Camera Calibration from a Single Image based on Coupled Line Cameras and Rectangle Constraint

Joo-Haeng Lee
Robot and Cognitive Systems Research Dept., ETRI
joohaeng@etri.re.kr

Abstract

Given a single image of a scene rectangle of an unknown aspect ratio and size, we present a method to reconstruct the projective structure and to find camera parameters including focal length, position, and orientation. First, we solve the special case when the center of a scene rectangle is projected to the image center. We formulate this problem with coupled line cameras and present the analytic solution for it. Then, by prefixing a simple preprocessing step, we solve the general case without the centering constraint. We also provides a determinant to tell if an image quadrilateral is a projection of a rectangle. We demonstrate the performance of the proposed method with synthetic and real data.

1. Introduction

Camera calibration is one of the most classical topics in computer vision research. We have an extensive list of related works providing mature solutions. In this paper, we are interested in a special problem to calibrate a camera from a single image of an unknown scene rectangle. We do not assume any prior knowledge on correspondences between scene and image points. Due to the limited information, a simple camera model is used: the focal length is the only unknown internal parameter.

The problem in this paper has a different nature with classical computer vision problems. In the PnP problem, we find transformation matrix between the scene and camera frames with prior knowledge on the correspondences between the scene and image points as well as the internal camera parameters [1]. In camera resection, we find the projection matrix from known correspondences between the scene and image points without prior knowledge of camera parameters [2]. Camera self-calibration does not rely on a known Euclidean structure, but requires multiple images from camera motion [3].

Several approaches are based on geometric properties of a rectangle or a parallelogram. Wu et al. proposed a calibration method based on rectangles of known aspect ratio [4]. Li et al. designed a rectangle landmark to localize a mobile robot with an approximate rectangle constraint, which does not give an analytic solution [5]. Kim and Kweon propose a method to estimate intrinsic camera parameters from two or more rectangle of unknown aspect ratio [6]. A parallelogram constraint can be used for calibration, which generally requires more than two scene parallelograms projected in multiple-view images as in [7, 8, 9].

Our contribution can be summarized as follows. Based on a geometric configuration of coupled line cameras, which models a simple camera of an unknown focal length, we give an analytic solution to reconstruct a complete projective structure from a single image of an unknown rectangle in the scene: no prior knowledge is required on the aspect ratio and correspondences.

Then, the reconstruction result can be utilized in finding the internal and external parameters of a camera: focal length, rotation and translation. The proposed solution also provides a simple determinant to tell if an image quadrilateral is a projection of a scene rectangle.

In a general framework for plane-based camera calibration, camera parameters can be found first using the image of the absolute conic (IAC) and its relation with projective features such as vanishing points [2, 10]. Then, a scene geometry can be reconstructed using a non-linear optimization on geometric constrains such orthogonality, which cannot be formulated as a closed-form in general.

2. Problem Formulation

2.1. Line Camera

A line camera is a conceptual camera, which follows the same projection model of a standard pin-hole camera. (See Fig 1a.) Let \(v_0v_2\) be a line segment in the
scene, which will be projected as a line $uc0v2$ in the line camera $C0$. Especially, we are interested in the position $pc$ and the orientation $θ0$ of $C0$ when the principal axis passes through the center $vm$ of $v0v2$ and the center $um = (0, 0, 1)$ of image line.

To simplify the formulation, we assume a canonical configuration where $∥vmv0∥ = 1$, and $vm$ is placed at the origin of the world coordinate system: $vm = (0, 0, 0)$. For derivation, we define followings:

$$d = ∥vmv0∥, \quad l_i = ∥vmui∥, \quad ψ_i = ∠vmv_i, \quad and \quad s_i = ∥pmv0∥.$$  

In this configuration, we can derive the following relation:

$$l_2 \frac{l_0}{l_0} = \frac{d - \cos θ_0}{d + \cos θ_0} = \frac{d_0}{d_1},$$  

where $d_0 = d - \cos θ_0 = s_0 \cos ψ_0$ and $d_1 = d + \cos θ_0 = s_2 \cos ψ_2$. We can derive the relation between $θ_0$ and $d$ from (1):

$$\cos θ_0 = d (l_0 - l_2)/(l_0 + l_2) = d α_0$$  

where $α_0 = (l_i - l_{i+2})/(l_i + l_{i+2})$, which is solely derived from a image line $v_iu_{i+2}$. Note that $θ_0$ and $d$ are sufficient parameters to determine the exact position of $pc$ in 2D. When $α_0$ is fixed, $pc$ is defined on a certain sphere as in Fig 1b. Once $θ_i$ and $d$ are known, additional parameters can be also determined: $\tan ψ_i = \sin θ_i/d$ and $s_i = \sin θ_i/\sin ψ_i$.

### 2.2. Coupled Line Cameras

A standard pin-hole camera can be represented with two line cameras coupled by sharing the center of projection. (See Fig 2.) Let a scene rectangle $G$ have two diagonals $v0v2$ and $v1v3$, each of which follows the canonical configuration in Section 2.1: $∥vmv0∥ = 1$ where $vm = (0, 0, 0)$. Each diagonal $vmv_{i+2}$ is projected to an image line $vmv_{i+2}$ in a line camera $C_i$. When two line cameras $C_i$ share the center of projection $pc$, two image lines $vmv_{i+2}$ intersect at $um$ on the common principal axis, say $pmv_m$. The four vertices $ui$ form a quadrilateral $H$, which is the projection of the rectangle $G$ in a pin-hole camera with the center of projection at $pc$. Note that the principal axis passes through $vm, um$, and $pc$.

Using this configuration of coupled line cameras, we find the orientation $θ_i$ of each line camera $Ci$ and the length $d$ of the common principal axis from a given quadrilateral $H$. Using the lengths of partial diagonals, $l_i = ∥vmui∥$, we can find the relation between the coupled cameras $Ci$ from (1):

$$\sin θ_i(d - \cos θ_0) \sin θ_i(d - \cos θ_1)$$  

Manipulation of (2) and (3) leads to the system of nonlinear equations:

$$d = \frac{β \sin θ_0 \cos ψ_1 - \cos θ_0 \sin ψ_1}{β \sin θ_0 - \sin θ_1} = \frac{cos θ_0}{α_0} = \frac{cos θ_1}{α_1}$$  

where $β = l_1/l_0$. Using (4), the orientation $θ_0$ can be represented with coefficients, $α_0, α_1$, and $β$, that are solely derived from a quadrilateral $H$:

$$\tan θ_i/2 = \sqrt{\frac{A_0 + A_1 \pm 2\sqrt{A_0A_1}}{A_1 - A_0}} = \sqrt{D_±}$$  

where

$$A_0 = B_0 + 2B_1, \quad A_1 = B_0 - 2B_1 \quad B_0 = 2(α_0 - 1)^2(α_0^2 + α_1^2) - 4α_0^2(α_1 - 1)^2β^2 \quad B_1 = (α_0 - 1)^2(α_0 - α_1)(α_0 + α_1)$$

The actual value of $θ_0$ should be chosen to make $d > 0$ using (2). With known $θ_0$, we can compute values of $θ_i$ and $d$ using (4). Based on the result of this section, the explicit coordinates of $pc$ and the shape of $G$ are reconstructed in Section 3.1.

### 2.3. Rectangle Determination

In our configuration of coupled line cameras, the existence of $θ_i$ implies that $H$ is the projection of a canonical rectangle $G$. As the orientation angle $θ_0$ should be
a positive real number, the expression inside the outer square root of (5) also should be a positive real number:

\[ D_\pm = \frac{A_0 + A_1 \pm 2\sqrt{A_0 A_1}}{A_1 - A_0} > 0 \]  

(6)

The above condition guarantees that a quadrilateral \( H \) is the projection of a scene rectangle \( G \), which will be fully reconstructed in Section 3.1.

3. Reconstruction and Calibration

3.1. Reconstructing Projective Structure

We define a projective structure as a frustum with a rectangular base \( G \) and an apex in the center of projection \( p_c \). (See Fig 2.) Since \( G \) has a canonical form and parameterized by \( \phi \), its vertices can be represented as \( v_0 = (1, 0, 0), v_1 = (\cos \phi, \sin \phi, 0), v_2 = -v_0, \) and \( v_3 = -v_1 \) where \( \phi \) is the crossing angle of two diagonals. Since \(|v_i| = 1\), the crossing angle \( \phi \) is represented as

\[ \cos \phi = \langle v_0, v_1 \rangle \]  

(7)

Two diagonals of the quadrilateral \( H \) intersect at the image center \( u_m \) on the principal axis with the crossing angle \( \rho \).

To compute \( \phi \), we denote the center of projection as \( p_c = (0, 0, 0) \), the image center as \( u_m = (0, 0, -1) \), the first two vertices of \( H \) as \( u_0^c = (\tan \psi_0, 0, -1) \) and \( u_1^c = (\cos \rho \tan \psi_1, \sin \rho \tan \psi_1, -1) \), the center of \( G \) as \( v_m^c = (0, 0, d) \), and the vertices of \( G \) as \( v_{m_i} = s_i u_i^c / \| u_i^c \| \). Since \( v_i = v_i^c - v_m^c \), \( \phi \) can be computed using (7) with known values of \( \rho, d, s_i, \) and \( \psi_i \).

The coordinates of \( p_c = (x, y, z) \) can be found by solving following equations: \( d \cos \psi_0 = x, d \cos \psi_1 = x \cos \phi + y \sin \phi \), and \( x^2 + y^2 + z^2 = d^2 \), which are derived from the projective structure. Now, the projective structure has been reconstructed as a frustum with the base \( G(\phi) \) and the apex \( p_c \).

Note that the reconstructed \( G \) is guaranteed to be a rectangle satisfying geometric constraints such as orthogonality since the proposed method is based on a canonical configuration of coupled line camera of Section 2.

3.2. Camera Calibration

Since our reconstruction method is formulated using a canonical configuration of coupled line cameras, the derivation based on \( H \) can be substituted with a given image quadrilateral \( Q \). Now we can find a homography \( H \) using four point correspondences between \( G(\phi) \) and \( Q \). A homography is defined as \( H = sKW \) where \( s, K \)

and \( W \) denote a scalar, a camera and a transformation matrices, respectively. Note that we assume a simple camera model \( K \): the only unknown internal parameter is a focal length \( f \). Since the model plane is at \( z = 0 \), it is straightforward to derive \( K \) and \( W \). See [2] for details.

3.3. Off-Centered Quadrilateral

In a centered case above, the centers of a scene rectangle \( G \) and an image quadrilateral \( Q \) are both aligned at the principal axis, which is exceptional in reality. The proposed method, however, can be readily applied to the off-centered case by prefixing a simple preprocessing step.

Let \( Q^0 \) and \( Q^9 \) denote an off-centered image quadrilateral and a corresponding scene rectangle, respectively. In the preprocessing step, we will find a centered quadrilateral \( Q \) which will reconstruct the projective structure where the projective correspondences between \( Q^0 \) and \( Q^9 \) are preserved. We can show that such \( Q \) can be geometrically derived from \( Q^9 \) using the properties of parallel lines and vanishing points. (See Fig 3.)

We choose one vertex \( u_0^q \) of \( Q^0 \) as the initial vertex \( u_0^q \) of \( Q \), say \( u_0 = u_0^q \). First, we compute two vanishing points \( u_1^q \) from \( Q^0 \): \( u_0^q \) as intersection of \( \overrightarrow{u_0^q u_1^q} \) and \( \overrightarrow{u_0^q u_3^q} \) and \( u_1^q \) as intersection of \( \overrightarrow{u_0^q u_1^q} \) and \( \overrightarrow{u_1^q u_2^q} \). Let \( u_2 \) be the opposite vertex of \( u_0 \) on the diagonal passing through \( u_0 \) and \( u_m \). \( u_2 = a(u_m - u_0) + u_0 \) where \( a \) is an unknown scalar defining \( u_2 \). Then, we can make symbolic definitions of the other vertices of \( Q \): \( u_1 \) as intersection of two lines \( \overrightarrow{u_0 u_1} \) and \( \overrightarrow{u_2 u_0} \), and \( u_3 \) as intersection of \( \overrightarrow{u_0 u_3} \) and \( \overrightarrow{u_2 u_3} \).

Note that the above definitions of \( u_i \) guarantees that each pair of opposite edges are parallel before projective distortion. The symbolic definitions of \( u_i \) above can be combined in one constraint: the intersection of \( \overrightarrow{u_0 u_2} \) and \( \overrightarrow{u_1 u_3} \) equals to the image center \( u_m^q \), which is formulated as a single equation of one unknown \( a \). Since we can find the value of \( a \), all the vertices \( u_i \) of the centered image quadrilateral \( Q \) can be computed ac-
We also tested with real images. In Fig 7a, the off-centered quadrilateral \( Q^g \) is the image of an ISO A4. We demonstrated the experimental results for both synthetic and real data. For experiment, the proposed method was implemented in Mathematica, and a Logitech C910 webcam was used for real data with autofocus enabled and 1280 × 720 resolution.

4.1. Synthetic Data

We tested with a synthetic scene containing a reference rectangle \( G^r \) (with random values of crossing angle \( \phi^r \), size, position and orientation) and a corresponding image quadrilateral \( Q^g \) generated by a known camera in Fig 4a.

First, we find a centered quadrilateral \( Q \) using vanishing points defined by \( Q^g \) in Fig 4b. Then, the projective structure is reconstructed as a frustum with the centered scene rectangle \( G \) and the center of projection \( p_c \) in Fig 4c. Then, the homography \( H \) between \( G \) and \( Q \) are derived, which is further decomposed to projection \( K \) and transformation \( W \) matrices as in Section 3.2. Using the homography \( H \) and the given off-centered quadrilateral \( Q^g \), we can reconstruct the target scene rectangle \( G^g(\phi^g) \) in Fig 4d.

In this example, the overall error can be measured between \( G^r \) and \( G^g \), which is found to be negligible: \( e_\Delta = \|\phi^r - \phi^g\| < 10^{-14} \) and \( \|v_i^g v_{i+2}^g - v_i^r v_{i+2}^r\| < 10^{-13} \). (Note that these error metrics are sufficient since the geometric constraint such as orthogonality is always satisfied as in Section 3.1.) Then, the focal length can be correctly reconstructed: \( f = 1280 \) in this example.

Figure 5: Experiment for a near-infinite vanishing point.

As a remark, we need to specially handle singular cases where two vanishing points \( w_0 \) and \( w_1 \) are not defined [10]. However, experimental results show that the proposed method is numerically stable for near-singular cases.

4.2. Real Data

We also tested with real images. In Fig 7a, the off-centered quadrilateral \( Q^g \) is the image of an ISO A4.
paper (210mm × 297mm) on a desk. ($Q^g$ is represented with its vertices, which can be extracted using a standard image processing method after any prior knowledge on the scene.) We can find a centered quadrilateral $Q$ in Fig 7b. Fig 7c and 7d show the results of reconstruction and rectification. The aspect ratios of an ISO A4 paper and reconstructed $G^g$ are 1.41427 and 1.40283, respectively, which is in the error of about 0.8%. Note that the reconstructed geometry $G^g$ is always a rectangle satisfying geometric constraints such as orthogonality. (See Section 3.1.)

We tested for a moving A4-sized paper in the same environment. The reconstruction and calibration were performed individually for each of nine instances of scene rectangles. The reconstructed canonical frustums were geometrically merged to be aligned along the common principal axis and to have the same center of projection in Fig 7e. The slight mismatches of frustums are mainly caused by lens distortion and auto-focusing, which are not modeled in the current method. The mean error of aspect ratios of reconstructed rectangles is within the error of 0.7% in Fig 7f, which can be reduced to 0.06% by compensating the lens distortion. The mean error of the computed focal lengths when compared to the explicit calibration [11] is about 0.9%.

5. Conclusion

We have presented a novel reconstruct-to-calibrate method for calibrating a simple camera based on reconstruction of the projective structure using a single image of a scene rectangle of an unknown aspect ratio. The proposed method is based on a geometric configuration of coupled line cameras, which leads to derive an analytic solution including a simple condition for rectangle determination.

We expect that our approach can be quite practical since rectangles are ubiquitous in everyday life. Moreover, the proposed method can be compactly implemented based on the closed-form solution without relying on any non-linear optimization, which is general in a calibrate-to-reconstruct approach. Hence, the proposed method may be suitable for applying in a mobile application.

Although the proposed method shows reliable results in both synthetic and real scenes, we need to further evaluate in various aspects such as numerical stability, robustness to noises, rigorous comparison with existing methods, and application in more complete camera models.

References