Pose Estimation with Radial Distortion and Unknown Focal Length

Klas Josephson and Martin Byröd
Centre for Mathematical Sciences, Lund University, Lund, Sweden

LUND UNIVERSITY

Overview

This work is about finding the location of a camera with unknown focal length and radial distortion from a minimal number of point correspondences.

Problem: A camera can be assumed to be semi-calibrated (known skew, aspect ratio and principal point) but the focal length can not be easily determined from the image. Images from standard cameras also have a radial distortion with impact on the result of the localization.

Solution: Incorporate the focal length and the radial distortion in the core of a RANSAC engine used for localization.

Contribution:

- Derivation of a minimal case that uses four points to find the location of a camera when the focal length and radial distortion are unknown.
- A solution to the problem and code that can be downloaded.

Motivation

- Localization is a common problem in computer vision.
- Radial distortion is not used in todays localization algorithms.
- If an unknown focal length is modeled **uncalibrated** cameras can be used.

The Camera Model

The standard pinhole camera is used,

$$\lambda \mathbf{x} = P\mathbf{X}.$$

with,

$$P = K[R \mid \mathbf{t}].$$

R is a rotation matrix and t a translation. K is the calibration matrix that is represented by,

$$K = \begin{bmatrix} f & s & p_x \\ 0 & \gamma f & p_y \\ 0 & 0 & 1 \end{bmatrix}.$$

In this work s is assumed zero and γ one. This is more or less true for all digital cameras. Further on p_x and p_y, the principal point, is assumed to be in the middle of the image. These values can also be precalibrated.

Radial distortion is modeled by the division model [Fit01],

$$\mathbf{p}_u = \mathbf{p}_d/(1 + \mu r_d^2),$$

where μ is the distortion parameter and r_d the distance from the distortion center, assumed to be at the principal point, and $\mathbf{p}_u$, $\mathbf{p}_d$ are the undistorted and distorted image coordinates, respectively.

This leads to the final projection model,

$$\lambda \begin{bmatrix} x_1 \\ x_2 \\ 1 + \mu(x_1^2 + x_2^2) \end{bmatrix} = P\mathbf{X},$$

where x_1 and x_2 are the measured image coordinates, with the origin in the distortion center.

Parametrization

Initial coordinate transformations are performed both in the image plane and in the world coordinate frame. These transformations lead to the following result,

$$\mathbf{X}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{X}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{x}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

The rotation is parameterized by quaternions which gives the following rotation matrix,

$$R = \begin{bmatrix} a^2+b^2-c^2-d^2 & 2bc-2ad & 2ac+2bd \\ 2ad+2bc & a^2-b^2+c^2-d^2 & 2cd-2ab \\ 2bd-2ac & 2ab+2cd & a^2-b^2-c^2+d^2 \end{bmatrix}.$$

To remove the scale ambiguity of the camera a is fixed to 1. This paremetrization results in five unknown variables, (f, μ, b, c, d) and with four points this gives five linearly independent equations. The resulting system of polynomial equations is solved with Gröbner basis techniques [BJÅ09]. The system has 32 solutions but the equations are all quadratic in the focal length so at most 16 solutions can be in front of the camera.

Code available: `http://www.maths.lth.se/vision/downloads/`

Execution time for one problem instance is 67 ms in the available Matlab implementation.

Synthetical Experiments

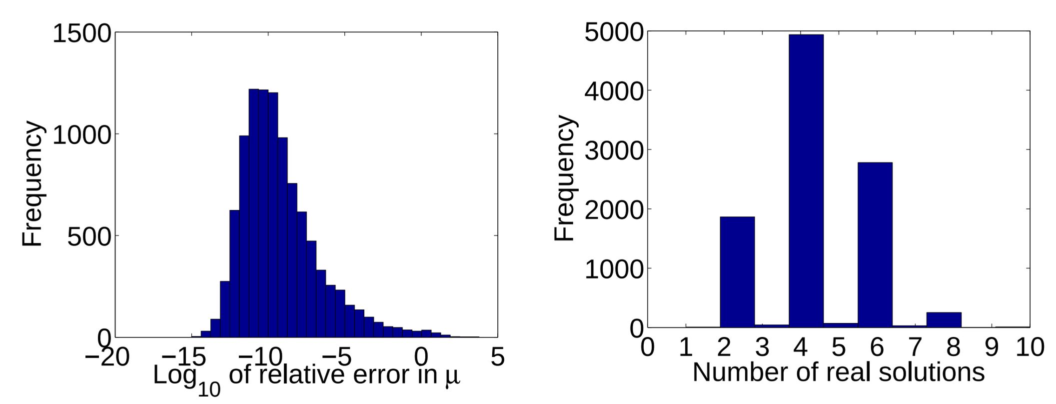

FIGURE 1: Left: Histogram of errors over 10000 runs on noise free data. Right: Histogram of the number of solutions with real positive focal length found on the same data.

Noise	Median	75th percentile
0.0	$1.5 \cdot 10^{-11}$	$5.1 \cdot 10^{-10}$
0.5	$1.4 \cdot 10^{-2}$	$4.1 \cdot 10^{-2}$
1.0	$2.3 \cdot 10^{-2}$	$6.8 \cdot 10^{-2}$
2.0	$5.2 \cdot 10^{-2}$	$1.5 \cdot 10^{-1}$
3.0	$6.7 \cdot 10^{-2}$	$1.5 \cdot 10^{-1}$

TABLE 1: The relative error of the focal length for different levels of noise. The noise is given in pixels.

Real Experiments

The real experiments were performed on a data set with approximately hundred images of a shopping street. These were used to build a 3D-model using the Photo Tourism code. Then localization was performed in a leave one out manner. The localization was performed with two different lenses on the camera, one regular lens and one fisheye lens. Comparisons were made with [BKP08] that solves the same problem but without modeling the radial distortion.

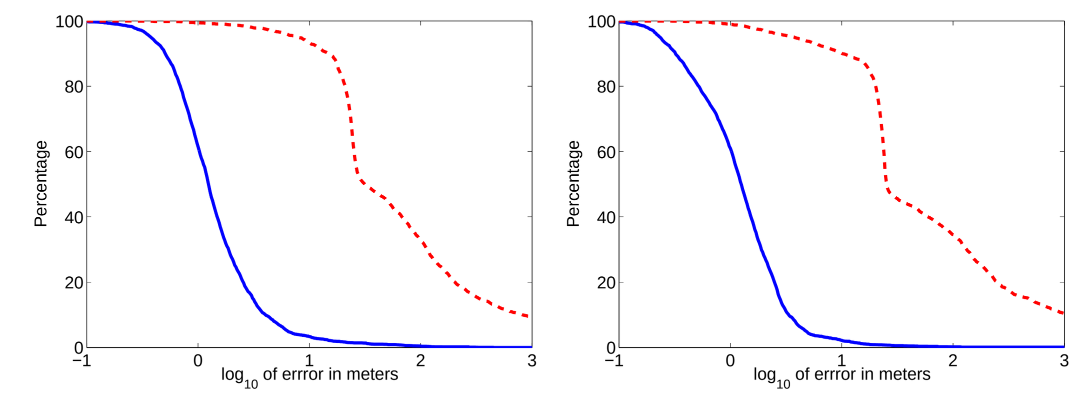

FIGURE 2: The percentage of images with an estimated position further away then a given distance to the position given by Photo tourism. The error is roughly given in meters. Notice the logarithmic scale. The blue solid line is for the proposed method and the red dashed represents the method without distortion. The left plot is for the fisheye lens and the right is for a regular lens.

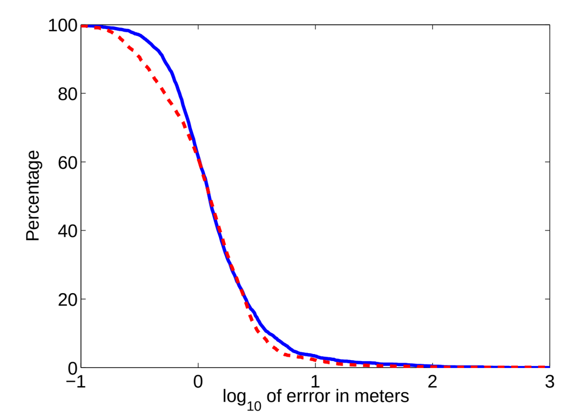

FIGURE 3: Error in meters for distorted image and non-distorted image using the proposed algorithm on a logarithmic scale. The blue line represents the distorted images and the red show the result for images taken with a regular lens.

Kernel Voting

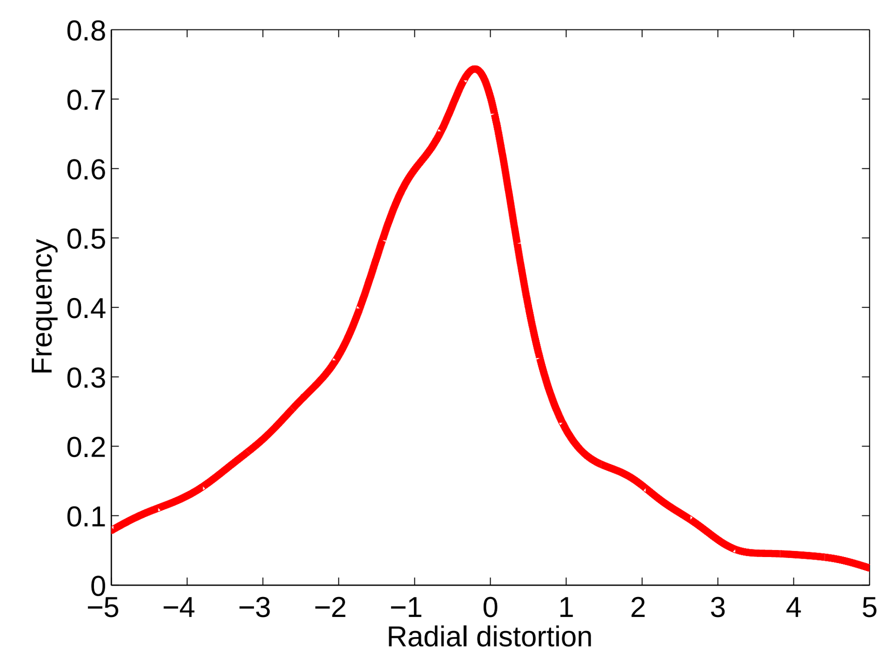

FIGURE 4: Result of kernel voting for radial distortion. The standard deviation of the Gaussian kernel was fixed to 1/3 and the peak of the curve is at $\mu = -0.20$. This value was used to remove distortion from the image. In Figure 5 the result after rectification is shown.

FIGURE 5: Left: An image taken with a fisheye lens. Right: The same image rectified after kernel voting was used to determine the radial distortion

References

[BJÅ09] Martin Byröd, Klas Josephson, and Kalle Åström. Fast and stable polynomial equation solving and its application to computer vision. *Int. Journal of Computer Vision*, 84(3):237–255, 2009.

[BKP08] Martin Bujnak, Zuzana Kukelova, and Tomas Pajdla. A general solution to the p4p problem for camera with unknown focal length. In *Proc. Conf. Computer Vision and Pattern Recognition, Anchorage, USA*, 2008.

[Fit01] A. W. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *Proceedings of Computer Vision and Pattern Recognition Conference (CPVR)*, pages 125–132, 2001.

Whiteboard