Unsupervised intrinsic calibration from a single frame using a "plumb-line" approach

R. Melo[1], M. Antunes[1], J. P. Barreto[1], G. Falcão[2] and N. Gonçalves[1]

[1] Institute for Systems and Robotics, University of Coimbra

[2] Instituto de Telecomunicações, University of Coimbra

Abstract

Estimating the amount and center of distortion from lines in the scene has been addressed in the literature by the so-called "plumb-line" approach. In this paper we propose a new geometric method to estimate not only the distortion parameters but the entire camera calibration (up to an "angular" scale factor) using a minimum of 3 lines. We propose a new framework for the unsupervised simultaneous detection of natural image of lines and camera parameters estimation, enabling a robust calibration from a single image. Comparative experiments with existing automatic approaches for the distortion estimation and with ground truth data are presented.

1. Introduction

We investigate the problem of fully calibrating an image with significant distortion without requiring any type of manual supervision. A solution for automatic, single frame calibration is specially relevant in the case of images mined from the internet, for which knowing the camera parameters can be useful for multiple tasks. Possible applications include the distortion correction via image warping for visualization purposes [20], the normalization of images for subsequent use in content retrieval frameworks [19], the estimation of camera rotation by aligning vanishing directions with a manhattan world scene [6], or the inference of 3D metric information from the image [22].

The article considers the case of cameras with distortion that can be described by the 1-parameter division model (DM) [7, 11], and assumes that the imaged scene has a reasonable number of straight lines. We propose for the first time a calibration algorithm that, given the image of 3 lines, it estimates the distortion, principal point, aspect ratio, and skew. Such result is not surprising if we consider that the division model has obvious resemblances with the stereographic projection used to describe the para-catadioptric sensor [3], and that para-catadioptric cameras can be fully
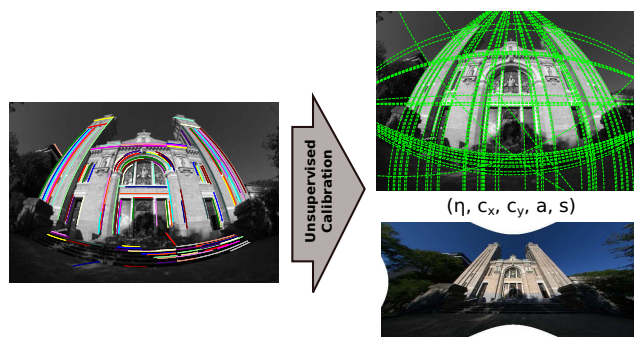

Figure 1: Proposed unsupervised calibration result.

calibrated from a minimum of 3 line images [4]. Nevertheless, and to the best of our knowledge, the possibility of calibrating the intrinsics of dioptric camera with distortion from 3 lines has never been reported. It is also shown that there are some differences with respect to the para-catadioptric case, namely the possibility of recovering aspect-ratio and skew from a single line projection, and the fact that the distortion can only be computed in pixels, being proved that line information is insufficient to decouple the focal length parameter from the distortion parameter in mm. Such decoupling requires additional information that is often available (e.g. the EXIF tag, the nominal field-of-view) and only knowing the distortion in pixels still enables accurate distortion compensation with the focal length being chosen as a function of the desired resolution for the output perspective [20].

The main contribution of the paper is a processing pipeline that receives as input a set of image contours, selects the ones that are likely to be projections of lines, and outputs both the detected lines and the camera calibration parameters (see Fig. 1). While standard "plumb-line" calibration requires user intervention for selecting the image edges that are projection of lines, our method carries this operation in a fully automatic manner. The line contours are detected in the clutter by evaluating their geometric consistency with putative calibrations. This leads to a complex

problem of multi-model fitting that is difficult to solve in practice. We start by showing that, if the camera is calibrated, then the line detection can be conveniently casted as an Uncapacited Facility Location (UFL) problem [17] that enables correct selection and clustering of contours. For the uncalibrated case we use contour triplets to establish different calibration hypothesis that give raise to different UFL instances. These multiple (UFL) instances can be combined in a large Hierarchical Facility Location (HFL) [12] problem for whose solution is the solution of the UFL instance with lowest energy. This provides an efficient, robust manner of simultaneously detecting the line contours and finding the camera calibration.

The approach is validated through extensive experiments using real images that prove the feasibility of unsupervised "plumb-line" calibration. It is important to mention that, although the algorithm has been originally designed for handling fully uncalibrated images, we verify that in practice weak assumptions about the camera aspect ratio and skew dramatically improve the stability and robustness of results. The reason for this is that the detected arc contours are usually small and, under strong occlusion, it is difficult to obtain plausible conic estimation [4] and, consequently, plausible calibration hypothesis to be used in the (UFL)-(HFL) framework. Nevertheless, this barely limits the usefulness of the algorithm because the assumption of zero skew and square pixels is verified by the vast majority of the images available in the internet.

1.1. Related work

The geometric calibration of cameras with distortion is a well-studied topic, with several methods and approaches being described in the literature [22]. However, none of these solutions is well suited for the automatic calibration of images mined from the internet. Auto-calibration techniques rely in point correspondences across views for determining both motion and camera parameters [5, 16], but it does not apply to our case where we have in general a single frame. Barreto et al. showed in [1] that it is possible to fully calibrate a camera with distortion using a single image of a chessboard pattern. Since we are addressing the calibration of images of natural scenes, this approach is also no solution for our problem. Very recently, Zhang et al. presented a technique for calibrating a camera using one or more images of a pattern that is unknown but sufficiently structured to be considered a low rank texture matrix [25]. Given a single image, the algorithm enables to recover the distortion parameters and the principal point whenever the scene has two patterns orthogonal to each other. This requirement considerably limits the number of cases where the camera parameters can be recovered. Moreover the approach requires the user to roughly indicate the localization of the patterns which precludes fully automatic applications.

Contrary to what happens with conventional perspective cameras, in the case of cameras with distortion it is possible to recover calibration information from the projection of 3D lines in random position [22]. Since lines are features that often appear in natural images, with special relevance in the case of man-made environments, line-based calibration is an appealing proposition. The first contributions in camera calibration using the so-called "plumb-line" constraint go back to the 70's when Brown suggested to model the distortion by a polynomial and estimate its parameters by straightening up the lines in the image [8]. Latter on, Brauer-Burchardt et al. [7] and Fitzgibbon [11] simultaneously proposed to describe the image distortion by a 1-parameter rational function. The geometry of line projection considering the DM was investigated by Barreto in [3]. He concluded that, similarly to para-catadioptric cameras, the lines in 3D are projected into a family of conic curves that intersect in two points and satisfy an harmonic conjugate relation with two other points [4]. He also showed that the conic where a line is projected has only two independent degrees-of-freedom (DOF) and that, if the center is known, then it is possible to estimate and correct the image distortion using a single line. More recently Wang. et al. proposed an algorithm for computing both the distortion parameter and the principal point from an image of 3 lines [24] using an algebraic interpretation of the division model. In all these works the user is required to manually select the image contours corresponding to the projection of lines. Very recently Bukhari et al. [9] suggested an algorithm for automatically detecting lines and accomplishing the calibration following the methods of [24]. We advance the state-of-the-art in "plumb-line" calibration by showing that 3 lines enable to also recover the aspect ratio and skew and, more importantly, by providing an algorithm for unsupervised calibration that largely outperforms [9] in terms of accuracy, robustness, and computational efficiency.

The structure of the paper is as follows: section 2 introduces some background notions, section 3 addresses the problem of camera calibration from 3 lines and section 4 addresses the problem of line extraction from calibrated images. In section 5 we present the unsupervised calibration algorithm and section 6 shows the experimental results. Finally, conclusions are drawn in section 7.

1.2. Notation

Vectors and vector functions are represented by bold symbols, e.g $\mathbf{x}$, $\mathbf{F}(\mathbf{x})$, scalars and scalar functions are indicated by plain letters, e.g. r, $f(x)$, $g(r)$, matrices and image signals are respectively denoted by capital letters in sans serif and typewriter fonts, e.g. the matrix M and the image I. Points, lines and conics are represented in homogeneous coordinates. We do not distinguish between a projective transformation and the matrix representing it.

2. Background concepts

Throughout this article we will model the camera distortion using the so called division model [7,11], where ξ is the negative parameter that quantifies the amount of distortion. $\mathbf{h}()$ is the radial distortion function that maps undistorted points $\mathbf{u}$ in $\mathbb{P}^2$ into distorted points $\mathbf{d}$ in $\mathbb{P}^2$:

$$\mathbf{d} \sim \mathbf{h}(\mathbf{u}) \sim \begin{pmatrix} 2u_1 & 2u_2 & u_3 + \sqrt{u_3^2 - 4\xi(u_1^2 + u_2^2)} \end{pmatrix}^T \tag{1}$$

Let $\mathbf{n} \sim \begin{pmatrix} n_1 & n_2 & n_3 \end{pmatrix}$ be the projection of a 3D line according to the conventional pinhole model. The distortion function 1 transforms a line $\mathbf{n}$ into the conic Ω given by [3]:

$$\Omega = \begin{pmatrix} \xi n_3 & 0 & \frac{n_1}{2} \\ 0 & \xi n_3 & \frac{n_2}{2} \\ \frac{n_1}{2} & \frac{n_2}{2} & N_3 \end{pmatrix}. \tag{2}$$

It has been shown in [3] that Ω is the distorted image of a world line *iff* it passes through the circular points $\mathbf{I}$ and $\mathbf{J}$, and points $\mathbf{r}_+$ and $\mathbf{r}_-$ are harmonic conjugates [21] with respect to Ω:

$$\begin{cases} \mathbf{I}^T \Omega \mathbf{I} = 0 & \text{with} & \mathbf{I} = \begin{pmatrix} 1 & i & 0 \end{pmatrix}^T \\ \mathbf{J}^T \Omega \mathbf{J} = 0 & \text{with} & \mathbf{J} = \begin{pmatrix} 1 & -i & 0 \end{pmatrix}^T \\ \mathbf{r}_+^T \Omega \mathbf{r}_- = 0 & \text{with} & \mathbf{r}_\pm = \begin{pmatrix} 1 & 0 & \pm\sqrt{(\xi)} \end{pmatrix}^T \end{cases} \tag{3}$$

For convenience we will use, whenever needed, lifted representations of points and conics:

$$\mathbf{q}^t \underbrace{\begin{bmatrix} \omega_1 & \frac{\omega_2}{2} & \frac{\omega_4}{2} \\ \frac{\omega_2}{2} & \omega_3 & \frac{\omega_5}{2} \\ \frac{\omega_4}{2} & \frac{\omega_5}{2} & \omega_6 \end{bmatrix}}_{\Omega} \mathbf{q} = 0 \Leftrightarrow$$

$$\Leftrightarrow \underbrace{\begin{pmatrix} \omega_1 & \omega_2 & \omega_3 & \omega_4 & \omega_5 & \omega_6 \end{pmatrix}}_{\boldsymbol{\omega}^T} \hat{\mathbf{q}}^T = 0 \tag{4}$$

where $\hat{\mathbf{q}}$ denotes the lifted coordinates of $\mathbf{q}$ according to a second order Veronese map [3].

3. Calibration from 3 lines

The division model of equation 1 is usually studied as a bijective mapping in image coordinates. In this work we assume the mapping to be before the intrinsics. In this case, a point $\mathbf{Q}$ in 3D is projected onto a point $\mathbf{q}$ in the image by:

$$\mathbf{q} \sim \mathsf{K}\mathbf{h}(\mathsf{P}\mathbf{Q}),$$

with K being the matrix of intrinsic parameters of the camera:

$$\mathsf{K} = \begin{pmatrix} af & sf & c_x \\ 0 & a^{-1}f & c_y \\ 0 & 0 & 1 \end{pmatrix},$$

P being the standard 3×6 projection matrix [14], and $\mathbf{h}()$ being the radial distortion function with ξ being now expressed in millimetres instead of pixels.

3.1. Condition for a conic to be the image of a line

The conic Ω where a line is imaged is now given by transforming the result of equation 2 by the intrinsic parameters K, as shown in Fig. 2a. Since projective transformations preserve incidence and cross-ratio relations, the conic Ω must intersect the line at infinity in points $\mathbf{I}' \sim \mathsf{K}\mathbf{I}$ and $\mathbf{J}' \sim \mathsf{K}\mathbf{J}$, and must verify an harmonic relation with respect to points $\mathbf{r}'_\pm \sim \mathsf{K}\mathbf{r}_\pm$. Therefore, a conic is the image of a line *iff* it verifies $\Phi\boldsymbol{\omega} = 0$, with $\boldsymbol{\omega}^T$ being its representation in $\mathbb{P}^5$ and Φ being the 3×6 matrix:

$$\Phi = \underbrace{\begin{bmatrix} (as - ia^2)^2 & as - ia & 1 & 0 & 0 & 0 \\ (as + ia^2)^2 & as + ia & 1 & 0 & 0 & 0 \\ c_x^2 - \frac{a^2}{\eta} & c_x c_y & c_y^2 & c_x & c_y & 1 \end{bmatrix}}_{\Lambda} \tag{5}$$

$$\text{with} \qquad \eta = \frac{\xi}{f^2}.$$

The two first rows are the lifted representation of $\mathbf{I}'$ and $\mathbf{J}'$ respectively, and the last row is the joint lifted representation of $\mathbf{r}_+$ and $\mathbf{r}_-$ [2].

If the calibration parameters of the camera are known, then the conic Ω can be estimated from $N \geq 2$ image points using constrained least squares [13], with equation 5 giving the set of 3 linear constraints.

3.2. Minimal solution for the calibration

From equation 5, we observe that the images of lines lie in a 2D subspace $\mathbb{S}$ of $\mathbb{P}^5$ that encodes the calibration. We now show how to recover the calibration parameters from 3 line images $\boldsymbol{\omega}_1$, $\boldsymbol{\omega}_2$ and $\boldsymbol{\omega}_3$ (Fig. 2b). If the projection of a 3D line is correctly estimated in the image plane, intersecting it with the line at infinity defines points $\mathbf{I}'$ and $\mathbf{J}'$, whose coordinates encode a and s as shown in Fig.2a. Solving with respect to these parameters it can be shown that:

$$a = \sqrt{\frac{\omega_3}{\omega_1} - \frac{1}{4}\frac{\omega_2^2}{\omega_1^2}} \quad \text{and} \quad s = -\frac{\omega_2}{\omega_1}a^{-1}, \tag{6}$$

where ω_i are the entries of $\boldsymbol{\omega}$. We can therefore retrieve a and s from the image of a single line.

The principal point (c_x, c_y) and distortion parameter are encoded in the third orthogonal vector to the subspace of
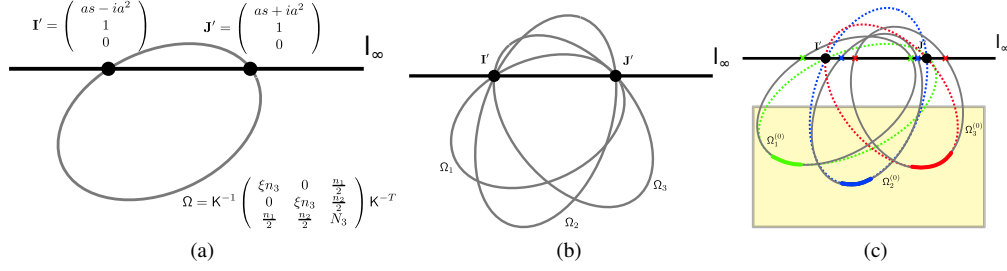
Figure 2: Intersecting projections of 3D lines with the line at infinity.

lines, $\mathbb{S}$. Given tree line images, we can determine this subspace and compute $\mathbf{\Lambda}$ by parametrizing the null space of the lines as follows:

$$\mathbb{N}(\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \boldsymbol{\omega}_3) = K_1\mathbf{V}_1 + K_2\mathbf{V}_2 + K_3\mathbf{V}_3 \qquad (7)$$

with $\Phi \in \mathbb{N}(\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \boldsymbol{\omega}_3)$. Considering the third row of Phi, $\mathbf{\Lambda}$, the following conditions must hold:

$$\begin{cases} \Lambda_6 = 1 \\ \Lambda_2 - \Lambda_5\Lambda_4 = 0 \\ \Lambda_5^2 - \Lambda_3 = 0 \end{cases}$$

After solving the first constraint in order to K_3 and replacing it in the other two, the parameters K_1 and K_2 can be computed by intersecting the two conic curves given by the 2nd and 3rd constraints (system of two second order homogeneous polynomials). From the 4 possible solutions we choose the one with physical meaning.

In summary, we have shown that tree world lines are sufficient to calibrate a camera. However, two important remarks are done:

1. We are only able to determine the ratio η between ξ and f^2, that can be understood as the distortion parameter expressed in pixels rather than in mm. Nevertheless this coupled parameter enables to rectify the image distortion (as shown in the experiments).

2. We can verify that, considering a fourth line projection $\boldsymbol{\omega}_4$, puts no further constraints to the calibration problem. The conic curve must satisfy two linear constraints since it must pass by points $\mathbf{I}'$ and $\mathbf{J}'$, and must belong to the subspace orthogonal to $\mathbf{\Lambda}$. This means that only 2 of the 5 DOF of the conic curve are really independent and they refer to the orientation of the plane containing the original line in 3D (see [3]). Thus, we conclude that line images $\boldsymbol{\omega}_i$, with $i > 3$, bring no additional information about the camera calibration and it is impossible to decouple the focal length f from the distortion parameter ξ using exclusively line features.

The calibration solution demonstrated above enables to determine the back-projection directions up to an angular multiplicative factor. Determining this multiplicative factor requires knowing, directly or indirectly, the absolute angle between two back-projection rays (e.g. know 'a priori' the field-of-view (FOV) of the camera or extract focal distance from EXIF tag).

3.3. Calibration algorithm

The previous section demonstrates how to calibrate a camera from a minimum of 3 imaged lines. In practice, for each conic, we are only able to extract points belonging to a small arc of the conic. The joint effect of noise and strong partial occlusion makes the estimation of the conic very uncertain [2]. In Fig.2c we can see that the arrangement of the initial conics $\Omega_i^{(0)}$ does not comply with the constrains derived in section 3.1. In this case the conics do not intersect the line at infinity in two unique points and the harmonic relations with respect to points $\mathbf{r}'$ and $\mathbf{t}'$ are not consistent.

Table 1 summarizes the unsupervised calibration algorithm. We start by estimating the likely location of the conics intersection with the line at infinity (steps 1 to 3) and then the conics a re-estimated from the corresponding image points enforcing the incidence with points $\mathbf{I}'$ and $\mathbf{J}'$ (step 4). The principal point (c_x, c_y) and η are estimated by computing the vector $\mathbf{\Lambda}$ lying in the null space $\mathbb{N}(\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \boldsymbol{\omega}_3)$ (steps 5 to 6).

The calibration estimation of steps 1 to 6 is sub-optimal and is used as initialization for a final iterative optimization step. From the analytical form of Ω (see Fig:2a), and after some algebraic manipulation, we derive the following equation:

$$
\underbrace{\begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ \omega_4 \\ \omega_5 \\ \omega_6 \end{pmatrix}}_{\boldsymbol{\omega}} \sim \underbrace{\begin{pmatrix} 0 & 0 & \eta a^{-2} \\ 0 & 0 & -2\eta a^{-1}s \\ 0 & 0 & \eta(a^2+s^2) \\ a^{-1} & 0 & -2\eta c_x a^{-1} \\ -s & a & 2\eta(c_x s + c_y a) \\ -c_x & -c_y & 1+\eta(c_x^2+c_y^2) \end{pmatrix}}_{\mathsf{G}} \cdot \underbrace{\begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}}_{\mathbf{m}}
$$
$$(8)$$

with G being a 6×3 matrix that encodes the calibration parameters and $\mathbf{m}$ being the 3×1 vector encoding the orientation of the plane that contains the line [3].

Given the conics $\boldsymbol{\omega}_i$ and the matrix G, computed with the calibration initialization, the corresponding vector $\mathbf{m}_i$ is determined linearly. Let $\mathbf{q}_j^{(i)}$ be contour point $j = 1 \cdots N_i$ belonging to the ith conic $\boldsymbol{w}_i$. The bundle adjustment of the calibration parameters is carried by minimizing the function of equation 9:

$$
f = \min_{a,s,c_x,c_y,\eta,\mathbf{m}_i} \sum_{i=1}^{3} \sum_{j=0}^{N_i} \left(\mathbf{m}_i^T \mathsf{G}^T \hat{\mathbf{q}}_j^{(i)} \right)^2, \quad (9)
$$

Note that, if there is no prior knowledge about a and s, it is possible to estimate $\mathbf{I}'$ and $\mathbf{J}'$ from steps 1 to 3 in Table 1. However, if there is some prior knowledge such that reasonable assumptions can be made about a and s (e.g. the camera is skewless and has square pixels), then the 3 first steps can be skipped and the calibration carried trough 4 to 7.

4. Line extraction from a calibrated image

Let us assume a calibrated image with distortion for which we want to detect projections of world straight lines. We start by applying a standard edge detector [10], followed by a connected components algorithm in order to obtain several contours e_i that are line projection candidates. We aim at identifying the contours e_i that support conics $\boldsymbol{\omega}_j$ lying on the 2D subspace $\mathbb{S} \in \mathbb{P}^5$ defined by the calibration parameters. This can be seen as a multi-model fitting problem where the models are the conics $\boldsymbol{\omega}_j$ consistent with the calibration and we want to assign to each contour e_i a model (or discard the contour in case it does not fit any model). Following this, we formulate the detection of lines as an optimal labelling problem that is cast as an UFL instance [17].

4.1. Uncapacited Facility Location (UFL) problem

Suppose that we need to open a set of facilities ω_j^0 to serve N customers $\mathsf{e}_i \in \mathcal{E}$ whose locations are known. Consider a set $\mathcal{V}_0$ comprising M possible facility locations, the cost c_{ij}^0 for assigning the facility ω_j^0 to the customer e_i and the cost v_j^0 for opening the particular facility ω_j^0. The

Table 1: Unsupervised calibration algorithm. Steps 1 to 3 apply when there is no prior knowledge of the camera aspect ratio and skew angle.

1. Given 3 contours $\mathsf{e}_{i=1\ldots3}$, estimate the conics $\Omega_{i=1\ldots3}^{(0)}$ using a standard conic fitting algorithm. For our experiments we use Taubin's method [23] due to its better performance with small arcs [2].

2. Intersect each conic $\Omega_i^{(0)}$ with the line at infinity and obtain $\mathbf{I}_i'$ and $\mathbf{J}'$ estimates [2].

3. Estimate the pair of complex conjugate points $\mathbf{I}'$, $\mathbf{J}'$ from the pairs $\mathbf{I}_i'$, $\mathbf{I}_i'$ by averaging the real and imaginary parts.

4. (Re)-estimate the conics using constrained least squares [13], forcing them to intersect $\mathbf{I}'$ and $\mathbf{J}'$.

5. Given the conics $\Omega_{i=1\ldots3}$ compute a basis for the null space $\mathbb{N}$ and determine the $\boldsymbol{\Lambda}$ vector by solving equation 7.

6. Extract the principal point (c_x, c_y) and η from vector $\boldsymbol{\Lambda}$.

7. Refine the calibration result by minimizing equation 8 using iterative optimization (equation 9).

goal of the UFL problem is to select a subset of $\mathcal{V}_0$ such that each customer is served by one facility and the sum of the customer-facility costs plus the sum of facility opening costs is minimized. This leads to an integer programming problem that is usually formulated using unary indicator variables y_j^0 and binary indicator variables x_{ij}^0, and whose objective is to find the vector $\mathbf{x}^0 = \{x_{11}^0 ... x_{ij}^0 ... x_{NM}^0\}$ such that :

$$
\min_{\mathbf{x}^0} \sum_{i=1}^{N} \sum_{j=1}^{M} c_{ij}^0 x_{ij}^0 + \sum_{j=1}^{M} v_j^0 y_j^0 \quad (10)
$$

$$
\text{subject to} \quad \begin{cases} x_{ij}^0, y_j^0 \in \{0,1\}, \forall_{i,j} \\ \sum_{j=1}^{M} x_{ij}^0 = 1, \forall_i \\ y_j^0 \geq x_{ij}^0, \forall_{i,j} \end{cases} \quad (11)
$$

The second constraint in equation 11 ensures that each customer is assigned to one facility, while the last constraint guarantees that each customer is only served by open facilities. For solving the UFL problem, we use the local message passing approach proposed by Lazic et al [17, 18].

4.2. Line detection as a UFL problem

Let $e_i \in \mathcal{E}$ with $i = 1...N$ be the ith connected component in the image. The objective is to assign to each segment e_i an image conic $\omega_j^0 \in \mathcal{V}_0$ using as few unique models as possible. Consider that the segments e_i are the customers and the putative conics ω_j^0 are the facilities. Let the cost c_{ij}^0 be the root mean square geometric distance between points of e_i and conic ω_j^0. Let v_j^0 be a constant cost for adding ω_j^0 in the final models assignment. The goal is to select a subset of conics in $\mathcal{V}_0$ such that sum of the consistency measures c_{ij}^0 and the costs $\mathbf{v}_j^0$ is minimized, which corresponds to the minimization of Eq.10. Fig. 3 shows the result of the UFL formulation applied to an image with considerable amount of clutter and where straight lines projections are often separated. It can be seen that the line extraction algorithm successfully identifies the correct lines and clusters disconnected segments in the same line.

5. Unsupervised Plumb-line calibration using RANSAC-UFL

In the previous section we presented an algorithm that, given the calibration, detects and estimates distorted world lines projections. This section considers the unsupervised calibration of the camera, which consists in simultaneously determining a suitable set of calibration parameters along with the corresponding world line projections in the image. This problem can be seen as a HFL instance.

5.1. Hierarchical Facility Location (HFL) problem

Consider that the facilities ω_j^0 described previously need to be stocked by additional storage facilities Γ_k^1. Consider a set of M facility locations $\mathcal{V}_0$ and L storage facilities $\mathcal{V}_1$. In addition to the costs v_j^0 and c_{ij}^0 described in the previous section, we now add the cost v_j^1 for opening the storage facility Γ_k^1, and the cost c_{jk}^1 associated with the facility Γ_k^1 supplying the facility ω_j^0. The goal of this two layer HFL problem is to find the vector $\mathbf{x} = \{\mathbf{x}^0, \mathbf{x}^1\}$ that minimizes the following function:

$$\min_{\mathbf{x}} \quad \sum_{i=1}^{N}\sum_{j=1}^{ML} c_{ij}^0 x_{ij}^0 + \sum_{j=1}^{ML}\sum_{k=1}^{L} c_{jk}^1 x_{jk}^1 + \sum_{j=1}^{ML} f_j^0 y_j^0 + \sum_{k=1}^{L} f_k^1 y_k^1$$

$$\text{s.t.:} \quad \begin{cases} x_{ij}^0, x_{jk}^1, y_j^0, y_k^1 & \in \quad \{0,1\} \\ \sum_{j=1}^{M} x_{ij}^0 = 1, \forall_i & \bigwedge \quad \sum_{k=1}^{L} x_{jk}^1 = y_j^0, \forall_j \\ y_j^0 \geq x_{ij}^0, \forall_{i,j} & \bigwedge \quad y_k^1 \geq x_{jk}^1, \forall_{j,k} \end{cases}$$

The additional restrictions compared to Eq. 11 are that if a facility ω_j^0 is closed in layer 0, then ω_j^0 will not need to be stocked by a storage facility Γ_j^1, whereas if a facility ω_j^0 is open, then it must be stocked by a facility in layer 1.
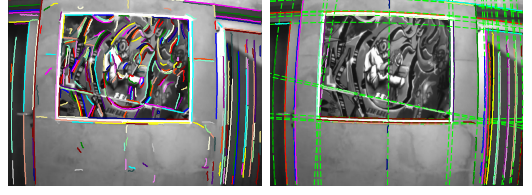

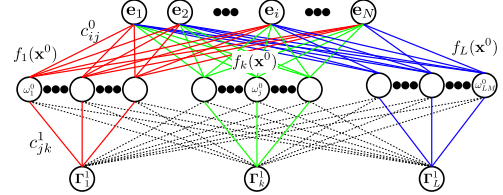Figure 3: UFL for finding lines according to calibration.

Figure 4: HFL formulation for the unsupervised calibration. Γ_k^1 are calibration hypothesis and ω_j^0 are conics estimated from segments e_i constrained by the associated calibration. The dotted lines indicates infinite costs between nodes.

5.2. Plumb-line calibration as a HFL problem

Given a set of connected components $e_i \in \mathcal{E}$, a set of image conics $\omega_j^0 \in \mathcal{V}^0$, and a set of calibration hypotheses $\Gamma_k^1 \in \mathcal{V}^1$, the objective is to assign an image conic to each segment e_i, minimizing the number of conics as well as the number of calibration hypotheses. This problem is cast as a HFL instance with two different layers (Fig.4). In addition to the costs c_{ij}^0 and v_j^0 described in Sec. 4, we add a new penalization v_k^1 for every Γ_k^1 contained in the solution. Since only one calibration Γ_k^1 is desirable, the penalization v_k^1 should be very high.

Calibration hypotheses Γ_k^1 are generated from segment e_i triplets using the method described in Table 1. For each generated Γ_k^1 we compute the M conics ω_j^0 that are consistent with the calibration Γ_k^1 and minimize the geometric distance of e_i to ω_j^0 (the method described in section 3.1), with:

$$j \in \{..., \underbrace{M(k-1)+1, ..., kM}_{j \in \mathbf{j}_k}, ...\}.$$

where $\mathbf{j}_k$ contains the indices of the conics that were generated from a particular calibration Γ_k^1. Our HFL formulation retrieves a single calibration by setting the connection costs c_{jk}^1 between ω_j^0 and Γ_k^1 as

$$c_{jk}^1 = \begin{cases} 0 & \text{if } j \in \mathbf{j}_k \\ \infty & \text{otherwise} \end{cases}$$

5.3. RANSAC-UFL

The high combinatorial nature of calibration models and line segments, along with the fact that we aim at jointly detecting+clustering the contours which are projection of

lines and finding the calibration parameters, motivate the use of a HFL approach over other multi-model fitting approaches [15]. Being formulated as a HFL problem, the unsupervised calibration algorithm can be computationally intensive if the number of segments e_i and/or the number of calibration hypothesis Γ_k is high. We show that our HFL problem can be efficiently solved as a minimization over a calibration dependent function $f_{\Gamma_k^1}(\mathbf{x}^0)$, which in turn is the result of solving the UFL problem (please refer to the supplementary material for details):

$$f_{\Gamma_k^1}(\mathbf{x}^0) = \min_{\mathbf{x}^0} \sum_{i=1}^{N} \sum_{j \in \mathbf{j}_k} c_{ij}^0 x_{ij}^0 + \sum_{j \in \mathbf{j}_k} f_j^0 y_j^0 + f_k^1 \quad (12)$$

subject to the constraints in Eq. 11. Following this, we propose a RANSAC-based approach for unsupervised calibration: the **RANSAC-UFL**. As the name suggests, this framework consists in subdividing the general HFL problem into various UFL problems, therefore greatly reducing the computational cost without changing the optimality bounds. The RANSAC-UFL randomly samples triplets of connected components, generating calibration hypothesis Γ_k^1. Then, each Γ_k^1 is evaluated separately using Eq. 12. The calibration with the lowest UFL energy $f_{\Gamma_k^1}(\mathbf{x}^0)$ is the output of the unsupervised calibration.

6. Experimental results

To evaluate our unsupervised calibration accuracy we compared the camera parameters estimation in 8 image of a cluttered environment against ground truth calibration obtained with [1]. In Fig. 5, we show results of: (i) using 3 manually selected lines per image and calibrating with the minimal solution of section 3.2; and (ii) using the completely unsupervised algorithm described in this paper. It can be seen that the estimation is accurate and robust even in a heavily cluttered environment.

In Fig. 7, we compare our approach with [9] by correcting the radial distortion in some images dominated by straight world lines, showing that out approach outperforms [9] for the estimation of the center and amount of distortion in both robustness and accuracy. In Fig. 6 we present results in challenging scenarios mined from the internet (where the method in [9] fails to obtain a valid calibration).

Note that, in the absence of large segments in the image, the estimation of a and s (Table 1, steps 1 to 3) becomes very unstable (see Fig. 2c). To overcome this we set a and s to reasonable values in images with small segments and proceed with the estimation from steps 4 to 7.

Further results are shown in the supplementary material, including the complete set of images used in Fig. 5. The code is implemented in MATLAB and is publicly available for download at the author's website.
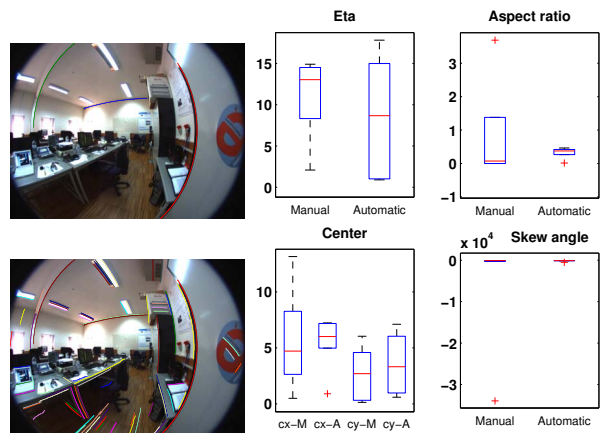

Figure 5: Calibration percentage deviation from ground truth. The top left image shows an exemplar of the manually selected lines. The bottom right image shows the automatically selected lines during unsupervised calibration.
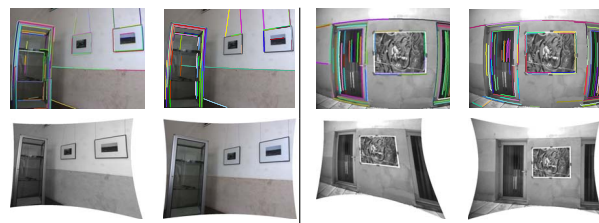

Figure 7: Comparison in distortion correction between Bukhari et al. [9] (leftmost image) and our method (rightmost image) in 2 scenes. For each scene we show the segment e_i on the top and the resulting distortion correction in the bottom.

7. Conclusion

In this work we have proposed a new method for the calibration of a camera using a minimum of 3 natural lines in a single image. Our work is based on a solid geometric interpretation of the line projection under the division model in perspective cameras and is able to estimate the principal point, aspect ratio, skew angle and a coupled parameter of the distortion and focal distance. For unsupervised camera calibration, we devised a framework for the joint line detection and calibration parameters estimation from a single image, that has been tested in challenging situations.
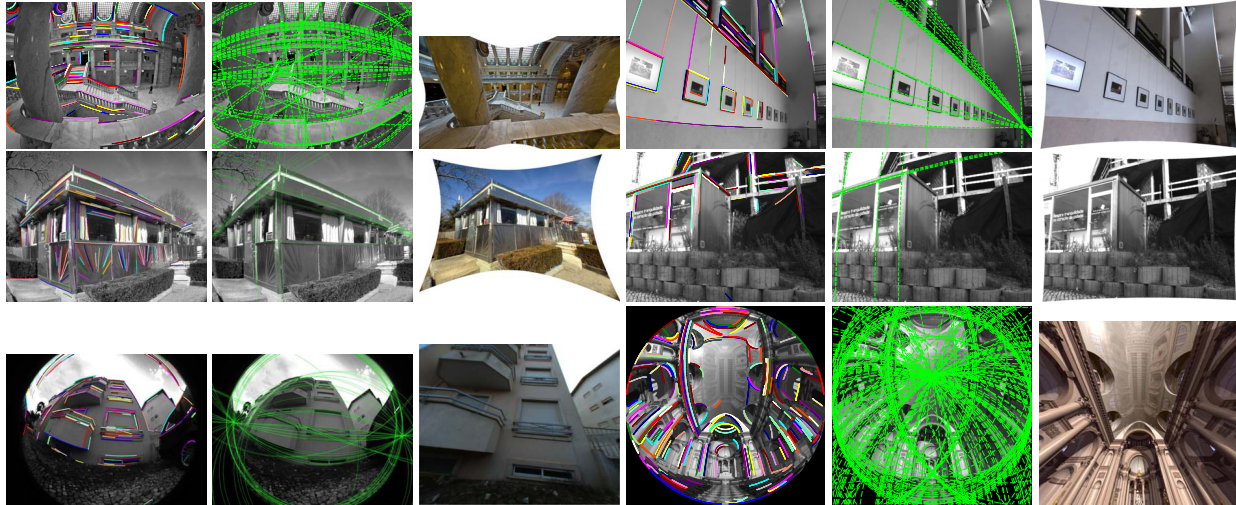
Acknowledgments

Figure 6: Results of the unsupervised calibration in images mined from the internet. The first column shows the segments e_i highlighted in different colors. The second column shows the detected lines consistent with the calibration. The third column shown the distortion correction. The next 3 columns have the same meaning.

References

[1] J. Barreto, J. Roquette, P. Sturm, and F. Fonseca. Automatic Camera Calibration Applied to Medical Endoscopy. In *BMVC*, 2009. 2, 7

[2] J. P. Barreto. General central projection systems, modeling, calibration and visual servoing. *PhD thesis*, 2003. 3, 4, 5

[3] J. P. Barreto. A unifying geometric representation for central projection systems. *CVIU*, 2006. 1, 2, 3, 4, 5

[4] J. P. Barreto and K. Daniilidis. Unifying image plane liftings for central catadioptric and dioptric cameras. *OMNIVIS*, 2004. 1, 2

[5] J. P. Barreto and K. Daniilidis. Fundamental matrix for cameras with radial distortion. In *ICCV*, 2005. 2

[6] J. Bazin, Y. Seo, C. Demonceaux, P. Vasseur, K. Ikeuchi, I. Kweon, and M. Pollefeys. Globally optimal line clustering and vanishing point estimation in manhattan world. In *CVPR*, 2012. 1

[7] C. Brauer-Burchardt and K. Voss. A new algorithm to correct fish-eye and strong wide-angle-lens-distortion from single images. In *ICIP*, 2001. 1, 2, 3

[8] D. C. Brown. Close-range camera calibration. *Photogrammetric Engineering*, 1971. 2

[9] F. Bukhari and M. Dailey. Automatic radial distortion estimation from a single image. *Journal of Mathematical Imaging and Vision*, 2012. 2, 7

[10] J. Canny. A computational approach to edge detection. *PAMI*, 1986. 5

[11] A. W. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *CVPR*, 2001. 1, 2, 3

[12] I. Givoni, C. Chung, and B. J. Frey. Hierarchical affinity propagation. *CoRR*, 2012. 2

[13] G. H. Golub and C. F. Van Loan. *Matrix computations*, volume 3. JHUP, 2012. 3, 5

[14] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. 2004. 3

[15] H. Isack and Y. Boykov. Energy-based geometric multi-model fitting. *IJCV*, 97(2):123–147, 2012. 7

[16] Z. Kukelova and T. Pajdla. A minimal solution to radial distortion autocalibration. *PAMI*, 2011. 2

[17] N. Lazic, B. J. Frey, and P. Aarabi. Solving the uncapacitated facility location problem using message passing algorithms. *Journal of Machine Learning Research*, 2010. 2, 5

[18] N. Lazic, I. E. Givoni, B. J. Frey, and P. Aarabi. Floss: Facility location for subspace segmentation. In *ICCV*, 2009. 5

[19] M. Lourenco, V. Pedro, and J. P. Barreto. Localization in large-scale indoor environments by querying omnidirectional visual maps using perspective images. *IEEE International Conference on Robotics and Automation*, 2012. 1

[20] R. Melo, J. Barreto, and G. Falcao. A new solution for camera calibration and real-time image distortion correction in medical endoscopy - initial technical evaluation. *Biomedical Engineering, IEEE Transactions on*, 2012. 1

[21] J. Semple and G. Kneebone. Algebraic Projective Geometry, 1952. 3

[22] P. Sturm, S. Ramalingam, J.-P. Tardif, S. Gasparini, and J. a. Barreto. Camera models and fundamental concepts used in geometric computer vision. *Found. Trends. Comput. Graph. Vis.*, 2011. 1, 2

[23] G. Taubin. Estimation of planar curves, surfaces, and non-planar space curves defined by implicit equations with applications to edge and range image segmentation. *IEEE TPAMI*, 13(11):1115–1138, 1991. 5

[24] A. Wang, T. Qiu, and L. Shao. A simple method of radial distortion correction with centre of distortion estimation. *J. of Math. Imag. and Vis.*, 2009. 2

[25] Z. Zhang, Y. Matsushita, and Y. Ma. Camera calibration with lens distortion from low-rank textures. In *CVPR*, 2011. 2