Robust Camera Calibration from Images and Rotation Data

Jan-Michael Frahm and Reinhard Koch

Institute of Computer Science and Applied Mathematics
Christian Albrechts University Kiel
Herman-Rodewald-Str. 3, 24098 Kiel, Germany
{jmf,rk}@mip.informatik.uni-kiel.de

Abstract. The calibration of cameras from external orientation information and image processing is addressed in this paper. We will show that in the case of known rotation the calibration of rotating cameras is linear even in the case of fully varying parameters. For freely moving cameras the calibration problem is also linear but underdetermined for fully varying internal parameters. We show one possible set of contraints to reach a fully determined calibration problem. Furthermore we show that these linear calibration techniques tend to fit to noise for some of the intrinsics. To avoid this fit to noise we introduce a statistical calibration technique which uses the robust components of linear calibration and prior knowledge about cameras. This statistical calibration is fully determined even for freely moving cameras.

1 Introduction

We have seen a lot of research on camera calibration from image sequences over the last decade. These approaches calibrate the cameras by observing unknown scenes and therefore they may suffer under degeneracies caused by the scenes respectively the image information. We will introduce a technique for selfcalibration from image sequences together with external orientation information. This information is available in many applications. Today's cars are already equipped with orientation sensors for Electronic Stability systems (ESP) for example. Future cars will also have smart cameras. Another popular application is the surveillance with rotating and zooming cameras. In this case we have rotation information of the camera but normally lack correct zoom data.

In this contribution we will discuss the possibilities to use this external orientation information for selfcalibration of arbitrary moving and zooming cameras. We will first review the literature in section 2. Selfcalibration from image and rotation data will be discussed in detail in section 3. Finally we will discuss some experiments and conclude.

2 Previous work

Camera calibration has always been a subject of research in the field of computer vision. The first major work on selfcalibration of a camera by simply observing an unknown scene was presented in [9]. Since that time various methods have been developed. Methods for the calibration of rotating cameras with unknown but constant intrinsics were first developed in [11]. The approach was extended for rotating cameras with partially varying intrinsic parameters in [5]. This work uses the infinite homography constraint and has the disadvantage that not all parameters are allowed to vary. The calibration process has three major steps: linearized calibration, nonlinear calibration, and statistical calibration. Sometimes the first calibration step may fail due to noisy data.

Camera selfcalibration from unknown general motion and constant intrinsics has been dicussed in [12]. For varying intrinsics and general camera motion the selfcalibration was proved by [8]. All these approaches for selfcalibration only use the images of the cameras themselves for the calibration.

Only few approaches exist to combine image analysis and external rotation information for selfcalibration. In [10] cameras with constant intrinsics and known rotation were discussed. They use unconstrained nonlinear optimization to estimate the camera parameters. This lack of attention is somewhat surprising since this situation occurs frequently in a variety of applications: cameras mounted in cars for driver assistence, robotic vision heads, surveillance cameras or PTZ-cameras for video conferencing often provide rotation information.

In this paper we will address one of the few cases which have not yet been explored, that of a rotating camera with varying intrinsics and known rotation information. We will show that orientation information is helpful for camera calibration. Furthermore it is possible to detect degenerate cases for calibration like rotation about only one axis or about the optical axis.

3 Selfcalibration with known rotation

In this section we will develop novel techniques to use available external orientation information for camera selfcalibration. We will address both cases of purely rotating and arbitrarily moving cameras.

3.1 Rotating cameras

We can exploit given rotational information to overcome the limitations on the number of varying intrinsics and the problems caused by noise during computation in [5]. The homography $H^{\infty}_{j,i}$ between two images i and j of a rotating camera is given by

$$H^{\infty}_{j,i} = K_i R_{j,i} K_j^{-1} \text{ with } K = \begin{bmatrix} f & s & c_x \\ 0 & a \cdot f & c_y \\ 0 & 0 & 1 \end{bmatrix}, \qquad (1)$$

where f is the focal length of the camera expressed in pixel units. The aspect ratio a of the camera is the ratio between the size of a pixel in x-direction and the size of a pixel in y-direction. The principal point of the camera is (c_x, c_y) and s is a skew parameter which models the angle between columns and rows of the CCD-sensor. $R_{j,i}$ is the relative rotation of the camera between image j and i. If $R_{j,i}$ is known from an external orientation sensor, then equation (1) can be rewritten as

$$K_i R_{j,i} - H^\infty_{j,i} K_j = 0_{3x3} \quad \text{or} \quad K_i^{-1} m_{i,k} - \alpha(m_{j,k}) R_{j,i} K_j^{-1} m_{j,k} = 0_{3x3}, \quad (2)$$

where $\alpha(m_{j,k})$ is the factor to homogenize $R_{j,i} K_j^{-1} m_{j,k}$. $(m_{i,k}, m_{j,k})$ is in the set of point correspondences between a point m_k in image j and a point m_k in image i. The homography $H^\infty_{j,i}$ can be estimated from the image point correspondences [3]. Therefore, (2) is linear in the components of K_i and K_j and provides nine linear independent contraints on the intrinsics of the cameras.

The estimated homographies are determined only up to scale $\rho_{j,i}$. Therefore we can estimate only $\tilde{H}^\infty_{j,i} := \rho_{j,i} H^\infty_{j,i}$ from the images. For the estimated homographies $\tilde{H}^\infty_{j,i}$ equation (2) is modified to

$$0_{3x3} = \tilde{K}_i R_{j,i} - H^\infty_{j,i} K_j \text{ with } \tilde{K}_i = \rho_{j,i}^{-1} K_i, \quad (3)$$

which is also linear in the intrinsics of the camera j and linear in the elements of $\tilde{K}_i$. Note that due to the unknown scale we have now six unknowns in $\tilde{K}_i$. Eq. (2) provides nine linearily independent equations for each camera pair for the five intrinsics contained in K_j and the five intrinsics of K_i plus the scale $\rho_{j,i}^{-1}$ contained in $\tilde{K}_i$. If there are no constraints available for the intrinsics, (2) has no unique solution for a single camera pair. With two constraints for the intrinsics or the scale $\rho_{j,i}^{-1}$ the solution is unique. Alternatively, if we consider a camera triplet (i, j, k) with estimated homographies $\tilde{H}^\infty_{j,i}$ and $\tilde{H}^\infty_{j,k}$, (3) provides

$$\tilde{K}_i R_{j,i} - H^\infty_{j,i} K_j = 0_{3x3} \quad \text{and} \quad \tilde{K}_k R_{j,k} - H^\infty_{j,k} K_j = 0_{3x3}, \quad (4)$$

with 17 unknowns and up to 9 independent equations for each camera pair. Therefore, for each camera triplet the solution for the intrinsics and scales is unique and can be solved even for fully varying parameters. In contrast to the approach in [5] this calibration can always be computed even in the case of strong noise.

Evaluation for rotating cameras: To measure the noise robustness of the calibration we test the approach with synthetic data. The center of the rotating camera is at the origin , the image size is 512x512. The camera rotates about x-axis and y-axis with up to six degrees and observes a scene in front of the camera. The location of the projected points is disturbed by Gaussian noise with variance of 2 pixel. The known camera orientation is also disturbed by noise of up to 2 degrees per axis. We varied both pixel and rotational noise. The homographies $\tilde{H}^\infty_{j,i}$ are estimated from point correspondences by least squares estimation. The measurements for the first camera with focal length $f = 415$ and $c_y = 201$ are shown in figure 1. The measured errors for the other images are similar to these results.

It can be seen from figure 1 that the estimated focal length f is rather stable if the pixel noise is less than one pixel and the orientation data are noisy by angular errors of less than one degree. The measured variance of the linear estimation is below 10% for angle noise of up to 1 degree. The estimation for the aspect ratio shows similar stability. The estimated principle point component c_y is not as stable as focal length f. It fits to noise for angular noise greater than 0.5 degrees. The estimation for the other principal point component c_x and the skew s is similar to the values of the principal point component c_y. Furthermore the influence of the orientation noise is much larger since the absolute rotation angle between the cameras is in the range of the noise (6 degree camera rotation with up to 2 degree noise). In the next subsection we will introduce a statistical calibration method for robust calibration of all intrinsics.

3.2 Statistical calibration with known rotation

The above linear approach (4) is able to robustly estimate the focal length and the aspect ratio. The estimation of the principal point is an ill posed problem [5]. For the most cameras the principal point is located close to the image center and the skew is zero. Therefore the use of prior knowledge about the distribution of the principal point and the skew can be used to reduce the estimation problems of the linear approach (4).

Let us consider that the noise n on the measured image positions is additive and has a Gaussian distribution with mean zero and standard deviation σ. Then an approximation of the Maximum Likelihood estimation is given by:

$$\mathrm{MLE} = \arg\min_{K_i,R_i} \sum_{i=1}^{\#cameras} \sum_{k=1}^{\#points} \|K_i^{-1} m_{i,k} - \alpha(m_{j,k}) R_{j,i} K_j^{-1} m_{j,k}\|^2 \qquad (5)$$

To compute an exact Maximum Likelihood we have to weight the backprojection error with the inverse variance of the image measurement. The approximation error is small because we use normalized coordinates [2] for the computation. If we model the expectation pp_{prior} that the principal point probably lies close to the center of the camera and has also a Gaussian distribution whose mean is the image center, a Maximum a Posterori estimation of the intrinsics is simply

$$\mathrm{MAP}_{pp} = \mathrm{MLE} + \lambda_{pp} \sum_{i \in cameras} (c^i - pp_{prior})^T \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} (c^i - pp_{prior}). \qquad (6)$$

where λ_{pp} is the weight of the prior knowledge and σ_x^2, σ_y^2 the distribution parameters for the components of the principal point. Furthermore we are able to use the given sensor orientation as a prior knowledge:

$$\mathrm{MAP}_{ori} = \mathrm{MAP}_{pp} + \lambda_{ori} \sum_{i \in cameras} (1 - < r_{j,i}, r_{est} >) + |\phi_{i,j} - \phi_{est}|. \qquad (7)$$

where $r_{j,i}$ is the rotation axis of $R_{j,i}$ and $\phi_{j,i}$ is the rotation angle about $r_{j,i}$. The estimated rotation axis is r_{est} and ϕ_{est} is the estimated rotation angle about r_{est}. Now we are able to optimize the orientation information concurrently with the calibration. This can be used to improve the orientation data. The statistical optimization is started with the linearly estimated focal length and aspect ratio and the prior knowledge about principal point and skew.

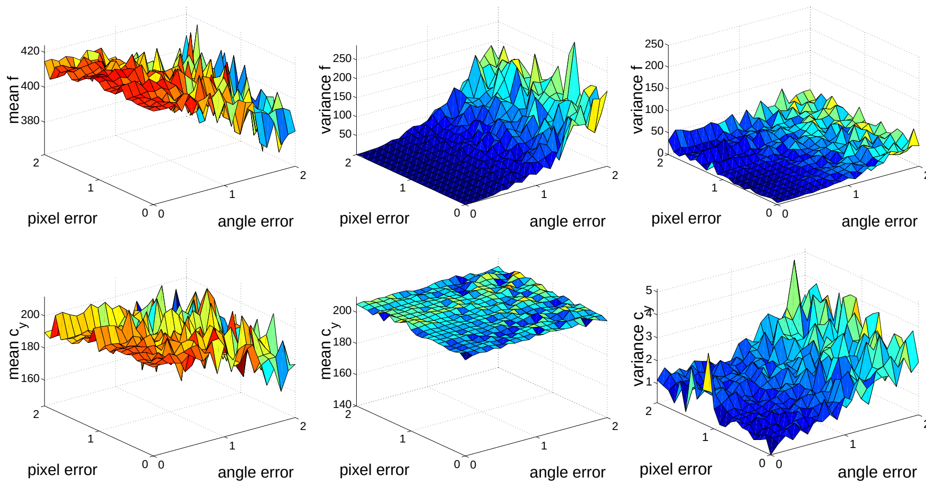

Fig. 1. Noise robustness measurements. Top from left to right: mean of linear estimated focal length f, variance of linear estimated focal length f, variance of MAP estimated focal length f. Bottom from left to right:mean of linear estimated principal point component c_y, mean of MAP estimated c_y, variance of MAP estimation

Evaluation for rotating cameras: To measure the noise robustness of the calibration technique (7) we test the approach with the above described synthetic data. The measurements are shown in figure 1. The principal point varies with about eight percent around the image center. As prior knowledge we use the principal point at the image center. It can be seen that the statistical estimation is more robust if the data is noisy. The variance of the focal length is much better than in the linear case. The estimation of the principal point is much more stable than in the linear case. The results for the other intrinsics are very similar. Since the error of orientation sensors like the InertiaCube[2] from InterSense is normally in the range below one degree, we can rely on the rotation information. The homography estimation can also be estimated with an error of less than 1 pixel for the features' positions in most situations. This shows that the proposed calibration with (7) is robust for most applications.

3.3 Calibration for freely moving cameras

We will investigate how to combine rotational information and the Fundamental matrix $F_{j,i}$ in the general motion case. The Fundamental matrix as opposed to projection matrix is not affected by projective skew, therefore we will use $F_{j,i}$ in the following to calibrate the cameras.

Without loss of generality[3] each Fundamental matrix $F_{j,i}$ can be decomposed to

$$F_{j,i} = [e]_x K_i R_{j,i} K_j^{-1} \Leftrightarrow [e]_x K_i R_{j,i} - F_{j,i} K_j = 0_{3\times3}. \quad (8)$$

This is linear in the intrinsics of camera i and camera j. Please note the relationship to Eq. (2). One can see that (8) is an extension of (2) which contains the unknown camera translation t in the epipole. Equation (8) provides six linear independent equations for the intrinsics of the cameras. So we need five image pairs to compute the camera calibration in case of fully varying intrinsics.

The Fundamental matrices $\tilde{F}_{j,i}$ that have to be estimated from the images are scaled by an arbitrary scale $\rho_{j,i}$

$$\tilde{F}_{j,i} = \rho_{j,i} F_{j,i}. \tag{9}$$

For these estimated Fundamental matrices $\tilde{F}_{j,i}$ (8) is

$$0_{3\times 3} = [e]_x K_i R_{j,i} - \tilde{F}_{j,i} K_j = [e]_x \tilde{K}_i R_{j,i} - F_{j,i} K_j \text{ with } \tilde{K}_i = \rho_{j,i}^{-1} K_i, \tag{10}$$

which is also linear in the intrinsics of camera j and the scaled intrinsics of camera i in conjunction with the scale $\rho_{j,i}^{-1}$. It provides six linear independent equations for the scale and the intrinsics of the cameras. From the counting argument follows that the solution is never unique if no constraints for the scales $\rho_{j,i}^{-1}$ or the intrinsics of the cameras are available.

If we use prior knowledge about the principal point of the cameras we are able to compute the camera calibration from an image triplet (j, i, k) with (10). To get a full camera calibration we use an approach similar to (7).

Evaluation for freely moving cameras: To measure the noise robustness of the proposed calibration for arbitrarily moving cameras we use synthetic data with known noise and ground truth information. Six cameras are positioned on a sphere, observing the same scene as used before in case of purely rotated cameras. The cameras also have a resolution of 512x512 pixels. The noise is the same as above. We calculate the Fundamental matrices $\tilde{F}_{j,i}$ for the image pairs by least squares estimation. The computed Fundamental matrices $\tilde{F}_{j,i}$ are used for the robustness measurements. The results for the case of known principal point (c_x, c_y) and known skew s are shown in figure 2 for the first camera with focal length $f = 415$. The errors and variances for the other images are very similar to these measurements.

It can be seen that for orientation noise of up to 1 degree and pixel noise of up to 1 pixel the calibration is rather stable. The noise sensitivity for this calibration is very similar to the rotational case, but one can see a slightly larger influence of pixel noise for F-estimation.

4 Experiments

In this section we show some experiments on real data for rotating cameras and for simulator scenes for fundamental matrix calibration.

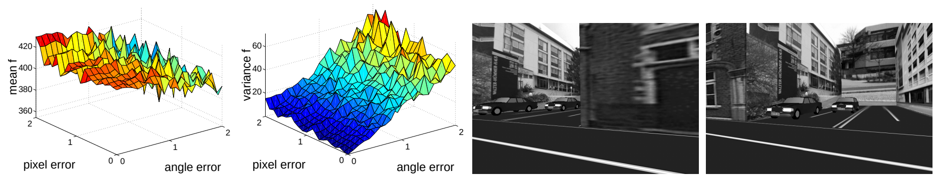

Fig. 2. Noise robustness measurements. Left: mean of estimated focal length f, and variance of estimated focal length f. Right: images from the sequence for Fundamental matrix calibration.

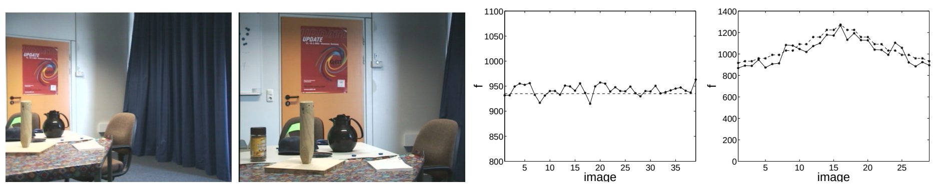

Fig. 3. Left:Images from the zoom-pan sequence for rotation calibration, right: calibration results for constant and varying focal length

4.1 Calibration of rotating camera

We tested the calibration techniques for rotating cameras with a sequence taken by a consumer pan-tilt-zoom camera as used in video conferencing (Sony DV-31). The camera is panning, and zooming during the sequence. Some frames of the sequence are shown in figure 3. The camera rotation is taken from the camera control commands, which means that we used the angles which are sent to the camera. Therefore the rotation error depends on the positioning accuracy of the pan-tilt head which is in the range of below 0.5 degrees for each axis. As reference for the zoom we manually measured the focal length of the different zoom positions to calculate approximate ground truth. The focal length of the camera varied between 875-1232 (in pixel). We also compensated the zoom-dependent radial distortion beforehand. This can be done for the different zooming steps of the camera without knowledge of the correct zoom.

The sequence was processed by tracking feature points with a KLT-tracker [13]. From these tracks we calculated the homographies for the sequence with RANSAC and least-squares-estimation over the inliers. The reprojection error gave a mean pixel error of 0.8 pixel. Calibration estimates for the focal length were computed from triples of images.

Figure 3 shows results for focal length estimation. The dashed line gives the true values, the solid line the estimated values. The left chart shows the estimated focal length (in pixel) for constant focal length f_{true} =940 pixel, the right chart contains a zooming camera. The average relative estimation error is around 3% for fixed zoom and 7% for changing zoom.

We tested the calibration of a moving and rotating camera by using images rendered from a photorealistic car driving simulator. A camera in the car is looking sideways and is panning while the car is driving forward (see figure 2).

The focal length was fixed to 415 (in pixel). From this sequence we estimated the fundamental matrix with RANSAC. The rotation is the given rotation of the ground truth data. However, we were able to detect this situation easily due to the known rotation information. The estimated focal length has a relative error of 3% w.r.t. the true focal length.

5 Conclusions

We introduced a novel linear calibration technique for rotating and moving cameras which uses external orientation information. This orientation information is already avaible in many applications. Furthermore the robustness of this calibration approach was discussed.

The analysis of the linear calibration technique leads to a statistical approach for calibration. We showed that the statistical approach is more robust and can be used for a wide range of applications.

References

1. R. Franklin, "Efficient Rotation of an Object", IEEE Transactions on Computing, 1983.
2. R. Hartley, " In defence of the 8-Point-Algorithm", *ICCV95*
3. R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision" *Cambridge university press, Cambrige, 2000*
4. H. Shum and R.Szeliski, "Panoramic Image Mosaics" *Microsoft Research*, Technical Report MSR-TR-97-23, 1997.
5. L. de Agapito and E. Hayman and I. Reid, "Self-calibration of a rotating camera with varying intrinsic parameters" *British Machine Vision Conference 1998*
6. H. Sawhney, S.Hsu and R. Kumar, "Robust Video Mosaicing through Topology Inference and Local to Global Alignment" *ECCV*, 1998.
7. C. E. Pearson, *Handbook of Applied Mathematics*, S.898, Second Edition, Van Nostrand Reinhold Company, 1983.
8. B. Triggs,"Autocalibration and the Absolute Quadric", Proceedings Conference on Computer Vision and Pattern Recognition, pp. 609-614, Puerto Rico, USA, June 1997.
9. O. D. Faugeras and M. Herbert, "The representation, recognition and locating of 3-D objects," *Intl. J. of Robotics Research*, 1992.
10. G. Stein, "Accurate internal camera calibration using rotation, with analysis of sources of error," *ICCV*, 1995.
11. R. I. Hartley, "Self-calibration from multiple views with a rotating camera" *ECCV*, 1994.
12. S.J. Maybank and O. Faugeras, "A therory of self-calibration of a moving camera," *Int. J. of Computer Vision*, 1992.
13. Bruce D. Lucas and Takeo Kanade, " An Iterative Image Registration Technique with an Application to Stereo Vision," *International Joint Conference on Artificial Intelligence, pages 674-679*, 1981.