# Unsupervised Vanishing Point Detection and Camera Calibration from a Single Manhattan Image with Radial Distortion

Michel Antunes[1], João P. Barreto[2], Djamila Aouada[1], and Björn Ottersten[1]

[1]Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg,
{michel.antunes,djamila.aouada,bjorn.ottersten}@uni.lu
[2]Institute of Systems and Robotics (ISR), University of Coimbra,
jpbar@deec.uc.pt

## Abstract

*The article concerns the automatic calibration of a camera with radial distortion from a single image. It is known that, under the mild assumption of square pixels and zero skew, lines in the scene project into circles in the image, and three lines suffice to calibrate the camera up to an ambiguity between focal length and radial distortion. The calibration results highly depend on accurate circle estimation, which is hard to accomplish, because lines tend to project into short circular arcs. To overcome this problem, we show that, given a short circular arc edge, it is possible to robustly determine a line that goes through the center of the corresponding circle. These lines, henceforth called Lines of Circle Centres (LCCs), are used in a new method that detects sets of parallel lines and estimates the calibration parameters, including the center and amount of distortion, focal length, and camera orientation with respect to the Manhattan frame. Extensive experiments in both semi-synthetic and real images show that our algorithm outperforms state-of-the-art approaches in unsupervised calibration from a single image, while providing more information.*

## 1. Introduction

The 3D reconstruction of a scene from multiple images is a widely investigated topic in computer vision [1, 16]. The abundance of imagery in the internet has motivated recent efforts toward single view reconstruction. The works in single view reconstruction can be broadly divided into learning, e.g. [10], or geometrical based, e.g. [14]. The first category usually requires 2D or 3D data for the learning task, while the second uses assumptions about the layout of the scene for restricting the reconstruction process. Learning based approaches are nowadays showing very promising re-
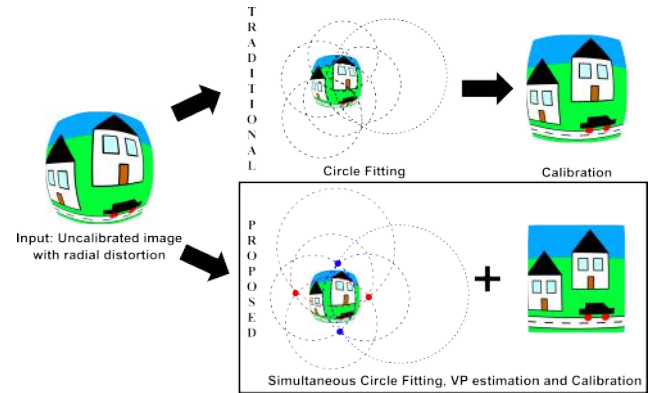


Figure 1. Unsupervised calibration from a single image with radial distortion. Traditional approaches (top) start by estimating circles from edges, and use these circles for performing the calibration. The quality of the estimated circles is usually low, and this has significant impact in the calibration accuracy. We propose a new framework (bottom) that simultaneously performs circle fitting and VP estimation for accurate calibration.

sults. However, they usually only provide an approximation of the scene being perceived [15, 10, 22]. The actual 3D scene can only be recovered using geometric approaches.

Many geometric based approaches use the assumption of orthogonality in man-made environments for computing a 3D model of the scene [6, 11, 14, 23]. They usually require that the intrinsic camera calibration information is available [14, 23], or the images do not exhibit significant radial distortion (RD) [11]. Nowadays, wide-angle-lenses incorporated into small cameras, e.g. GoPro, are ubiquitous, and single image reconstruction techniques handling also this type of imagery would be desirable. There are several methods and approaches for the calibration of cameras with RD [25, 18]. However, they require multiple images and/or the acquisition of a known object (e.g. checkerboard), which is,

in many cases, not possible.

Table 1. Computed calibration parameters.

| Solver | $\mathbf{c}_\eta$ | $\eta$ | $f$ | R | Assumptions |
|---|---|---|---|---|---|
| **R3l** [13, 24] | ✓ | ✓ | ✗ | ✗ | no |
| **R5l** [21] | ✗ | ✓ | ✓ | ✓ | Manhattan |
| **R7C** (ours) | ✓ | ✓ | ✓ | ✓ | Manhattan |

Promising works for unsupervised single image calibration with RD were recently introduced [4, 13, 21, 24]. As depicted in Fig. 1, all these works start by detecting circles[1] in the image that are likely to be projections of scene lines. They use minimal solvers in a consensus maximization [4, 21] or energy minimization [13, 24] frameworks for computing some of the calibration parameters (refer to Table 1). It is rather obvious that the quality of the circle estimation in images with RD has a direct impact in the accuracy of camera calibration. The major problem of scene lines in images with RD is that the perceived arc is small, and it is difficult to accurately estimate the circle parameters even for low noise levels [2]. The works [13, 24] try to overcome this issue by clustering circular arc edges that are spatially separated. Nevertheless, dependence on the quality of the estimated circles for the calibration still remains.

## 1.1. Contributions

This paper proposes a new framework for simultaneously performing the circle fitting and vanishing point (VP) estimation for accurate camera calibration using a single image with RD (refer to Fig. 1). The contributions can be summarized as follows:

1. *Line of circle centers (LCC)*: Given a circular arc edge, we show that the minor axis of the scattering ellipse, herein called the LCC, provides a robust 1D restriction for the location of the circle center. The LCC can be understood as a circle pencil, and can also be represented by a line in the circle space $\mathbb{P}^3$. We propose to use the LCCs for the calibration;

2. *VPs in images with RD*: We analyse the geometry of projection of parallel lines in images with RD, and show that it defines a line in the circle space $\mathbb{P}^3$. This line can be efficiently computed from 4 LCCs belonging to parallel lines in 3D;

3. *Camera calibration using VPs and LCCs*: We present a method based on a minimal solution that uses 7 LCCs for extracting the calibration parameters from the detection of two orthogonal VP models. In contrast to

---

[1]As mentioned in [13], lines in 3D are projected into conic curves. The weak assumption about the camera aspect ratio and skew improves the results, and the standard values (one and zero, respectively) are adequate for obtaining an initial estimation that can be refined using iterative schemes. In this case, the line projections can be represented by circles.

existing solutions (refer to Table 1), our method provides an estimate of the center $\mathbf{c}_\eta$ and amount $\eta$ of distortion, the focal length $\mathbf{f}$, and the camera rotation R with respect to the Manhattan frame.

## 2. Background

This section briefly reviews background concepts that are used throughout the article.

### 2.1. Circle fitting

We detect circular arcs using the approach proposed in [21], and denote a circular arc segment by $\mathbf{e}$. The canonical equation of a circle fitted to $\mathbf{e}$ is given by

$$(x - a)^2 + (y - b)^2 = R^2,$$

where $(x, y)$ is a generic data point in $\mathbf{e}$, $(a, b)$ is the center and $R$ is the radius of the circle. The standard circle parameters $\mathbf{a} = (a, b, R)$ have several drawbacks [5], and usually the algebraic parameters $\mathbf{A} = \begin{bmatrix} A & B & C & D \end{bmatrix}^\mathsf{T}$ are applied:

$$\mathbf{A}^\mathsf{T} \begin{bmatrix} z & x & y & 1 \end{bmatrix}^\mathsf{T} = 0,$$

where $z = x^2 + y^2$, and with the constraints of $A \neq 0$ and $B^2 + C^2 - 4AD > 0$. The conversion between algebraic and geometric parameters can be consulted in [5]. For a given circular arc, the standard parameters $\mathbf{a}$ will be used for reasoning in the image space, while the algebraic parameters $\mathbf{A}$ will be used for geometric analyses in the circle space $\mathbb{P}^3$. The distance for measuring how well $\mathbf{e}$ containing $n$ points agrees with $\mathbf{a}$ is given by:

$$d(\mathbf{e}, \mathbf{a}) = \left( \frac{\sum_{k=1}^{n} \left( \sqrt{(x_k - a)^2 + (y_k - b)^2} - R \right)^2}{n} \right)^{1/2}. \tag{1}$$

### 2.2. Plücker coordinates

The geometric analysis of the projection of lines in images with RD will be partially carried in the circle space $\mathbb{P}^3$. In this regard, we need to represent lines in 3D. As discussed in [17], a 3D line can be represented by a 6D vector of Plücker coordinates. Given two 3D points $\mathbf{P}$ and $\mathbf{Q}$, the plücker coordinates are given by:

$$\mathbf{\Pi} = \begin{bmatrix} \Pi_1 \\ \vdots \\ \Pi_6 \end{bmatrix} \sim \begin{bmatrix} \mathbf{D} \\ \mathbf{M} \end{bmatrix} \sim \begin{bmatrix} \mathbf{P} - \mathbf{Q} \\ \mathbf{P} \times \mathbf{Q} \end{bmatrix}.$$

Two different lines $\mathbf{\Pi}$ and $\mathbf{\Pi}'$ intersect if and only if:

$$\mathbf{\Pi} \odot \mathbf{\Pi}' = \mathbf{D}^\mathsf{T}\mathbf{M}' + \mathbf{D}'^\mathsf{T}\mathbf{M} = 0,$$

and the lines need to satisfy the Plücker constraint

$$\mathbf{\Pi} \odot \mathbf{\Pi} = 0. \qquad (2)$$

The dual of the Plücker line $\mathbf{\Pi}$ is given by

$$\mathbf{\Pi}^* = \begin{bmatrix} \mathbf{M} \\ \mathbf{D} \end{bmatrix}.$$

## 2.3. Lines through four lines

In [20], Teller and Hohmeyer describe how to compute the lines that intersect four lines in 3D. The authors represent the 3D lines using Plücker coordinates, and explain that each 3D line is mapped to a hyperplane in $\mathbb{P}^5$. Four such hyperplanes in general position intersect a line in $\mathbb{P}^5$, and intersecting this line with the Plücker quadric determines the incident lines. Since we use a different Plücker representation from the one used in [20], and also because we will need to modify the solution presented in [20] for adding a particular restriction, we briefly describe how the lines intersecting four lines $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{C}_3$ and $\mathbf{C}_4$ represented in Plücker coordinates can be computed.

First, the following matrix is constructed

$$\mathsf{M} = \begin{bmatrix} \mathbf{C}_1^* & \mathbf{C}_2^* & \mathbf{C}_3^* & \mathbf{C}_4^* \end{bmatrix}^\mathsf{T}.$$

Assuming that $\mathsf{M}$ is full rank, then the null-space of $\mathsf{M}$ has two dimensions and can be parametrized as

$$\mathbf{I} = t\mathbf{F} + \mathbf{G}, \qquad (3)$$

where $\mathbf{F}$ and $\mathbf{G}$ are unitary vectors, and $t$ is a scalar. The condition of Eq. 2 needs to be verified, meaning that $\mathbf{I} \odot \mathbf{I} = 0$. Using Eq. 3 comes that

$$(\mathbf{F} \odot \mathbf{F})\,t^2 + 2\,(\mathbf{F} \odot \mathbf{G})\,t + \mathbf{G} \odot \mathbf{G} = 0,$$

from which $t$ can be computed, and by substituting $t$ in Eq. 3, the incident lines $\mathbf{I}$ can be recovered.

## 3. Line of circle centers (LCC)

In this section, it is shown that the circle center of noise free arc edge points is located on the minor axis of the scattering ellipse, and that in case of noisy edge points, this line, called LCC in this article, provides an accurate and robust 1D restriction about the location of the circle center.

## 3.1. Chord method for circle fitting

As explained in [5], a simple circle fitting approach is the chord method. For any two points that lie on a circle, the perpendicular bisector of the line connecting these two points, called the chord, intersects the circle center. For estimating the circle center, the chord method minimizes the distance of the circle center to all perpendicular bisectors, usually using a certain weighting factor [5]. The chord
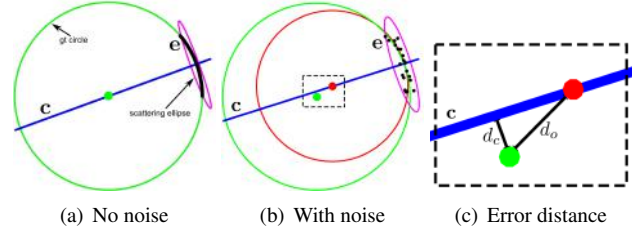


| (a) No noise | (b) With noise | (c) Error distance |

Figure 2. Circle fitting and LCC. (a) The center of the ground truth (gt) circle (green), related to the edge points in $\mathbf{e}$, is located on the minor axis (blue) of the scattering ellipse (magenta). We call $\mathbf{c}$ the LCC. (b) In case of noise, the circle obtained using the Taubin fit (red) is different from the gt, and $\mathbf{c}$ does not intersect the gt center. (c) In Fig. 3, we compare the distance $d_o$ between the gt and Taubin centers, and the distance $d_c$ from the gt center to $\mathbf{c}$.

method is not efficient and rarely used, but will give insights that are relevant for constraining the circle fitting procedure.

Given a noise free circular arc, there is a particular bisector whose sum of corresponding chord weights is the highest (refer to the supplementary material). This is the line that best divides the circular points in two parts. In case of noise, there is a unique perpendicular bisector for each point chord. Nevertheless, we show in the next section that it is possible to compute a line that best divides the circle points even under noise.

## 3.2. 1D restriction for circle fitting

As discussed previously, the problem of the projection of scene lines in images with RD is that the perceived arc is small, and circle fitting methods are inaccurate even for small noise magnitudes [2]. To avoid the estimation of all circle parameters for each arc edge a priori, we aim at constraining the circle search space for each edge, and use global information for the actual circle fitting.

Inspired by the chord method described in the previous section, our objective is to compute a line that constrains the location of the circle center. A cue is to use the line that best divides the edge points in two parts. In case of noise, this can be achieved using the minor axis of the scattering ellipse (refer to Fig. 2). The scattering ellipse is defined by the scatter matrix:

$$\mathsf{S} = \begin{bmatrix} s_{xx} & s_{xy} \\ s_{xy} & s_{yy} \end{bmatrix},$$

where $s_{xx} = \sum_{k=1}^{n}(x_k - \bar{x})^2$, $s_{yy} = \sum_{k=1}^{n}(y_k - \bar{y})^2$, $s_{xy} = \sum_{k=1}^{n}(x_k - \bar{x})(y_k - \bar{y})$, $\bar{x}$ and $\bar{y}$ are the coordinates of the centroid. The scatter matrix $\mathsf{S}$ defines the spread of the points around the centroid. The minor axis of the scattering ellipse intersects the centroid of the edge points $(\bar{x}, \bar{y})$ and the angle can be computed as follows:

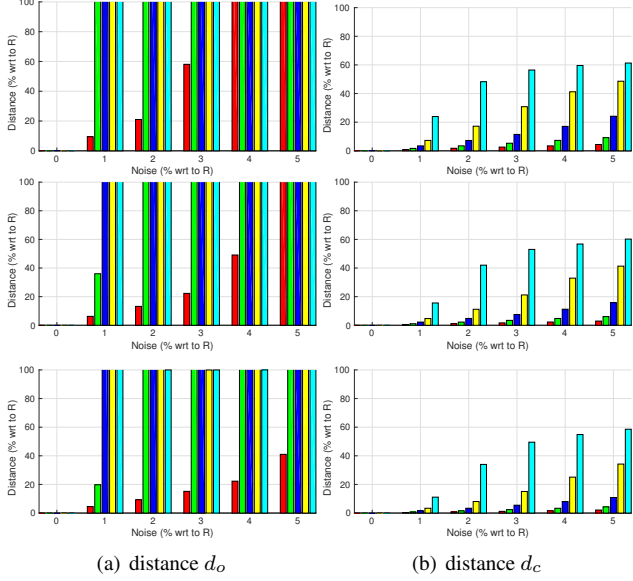$$\theta = \frac{1}{2} \operatorname{atan}\left( \frac{2s_{xy}}{s_{xx} - s_{yy}} \right).$$

| Arc $= 40°$ | Arc $= 20°$ | Arc $= 10°$ | Arc $= 5°$ | Arc $= 2°$ |

(a) distance $d_o$      (b) distance $d_c$

Figure 3. Taubin fit vs. LCC. Referring to Fig. 2, (a) distance of gt and estimated centers; (b) distance of gt centers and LCCs **c**. The analysis is carried for different noise levels (x-axis); number of arc points (rows), from top to bottom: 20, 50 and 100 points, respectively; and arc lengths (colors). For each combination of noise level, number of points and arc length, $50k$ runs were realized and the mean values are shown.

Refer to [5] where this equation is used for estimating a line fitting a set of points. We call the minor axis of the scattering ellipse the LCC, because we will use it as a 1D restriction for the location of the circle centers to be estimated.

### 3.3. Robustness of the LCC

This section analyses the suitability of using the LCC for constraining the location of the circle center. This is done by comparing the accuracy and robustness with the popular circle fitting method called the Taubin fit [19]. Refer to Fig. 3 for synthetic comparison experiments. A careful analysis of the graphics shows that the distance $d_c$ of the ground truth center to the LCCs is always smaller than the distance $d_o$ between the ground truth and estimated centers, and that the LCCs are much more robust to increasing noise magnitude and decreasing arc length. The next sections show how global information can be used for determining a circle on this pencil for fitting a circular arc edge.

## 4. VPs in images with radial distortion

This section concerns the geometric analysis of the projection of parallel scene lines in images with RD. We show that a VP defines a line in the circle space $\mathbb{P}^3$, and that it can be estimated using 4 LCCs.
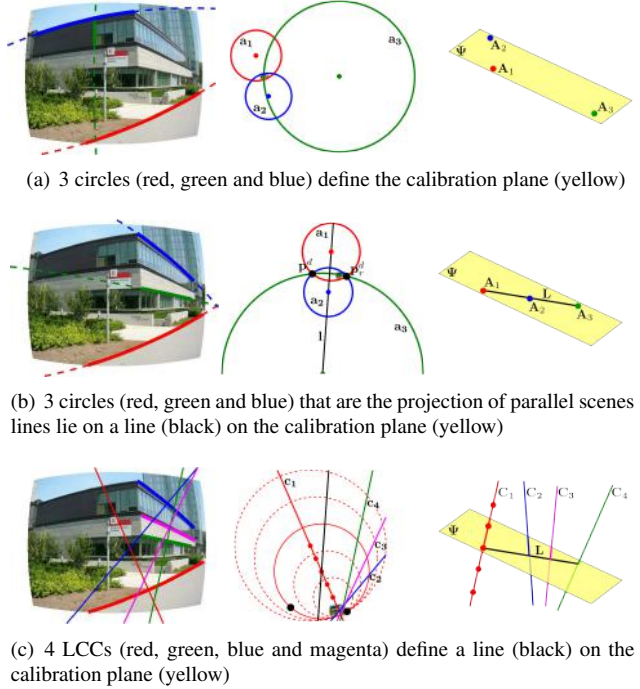


(a) 3 circles (red, green and blue) define the calibration plane (yellow)



(b) 3 circles (red, green and blue) that are the projection of parallel scenes lines lie on a line (black) on the calibration plane (yellow)



(c) 4 LCCs (red, green, blue and magenta) define a line (black) on the calibration plane (yellow)

Figure 4. Line projections in images with radial distortion. The calibration plane $\mathbf{\Psi}$ and VP model $\mathbf{L}$ in the circle space $\mathbb{P}^3$ can be computed using circles or LCCs.

### 4.1. Calibration plane

We consider the case of cameras with zero skew, unitary aspect ratio and RD that can be described by the 1-parameter division model [8]. In this case, the matrix of intrinsic parameters is:

$$\mathsf{K} = \begin{bmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix}, \tag{4}$$

where $f$ is the focal length, and $\mathbf{c}_\eta = \begin{bmatrix} c_x & c_y \end{bmatrix}^\mathsf{T}$ is the distortion center and principal point. An undistorted point $\mathbf{u}$ is distorted into point $\mathbf{d}$ as

$$\mathbf{d} = h(\mathbf{u}) = \mathbf{c}_\eta - \frac{2\mathbf{c}_\eta - 2\mathbf{u}}{1 + \sqrt{1 - 4\eta r^2}}, \tag{5}$$

where $r^2 = (u_x - c_x)^2 + (u_y - c_y)^2$, and $\eta$ quantifies the amount of distortion in pixels. As discussed in the seminal paper of Barreto and Daniilidis [3], a line $\mathbf{n} = \begin{bmatrix} n_x & n_z & n_y \end{bmatrix}^\mathsf{T}$ in the scene is projected into a circle that is given by the following equation

$$\Lambda \underbrace{\begin{bmatrix} \eta \\ n_x - 2c_x\eta \\ n_y - 2c_y\eta \\ n_z + (c_x^2 + c_y^2)\eta \end{bmatrix}}_{\mathbf{A}}^{\mathsf{T}} \begin{bmatrix} (x^2 + y^2) \\ x \\ y \\ 1 \end{bmatrix} = 0, \tag{6}$$

where $\Lambda = (n_z + c_x n_x + c_y n_y)$. Note that the circle $\mathbf{A}$ of Eq. 6 defines a point in the circle space $\mathbb{P}^3$. Given that there are three unknowns, from three circles in general position, the calibration parameters $c_x$, $c_y$ and $\eta$ can be recovered. Referring to Fig. 4(a), this can be geometrically interpreted as the three circles defining a plane, henceforth called the calibration plane, that is given by

$$\mathbf{\Psi} = \begin{bmatrix} c_x^2 + c_y^2 - 1/\eta & c_x & c_y & 1 \end{bmatrix}^\mathsf{T}. \qquad (7)$$

## 4.2. VPs in images with radial distortion

As mentioned previously, a set of parallel lines in the scene project into a pencil of lines intersecting a VP. In case of images with RD and with intrinsic calibration as in Eq. 4, the projection of 3D lines are given by circles. In this case, the geometry of VPs is slightly different.

Let us consider the projection of two parallel scene lines $\mathbf{n}_1$ and $\mathbf{n}_2$ that intersect the VP $\mathbf{p} = \begin{bmatrix} p_x & p_y & 1 \end{bmatrix}^\mathsf{T}$ in the undistorted image plane. Consider the projections in the distorted image plane as $\mathbf{a}_1$ and $\mathbf{a}_2$. The corresponding circle centers $(a_1, b_1)$ and $(a_2, b_2)$, respectively, define a line (refer to Fig. 4(b)):

$$\mathbf{l} = \begin{bmatrix} a_1 \\ b_1 \\ 1 \end{bmatrix} \times \begin{bmatrix} a_2 \\ b_2 \\ 1 \end{bmatrix} = \begin{bmatrix} -2\eta(c_x - p_x) \\ -2\eta(c_y - p_y) \\ 2\eta(-c_x^2 + p_x c_x - c_y^2 + p_y c_y) + 1 \end{bmatrix}.$$

Let us now consider a third projection of a scene line $\mathbf{n}_3$ that also intersects $\mathbf{p}$. It can be shown that $\mathbf{l}^\mathsf{T} \begin{bmatrix} a_3 & b_3 & 1 \end{bmatrix}^\mathsf{T} = 0$. This means that the projection of parallel lines in 3D will have the corresponding circle centers on a particular line $\mathbf{l}$, which depends on the center and the amount of distortion, and the direction of the lines in 3D.

Let us now analyse the intersection points of $\mathbf{a}_1$ and $\mathbf{a}_2$. There is the distorted VP given by

$$\mathbf{p}^d = h(\mathbf{p}),$$

where $h()$ is the RD function in Eq. 5. Computing the distance $d$ from $\mathbf{p}^d$ to $\mathbf{l}$ comes that

$$d = \mathbf{l}^\mathsf{T} \mathbf{p}^d = \frac{\sqrt{1 - 4\eta r^2}}{2\eta(-c_x^2 + p_x c_x - c_y^2 + p_y c_y) + 1}.$$

Note that there is only a particular combination of $r$ and positive $\eta$ (pincushion distortion) for which there is a single intersection point ($d = 0$), in all other cases there are two ($d > 0$). We are concerned with wide-angle lenses (negative $\eta$), and henceforth assume that there are always two intersection points. Given that the circle centers of $\mathbf{a}_1$ and $\mathbf{a}_2$ lie on the same line $\mathbf{l}$, then the second intersection point is the reflection of $\mathbf{p}^d$ about $\mathbf{l}$:

$$\mathbf{p}_r^d = 2\mathbf{c}_\eta - \mathbf{p}_d + \mathbf{D},$$

where

$$\mathbf{D} = -\frac{1}{\eta r^2} \begin{bmatrix} c_x - p_x \\ c_y - p_y \end{bmatrix}.$$

Remark that $\mathbf{p}_r^d$ can also be obtained using the backward projection function. As conclusion, the projection of parallel lines in images with RD defines a line $\mathbf{l}$ where the circle centers are located, and two intersection points $\mathbf{p}^d$ and $\mathbf{p}_r^d$ that are reflected with respect to $\mathbf{l}$.

The previous analysis also means that the images of a set of lines with the same direction in 3D is a circle pencil. Using Eq. 6, the family of line images with the same VP defines a line in the space of circles $\mathbb{P}^3$ that is computed using Plücker coordinates (refer to Fig. 4(b)):

$$\mathbf{L} = \begin{bmatrix} c_x p_y - c_y p_x \\ (c_x^2 - c_y^2)(c_y - p_y) - 2c_x^2 c_y + p_y \eta^{-1} + 2c_x c_y p_x \\ c_x^3 - p_x c_x^2 + c_x c_y^2 - 2p_y c_x c_y + p_x c_y^2 - p_x \eta^{-1} \\ 2c_x^2 - 2p_x c_x + 2c_y^2 - 2p_y c_y - \eta^{-1} \\ c_x - p_x \\ c_y - p_y \end{bmatrix}. \qquad (8)$$

It can be shown that the line $\mathbf{l}$, and the points $\mathbf{p}^d$ and $\mathbf{p}_r^d$ can be recovered from $\mathbf{L}$.

## 4.3. Four LCCs define a VP

We have discussed in the previous section that the projection of parallel lines define a line $\mathbf{L}$ in the calibration plane $\mathbf{\Psi}$. In this regard, the projection of two parallel scene lines (circles) is enough to compute $\mathbf{L}$ and the VP geometry (refer to Fig. 4(b)).

Let us now consider the LCCs. Referring to Fig. 4(c) and as discussed in Sec. 3, from each circular arc edge it is possible to compute a LCC, which serves as a 1D restriction for the location of the circle center. This line $\mathbf{c}$ can also be understood as a circle pencil, and defines a 3D line $\mathbf{C}$ in the circle space $\mathbb{P}^3$. As shown in Fig. 4(c), every point on $\mathbf{C}$ defines a different circle in the image space. From the paper of Teller and Hohmeyer [20] briefly discussed in Sec. 2.3, we know that four generic 3D lines are required for determining incident lines. Following this, using four LCCs it is possible to compute $\mathbf{L}$ and the complete VP geometry. Note that the method of Teller and Hohmeyer [20] usually provides two solutions. The one that minimizes the consistency function described in Sec. 5.3 is selected. Finally, the intersection of $\mathbf{L}$ with the LCCs defines the circle parameters fitting the corresponding edge points.

## 5. Camera Calibration using LCCs

This section presents a method for detecting VPs and calibrating a camera from a single image using LCCs.

## 5.1. Calibration plane from two VPs

We have seen in Sec. 4.3 that from four LCCs belonging to parallel scene lines, the VP geometry represented by the line $\mathbf{L}$ can be estimated (refer to Eq. 8). In order to estimate the calibration plane, we need two such lines. This section shows how the calibration plane $\mathbf{\Psi}$ can be computed from two sets of arc edges that belong to parallel lines in 3D.

Let us suppose we are given four LCCs $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{C}_3$ and $\mathbf{C}_4$ that belong to parallel scene lines, from which $\mathbf{L}_1$ is computed using the method described in Sec. 4.3. For obtaining a second VP model $\mathbf{L}_2$, we could use again four LCCs belonging to a different VP. Note however that this would not be a minimal solution for computing $\mathbf{\Psi}$, because we would have 4 LCCs plus $\mathbf{L}_1$, which makes in total five lines, and only four are required (refer to Sec. 4.3). Following this, we propose to compute $\mathbf{L}_2$ using three LCCs $\mathbf{C}_5$, $\mathbf{C}_6$ and $\mathbf{C}_7$ and the (co-planarity) restriction that it should also intersect $\mathbf{R} = \mathbf{L}_1$. This is done by modifying the solution of Teller and Hohmeyer in Sec. 2.3 as follows:

1. Construct the matrix $\mathsf{M} = \begin{bmatrix} \mathbf{C}_5^* & \mathbf{C}_6^* & \mathbf{C}_7^* \end{bmatrix}^\mathsf{T}$;

2. Define $\mathbf{I} = u\mathbf{E} + v\mathbf{F} + \mathbf{G}$, since the null-space of $\mathsf{M}$ has now three dimensions;

3. By defining $\mathbf{I} \odot \mathbf{R} = 0$, it can be shown that

$$u = X + Yv, \tag{9}$$

    where

$$X = -\frac{(G_3R_1 + G_1R_6 - G_2R_5 + G_4R_4 + G_5R_3 - G_6R_2)}{(E_3R_1 + E_1R_6 - E_2R_5 + E_4R_4 + E_5R_3 - R_6R_2)}$$
$$Y = -\frac{(F_3R_1 + F_1R_6 - F_2R_5 + F_4R_4 + F_5R_3 - F_6R_2)}{(E_3R_1 + E_1R_6 - E_2R_5 + E_4R_4 + E_5R_3 - R_6R_2)}$$

4. By substituting Eq. 9 in 2., and considering $\mathbf{I} \odot \mathbf{I} = 0$, the lines piercing $\mathbf{C}_5$, $\mathbf{C}_6$, $\mathbf{C}_7$ and $\mathbf{R}$ can be computed.

Using $\mathbf{L}_1$ and $\mathbf{L}_2$, the plane $\mathbf{\Psi}$ can be estimated, and the parameters $c_x$, $c_y$ and $\eta$ can be extracted using Eq. 7.

## 5.2. Focal length from the Manhattan assumption

Given $\mathbf{L}_1$ and $\mathbf{L}_2$, it is possible to calculate the two VPs $\mathbf{p}^1$ and $\mathbf{p}^2$ in the undistorted image plane using Eq. 8. Assuming that $\mathbf{p}^1$ and $\mathbf{p}^2$ are orthogonal, then the focal length $f$ can be determined [9]:

$$f = \sqrt{\frac{-p_x^1 p_x^2 - p_y^1 p_y^2}{p_z^1 p_z^2}}.$$

Finally, it can be shown that the third orthogonal VP belonging to the Manhattan frame is given by [9]:

$$\mathbf{p}^3 = \begin{bmatrix} f^2 & f^2 & 1 \end{bmatrix} \left( \mathbf{p}^1 \times \mathbf{p}^2 \right).$$
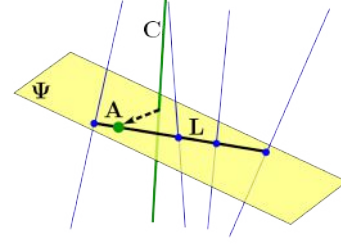


Figure 5. The function $f_\mathbf{C}(\mathbf{L})$ of Eq. 10 finds the 3D point $\mathbf{A}$ on $\mathbf{L}$ that minimizes the orthogonal distance from the line $\mathbf{C}$.

## 5.3. Consistency function

In order to measure how well a VP model $\mathbf{L}$ fits the points of an arc edge $\mathbf{e}$ with LCC $\mathbf{C}$, we need to compute a circle from $\mathbf{C}$ that agrees with $\mathbf{L}$ (refer to Fig. 5). As shown in [12], the point on $\mathbf{L}$ that minimizes the orthogonal distance from $\mathbf{C}$ can be computed using the function

$$f_\mathbf{C}(\mathbf{L}) = \frac{\mathbf{M}^L \times (\mathbf{D}^C \times \mathbf{D}_\times) + (-(\mathbf{M}^C)^\mathsf{T}\mathbf{D}_\times)\mathbf{D}^L}{||\mathbf{D}_\times||^2}, \tag{10}$$

where $\mathbf{D}_\times = \mathbf{D}^L \times \mathbf{D}^C$. The consistency function between $\mathbf{e}$ and $\mathbf{L}$ is then given by

$$D(\mathbf{e}, \mathbf{L}) = d(\mathbf{e}, f_\mathbf{C}(\mathbf{L})), \tag{11}$$

where $d()$ is defined in Eq. 1.

## 5.4. Method for camera calibration

We propose a RANSAC based framework for camera calibration by extracting dominant VPs using LCCs. We start by detecting circular arc edges, and for each arc edge compute the corresponding LCC, as discussed in Sec. 3. Then, a minimal sample set of 7 LCCs is drawn, of which 4 are taken to compute the VP model $\mathbf{L}_1$ and the remaining 3 are used for computing $\mathbf{L}_2$. From $\mathbf{L}_1$ and $\mathbf{L}_2$, we can compute the calibration parameters and Manhattan frame, as described in Sec. 5.1 and Sec. 5.2. From each RANSAC iteration, a hypothesis $\mathcal{H} = \{c_x, c_y, \eta, f, \mathsf{R}\}$ is obtained, where $\mathsf{R}$ represents the camera rotation. The hypothesis that maximizes the inlier consensus is selected. An edge $\mathbf{e}$ is considered as inlier of $\mathcal{H}$ if the consistency error of Eq. 11 is below a pre-defined threshold for $\mathbf{L}_1$, $\mathbf{L}_2$ or $\mathbf{L}_3$. After the RANSAC selection, we perform a non-linear refinement of the calibration parameters. The idea is to simultaneously optimize the calibration plane of Eq. 7, the focal length $f$, and the rotation $\mathsf{R}$ parametrized by the exponential map.

## 6. Experimental comparisons

This section concerns the experimental comparison of our method, described in the previous section, with state-of-the-art approaches. We denote our RANSAC based
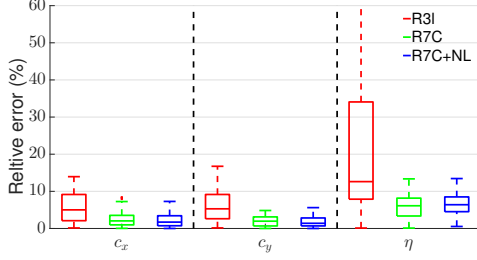
Figure 6. Comparison of **R7C** and **R7C+NL** (ours) with **R3l**.
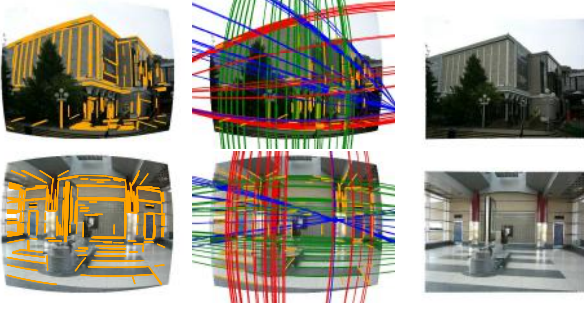


Figure 7. Calibration of images from YUD using **R7C+NL**. 2 examples are shown (rows); (left) edges; (middle) VP labelling, different colors identify different VPs; (right) undistorted image.
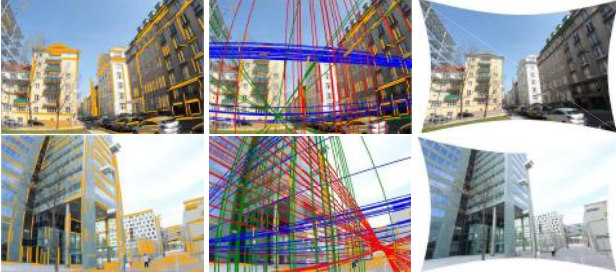


Figure 8. Calibration of images from [21] using **R7C+NL**. 2 examples are shown (rows); (left) edges; (middle) VP labelling; (right) undistorted image.

calibration method using a minimal sample of 7 LCCs by **R7C** and the variant using non-linear optimization as **R7C+NL**. Our method provides an estimate of $c_x$, $c_y$, $\eta$, $f$ and R. We use an inlier threshold of 2 pixels and run RANSAC until 4000 valid solutions are found. A valid solution verifies the following: $\eta < 0, f > 0$, and the distortion center is located at most 200 pixels from the image center.

## 6.1. Comparison with R3l

In this section, we compare our method against the solver described in [13] that uses 3 line projections for estimating the calibration parameters $c_x$, $c_y$ and $\eta$. Note that $f$ and R are not computed using this approach. We fit a circle using the Taubin fit [19] to each arc edge, and run RANSAC until 4000 valid solutions are found. The solution that maximizes the number of inliers is selected. Remark that the

consistency function in Eq. 11 is modified, and instead of using Eq. 10 for finding a circle that agrees with a VP model **L**, we project the estimated circle onto the calibration plane hypothesis **Ψ** and use this projection as the new circle. We refer to this approach as **R3l**. The methods are tested using the York Urban Database (YUD) [7] that consists of 102 images of man-made environments, to which we artificially added RD. The experimental comparison is presented in Fig. 6. By avoiding hard decision on the circle parameters using LCCs, and by estimating the calibration using global VP information, we improve the quality of the estimation of the calibration parameters. Fig. 7 shows two cases of VP detection and image undistortion.

## 6.2. Comparison with R5l [21]

In this section, we compare our method against the solver described in [21] that uses 5 line projections for computing $\eta$, $f$ and R. Remark that this approach assumes that the distortion center is the center of the image. We refer to this method as **R5l** and to the variant using a maximum likelihood estimator **R5l+MLE**. The methods are compared using the dataset presented in [21] that consists of 102 images with RD. The images were pre-calibrated by placing the distortion center and principal point fixed at the center of the image, obtaining a focal length of 446 pixels and distortion of $-1.5909 \times 10^{-6}$ pixels. Fig. 8 shows two cases of VP detection and image undistortion, while the quantitative experimental comparison is presented in Fig. 9. As can be seen, our method provides better estimations of the amount of RD, and slightly better estimation of the focal length. In contrast to **R5l** and **R5l+MLE**, we achieve this without any constraint about the location of the distortion center. Finally, we also show in Fig. 9 the distribution of our estimation of the distortion center. It can be seen that $c_y$ has an offset with respect to the image center location. .

The current drawback of our approach with respect to [21] is the computational complexity. Their minimal solution is of dimension 5, and the authors suggest to run RANSAC until 4000 valid solutions are found, which on average involves the evaluation of $120k$ hypotheses. In our case, the minimal sample dimension is 7. We also run RANSAC until 4000 valid solutions are found, which involves on average $650k$ iterations.

## 7. Experiments with images from the Internet

Fig. 10 shows experiments on images of man-made environments mined from the internet. These promising results show that it is possible to accurately detect VPs and calibrate a camera from a single image with RD.
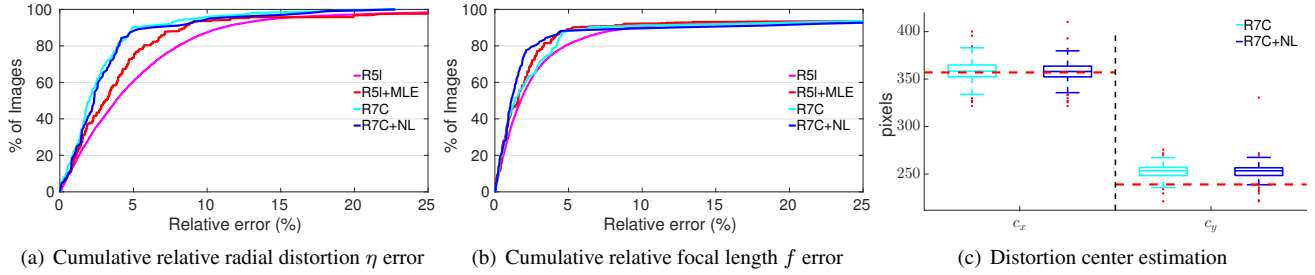
(a) Cumulative relative radial distortion $\eta$ error     (b) Cumulative relative focal length $f$ error     (c) Distortion center estimation

Figure 9. Comparison of **R5l** and **R5l**+**MLE** with our methods **R7C** and **R7C**+**NL**. (a) and (b) Cumulative histogram of the relative errors; (c) our distortion center estimation compared to the center of the image assumption used in **R5l** and **R5l**+**MLE** [21] (red lines).
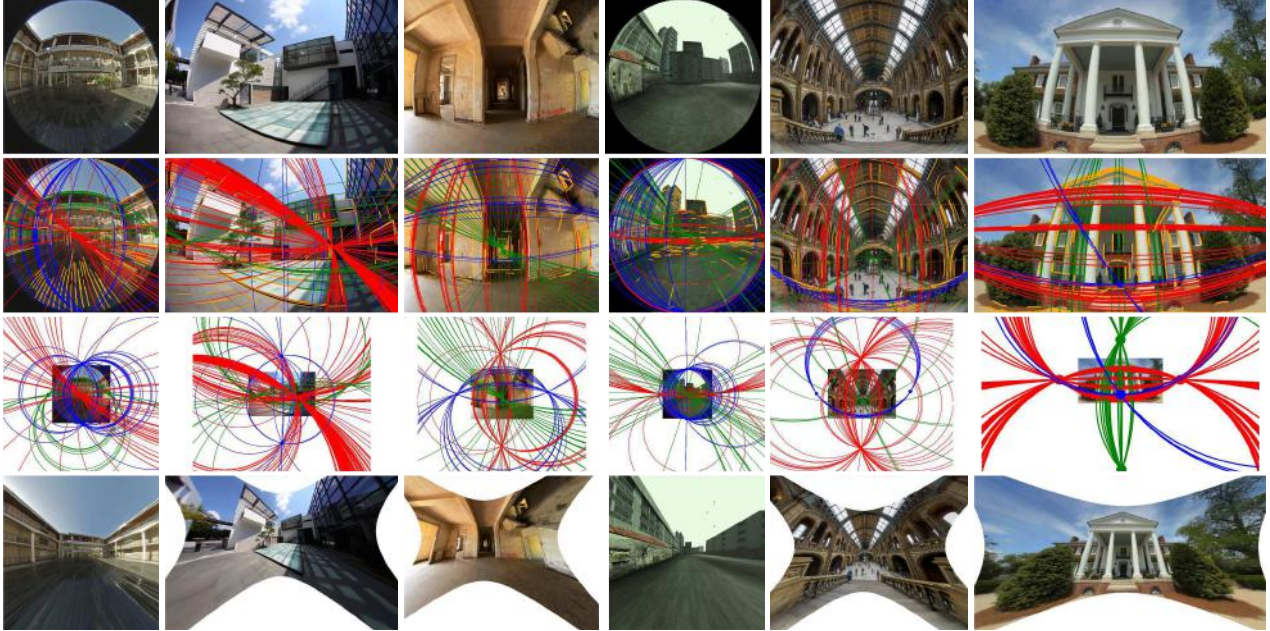


Figure 10. Detection of VPs and calibration of images mined from the internet. 6 examples are shown (columns); (top) input, (second row) VP labeling; (third row) zoom out for better visualizing the line projections and VP locations; and (last row) undistorted images.

## 8. Conclusions

We presented an automatic approach for the detection of VPs and camera calibration from a single image with RD. The core of the framework is (1) a theory that shows how to compute LCCs for circular arc edges, which are robust to noise and small arcs; (2) a theoretical analysis of the geometry of projection of parallel lines in images with RD; and (3) the use of LCCs for estimating camera parameters using the Manhattan prior. The effectiveness of the approach is proved using images downloaded from the Internet.

## Acknowledgements

## References

[1] S. Agarwal, N. Snavely, I. Simon, S. M. Seitz, and R. Szeliski. Building rome in a day. In *ICCV*, 2009.

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

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

[4] F. Bukhari and M. Dailey. Automatic radial distortion estimation from a single image. *Journal of Mathematical Imaging and Vision*, 2013.

[5] N. Chernov. *Circular and Linear Regression: Fitting Circles and Lines by Least Squares*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, 2010.

[6] E. Delage, H. Lee, and A. Ng. Automatic single-image 3d reconstructions of indoor manhattan world scenes. In *Robotics Research*, Springer Tracts in Advanced Robotics, 2007.

[7] P. Denis, J. H. Elder, and F. J. Estrada. Efficient edge-based methods for estimating manhattan frames in urban imagery. In *ECCV*, 2008.

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

[9] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.

[10] A. Kar, S. Tulsiani, J. Carreira, and J. Malik. Category-specific object reconstruction from a single image. In *CVPR*. 2015.

[11] D. C. Lee, M. Hebert, and T. Kanade. Geometric reasoning for single image structure recovery. In *CVPR*, 2009.

[12] M. T. Mason. *Mechanics of Robotic Manipulation*. MIT Press, Cambridge, MA, USA, 2001.

[13] R. Melo, M. Antunes, J. P. Barreto, G. Falcao, and N. Goncalves. Unsupervised intrinsic calibration from a single frame using a "plumb-line" approach". 2013.

[14] S. Ramalingam and M. Brand. Lifting 3d manhattan lines from a single image. In *ICCV*, 2013.

[15] A. Saxena, S. H. Chung, and A. Y. Ng. 3d depth reconstruction from a single still image. *IJCV*, 2008.

[16] N. Snavely, S. M. Seitz, and R. Szeliski. Modeling the world from internet photo collections. *IJCV*, 2008.

[17] 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.

[18] P. Sturm, S. Ramalingam, J. P. Tardif, S. Gasparini, and J. Barreto. Camera Models and Fundamental Concepts Used in Geometric Computer Vision. *Foundations and Trends in Computer Graphics and Vision*, 2010.

[19] G. Taubin. Estimation of planar curves, surfaces, and non-planar space curves defined by implicit equations with applications to edge and range image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 1991.

[20] S. Teller and M. Hohmeyer. Determining the lines through four lines. *J. Graph. Tools*, 1999.

[21] H. Wildenauer and B. Micusik. Closed form solution for radial distortion estimation from a single vanishing point. In *Proceedings of the British Machine Vision Conference*, 2013.

[22] J. Wu, T. Xue, J. J. Lim, Y. Tian, J. B. Tenenbaum, A. Torralba, and W. T. Freeman. Single image 3d interpreter network. In *ECCV*, 2016.

[23] H. Yang and H. Zhang. Efficient 3d room shape recovery from a single panorama. In *CVPR*, 2016.

[24] M. Zhang, J. Yao, M. Xia, K. Li, Y. Zhang, and Y. Liu. Line-based multi-label energy optimization for fisheye image rectification and calibration. In *CVPR*, 2015.

[25] Z. Zhang. A flexible new technique for camera calibration. *PAMI*, 2000.