Calibration harp

April 15, 2011

Abstract

A tool of lens distortion measurement is introduced in this chapter. Lens distortion is a non-linear deformation which deviates a pinhole camera from central projection. The alignement is the only property preserved in the central projection. So it is reasonable to measure the straightness of the projection of 3D straight lines to evaluate the lens distortion. To have a precise evaluation in practice, we need some very straight strings of good quality. It is relatively easy to ensure the straightness by tightly strecthing the strings and twisting them on a frame, while it is more delicate to choose an appropriate type of string. We tried four types of strings and found that the fishing string is the best choice for our purpose. An evaluation pattern made up of several parallel tightly stretched fishing strings, called a "calibration harp" is thus built. The Devernay sub-pixel precision edge detector is used to extract the edge points in image, which are then associated to the line segments detected by LSD (Line Segment Detector). Finally, the distortion is evaluated as the root-mean-square (RMS) distance from the edge points belonging to a same line segment to their corresponding linear regression line.

1 Introduction

For precise 3D stereo applications, lens distortion correction is a crucial step. Once the camera is calibrated, it is relatively easy to use the techniques like triangulation to reconstruct the 3D scene. But if an imprecise distortion model is used to correct the images, the residual distortion will be directly back-projected to the reconstructed 3D scene and make the scene distorted. This can be very harmful in the applications like remote sensing of early warning of geology disasters or topographic maps making from stereographic pairs of aerial photographs. Surprisingly, in despite of its importance, the precision of distortion correction is not explicitly verified before.

In the literature, three kinds of distortion correction methods coexist by minimizing different error:

- classic pattern-based methods, by minimizing the re-projection error;

- plumb-line methods, by minimizing the straightness error of corrected lines;
- enlarged epipolar methods, by minimizing the algebraic error in the estimate of enlarged fundamental matrix.

Traditionally, in classic pattern-based methods, the lens distortion is estimated together with the camera internal and external parameters [25, 30, 33, 19, 31]. So we call these methods global camera calibration methods. These methods usually are not blind and need a known planar or non-planar pattern which contains simple geometric shapes. The corners or the centroid of these shapes are used as accurate control points. The global process finds the camera parameters minimizing the distance between the observed position of these points in the real image, and their position in the image simulated by retro-projection of the pattern model using the camera model. This is a non-linear problem with many parameters. So the result will be precise only if the model parameters capture the correct physical camera properties of cameras, and if the minimization algorithm finds a global minimum.

The second method is called "plumb-line" methods, which rectify the distorted lines in images which are the projection of 3D straight lines. The first paper about "plumb-line" method is written by Brown in 1971 [6]. The similar idea has been used by integrating differet distortion models, like radial model [3, 23, 29], FOV (Field Of View) model [13] or rational function model [9]. The error to be minimized is the straightness error of corrected lines.

Recently more attention has been paid to estimate the distortion without specific patterns. The distortion is estimated from the correspondences between two or several images without knowing any camera information. The main tool used here is enlarged epipolar constraints, which incorporate lens distortion into the epipolar geometry. Some iterative [26, 32] or non-iterative method, for example, quadratic-eigenvalue problem (QEP) in [4, 14], lifting method in [5], companion matrix method in [20], radial trifocal tensor in [28], quadrifocal tensor in [27], ratio function model in [10], Gröbner basis [18, 22, 7, 17, 16], are used to estimate the distortion and correct it. These methods minimize the algebraic error in the estimate of enlarged fundamental matrix.

Differet types of error mean that there is no common evaluation for the correction precision of different methods. That is why we want to propose an absolute measurement. The proposed evaluation is very simple and evident, which is based on the theorem in section 2. In section 3, we discuss how to build a pattern called "calibration harp" for precision verification. Section 4 shows in practice how to use the photos of the calibration harp to compute the correction precision.

2 From straight lines to straight lines

In this section, we want to present and prove the theorem on which the evaluation tool is based. It is know to all that the alignement is

the only property preserved in the central projection. And it can be presented in a more formal way by the following theorem:

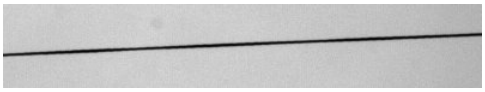
Theorem 1 *A camera follows the pinhole model if and only if the projection of every straight line in space onto the camera is a line.*

The aim of distortion correction is exactly to bring a real camera back to a pinhole camera. Since the above theorem presents a sufficient and necessary condition, it is enough to rectify the distorted lines in the image to obtain a pinhole camera. So compared to other methods, plumb-line methods seem to minimize the approriate error. But in the literature of plumb-line methods, there is no details about experiment setups, including the type of lines, the photographe condition, the line/edge detection, etc. Since all of these are important to devise a precise evaluation tool, we will present them in the following sections.

Even though this theorem has been widely cited in computer vision papers [6, 3, 23, 29, 13], there is no proof for it in the literature. Due to this reason, we want to prove it. **lack of proof. Insert later.**

3 Build the pattern of verification: a calibration harp

According to the above theorem, it is necessary to find some absolutely straight lines in the scene. But in the natural scene, it is usually difficult to find the absolutely straight lines. The only thing we have in mind is the cables on the cable-stayed bridge, which support the weight of the bridge, thus are stretched very tightly. Nonetheless, it is still more convenient if we have our own pattern containing very straight lines. In practice, it is relatively easy to build this kind of pattern. We first find a solid frame, like a wood frame we found. Then we screwed some screws on two sides of the wood frame. Finally we stretched the strings very tightly and fixed them with the aid of the screws to ensure the straightness of the strings. This pattern looks like the musical instrument harp, where comes from the name "calibration harp". The quality of strings in fact plays also an important role in building a good calibration harp. Four different strings were tried: the sewing string, the tennis racket string, the transparent fishing string and the opaque fishing string. The sewing strings are not very smooth and their thickness oscillates in a braid pattern along the strings, due to their twisted structure (see Fig. 1a). Among the four types of strings, the tennis racket string (Fig. 1b), the transparent fishing string and the opaque fishing string (Fig. 1c) are apparently more smooth than the sewing string and have a more uniform thickness. But, as we shall see, tennis racket strings are rigid and would require an extreme tension to become straight. The fishing strings are both smooth and flexible, thus can be easily stretched very straight on the wood framework (see Fig. 2). But the transparent fishing string is like a lens which can introduces makes ..... So finally we think that the opaque fishing string is the best choice to build the calibration harp.



(a) The sewing line

(b) The tennis racket line

(c) The fishing line

Figure 1: The quality of lines. (a) sewing line. (b) tennis racket line. (c) fishing line.

To ensure the extraction precision of the edge points from the string images, a uniform background with contrast to the string color must be preferred. The first idea was to use an uniform wall as background. However, the projected shadows of the strings on the wall are a nuisance which can cause an edge detector failure (see Fig. 2a and 2c). The uniform wall should therefore be far away from the harp, which is not easy to realize. Next to a uniform wall, the sky is the only distant and uniform background that comes into mind. Yet, in the experiments, we found that it was difficult to take photos of the harp against the sky. When the angle of view of the camera is large, it is difficult to photograph only the sky and to avoid the interference of buildings, trees, etc. in the photos. In addition, even if at first sight the sky looks uniform, it turns out to be often inhomogeneous: see Fig. 3a). The final solution was to fix a translucent paper on the back of the harp and to use back lighting (see Fig. 2b and 2d for the harp with the translucent paper). This setup allows us to take photos anywhere, provided the back of the harp is sufficiently lit.

Remark that the above theorem only concerns about the geometric aspect by assuming that there is no optic effect, like blur, aliasing, vignetting, etc. But in reality, the photos are always disturbed more or less by the optic effect. To lessen the optic aberration as much as possible, the photo must be taken under a controlled condition. A Canon EOS 30D reflex camera was installed on a tripod with 10 seconds timer to avoid hand shakes and motion blur. The camera can be rotated on the tripod to take photos of different orientations, by keeping the the camera parallel and the same distance to the harp. Compared to the photos of sewing strings taken by hand against the

(a) The harp with an uniform opaque object as background

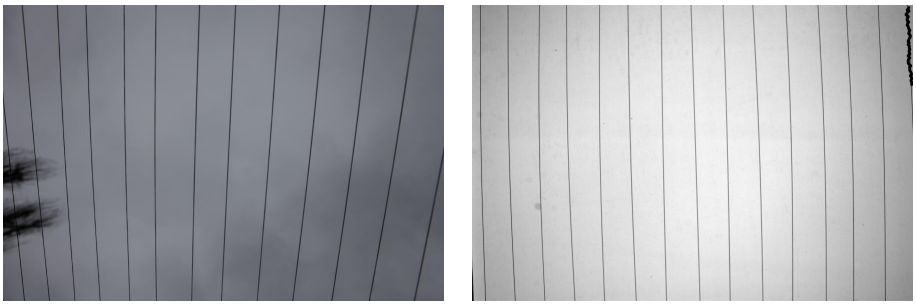
(b) The harp with a translucent paper as background

(c) A close-up of the harp with an uniform opaque object as background

(d) A close-up of the harp with a translucent paper as background

Figure 2: The harp with an opaque object or a translucent paper as background. (a) The harp with an uniform opaque object as background (see a close-up in (c)). (b) The harp with a translucent paper as background (see a close-up in (d)). The shadow can be seen in (a) and (c), while there is no shadow in (b) or (d).

sky (Fig. 3a), the photos of fishing strings with a translucent paper (Fig. 3b) have a more uniform background. In addition, the images taken by hand (Fig. 3a) suffer from inhomogeneous blur or variation of strings thickness caused by the inconstant distance from camera to the harp or the hand motion, while the images taken by tripod (Fig. 3b) have better quality.



(a) The photo of the harp is taken by hand against the sky

(b) The photo of the harp is taken against a translucent paper by a tripod

Figure 3: The quality of photos depends on the harp, its background and also the stability of camera for taking photos.

4 Line segment detection and edge points extraction

In this section, we show how to evaluate the correction precision in practice. Assume the distortion in the photos of calibration harp is estimated and corrected by a certain distortion correction method. According to the above theorem, whether a camera is a pinhole camera or not can be evaluated by judging whether the corrected lines are straight or not. The straightness of a line is defined as the root-mean-square (RMS) distance from the edge points to their corresponding linear regression line. To compute the RMS distance, we need to extract the edge points of the distorted lines from the images in sub-pixel precision. Briefly, the lines are first detected by the LSD algorithm which groups the pixels having coherent gradient direction into line support regions [24]. In each validated line support region, Devernay's algorithm [12] is used to extract the edge points at sub-pixel precision. Finally, a 1D Gaussian convolution followed by a sub-sampling is performed on the extracted edge points to reduce the detection and aliasing noise left by this detection.

4.1 Line detection

LSD is a linear-time line segment detector that gives accurate results, controls its own false detection rate, and requires no parameter tuning

[24]. The algorithm starts by computing the gradient direction at each pixel to produce a level-line field, i.e., a unit vector field such that all vectors are tangent to the level line going through their base point. Then, this field is segmented into connected regions of pixels that share the same level-line angle up to a certain tolerance (see Fig. 4). These connected regions are called line support regions. Each line support region (a set of pixels) is a candidate for a line segment, which is then validated by *a contrario* approach and the Helmholtz principle proposed in [1, 2]. **More detail will be added by Rafael.**

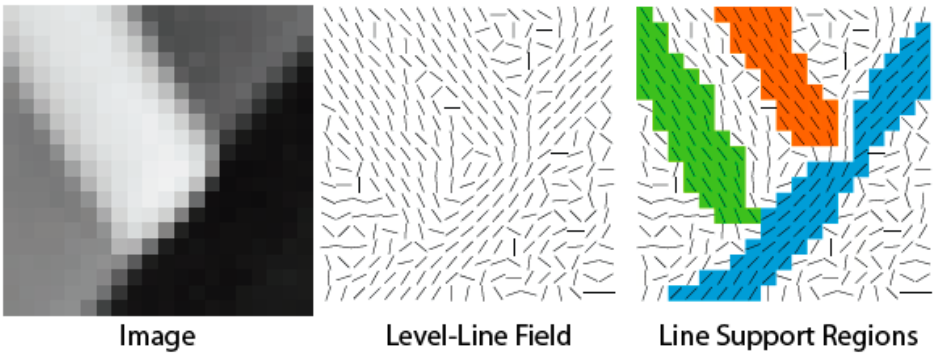


Figure 4: LSD algorithm.

4.2 Devernay's detector

The LSD algorithm gives a validated line support region associated to a line segment, which groups a set of pixels sharing the same gradient orientation up to some toleration. Devernay's detector [12] is then used to extract the edge points of the line segments with sub-pixel precision in each validated line support region. On good quality images (SNR larger than 100), Devernay's detector can attain a precision of about 0.05 pixels. The implementation of Devernay's detector is very simple since it is derived from the well-known Non-Maxima Suppression method [8, 11]. It can be recapitulated in the following steps.

1. Let a point (x, y), where x and y are integers and $I(x, y)$ the intensity of pixel (x, y).
2. Calculate the gradient of image intensity and its magnitude in (x, y).
3. Estimate the magnitude of the gradient along the direction of the gradient in some neighborhood around (x, y).
4. If (x, y) is not a local maximum of the magnitude of the gradient along the direction of the gradient then it is not an edge point.
5. If (x, y) is a local maximum then estimate the position of the edge point in the direction of the gradient as the maximum of an interpolation on the values of the gradient norm at (x, y) and the neighboring points.

In step 3, the magnitude of the gradient along the direction of the gradient at points (x, y) is computed by linearly interpolating the closest points in the 3×3 neighborhood of the point (x, y) (see Fig. 5). In step 5, if (x, y) is a local maximum, then its position is refined by a simple quadratic interpolation of the values of the gradient magnitude between the 3 values in the gradient direction. The quadratic function of gradient magnitude along the gradient direction can be written as:

$$f(l) = al^2 + bl + c \tag{1}$$

with l the distance to the point (x, y) and a, b and c unknown parameters. In Fig. 5, three points A, B and C is sufficient to solve a, b and c. Then the offset l_0 of the refined edge point to the point (x, y) can be obtained by computing the derivative of $f(l)$ and setting it zero: $\frac{df(l)}{dl} \mid_{l=l_0} = 0$.

Remark that the sub-pixel refinement of Devernay's detector is similar to the one of the SIFT method [21] except that SIFT works on the Laplacian value and uses a two-dimension quadric interpolation, while Devernay's detector works on the magnitude of gradient and uses a one-dimensional quadric interpolation in the gradient direction.

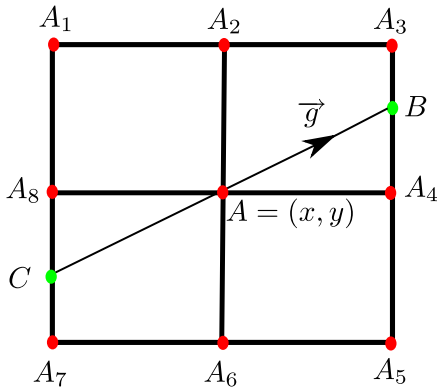


Figure 5: Devernay sub-pixel precision edge detector. The vector $\overrightarrow{g}$ is the gradient direction at the point (x, y). The gradient magnitude at point B is linearly interpolated by the gradient magnitude at point A_3 and A_4. Similarily, the gradient magnitude at point C is linearly interpolated by the gradient magnitude at point A_7 and A_8. If point A has the gradient magnitude bigger than B and C, then its position is refined by computing an offset through a quadric interpolation along the direction of $\overrightarrow{g}$.

4.3 Convolution and sub-sampling of edge points

For the photos of strings, almost every pixel along each side of one string is detected as edge point in sub-pixel precision. So there are

about 1000 edge points detected for a line of length about 1000 pixels. This large number of edge points give the possibility to further reduce the detection and aliasing noise left by the detection through a convolution followed by a sub-sampling. This process is similar to the two-dimension case where the Gaussian blur about $0.8 \times \sqrt{t^2 - 1}$ is needed before a t-subsampling to avoid the aliasing [15]. A similar rule can be applied here. But instead of a two-dimensional signal, we have two one-dimension signals (x-coordinate and y-coordinate of edge points) along the length of the line. The Gaussian convolution is performed on two one-dimension signals separately. Since the edge points are not regularly sampled along the line, a preliminary re-sampling step is needed. This re-sampling is made along the length of the line by taking a uniform sampling step, which is set to be m times smaller than the average distance between two adjacent edge points. A linear interpolation is used here to accelerate the re-sampling (see Fig. 6). Assume the distance between two adjacent edge ponits (x_1, y_1) and (x_2, y_2) is l and the re-sampling step is d. Then the re-sampled point (x', y') can be expressed as

$$\begin{aligned} x' &= \frac{d}{l}(x_2 - x_1) + x_1 \\ y' &= \frac{d}{l}(y_2 - y_1) + y_1. \end{aligned}$$

Once the line is re-sampled, the Gaussian blur $0.8 \times \sqrt{t^2 - 1}$ can be applied then followed by a sub-sampling with factor mt on the x and y coordinates separately (the re-sampling step is m times smaller than the average distance between two adjacent edge points).

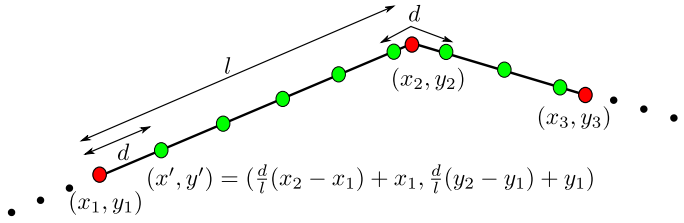


Figure 6: Line re-sampling. The red points (x_1, y_1), (x_2, y_2), $\cdots$ are the edge points extracted by Devernay's detector. They are irregularly sampled along the line. The re-sampling (in green) is along the length of the line with the uniform step d. The linear interpolation is used to compute the re-sampled point fast.

4.4 Computation of straightness

As have been presented before, the straightness of a line is defined as the root-mean-square (RMS) distance from its edge points to their corresponding linear regression line. Given a set of corrected edge

points $(x_{u_i}, y_{u_i})_{i=1,\cdots,N}$ which are supposed to be on a line, we compute the linear regression line:

$$\alpha x_{u_i} + \beta y_{u_i} - \gamma = 0 \tag{2}$$

with $\tan 2\theta = -\frac{2(A_{xy} - A_x A_y)}{V_{xx} - V_{yy}}$, $\alpha = \sin\theta$, $\beta = \cos\theta$, $A_x = \frac{1}{N}\sum_{i=1}^{N} x_{u_i}$, $A_y = \frac{1}{N}\sum_{i=1}^{N} y_{u_i}$, $A_{xy} = \frac{1}{N}\sum_{i=1}^{N} x_{u_i} y_{u_i}$, $V_{xx} = \frac{1}{N}\sum_{i=1}^{N}(x_{u_i} - A_x)^2$, $V_{yy} = \frac{1}{N}\sum_{i=1}^{N}(y_{u_i} - A_y)^2$ and $\gamma = A_x \sin\theta + A_y \cos\theta$. Then the straightness is computed as:

$$S = \sqrt{\frac{\sum_{i=1}^{N} \left(\alpha x_{u_i} + \beta y_{u_i} - \gamma\right)^2}{N}}. \tag{3}$$

5 Some preliminary results

In this section, we just show some results to support our argument that the opaque fishing string is more appropriate to evaluate the correction precision. A good string should not have other imperfection aspects which introduce some error susceptible to be mixed with the lens distortion. We hope that once the distorted line is ideally corrected by a certain correction method, the straightness only reflects the correction performance, but not affected by other factors.

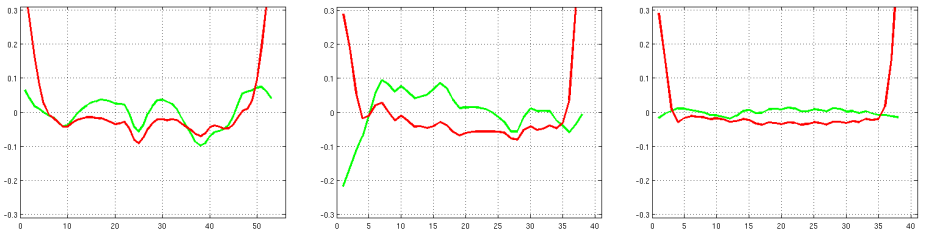
In Fig. 7, the high frequency of the distorted sewing string, the distorted tennis racket string and the distorted opaque fishing string are compared to the straightness error of their corresponding corrected strings. The almost superimposing high frequency oscillation confirms that the high frequency of the distorted strings is not changed by the lens distortion correction. In such case, the straightness error includes the high frequency of the distorted strings and does not really reflect the correction performance. So it is better to use the string which contains the high frequency oscillation as small as possible. Among the three types of strings, the fishing string shows the smallest such oscillation. The larger oscillation of the sewing string is due to a variation of the thickness related to its twisted structure, while the tennis racket string is simply too rigid to be stretched, even if this is not apparent in Fig. 1b).

References

[1] Jean-Michel Morel Agnès Desolneux, Lionel Moisan. Meaningful alignments. *International Journal of Computer Vision,*, 40(1):7–23, 2000.

[2] Jean-Michel Morel Agnès Desolneux, Lionel Moisan. *From Gestalt Theory to Image Analysis, a Probabilistic Approach.* Springer, 2008.

[3] Luis Alvarez and J. Rafael Sendra. An algebraic approach to lens distortion by line rectification. *Journal of Mathematical Imaging and Vision*, 35:36–50, 2009.



(a) The sewing string

(b) The tennis racket string

(c) The fishing string

Figure 7: The small oscillation of the corrected lines is related to the quality of the strings. The green curve shows the RMS distance (in pixels) from the edge points of a corrected line to its regression line. The red curve shows the high frequency of the corresponding distorted line. The corrected line inherits the oscillation from the corresponding distorted line. (a) the sewing string. (b) the tennis racket string. (c) the fishing line. The x-axis is the index of edge points. The range of y-axis is from -0.3 pixels to 0.3 pixels.

[4] T. Pajdla B. Micusik. Estimation of omnidirectional camera model from epipolar geometry. *CVPR*, 1:485, 2003.

[5] K. Barreto, J.P. Daniilidis. Fundamental matrix for cameras with radial distortion. *ICCV*, 1:625– 632, 2005.

[6] Duane C. Brown. Close-range camera calibration. *Photogrammetric Engineering*, 37:855–866, Brown.

[7] M. Byrod, Z. Kukelova, K. Josephson, T. Pajdla, and K. Astrom. Fast and robust numerical solutions to minimal problems for cameras with radial distortion. *Computer Vision and Image Understanding*, 114(2):1–8, 2008.

[8] J. F. Canny. Finding edges and lines in images. *Technical Report AI-TR-720, Massachusets Institute of Technology, Articial Intelligence Laboratory*, 1983.

[9] D. Claus and A. W. Fitzgibbon. A plumbline constraint for the rational function lens distortion model. *BMVC*, pages 99–108, 2005.

[10] D. Claus and A.W. Fitzgibbon. A rational function lens distortion model for general cameras. *CVPR*, 1:213–219, 2005.

[11] R. Deriche. Using cannys criteria to derive a recursively implemented optimal edge detector. *The International Journal of Computer Vision*, 1(2):167187, 1987.

[12] F. Devernay. A non-maxima suppression method for edge detection with sub-pixel accuracy. Technical Report 2724, INRIA rapport de recherche, 1995.

[13] F. Devernay and O. Faugeras. Straight lines have to be straight. *Mach. Vision Appli.*, 13:14–24, 2001.

[14] A. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. *ICPR*, 1:125–132, 2001.

[15] Guoshen Yu Jean-Michel Morel. On the consistency of the sift method. Technical report, CMLA, ENS-Cachan, 2008.

[16] K. Josephson and M. Byrod. Pose estimation with radial distortion and unknown focal length. *CVPR*, pages 2419–2426, 2009.

[17] Z. Kukelova and T. Pajdla. A minimal solution to the autocalibration of radial distortion. *CVPR*, page 17, 2007.

[18] Z. Kukelova and T. Pajdla. Two minimal problems for cameras with radial distortion. *OMNIVIS*, 2007.

[19] Dhome M. Lavest J., Viala M. Do we really need accurate calibration pattern to achieve a reliable camera calibration. *ECCV*, 1:158–174, 1998.

[20] Hongdong Li and Richard Hartley. A non-iterative method for correcting lens distortion from nine point correspondences. *OmniVis*, 2005.

[21] David G Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91110, 2OO4.

[22] T. Pajdla, Z. Kukelova, and M. Bujnak. Automatic generator of minimal problem solvers. *ECCV*, pages 302–315, 2008.

[23] B. Prescott and G. F. Mclean. Line-based correction of radial lens distortion. *Graphical Models and Image Processing*, 59:39–47, 1997.

[24] J.-M. Morel G. Randall R. Grompone von Gioi, J. Jakubowicz. Lsd: A fast line segment detector with a false detection control. *IEEE Trans. on PAMI*, 99, 2008.

[25] C. C Slama. *Manual of Photogrammetry, 4th edition.* Falls Church, American Society of Photogrammetry, Virginia, 1980.

[26] Gideon P. Stein. Lens distortion calibration using point correspondences. *CVPR*, 602–608, 1997.

[27] M. Thirthala and S. Pollefeys. Multi-view geometry of 1d radial cameras and its application to omnidirectional camera calibration. *ICCV*, pages 1539–1546, 2005.

[28] M. Thirthala, S. Pollefeys. Multi-view geometry of 1d radial cameras and its application to omnidirectional camera calibration. *ICCV*, 2:1539–1546, 2005.

[29] Václav Hlaváč Tomáš Pajdla, Tomáš Werner. Correcting radial lens distortion without knowledge of 3-d structure. *Research Report, Czech Technical University*, 1997.

[30] Roger Y. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, Vol. RA-3, 1987.

[31] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *TPAMI*, 14(10):965–980, 1992.

[32] Z. Zhang. On the epipolar geometry between two images with lens distortion. *Proceedings of the 1996 International Conference on Pattern Recognition*, 7270:407, 1996.

[33] Z. Zhang. A flexible new technique for camera calibration. *ICCV*, pages 663–673, September 1999.