Estimating Intrinsic Parameters of Cameras using Two Arbitrary Rectangles

Jun-Sik Kim and In So Kweon
Dept. of EECS, KAIST
373-1 Kusong-dong, Yuseong-Gu, Daejeon, Korea
jskim@rcv.kaist.ac.kr, iskweon@kaist.ac.kr

Abstract

In this paper, we propose new camera calibration methods assuming a static camera. Two corresponding imaged rectangles whose aspect ratios are unknown are sufficient to calibrate a camera. By warping the images properly, we show that the information from the imaged rectangles can be transformed to the form of camera constraints. Based on this results, we propose two methods, one for three or more images and the other for only two images. The proposed methods are verified with synthetic and real images, and the results are comparable with less assumptions on cameras and on scenes.

1. Introduction

Estimating the intrinsic parameters of cameras is an important procedure in many vision-based methods, including three-dimensional measurements, rendering, robot localization, etc. The approaches to estimating the intrinsic parameters of cameras are summarized in three categories. The first one uses known targets, called calibration objects, which are three-dimensional [3, 10], or two-dimensional [7, 9, 12], or even one-dimensional [13]. In these methods, camera calibration works quite robustly, but the physical measurements of the artificial objects are needed. The second approach uses only some camera assumptions without calibration targets, called autocalibration [2, 4, 8]. To autocalibrate cameras, there must be much longer image sequences and robustly matched feature sets. The final approach is in-between the first two approaches. There is some knowledge of both the scenes and the cameras. By combining the various constraints on the cameras and the scenes, camera calibration is possible [1, 7, 11]. In these methods, the constrained camera models are used, such as zero-skewed or square-pixel cameras. However, many industrial cameras do not suffice under these assumptions.

In [6], we propose a method to estimate infinite homography between two views using two arbitrary rectangles. In this paper, we propose new camera calibration methods based on [6]. In our proposed methods, we assume that the camera used is static. We only have to find two corresponding imaged rectangles whose aspect ratios are not known. The images are warped so that the projected rectangle is to be a real rectangle. We imagine a virtual camera, called a fronto-parallel camera, that can capture a warped image, and show that there are some constraints on the imagined camera. Based on the properties, we propose two methods, one for three or more images and the other for only two images. Both methods are tested with synthetic and real image sets, and the results are comparable to the well-known metric measurement based calibration methods.

This paper is organized as follows. In Section 2, we briefly introduce warping based on the imaged rectangles and the concept of the fronto-parallel camera, which may capture the warped image. Section 3 gives sketches about the computation of infinite homography between two views based on the virtual fronto-parallel cameras. We propose two algorithms, one for three or more images and the other for just two images in Section 4. Section 5 gives an analysis of the proposed algorithm and the results from real input images. We conclude this paper in Section 6.

2. Fronto-parallel camera

Assume that there is a rectangle in 3D, whose aspect ratio is unknown, denoted as R_m, and we have a view capturing the rectangle in a general position. In this case, we can find a homography to make the projectively distorted rectangle to align with the orthogonal axis of the rectangle. The warped image is called a semi-metric image [6]. Figure 1 shows an example of a semi-metric image. In the semi-metric image, the projected rectangle in the input image becomes a rectangle whose aspect ratio is R_{sm}, which can be measured.

This research has been supported by the Korean Ministry of Science and Technology for NRL Program(Grant number M1-0302-00-0064).

Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE

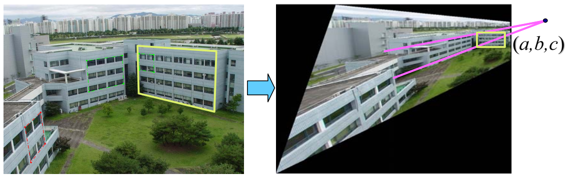


Figure 1: Example of a semi-metric image

Assume that there is a virtual camera to take the semi-metric image. We call the camera as the *Fronto-Parallel (FP) camera*, because the image plane of the virtual camera is parallel to the scene plane containing the rectangle [6].

The camera matrix of a FP camera is given as

$$\mathsf{K}_{FP} = \begin{bmatrix} 1/R_{sm} & 0 & a \\ & 1/R_m & b \\ & & c \end{bmatrix} \tag{1}$$

up to scale. As you can easily see, the camera matrix K_{FP} expresses a camera whose skew is zero, and its pixel aspect ratio is equal to a ratio between an aspect ratio of the reference rectangle R_m and the corresponding semi-metric aspect ratio R_{sm}. The principal point of the camera $(a, b, c)^\top$ is expressed with the scaled vanishing point orthogonal to the reference plane, which appears in Figure 1; the scale has the role of a focal length. To summarize, the FP camera matrix is determined with scene information and a camera pose.

Naturally, the relationship between the IAC in the projective space $\boldsymbol{\omega}$ and the IAC of a FP camera $\boldsymbol{\omega}_{FP}$ is obtained from such basic conic transformation as

$$\mathsf{H}_{sm}^{-\top}\omega\mathsf{H}_{sm}^{-1} = \omega_{FP} \tag{2}$$

where H_{sm} is a plane homography from projective space to semi-metric space.

The FP cameras have some motion constraints.

1. The FP cameras rotate to the physical camera of the original image.
2. The FP cameras that are derived from an identical rectangle in 3D from different viewpoints are purely translating to each other.

Because two FP cameras derived from an identical rectangle are purely translating, the infinite homography between the corresponding FP cameras is given as

$$\begin{aligned} \mathsf{T} &= \mathsf{K}_{FP2}\mathsf{R}_{21}\mathsf{K}_{FP1}^{-1} \\ &= \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & t_z \end{bmatrix} \end{aligned} \tag{3}$$

because K_{FP1} and K_{FP2} have the same pixel aspect ratios with zero skew from Eq. (1), and the relative rotation is $\mathtt{I}_{3\times 3}$.

3. Linear estimation of infinite homography from two arbitrary rectangles

If a captured scene contains two arbitrary rectangles whose aspect ratios are unknown, the infinite homography is estimated linearly using the parametrization, as in Eq. (3).

Assume that there are two views, view 1 and view 2 which contain two arbitrary rectangles named rectangles i and j. In that case, we can find two infinite homographies with respect to two rectangles as

$$\begin{aligned} \mathsf{H}_{\infty,i}^{12} &= \mathsf{H}_{sm2,i}^{-1}\mathsf{T}_i\mathsf{H}_{sm1,i} \\ \mathsf{H}_{\infty,j}^{12} &= \mathsf{H}_{sm2,j}^{-1}\mathsf{T}_j\mathsf{H}_{sm1,j} \end{aligned} \tag{4}$$

where $\mathsf{H}_{sm1,i}$ means a semi-metric warping matrix of view 1 w.r.t. the rectangle i.

However, the infinite homography is dependent only on the intrinsic parameters of cameras and a relative rotation between two views [5] such as

$$\mathsf{H}_{\infty}^{12} \sim \mathsf{K}_2\mathsf{R}_{21}\mathsf{K}_1^{-1}.$$

This means that the infinite homography is determined identically regardless whether rectangle is selected as a reference. This gives us a constraint equation

$$\rho\mathsf{H}_{sm2,i}^{-1}\mathsf{T}_i\mathsf{H}_{sm1,i} = \mathsf{H}_{sm2,j}^{-1}\mathsf{T}_j\mathsf{H}_{sm1,j} \tag{5}$$

where ρ is a proper scale factor.

The unknowns are the parameters of T_i and T_j and a scale factor ρ. The number of unknowns is seven and we have nine equations, so we can easily solve the equation linearly. Note that we do not use any metric measurements such as lengths and aspect ratios of scene rectangles.

4. Estimating intrinsic parameters

4.1. Using three or more images

If we know the infinite homographies between views captured by a static camera with three or more views, the camera calibration is linearly possible [5]. The algorithm is as follows:

1. Track two arbitrary rectangles.
2. Find semi-metric warping matrices in all views w.r.t. the two rectangles.
3. Estimate proper transformations T_i and T_j using Eq.(5).
4. Calculate the infinite homography H_{∞}^{12} using Eq.(4).
5. Find the image of absolute conic (IAC) $\boldsymbol{\omega}$ such that $\boldsymbol{\omega} = (\mathsf{H}_{\infty}^{12})^{-\top}\boldsymbol{\omega}(\mathsf{H}_{\infty}^{12})^{-1}$.

Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE

6. Determine the camera matrix K from IAC $\boldsymbol{\omega}$ using Cholesky decomposition $\boldsymbol{\omega} = (\mathsf{K}\mathsf{K}^\top)^{-1}$.

This algorithm can be compared with previous works that use information on scene geometry and proper camera assumptions [1, 7, 11]. The key difference is that we do not assume that there is important metric information of the scene such as line lengths or aspect ratios of the rectangles; we have no camera assumptions such as zero-skew and the known aspect ratio of the pixels. Because finding some rectangles in images is much easier than finding some metrics only with images, the proposed method is much more flexible than the previous ones.

4.2. Using only two images

If there are no further assumptions about the cameras, the camera calibration with only two views cannot generally be possible. However, we show that it is possible if we have two arbitrary corresponding rectangles in each view. That is because the "virtual" FP cameras can be treated as constrained physical cameras.

Assume that there are two views captured by a static camera, and we have infinite homography between the two views. From the infinite homography H_∞, the IAC ω is given as

$$\omega = \mathsf{H}_\infty^{-\top} \omega \mathsf{H}_\infty^{-1},$$

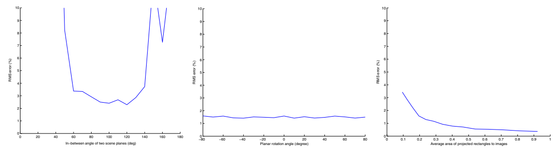
and we can make a linear equation $\mathsf{A}\bar{\omega} = 0$, as we did. However, the rank of A is only four, so we need another constraints of the static camera. By introducing FP cameras, we have two camera constraints of the two fronto-parallel cameras as in Eq. (1). The constraints on FP cameras can be transformed to ones of physical cameras using Eq. (2), and calibration of the camera is possible.

5. Experiments

5.1. With synthetic data

We analyze the performance of the algorithm in various situations. We generated three views having two arbitrary rectangles in general poses. Gaussian noises whose standard deviation is 0.5 pixels were added on the corners of the rectangles.

First, we test the effect of angles between the model plane and the image plane of the camera. In Figure 2, RMS errors of focal length estimation are depicted, with 500 iterations. Figure 2a shows the performance to rotation of the plane in 3D along the x-axis. In 40 degrees, the plane is orthogonal to the image plane, and all the features lie on a line. This is a singular case, and the calibration is not much degraded in general conditions.



(a) Pose (b) Planar rotation (c) Area

Figure 2: Simulated performance of the proposed algorithm using three images

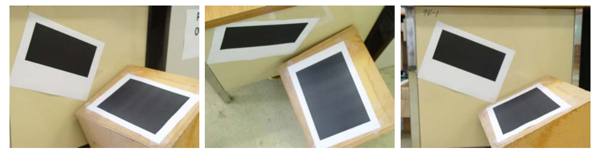


Figure 3: Real input images for camera calibration

Second, we analyze the effects of planar rotation of the model plane. Figure 2b shows the effects of planar rotation of the world plane. We tested the performance of the differences of the directions of the orthogonal axis of the two rectangles. We conclude that the directions of the model axis do not affect the performance of the algorithm.

Next, we test the effects of the area of the projected rectangles in input images. Figure 2c shows the performance of the area of the rectangles used in images. The performance is degraded exponentially when the area becomes smaller. This is natural because it is equivalent to estimating vanishing points from the four corner points of the rectangles. However, the algorithm works robustly if we have projected rectangles larger than 10% of the average of the whole images, as shown in Figure 2c.

5.2. With real images

We test our algorithm with real images. Figure 3 shows some input images containing two arbitrary rectangles. The images are captured with the SONY DSC-F717 in 640 × 480 resolution. We do not know the exact value of the aspect ratios of the rectangles in the metric world. The rectangles are placed arbitrarily, so we cannot use the pose relation between the two planes. Note that some imaged rectangles are rarely distorted projectively.

The calibration result is given as

$$\mathsf{K}_{estimated} = \begin{bmatrix} 899.4727 & 20.9762 & 322.9044 \\ 0 & 913.2549 & 297.9821 \\ 0 & 0 & 1.0000 \end{bmatrix}.$$

For comparison, we calibrated the camera with the well-known Zhang's plane based calibration method [12] using

Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE

six metric planes, as

$$\mathsf{K}_{Zhang} = \begin{bmatrix} 888.5763 & 14.3200 & 269.8877 \\ 0 & 887.2853 & 243.0086 \\ 0 & 0 & 1.0000 \end{bmatrix}.$$

Note that the proposed algorithm does not need any kind of metric measurements such as metric coordinates or line lengths. Also although we did not apply any robust method or refinement techniques such as Gauss-Newton manner non-linear minimization, the estimated camera parameters are comparable only with three images.

5.3. Two view cases

We test the method for only two views under the same condition as the three view cases. As you can see in Figure 4, calibration is not possible just with two views from a static camera, although the infinite homography is known. However with two imaged rectangles, the intrinsic camera parameter can be estimated.

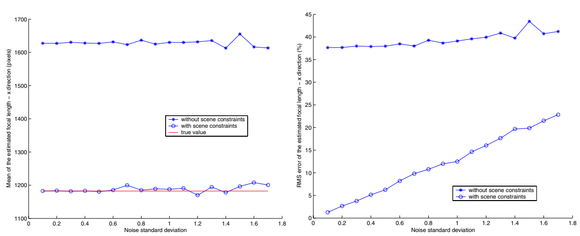


Figure 4: Analysis of the effects of adding two scene constraints to autocalibration of static cameras based on IAC

Applying this method to real images, we first used two real images in Figure 3. The intrinsic parameters of the cameras with two added scene constraints as in Figure 3 is given as

$$\mathsf{K}_{estimated} = \begin{bmatrix} 916.3691 & 107.5102 & 457.3421 \\ 0 & 877.6347 & 343.6547 \\ 0 & 0 & 1.0000 \end{bmatrix}.$$

Compared with the three image case in Figure 3, the accuracy is degraded. The accuracy for the focal length estimation is less 5% in this results.

6. Conclusion

In this paper, we propose new camera calibration methods for static cameras. To calibrate the cameras, we have to find only two corresponding imaged rectangles whose aspect ratios are not known. To solve the problem, we generate warped images so that the projectively transformed rectangles become real rectangles after warping, and imagine there is a camera that may capture the warped image physically. We call the virtual camera the fronto-parallel camera and show that there are some constraints on the camera. By analyzing the motion of the fronto-parallel camera, the infinite homography between two cameras can be retrieved easily with a simple linear method. Based on this, we proposed two methods to calibrate cameras, one for three or more images, and the other for only two images. Both methods are tested with synthetic and real image sets; the results are comparable to the classical calibration methods based on metric measurements, although we have fewer measurements of the scene and constraints on the cameras.

References

[1] B. Caprile and V. Torre. Using vanishing points for camera calibration. *International Journal of Computer Vision*, 4(2):127–140, March 1990.

[2] O. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In *European Conference on Computer Vision*, pages 563–578, 1992.

[3] O. Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press, 1993.

[4] O. Faugeras, Q. Luong, and S. Maybank. Camera self-calibration: Theory and experiments. In *European Conference on Computer Vision*, pages 321–334, 1992.

[5] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge, 2003.

[6] J.-S. Kim and I.-S. Kweon. Infinite homography estimation using two arbitrary planar rectangles. In *Proc. of Asian Conference on Computer Vision*, pages II:1–10, 2006.

[7] D. Liebowitz and A. Zisserman. Combining scene and auto-calibration constraints. In *Proceedings of the 7th International Conference on Computer Vision, Kerkyra, Greece*, Sept. 1999.

[8] M. Pollefeys and L. Van Gool. Stratified self-calibration with the modulus constraint. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 21(8):707–724, August 1999.

[9] P. Sturm and S. Maybank. On plane-based camera calibration: A general algorithm, singularities, applications. In *IEEE Computer Vision and Pattern Recognition or CVPR*, pages I: 432–437, 1999.

[10] R. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Trans. Robotics and Automation*, 3(4):323–344, 1987.

[11] M. Wilczkowiak, E. Boyer, and P. Sturm. Camera calibration and 3d reconstruction from single images using parallelepipeds. In *International Conference on Computer Vision*, pages I: 142–148, 2001.

[12] Z. Zhang. A flexible new technique for camera calibration. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, November 2000.

[13] Z. Zhang. Camera calibration with one-dimensional objects. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 26(7):892–899, July 2004.

Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)
0-7695-2521-0/06 $20.00 © 2006 IEEE