Self-Calibration from the Absolute Conic on the Plane at Infinity

Marc Pollefeys and Luc Van Gool
K.U.Leuven - ESAT-MI2
Marc.Pollefeys@esat.kuleuven.ac.be

Abstract

To obtain a Euclidean reconstruction from images the cameras have to be calibrated. In recent years different approaches have been proposed to avoid explicit calibration. In this paper a new method is proposed which is closely related to some of the existing methods. Some interesting relations between the methods are uncovered. The method proposed in this paper shows some clear advantages. Besides some synthetic experiments a metric model is extracted from a video sequence to illustrate the feasibility of the approach.

1 Introduction

Since a few years it has been shown that it is possible to recover constant intrinsic camera parameters from an uncalibrated image sequence. Translating this theoretical possibility into a working implementation proved to be difficult and several methods emerged. Most of them try to recover geometric entities whose projection stays fixed throughout the sequence. These projections are directly related to the camera intrinsic parameters.

For example Faugeras *et al* [?] and later on Zeller *et al* [?] used the absolute conic. In their approach the supporting plane of this conic (i.e. the plane at infinity) is eliminated from the equations. Heyden and Åström [?] and Triggs [?] proposed methods based on the absolute quadric. Other methods were proposed by Hartley [?] and Pollefeys *et al* [?].

The approach followed in this article is based on the absolute conic and the plane at infinity. Some nice relationships between this method and the methods based on the absolute quadric will be uncovered. It seems that all these methods are very similar. Our approach naturally results in a different way of dealing with the scale factors which appear in the equations. This is one of its main advantages.

Heyden and Åström [?] consider these scale factors as additional unknowns resulting in convergence problems with longer image sequences (i.e. more unknown scale factors). Triggs [?] eliminates them by cross-multiplying components, thereby introducing other disadvantages.

2 Basic principles

In this paper a pinhole camera model will be used. The following equation expresses the relation between image points and world points.

$$\lambda_{ik}\mathbf{m}_{ik} = \mathbf{P}_k M_i \tag{1}$$

Here $\mathbf{P}_k$ is a 3×4 camera matrix, $\mathbf{m}_{ik}$ and M_i are column vectors containing the homogeneous coordinates of the image points resp. world points, λ_{ik} expresses the equality up to a scale factor, in the remainder of the text this is replaced by $\simeq$.

2.1 Projective geometry

The projective calibration of a camera setup can be retrieved from correspondences between the images (see for example [**?**, **?**, **?**, **?**]). These projective cameras can be chosen as follow:

$$\begin{aligned} \mathbf{P}_1 &\simeq [\mathbf{I}\,|\,0] \\ \mathbf{P}_k &\simeq [\mathbf{H}_{1k}|\mathbf{e}_{1k}] \end{aligned} \tag{2}$$

with $\mathbf{H}_{1k}$ the homography for some reference plane (the same for all views) and the epipole $\mathbf{e}_{1k}$ the projection in image k of the first camera position[1] ($[0\,0\,0\,1]^\top$). This representation is not unique. In fact the homographies for any plane are valid in equation (**??**). The following transformation can be applied to equation (**??**) without altering $\mathbf{P}_1$:

$$\mathbf{T} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{a} & \sigma \end{bmatrix} \tag{3}$$

with $\mathbf{a} = [a_1\, a_2\, a_3]$ indicating the change in orientation of the reference plane and σ the change in scale. Therefore the homographies for different planes are related as follow:

$$\mathbf{H}'_{1k} \simeq \mathbf{H}_{1k} + \mathbf{e}_{1k}.\mathbf{a} \ . \tag{4}$$

The epipole stays of course unchanged (up to scale).

2.2 Affine geometry

The affine representation corresponds to a particular choice of the possible projective representations. In the affine case the reference plane has to be the plane at infinity:

$$\begin{aligned} \mathbf{P}_1 &\simeq [\mathbf{I}\,|\,\mathbf{0}] \\ \mathbf{P}_k &\simeq [\mathbf{H}_{\infty 1k}|\mathbf{e}_{1k}] \end{aligned} \tag{5}$$

From equation (**??**) it follows that the homography of the plane at infinity, $\mathbf{H}_{\infty 1k}$, can be written as follows:

$$\mathbf{H}_{\infty 1k} \simeq \mathbf{H}_{1k} + \mathbf{e}_{1k}.\mathbf{a} \tag{6}$$

for some $\mathbf{a}$ (which is in general unknown).

[1] Note that the relative scale of e_{1k} compared to $\mathbf{H}_{1k}$ is not free.

2.3 Euclidean geometry

In the Euclidean case the internal camera parameters have to be taken into account. This yields camera projection matrices of the following form:

$$\begin{aligned}\mathbf{P}_{E1} &\simeq \mathbf{K}\,[\,\mathbf{I}\,|\,\mathbf{0}\,]\\ \mathbf{P}_{Ek} &\simeq \mathbf{K}\,[\mathbf{R}_k| - \mathbf{R}_k\mathbf{t}_k]\end{aligned} \tag{7}$$

with $\mathbf{K}$ the calibration matrix, with $\mathbf{R}_k$ and $\mathbf{t}_k$ representing the orientation and position with respect to the first camera. $\mathbf{K}$ is an upper triangular matrix of the following form:

$$\mathbf{K} = \begin{bmatrix} f_x & s & u_x \\ & f_y & u_y \\ & & 1 \end{bmatrix} \tag{8}$$

with f_x and f_y the relative focal lengths, $\mathbf{u} = (u_x, u_y)$ the principal point and s a skew factor (see for example [?]).

Note that it follows from equation (??) that

$$\mathbf{H}_{\infty 1k} \simeq \mathbf{H}_{\infty k}\mathbf{H}_{\infty 1}^{-1} \simeq \mathbf{K}\mathbf{R}_k\mathbf{K}^{-1} \tag{9}$$

with $\mathbf{H}_{\infty k}$ the homography from the plane at infinity to the k^{th} image plane.

3 Self-calibration

Self-calibration methods try to solve for the intrinsic camera parameters which are contained in $\mathbf{K}$ under the assumption that these are constant throughout the sequence. Self-calibration is often based on fixed entities in the image or in the scene. Affine transformations always keep the plane at infinity Π_∞ fixed. This does not mean that every point in Π_∞ is mapped on the same point, but that they are all mapped in the same plane Π_∞. Euclidean transformations keep –in addition to Π_∞– also some special conic in Π_∞ fixed. This conic is called the absolute conic ω. There is also a degenerate quadric which is fixed under all Euclidean transformations, called the absolute quadric Ω[?].

Most self-calibration techniques try to retrieve one of these fixed entities (or its image) in the projective representation. From that point on metric measurements can be carried out. Often everything is transformed from these non-Euclidean frames to the usual Euclidean frame.

3.1 A fixed image for the absolute conic ω

Because the absolute conic ω is fixed under Euclidean transformations also its image ω_k will be fixed if the same camera is used. This is also valid for the dual of that conic ω_k^{-1}, which will be used here for convenience. Starting from $\omega = \mathbf{I}$, this can be proven as follows:

$$\begin{aligned}\omega_1^{-1} &\simeq \mathbf{H}_{\infty 1}\omega^{-1}\mathbf{H}_{\infty 1}^\top \simeq \mathbf{K}\mathbf{K}^\top\\ \omega_k^{-1} &\simeq \mathbf{H}_{\infty k}\omega^{-1}\mathbf{H}_{\infty k}^\top \simeq \mathbf{K}\mathbf{R}_k\mathbf{R}_k^\top\mathbf{K}^\top \simeq \mathbf{K}\mathbf{K}^\top\end{aligned} \tag{10}$$

This fact can be used to calculate the image of the dual of the absolute conic. Because ω lies in Π_∞ the following equations must hold:

$$\omega_k^{-1} \simeq \mathbf{H}_{\infty 1k} \omega_k^{-1} \mathbf{H}_{\infty 1k}^\top \tag{11}$$

Exact equality (not up to scale) can be obtained by scaling $\mathbf{H}_{\infty 1k}$ to obtain $\det \mathbf{H}_{\infty 1k} = 1$. In the case where the affine calibration was already obtained equation (**??**) results in a set of linear equations for the coefficients of ω_k^{-1}. Once ω_k^{-1} is retrieved $\mathbf{K}$ can be obtained from it by Cholesky factorization.

3.2 From projective to metric

If the affine calibration is not known then a_1, a_2, a_3 have to be retrieved in addition. In the general case the following equations are thus obtained (using an explicit scale factor and omitting the indices):

$$\lambda \mathbf{K}\mathbf{K}^\top = [\mathbf{H} + \mathbf{e}\mathbf{a}]\mathbf{K}\mathbf{K}^\top [\mathbf{H} + \mathbf{e}\mathbf{a}]^\top \tag{12}$$

The problem with these equations is that the scale factors are unknown. It is possible to consider these scale-factors as additional unknowns [**?**], but this poses additional problems and will make the scheme unworkable for longer image sequences (one additional scale factor per supplementary image).

It is possible to find an easy way of calculating these scale factors as a function of the 3 affine parameters. This is achieved by taking the determinant of the left- and right-hand side of Equation (**??**):

$$\lambda \det \mathbf{K}\mathbf{K}^\top = \det(\mathbf{H} + \mathbf{e}\mathbf{a}) \det \mathbf{K}\mathbf{K}^\top \det(\mathbf{H} + \mathbf{e}\mathbf{a})^\top \ . \tag{13}$$

Which gives us λ:

$$\lambda = \det(\mathbf{H} + \mathbf{e}\mathbf{a})^2 \tag{14}$$

which can be factorized as follow:

$$\lambda = (|\mathbf{e}\,\mathbf{h}_2\,\mathbf{h}_3|a_1 + |\mathbf{h}_1\,\mathbf{e}\,\mathbf{h}_3|a_2 + |\mathbf{h}_1\,\mathbf{h}_2\,\mathbf{e}|a_3 + |\mathbf{h}_1\,\mathbf{h}_2\,\mathbf{h}_3|)^2 \tag{15}$$

with $\det \mathbf{H} = |\mathbf{h}_1\,\mathbf{h}_2\,\mathbf{h}_3|$.

By filling in $\mathbf{K}$ as in equation (**??**) and λ as in equation (**??**) in equation (**??**) one obtain $5(n-1)$ equations in 8 unknowns (n being the number of images). Therefore at least 3 images are needed to obtain the calibration from correspondences alone.

These equations can be solved through a nonlinear minimization of the following criterion:

$$\sum_{k=2}^{n} \left(\frac{1}{\lambda} \mathbf{H}_{\infty 1k} \mathbf{K}\mathbf{K}^\top \mathbf{H}_{\infty 1k}^\top - \mathbf{K}\mathbf{K}^\top \right) \tag{16}$$

The implementation presented in this paper uses a Levenberg-Marquard algorithm. Results seem to be better when $\mathbf{K}$ is normalized to $||\mathbf{K}||_F = 1$ where $||.||_F$ denotes the Frobenius norm.

The advantage of Equation (**??**) is that it yields a simple closed form equation for λ. In practice during minimization it is more stable to use $\lambda = \frac{||\mathbf{K}\mathbf{K}^\top||_F}{||\mathbf{H}_{\infty 1k}\mathbf{K}\mathbf{K}^\top \mathbf{H}_{\infty 1k}^\top||_F}$. This avoids problems when $\mathbf{H}$ is badly conditioned.

4 Relation with other methods

The method presented in this paper is part of a family of methods which all try to retrieve the absolute entities in projective space (i.e. absolute conic ω or absolute quadric Ω). Once one of these entities is retrieved one can do metric measurements or transform to a metric frame.

Different methods will be discussed here. First the method based on the **Kruppa equations** proposed by Faugeras *et al* [?] and refined by Zeller [?]. In this method the affine parameters are eliminated from the equations. Only the fundamental matrices are needed, not a consistent projective frame for all cameras.

The second method was recently proposed by Heyden and Åström [?]. It is based on the **Kruppa constraints** which relate the dual of the image of the absolute conic to the absolute quadric. It will be shown that these constraints are equivalent with the constraints presented in this paper. Heyden introduced additional unknowns to cope with the scale factors. This strategy only works with a small number of images and is even then suboptimal.

Finally Triggs' method [?] to retrieve the **absolute quadric** is also reviewed. The principles are similar to Heyden's method, but the implementation is different. Scale factors are eliminated by doing cross-multiplication yielding 15 equations (from which only 5 are independent) per camera and doubling the order of the camera intrinsic parameters (i.e. 4^{th} order terms instead of 2^{nd}).

4.1 Kruppa equations

The Kruppa equation can be derived starting from equation (**??**):

$$\mathbf{K}\mathbf{K}^\top \simeq \mathbf{H}_\infty \mathbf{K}\mathbf{K}^\top \mathbf{H}_\infty^\top \tag{17}$$

There is an easy way of eliminating the affine parameters a_1, a_2, a_3 from these equations. They can be multiplied with $[\mathbf{e}]_\times$ to the left and $[\mathbf{e}]_\times^\top$ to the right:

$$[\mathbf{e}]_\times \mathbf{K}\mathbf{K}^\top [\mathbf{e}]_\times^\top \simeq \mathbf{F}\mathbf{K}\mathbf{K}^\top \mathbf{F}^\top \tag{18}$$

since the fundamental matrix $\mathbf{F} = [\mathbf{e}]_\times \mathbf{H}_\infty$. From the 5 equations obtained here only 2 are independent. Scale factors are eliminated by cross-multiplication. The disadvantage of this method is that a consistent supporting plane Π_∞ for ω is only indirectly enforced.

4.2 Kruppa constraints

It can be shown that the Kruppa constraints [?] are equivalent with the constraints used in this paper. Starting from equation (**??**),

$$\lambda \mathbf{K}\mathbf{K}^\top = [\mathbf{H} + \mathbf{e}\mathbf{a}]\mathbf{K}\mathbf{K}^\top [\mathbf{H} + \mathbf{e}\mathbf{a}]^\top \ , \tag{19}$$

and rewriting this equation the Kruppa constraints can easily be obtained (using $\tilde{\mathbf{a}} = \mathbf{aK}$):

$$\begin{aligned} \lambda \mathbf{KK}^\top &= [\mathbf{He}] \begin{bmatrix} \mathbf{KK}^\top & \mathbf{KK}^\top \mathbf{a}^\top \\ \mathbf{aKK}^\top & \mathbf{aKK}^\top \mathbf{a}^\top \end{bmatrix} \begin{bmatrix} \mathbf{H}^\top \\ \mathbf{e}^\top \end{bmatrix} \\ &= \mathbf{P} \begin{bmatrix} \mathbf{KK}^\top & \mathbf{K}\tilde{\mathbf{a}}^\top \\ \tilde{\mathbf{a}}\mathbf{K}^\top & ||\tilde{\mathbf{a}}||^2 \end{bmatrix} \mathbf{P}^\top \ . \end{aligned} \tag{20}$$

Equation (**??**) represents the Kruppa constraints like Heyden presented them in [**?**].

The problem is that in this form it does not seem possible to calculate the scale factors λ (the trick with the determinants does not work when non-square matrices are involved). Therefore Heyden and Åström [**?**] deal with them as additional unknowns. This renders this scheme unworkable for more than a few images because of the many additional unknowns.

4.3 The absolute quadric

Triggs' [**?**] equations are very similar to the Kruppa constraints (Eq. (**??**)) except that he does not assume $\mathbf{P}_1 = [\mathbf{I}\,|\mathbf{0}]$

$$\omega_{im}^{-1} \simeq \mathbf{P}\Omega\mathbf{P}^\top \ . \tag{21}$$

The consequence is that the absolute quadric Ω is not directly related to ω_{im}^{-1} through the parametrisation anymore. Therefore one has to cope with more unknowns. The constraint rank(Ω)=3 which also followed from the parametrisation in the previous methods now has to be enforced explicitly. The advantage is that all views are treated equally where previous methods implicitly favored the first view.

5 Experiments

Experiments have been done on both real and synthetic data. First the synthetic data give some insights in the behavior of the method depending on the number of views and the presence of noise. Then the feasibility of the method will be illustrated with some calibration/reconstruction work done on a real video sequence,

5.1 Simulations

The simulations were carried out on sequences of 3, 6 and 10 views. The scene consisted of 50 points uniformly distributed in a unit sphere with its center at the origin. For the calibration matrix the canonical form $\mathbf{K} = \mathbf{I}$ was chosen. The views were taken from all around the sphere and were all more or less pointing towards the origin. An example of such a sequence can be seen in figure **??**. The scene points were projected into the images. Gaussian noise with figure=../CVPR97/CVPR97seq.eps,width=6cm

Fig. 1. *Example of sequence used for simulations (the views are represented by the image axis and optical axis of the camera in the different positions.)*

an equivalent standard deviations of 0, 0.1, 0.2, 0.5, 1 and 2 pixels for 500×500 images was added to these projections. For every sequence length and noise level ten sequences were generated. The self-calibration method proposed in this paper was carried out on all these sequences. The results for the camera intrinsic parameters were compared with the real values and the RMS error is shown in table **??** for 6 view sequences.

	0.0	0.1	0.2	0.5	1.0	2.0
f_x	1.0000	0.9998	0.9992	0.9997	1.0014	0.9979
	±0.0000	±0.0005	±0.0017	±0.0019	±0.0030	±0.0170
f_y	1.0000	0.9999	0.9991	0.9999	1.0004	0.9993
	±0.0000	±0.0006	±0.0022	±0.0021	±0.0034	±0.0129
$\frac{f_x}{f_y}$	1.0000	1.0000	1.0001	0.9998	1.0010	0.9986
	±0.0000	±0.0003	±0.0009	±0.0007	±0.0022	±0.0052
u_x	0.0000	0.0001	0.0004	0.0019	-0.0005	0.0067
	±0.0000	±0.0002	±0.0009	±0.0017	±0.0029	±0.0108
u_y	0.0000	0.0002	0.0007	0.0014	0.0029	0.0032
	±0.0000	±0.0002	±0.0008	±0.0014	±0.0043	±0.0053
s	0.0000	-0.0001	0.0000	-0.0000	-0.0004	0.0023
	±0.0000	±0.0002	±0.0003	±0.0006	±0.0012	±0.0072

Table 1. Results of synthetic experiment for 6 view sequences

When 6 or 10 views were used the accuracy was very good, even for high amounts of noise (around 1% error for 2 pixels noise). The method almost always converges without problems. For sequences of only 3 views the method gives good results for small amounts of noise, but the error grows when more noise is added. This is due to the fact that in the 3 view case no redundancy is present anymore. In this case the method regularly ends up in an alternative solution.

5.2 A real video sequence

In this paragraph results obtained from a real sequence are presented. The metric qualities of the calibration can be appreciated by looking at the reconstruction. The sequence consists of a university building. These were recorded with a video camera. The images used for self-calibration are shown in figure **??**. The projective camera matrices were obtained following the method described in [**?**]. These camera matrices were upgraded to metric using the self-calibration method described in this paper and then a 3D model was generated using these cameras and a dense correspondence map obtained as in [**?**].

figure=back3d.o3.ps,width=6cm
figure=back3d.o1.ps,width=6cm
figure=back3d.o2.ps,width=2.5cm

Fig. 3. *Orthographic views of the reconstruction (notice parallelism and orthogonality)*

In figure **??** one can see 3 orthographic views of the scene. Parallelism and orthogonality relations clearly have been retrieved. Look for example at the right angles. Figure **??** contains some perspective views of the reconstruction. Because

figure=../vanguard/back3d.2.ps,width=4.66cm figure=../vanguard/back3d.4.ps,width=4cm figure=../vanguard/back3d.3.ps,width=3.66cm

Fig. 4. *Some perspective views of the reconstruction*

the dense correspondence map was only obtained from two images there are some inaccuracies left in the reconstruction. This however has nothing to do with the accuracy of the calibration.

6 Conclusion

In this article a new method was proposed for self-calibration. It is based on the explicit retrieval of the absolute conic and its supporting plane, the plane at infinity. It was shown that this is theoretically equivalent to solving the Kruppa constraints for the absolute quadric. The advantage of our formulation is that it gives a closed formula for the scale factors. Experiments on real and synthetic data illustrated the feasibility and the accuracy of the method.

Further research is done on the combination of our method with other methods to increase robustness. We believe that a stratified approach where the calibration and scene knowledge is gradually built up is the best way to robustly end up with an optimal calibration and reconstruction.

References

1. P. Beardsley, P. Torr and A. Zisserman 3D Model Acquisition from Extended Image Sequences *Proc. ECCV'96*, vol.2, pp.683-695
2. O. Faugeras, What can be seen in three dimensions with an uncalibrated stereo rig, *Proc. ECCV'92*, pp.563-578.
3. O. Faugeras, Q.-T. Luong and S. Maybank. Camera self-calibration: Theory and experiments, *Proc. ECCV'92*, pp.321-334.
4. R. Hartley, Estimation of relative camera positions for uncalibrated cameras, *Proc. ECCV'92*, pp.579-587.
5. R. Hartley, Euclidean reconstruction from uncalibrated views, Applications of invariance in Computer Vision, LNCS 825, Springer-Verlag, 1994.

6. A. Heyden, K. Åström, Euclidean Reconstruction from Constant Intrinsic Parameters *Proc. ICPR'96.*
7. R.Koch, Automatische Oberflächenmodellierung starrer dreidimensionaler Objekte aus stereoskopischen Rundum-Ansichten, PhD thesis, University of Hannover, 1996.
8. M. Pollefeys, L. Van Gool and A. Oosterlinck, The Modulus Constraint: A New Constraint for Self-Calibration, *Proc. ICPR'96*, pp. 349-353.
9. C. Rothwell, G. Csurka and O.D. Faugeras, A comparison of projective reconstruction methods for pairs of views, *Proc. ICCV'95*, pp.932-937.
10. J. G. Semple and G. T. Kneebone, Algebraic Projective Geometry, University Press, Oxford, 1952.
11. M. Spetsakis and Y. Aloimonos, A Multi-frame Approach to Visual Motion Perception *International Journal of Computer Vision*, 6:3, 245-255, 1991.
12. B. Triggs , Autocalibration and the Absolute Quadric, *submitted to CVPR'97.*
13. C. Zeller and O. Faugeras, Camera self-calibration from video sequences: the Kruppa equations revisited. Research Report 2793, INRIA, 1996.