $([x, z, 1]^T)$ and the pre-defined WCS $([x_s, z_s, 1]^T)$ can be computed as:

$$[x_s, z_s, 1]^T = \mathbf{H}_U [x, z, 1]^T, \quad \text{w } \mathbf{H}_U = \mathbf{H}_S^{-1} \mathbf{H} \sim \begin{bmatrix} c & \alpha & \alpha & \alpha & i-s_x & \alpha & \alpha - t_x & \alpha \\ -s & e\alpha & i & r\alpha & \alpha & s & \alpha - s t_z & c n\alpha \\ 0 & 0 & & & & & -t_y & \end{bmatrix} \quad (20)$$

According to the user-specified position of points O and P , it can be obtained that:

$$\begin{cases} [x_o, z_o, 1]^T = \mathbf{H}_U [0, 0, 1]^T \\ [x_p, z_p, 1]^T = \mathbf{H}_U [0, 1, 1]^T \end{cases}, \quad \text{w } \begin{cases} [x_o, z_o, 1]^T = \mathbf{H}_S^{-1} [u_o, v_o, 1]^T \\ [x_p, z_p, 1]^T = \mathbf{H}_S^{-1} [u_p, v_p, 1]^T \end{cases} \cdot (21)$$

All non-critical parameters can then be derived as:

$$\begin{cases} \alpha = t \ a^{-1} \left((z_p - z_o) (x_p - x_o)^{-1} \right) \\ t_x = (x_p z_o - x_o z_p) \left((x_p - x_o)^2 + (z_p - z_o)^2 \right)^{-1} \\ t_y = - \left((x_p - x_o)^2 + (z_p - z_o)^2 \right)^{-1/2} \\ t_z = (x_o^2 + z_o^2 - x_o x_p - z_o z_p) \left((x_p - x_o)^2 + (z_p - z_o)^2 \right)^{-1} \end{cases}, \quad (22)$$

where x_o, z_o, x_p and z_p can be computed using Equation (6).

4. Results

The proposed camera calibration algorithm has been tested on both simulated data and real data, including data from both controlled and uncontrolled environments.

4.1 Experiments on Simulated Data

Simulated data are used to test the accuracy of the proposed method under different circumstances. In the experiments, the image resolution is set to be 640 (pixel) \times 480 (pixel). The ground true values for the camera parameters are set as follows: the tilt angle β is 25° , the roll angle γ is 10° , the focal length f is 1000 pixel, and $\alpha=59.1^\circ$, $t_x=8.2$ mm, $t_y=-14.9$ mm, and $t_z=-12.7$ mm. Under this settings, the projection of the WCS origin is (280.0, 475.0) and the projection of a 3D point (1000mm, 0, 0) on X axis is (319.5, 460.0).

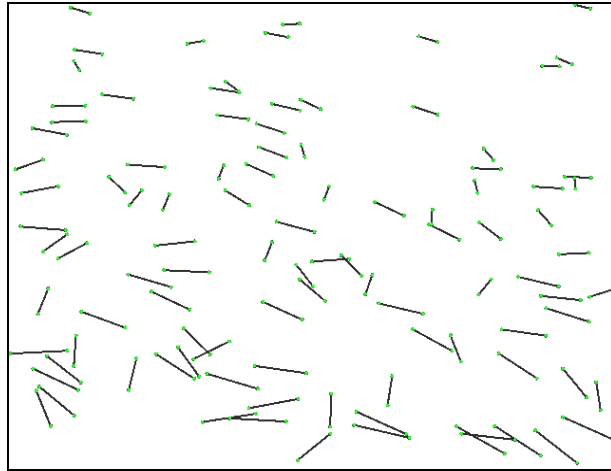


Figure 2: Image of 100 observations of a 1D object.

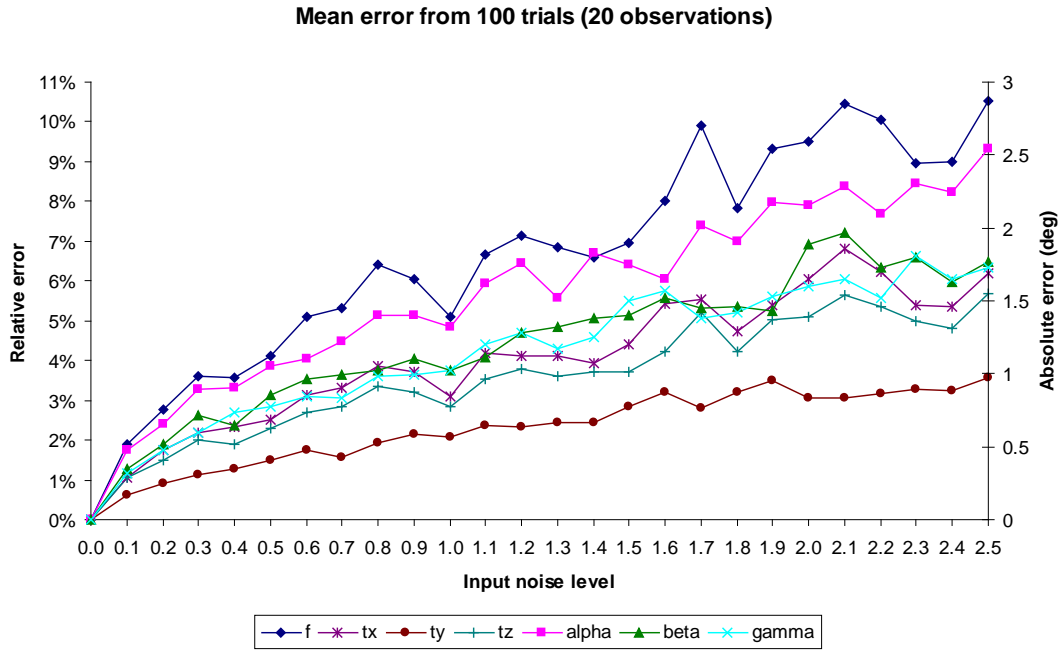
Two individual tests are conducted on the simulated data. The first one tests the calibration accuracy relative to the input noise level. The second one tests the calibration accuracy relative to the number of observation.

In Test 1, for each input noise level, 100 independent trials are performed. In each trial 20 observations of a 1D object is randomly generated. An example is shown in Figure 2 which blends all the observations together. Gaussian noise with 0 mean and σ standard deviation is added to the image points, where σ is varied from 0 to 2.5 pixels with the step length of 0.1 pixel. Based on the randomly generate video clips with the specified input noise, camera calibration is performed using the proposed technique. After calibrating the critical parameters, namely the focal length, the tilt angle and the roll angle, the non-critical parameters do not need to be calibrated if the pre-defined WCS is used.

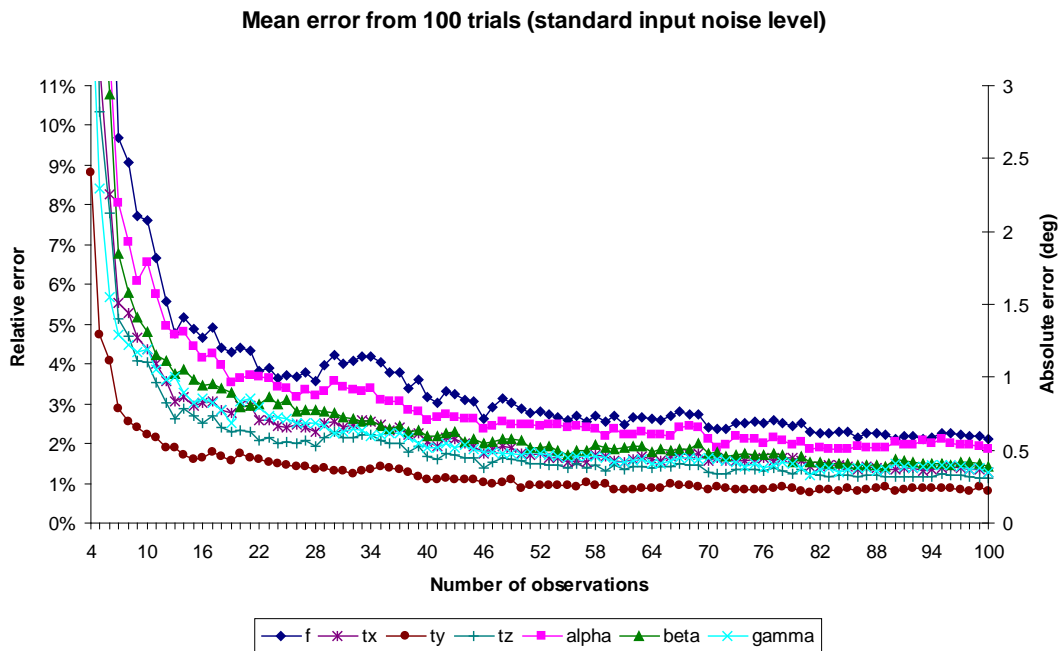
However, for the convenience of verifying the accuracy of non-critical parameter calibrations in the same scale as the ground truths, the user-specified WCS is used by giving the preference that a 2D vector from the image point (280.0, 475.0) to the image point (319.5, 460.0) is the projection of a 3D vector from the WCS origin to a point on X axis with the magnitude of 1000mm. The calibrated parameters are compared with the ground truths. The accuracies of the calibrated parameters are measured in different ways. For focal length, the relative error with respect to the ground truth ($|\Delta f|/f$) is measured; For pan (α), tilt (β), and roll (γ) angles, the absolute errors ($|\Delta\alpha|, |\Delta\beta|, |\Delta\gamma|$) are measured; For the translation parameters, the ground true magnitude of the vector from the camera center to WCS origin is first computed. The difference in each of the translational parameters is then computed in relation to this magnitude to obtain the relative error ($|\Delta t_x|/l, |\Delta t_y|/l, |\Delta t_z|/l$, where $l=(t_x^2+t_y^2+t_z^2)^{0.5}$).

Figure 3(a) shows the mean value of the relative errors of the focal length and the absolute errors of the tilt/roll angle from the 100 independent trials.

In Test 2, similar ground truth settings as Test 1 are used. The input noise level of 0.5 pixel (a.k.a. standard input noise level) is assumed. The number of observations varies, ranging from 4 to 100. For each number of observations, 100 independent trials are performed. The accuracy of the calibrated parameters is measured in the same way as in Test 1. The results of Test 2 are shown in Figure 3(b).



(a)



(b)

Figure 3: Experimental results on simulated data: (a) Calibration error with regards to input noise level; (b) Calibration error with regards to number of observations.

It can be seen from the results shown Figure 3(a) that the errors of all calibrated parameters are generally proportional to the input noise level. When the input noise level increases from 0 to 2.5 pixel, the mean relative error of focal length increases from 0 to about 11%, the absolute error of pan/tilt/roll angle increases from 0 to about 2.5° (shown in the secondary y-axis). The relative error of each translation parameter increases from 0 to less than 7%. From the results shown Figure 3(b) it can be seen that the errors of all calibrated parameters decreases as the number of observation increases. When the number of observations is 4, the relative error of every parameter is very large. When the number of observation is around 15, the relative error of each parameter is acceptable, with the mean relative error of the focal length at about 5%, the mean absolute error of rotation angles are less than 1.5° , and the mean relative error of each translation factor under 4%. When the number of observation reaches 100, the mean relative error of the focal length drops to 2%, the mean absolute error of rotational angles drops to under 0.5° and the mean relative error of each translation factor is decreased to be under 2%.

The experiments on simulated data demonstrate that the proposed calibration algorithm can produce very high accuracy on both intrinsic and extrinsic parameters under the standard input noise level (which sufficiently models the noise in most real data), if the number of observations can be more than 10.

4.2 Experiments on Real Data

After the successful experiments on simulated data, the proposed method is tested against some real video data, which are acquired from both controlled and uncontrolled environments. An explicit calibration object for the proposed method can be an explicit 1D object, or an object with an explicit 1D segment. These kinds of objects are very easy to find in the real world, which means that the proposed camera calibration method can be widely used for real videos. Since no ground true camera parameters are available in the tests for real videos, a popular plane-based camera calibration toolbox [26] is used to provide the reference value for all parameters. In the plane-based calibration cases, the user is required to specify the projection of the WCS origin and two axes on the plane with metric information for the purpose of constructing the WCS. Because the camera model defined in this camera

calibration toolbox is different from the one in this project, some parameter conversions are performed as follows. The reference focal length f is approximated by averaging the two calibrated focal lengths (represented in pixel width and pixel height respectively). The other parameters α , β , γ , t_x , t_y , and t_z are derived from the calibrated rotation matrix and translation vector. With these values derived, all camera parameters calibrated using the proposed method can be compared with the reference values.

For real data, if the user does not specify the WCS information, the non-critical parameters will be derived using the pre-defined WCS. The accuracy of the non-critical parameters derived this way cannot be compared with those calibrated using the plane-based method. In order to evaluate the accuracy of all camera parameters, the user must provide input for the specifications of the WCS. The same user input will be used for the plane-based calibration method and the proposed method with the exception that only one axis with metric information is required in the proposed method. Please note that in actual application of the proposed method, the users do not have to provide WCS specifications if they choose not to. In that case the non-critical camera parameters will be derived based on the pre-defined WCS, rather than calibrated (i.e. $\alpha=0$ and $\mathbf{T}_S=[0,-1,0]^T$ as described in Section 3.1). Again different schemes are used for measuring the accuracy of different parameters. For the focal length f , it is compared with respect to the reference focal length. For α , β , and γ , the absolute error is measured. For t_x , t_y , and t_z , they are measured relatively with respect to the distance between the origins of CCS and the reference WCS, calibrated using the plane-based method.

4.2.1 Real data from the controlled environment

In the controlled environment, an object with 1D information is directly used as the calibration object. First, the algorithm is tested using a chopstick– which is an explicit 1D calibration object; next, the algorithm is tested using a piece of paper, one edge of which is treated as a calibration object.

- Calibration using an explicit 1D object – a chopstick

A chopstick is an explicit 1D object, since it can be represented as a line segment with two end points. 22 observations of a single chopstick are recorded by randomly placing it on the ground. A sample

image is shown in Figure 4(a1). The chopstick is detected automatically from each image using simple image processing techniques, e.g., background subtraction and Hough transformation. From the 2D end points of a chopstick, three camera parameters are obtained through DHC optimization as described in Sec. 3. In order to evaluate the accuracy of the calibrated non-critical parameters, the same input for the plane-based calibration method, i.e., the 2D positions of O (for WCS origin) and P as shown in Figure 4(a4) and the measured distance ($28\text{mm}\times 7$ grids) between them, are used to compute the rest of the extrinsic parameters.

These errors are within very reasonable ranges, considering the existence of input noises in the real video. Other extrinsic parameters are calibrated based on the critical parameters and user's preference. It can be seen that the angle α can be quite accurately calibrated with error of about 1.5° , and the relative errors of t_x , t_y and t_z are under 5%.

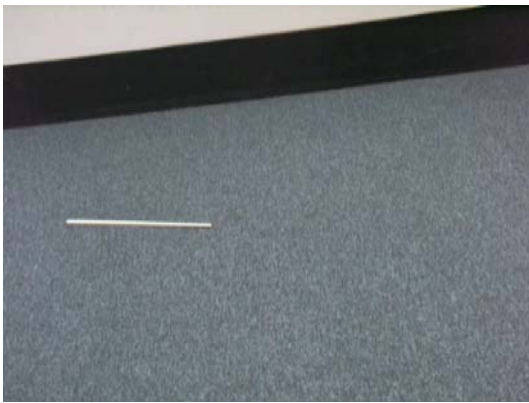
Based on the three calibrated critical camera parameters (f , β , and γ), the 2D homography matrix can be computed using Equation (3). In order to have an overview of all observations of the chopstick, we stacked all images into a single image shown in Figure 4(a2). The computed homography matrix is applied to this image and the recovered ground plane pattern is shown in Figure 4(a3). It can be seen that all appearances of the chopstick show a very similar length in the recovered image which confirm in another way the accuracy of the camera calibration.

- Calibration using an object with an explicit 1D segment - a paper edge

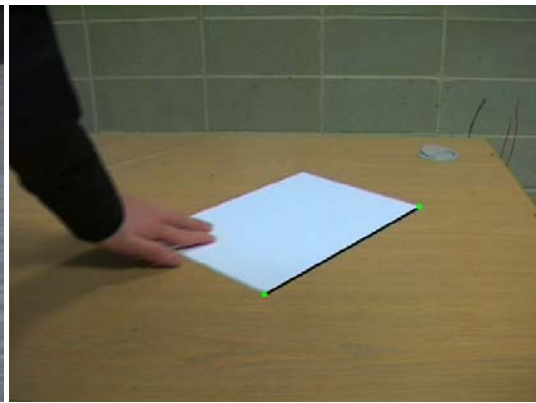
A piece of print/copy paper can be easily found in the office environment. It can be used straight away in the proposed algorithm, no matter whether it is blank or printed, without any specific patterns on it, because any edge of the paper is an explicit 1D segment. In this experiment a piece of A4 paper is moved randomly on a desk in front of the fixed camera. One of its long edges is chosen as the 1D calibration object, which is illustrated as a black line segment with green end points in Figure 4(b1). 20 observations of the A4 paper are used to calibrate the three critical camera parameters using the proposed method. The user's input is taken from the plane-based calibration method: the 2D positions of O and P as shown in Figure 4(b4) and the measured distance ($28\text{mm}\times 5$) between them. These inputs are used to compute the remaining extrinsic parameters. The experiment results in Table 1 show very

high accuracy with the relative differences for f at about 2%, for β , γ and α at under 2.5° , and for t_x , t_y and t_z at under 2%. Figure 4(b2) and Figure 4(b3) show the extracted paper edges and their positions after plane rectification.

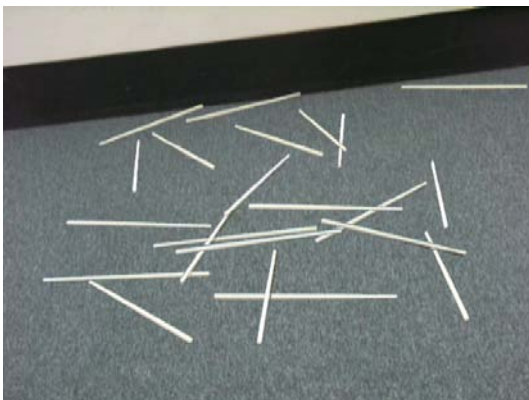
The experiments with an explicit calibration object have demonstrated the high accuracy and high flexibility of the proposed method. Compared with the traditional plane-based method, no specific preparation is required for the calibration object while the same level of calibration accuracy is achieved. Next, we will demonstrate even higher flexibility of the proposed method when the input data is captured from uncontrolled environment, where no specific object is prepared to be the calibration object.



(a1)



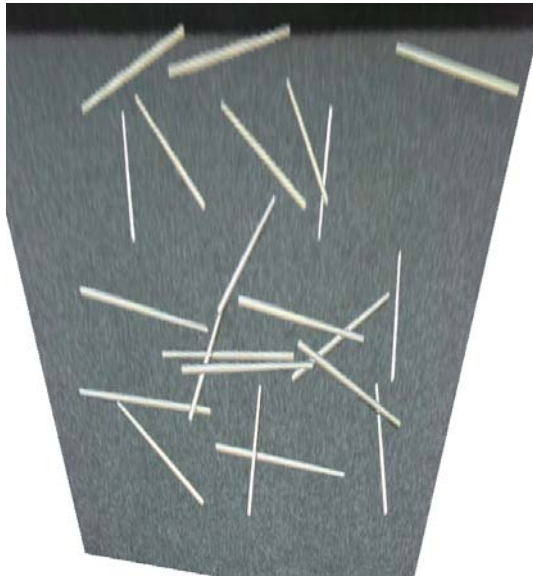
(b1)



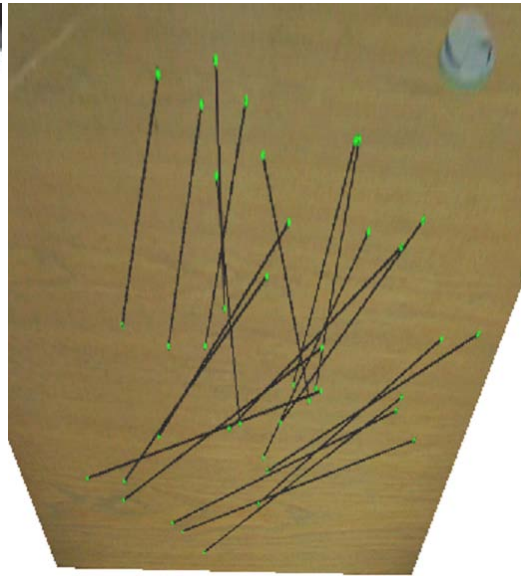
(a2)



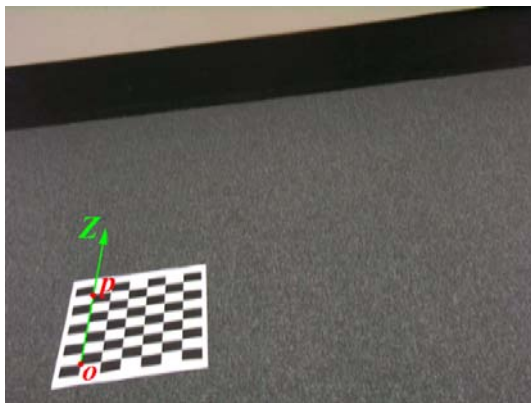
(b2)



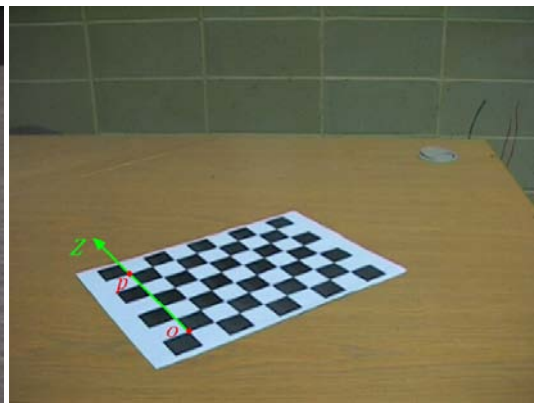
(a3)



(b3)



(a4)



(b4)

Figure 4: Real data with an explicit calibration object: (a) Explicit 1D object (chopstick), (b) Explicit 1D segment (paper edge): (1) Sample image; (2) Blended image; (3) Rectified plane; (4) Reference for extrinsic parameters.

4.2.2 Real data from the uncontrolled environment

In many circumstances, it is desirable to perform camera calibration from some pre-input data with no explicit calibration object present in them. In this kind of cases, if some information contained in the real data can be interpreted as a 1D object that satisfies the assumptions, the proposed method can be used for camera calibration. Three examples are presented here to show that the 1D calibration object satisfying our requirement can be either identified, or created in many uncontrolled environment. In the

first example the calibration object is identified from a cabinet door, which is an object with an explicit 1D segment. In the second example the calibration object is identified from a moving toy car, which is an object with no explicit 1D segment but containing explicit 1D information. The last example shows that it is even possible to create the calibration object from human walking motion, which contains implicit 1D information over time. In these examples, the plane-based calibration method is again used but is only for the purpose of providing reference values in each experiment. In these tests, only the critical parameters are compared because it is generally difficult for an user to specify preferences in pre-recorded video and thus the pre-defined WCS is used where the non-critical parameters are fixed ($\alpha=0$ and $\mathbf{T}_s=[0,-1,0]^T$).

- Calibration from an object with an explicit 1D segment – a cabinet door

Figure 5(a1) shows an image in which a person is opening the cabinet door in the office. The bottom edge of a cabinet door (illustrated in a red line segment) is an explicit 1D segment and is always parallel to the floor when moving. We can safely identify it as a 1D calibration object for the proposed method although it is not contained in a visible plane. 30 observations are obtained for the calibration and the corresponding “1D object” is illustrated in Figure 5(a2). The calibration results in Table 1 have shown that the relative error of focal length is about 10%, and the absolute error of tilt and roll angles are under 0.5° . It can be seen from Figure 5(a3) that on the reference plane rectified based on the calibrated camera parameters, all 1D segments have a consistent length validating that they are from the same object.

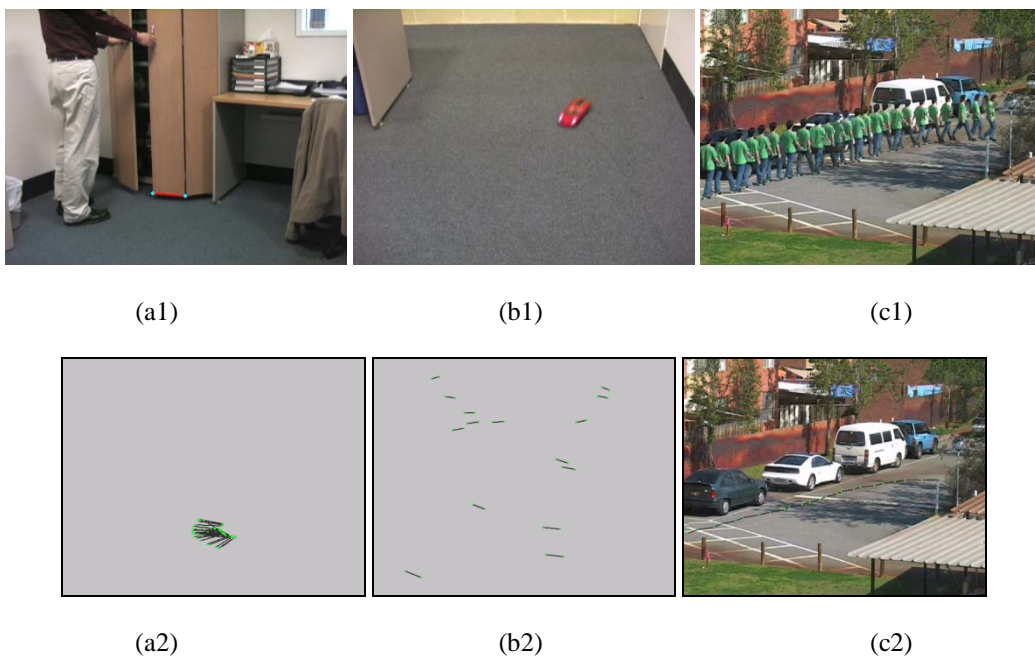
- Calibration from an object with no explicit 1D object but explicit 1D information – a moving toy car

The motion of a moving toy car with head lights on is captured. From the sample frame shown in Figure 5(b1), it can be noticed that a segment connecting the two head lights has fixed length and is always parallel to the floor. The two head lights can hence be treated as the two end points of a 1D calibration object, though there is not such an explicit 1D segment between them. Table 1 shows the calibration results (from 15 observations illustrated in Figure 5(b2)) for this experiment. The relative

error of focal length is about 8% and the absolute error of tilt and roll angles are under 2° . The rectified plane in Figure 5(b3) has shown the consistency.

- Calibration from implicit 1D information without an explicit object – human walking

Human walking is a very common scene in the daily life. It may not be possible to find any explicit information to represent a 1D object in images containing human walking. However, when a motion sequence is captured, the foot step (from left/right stance foot to right/left stance foot) can be treated roughly as a 1D calibration object if the person is walking on a relatively leveled ground, because the length of foot step is roughly constant for a uniformly moving person. Figure 5(c1) shows such a motion sequence, with all extracted footsteps shown in Figure 5(c2). Using the extracted 1D information (23 line segment observations), the proposed method can be applied. The calibration results are shown in Table 1. It can be seen that the relative error of the focal length is about 6%, while the absolute error of tilt/roll angle is about 1.5° . Similarly, Figure 5(c3) shows that the footsteps on the rectified plane (can be considered as viewing from the top) have consistent lengths.



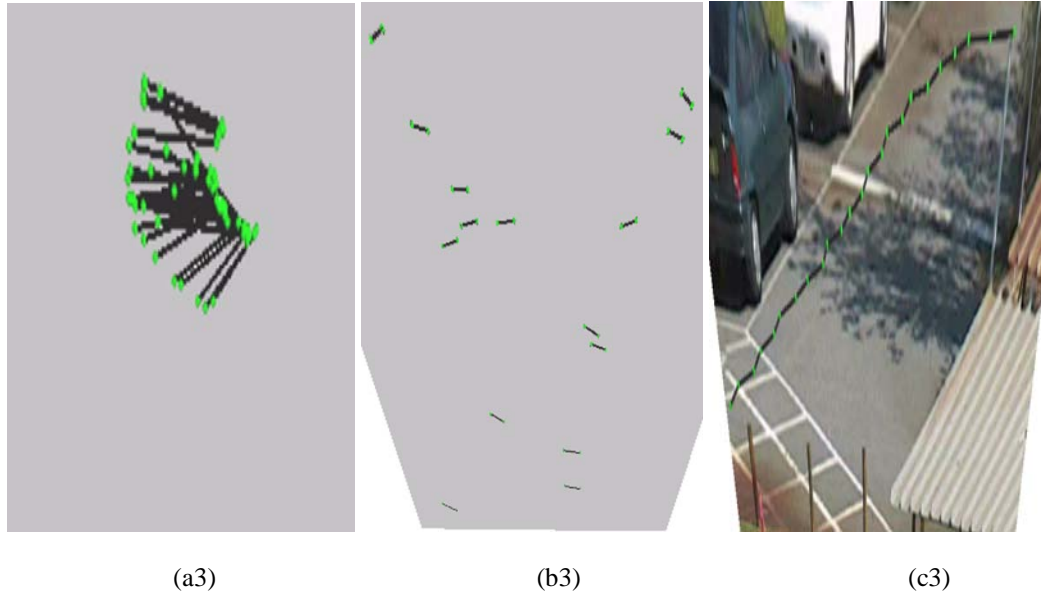


Figure 5: Real data without an explicit calibration object: (a) Explicit 1D segment (cabinet door edge), (b) Explicit 1D information (toy car headlights), (c) Implicit 1D information (footstep): (1) Sample image; (2) Blended image; (3) Rectified plane.

Table 1. Calibration results on real data

		f	β	γ°	α°	t_x	t_y	t_z
Controlled environment	Explicit 1D object (chopstick)	1332.4 pixel	22.9°	-6.7°	-2.1°	273.6 mm	-783.4 mm	-1242.2 mm
	Reference							
	Result	1273.8 pixel	25.1°	-7.3°	-0.6°	272.8 mm	-810.6 mm	-1311.1 mm
	Error	4.40%	2.2°	0.6°	1.5°	0.05%	1.82%	4.61%
	Explicit 1D segment (paper edge)	789.8 pixel	23.4°	-2.2°	-32.2°	-214.5 mm	-333.9 mm	-468.6 mm
	Reference							
	Result	773.9 pixel	22.5°	-1.1°	-34.7°	-225.4 mm	-342.0 mm	-470.9 mm
	Error	2.01%	0.9°	1.1°	2.5°	1.77%	1.32%	0.37%
	Explicit 1D segment (cabinet door edge)	831.5 pixel	12.6°	-0.7°	N/A	N/A	N/A	N/A
	Reference							
	Result	750.8 pixel	12.6°	-0.3°	0	0	-1 unit	0
	Error	9.70%	0.0°	0.4°	N/A	N/A	N/A	N/A
Uncontrolled environment	Explicit 1D information (toy car headlights)	749.8 pixel	33.5°	-2.5°	N/A	N/A	N/A	N/A
	Reference							
	Result	809.3 pixel	33.3°	-0.8°	0	0	-1 unit	0
	Error	7.94%	0.2°	1.7°	N/A	N/A	N/A	N/A
	Implicit 1D information (footstep)	2629.6 pixel	7.9°	-0.5°	N/A	N/A	N/A	N/A
	Reference							
Result	2779.1 pixel	8.3°	0.6°	0	0	-1 unit	0	

<i>(footstep)</i>	<i>Error</i>	5.69%	0.4°	1.1°	N/A	N/A	N/A	N/A
-------------------	--------------	-------	------	------	-----	-----	-----	-----

From the experiments in real data, great application potentials of the proposed method have been demonstrated. In Section 4.2.1, it has been demonstrated that the proposed calibration method can work well in controlled environments and can serve as an alternative to the plane-based method. In Section 4.2.2, it has been shown that the proposed calibration can also work in uncontrolled environments where other object-based methods would fail due to the absence of explicit calibration objects and the self-calibration methods would also fail due to insufficient number of point correspondences and the relatively static scene structure.

Through the experiments on both the simulated and real data, it can be seen that the proposed camera calibration method produce highly satisfactory results. One disadvantage of the proposed calibration method is that it could not recover comprehensive camera parameters, unlike some existing calibration methods. However, the benefits of the proposed method significantly outweigh this drawback. The proposed method can handle many cases where existing methods fail to work due to the absence of required information. In addition, the camera model calibrated using the proposed method is sufficient for most practical applications.

5. Conclusions and Future Work

In this paper, we proposed a novel approach to calibrate both the focal length and extrinsic parameters of a fixed ideal pin-hole camera from observations of a moving 1D object on a fixed plane. The proposed method does not require the traditional 3D object, 2D pattern or 1D object containing three or more collinear points, instead, only the simplest geometry of a 1D object with two end points is required. The object motion required by the proposed method is also less restricted than existing 1D object-based methods. The experimental results show that camera parameters can be calibrated with high accuracy from both simulated and real data. Another significant feature of the proposed algorithm is that it shows great potential on camera calibration based on a wide range of everyday videos, including those with no explicit 1D object. The proposed method is only a small step away from making object-based calibration employable in most uncontrolled environments. The motion required

for the 1D object still need to be further investigated so that more situations can be handled especially for those where self-calibration fails to work.

Reference

- [1] Hartley, R. Camera Calibration Using Line Correspondences. in DARPA93. 1993.
- [2] Maybank, S.J. and O.D. Faugeras, A Theory of Self-Calibration of a Moving Camera. *Int. J. Comput. Vision*, 1992. 8(2): p. 123-151.
- [3] Thirion, E. and C. Ronse. Self Calibration and 3D Reconstruction from Lines with a Single Translating Camera. in *The British Machine Vision Conference*. 1996. Universite Louis Pasteur, Strasbourg.
- [4] Faugeras, O.D., Q.-T. Luong, and S.J. Maybank. Camera Self-Calibration: Theory and Experiments. in *European Conference on Computer Vision*. 1992.
- [5] Hartley, R.I. Self-calibration from multiple views with a rotating camera. in *ECCV '94: Proceedings of the third European conference on Computer vision (vol. 1)*. 1994. Stockholm, Sweden: Springer-Verlag New York, Inc.
- [6] Luong, Q.-T. and O.D. Faugeras, Self-Calibration of a Moving Camera from Point Correspondences and Fundamental Matrices. *Int. J. Comput. Vision*, 1997. 22(3): p. 261-289.
- [7] Lv, F., T. Zhao, and R. Nevatia. Self-Calibration of a Camera from Video of a Walking Human. in *ICPR '02: Proceedings of the 16 th International Conference on Pattern Recognition (ICPR'02) Volume 1*. 2002. Washington, DC, USA: IEEE Computer Society.
- [8] Lv, F., T. Zhao, and R. Nevatia, Camera Calibration from Video of a Walking Human. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2006. 28(9): p. 1513-1518.
- [9] Grammatikopoulos, L., Karras, G.E., Pets, E., and Kalisperakis, I. A unified approach for automatic camera calibration from vanishing points, in *IEVM06*. 2006.
- [10] Faugeras, O., *Three-Dimensional Computer Vision: a Geometric Viewpoint*. 1993, Cambridge, MA, USA: MIT Press.
- [11] Grammatikopoulos, L., G.E. Karras, and E. Petsa. Camera Calibration Approaches Using Single Images of Man-Made Objects. in *CIPA 2003 XIXth International Symposium*. 2003. Antalya, Turkey.
- [12] Tsai, R., A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology using Off-the-Shelf TV Cameras and Lenses. *Robotics and Automation, IEEE Journal of* [legacy, pre - 1988], 1987. 3(4): p. 323-344.

- [13] Hartley, R.I. and A. Zisserman, Multiple View Geometry in Computer Vision. 2000: Cambridge University Press, ISBN: 0521623049.
- [14] Triggs, B. Autocalibration from Planar Scenes. in ECCV '98: Proceedings of the 5th European Conference on Computer Vision-Volume I. 1998. London, UK: Springer-Verlag.
- [15] Zhang, Z., A Flexible New Technique for Camera Calibration. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2000. 22(11): p. 1330-1334.
- [16] Cao, X. and H. Foroosh. Camera Calibration Without Metric Information Using 1D Objects. in Image Processing, 2004. ICIP '04. 2004 International Conference on. 2004.
- [17] Hammarstedt, P., P. Sturm, and A. Heyden. Degenerate Cases and Closed-Form Solutions for Camera Calibration With One-Dimensional Objects. in Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on. 2005.
- [18] He, X., et al. Estimation of Internal and External Parameters for Camera Calibration Using 1D Pattern. in Proc. IEEE International Conference on Video and Signal Based Surveillance (AVSS'06). 2006. Los Alamitos, CA, USA: IEEE Computer Society.
- [19] Zhang, Z., Camera Calibration with One-Dimensional Objects. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2004. 26(7): p. 892-899.
- [20] Wu, F.C., Z.Y. Hu, and H.J. Zhu, Camera Calibration with Moving One-Dimensional Objects. Pattern Recognition, 2005. 38(5): p. 755-765.
- [21] Qi, F., et al., Camera Calibration with One-Dimensional Objects Moving under Gravity. Pattern Recognition, 2007. 40(1): p. 343-345.
- [22] Qi, F., et al., Constraints on General Motions for Camera Calibration with One-Dimensional Objects. Pattern Recognition, 2007. 40(6): p. 1785-1792.
- [23] Bose, B. and Grimson, E. Ground Plane Rectification by Tracking Moving Objects, in Proceedings of the Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS). 2003.
- [24] Liebowitz, D. and Zisserman, A. Metric Rectification for Perspective Images of Planes, in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (ICPR). 1998.
- [25] Yuret, D. and M. Maza. Dynamic Hill Climbing: Overcoming the Limitations of Optimization Techniques, in the Second Turkish Symposium on Artificial Intelligence and Neural Networks. 1993.

[26] Bouguet, Jean-Yves. Camera Calibration Toolbox for Matlab, available online at http://www.vision.caltech.edu/bouguetj/calib_doc/, Oct 2008.

About the Author – EN PENG received his B.Sc. and Ph.D. degrees in Computer Science from South China Normal University in 2002 and Curtin University of Technology in 2008 respectively. He is now a research associate with Institute for Multi-sensor Processing and Content Analysis (IMPCA), Department of Computing, Curtin University of Technology, Australia.

About the Author – LING LI received her B.Sc., M.Eng. and Ph.D. degrees in computer science and engineering from Sichuan University, China Academy of Post and Telecommunication, and Nanyang Technological University (Singapore) respectively. She worked as an Assistant Professor and Associate Professor in Nanyang Technological University from 1994 to 2002. She is now an Associate Professor in the Department of Computing, Curtin University of Technology, Australia. Her main research interests include physically-based and video-based computer graphics and animation.