INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Calibration of Digital Amateur Cameras

Magdalena Urbanek — Radu Horaud — Peter Sturm

N° 4214

Juin 2001

THÈME 3

ISSN 0249-6399 ISRN INRIA/RR--4214--FR+ENG

Calibration of Digital Amateur Cameras

Magdalena Urbanek , Radu Horaud , Peter Sturm

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projet Movi

Rapport de recherche n° 4214 — Juin 2001 — 40 pages

Abstract: We introduce a novel outlook on the self-calibration task, by considering images taken by a camera in motion, allowing for zooming and focusing. Apart from the complex relationship between the lens control settings and the intrinsic camera parameters, a prior off-line calibration allows to neglect the setting of focus, and to fix the principal point and aspect ratio throughout distinct views. Thus, the calibration matrix is dependent only on the zoom position. Given a fully calibrated reference view, one has only one parameter to estimate for any other view of the same scene, in order to calibrate it and to be able to perform metric reconstructions. We provide a close-form solution, and validate the reliability of the algorithm with experiments on real images. An important advantage of our method is a reduced - to one - number of critical camera configurations, associated with it. Moreover, we propose a method for computing the epipolar geometry of two views, taken from different positions and with different (spatial) resolutions; the idea is to take an appropriate third view, that is "easy" to match with the other two.

Key-words: self-calibration, calibration, matching, 3D reconstruction, zoom.

Unité de recherche INRIA Rhône-Alpes
655, avenue de l'Europe, 38330 Montbonnot-St-Martin (France)
Téléphone : +33 4 76 61 52 00 — Télécopie +33 4 76 61 52 52

Calibration de Caméras Numériques Amateurs

Résumé : Nous introduisons un nouveau regard sur l'auto-calibration, en considérant des images saisies par une caméra en mouvement, et soumises à des changements de la focale et de la mise au point. A priori, les relations entre ces paramètres et les paramètres internes du modèle sténopé sont complexes. Pourtant, nous montrons qu'en pratique, une pré-calibration permet le plus souvent de négliger les effets de la mise au point et de fixer le point principal et le rapport d'échelle pour des vues différentes. Ainsi la matrice de calibration ne dépend plus que de la distance focale. Etant donné une vue de référence calibrée, il suffit d'estimer un seul paramètre associé à une vue différente de la même scène pour la calibrer et faire des reconstructions métriques. Nous proposons une solution explicite et validons la fiabilité de l'algorithme en l'expérimentant sur des images réelles. Un avantage important de notre méthode est de réduire à un le nombre des mouvements critiques de la caméra. Aussi, nous proposons une méthode pour calculer la géométrie épipolaire de deux vues, prises sous des angles différents et des resolutions spatiales distinctes ; l'idée est de prendre une troisième vue appropriée, telle que l'appariement avec les deux autres soit aisé.

Mots-clés : auto-calibration, calibration, appariement d'images, reconstruction 3D, zoom.

Table of Contents

1 Introduction **5**
1.1 Previous work . . . 5
1.2 Motivation . . . 5
1.3 Contribution . . . 6
1.4 Paper organization . . . 6
1.5 Notations . . . 7

2 Camera modeling **8**
2.1 The model . . . 8
2.2 Off-the-shelf digital camera . . . 8

3 Off-line stability study over calibration **8**
3.1 A way to calibrate . . . 9
3.2 Optical distortion . . . 9
3.3 Experiments . . . 9
3.4 Dependencies . . . 11
3.5 Reliability statistics . . . 14
3.5.1 Experimental data . . . 16
3.5.2 Is the distribution of α, u_0, v_0 Gaussian? . . . 16
3.5.3 Statistically estimated variations . . . 19
3.6 Final results to be used in self-calibration . . . 20
3.7 What about another digital camera? . . . 21

4 Self-calibration **21**
4.1 Kruppa's equations . . . 21
4.2 Outline of the algorithm . . . 24
4.3 Critical motions . . . 24

5 Matching **25**
5.1 Difficulties . . . 26
5.2 Our method . . . 26
5.3 Outline of the algorithm . . . 30

6 Experiments **30**
6.1 Image pairs . . . 30
6.2 Sequence of images . . . 31
6.3 Discussion . . . 31

7 3D reconstruction **34**

8 Conclusion **34**

A Link between images of the absolute conic and the fundamental matrix 38

B Affine transformation between two views taken from the same camera position but with different resolutions 39

1 Introduction

The problem of recovering the Euclidean structure of a scene is strongly associated with the estimation of the camera internal parameters, i.e. calibration. When no calibration knowledge provided, one can reconstruct only a projective model of the scene [6, 11].

1.1 Previous work

The most basic solution to compute the internal parameters employs a calibration grid or planes, and performs an off-line calibration. However the restriction of keeping an identical camera state (including zooming and focusing) while shooting subsequent images can hardly be fulfilled in practice.

Another idea is to self-calibrate an entire sequence. Existing approaches follow several directions. One is to assume invariance of unknown intrinsic parameters throughout distinct views [13, 18, 1, 9], thus not to allow for zooming/focusing, which is quite a strong constraint. Given a stereo pair of an arbitrary scene, one cannot vary but *the magnification parameter* (we use that term, to avoid confusion of associating different meanings to *the focal length*, in vision and optics), while having the other ones known [8]. Other methods [3, 14] allow the retrieval of varying magnification parameter and fixed principal point. Furthermore, if provided with at least 9 views, it is possible to fix only one camera internal parameter and let the other ones vary [14, 12].

In reality, such a general calibration problem cannot be solved reliably. On the other hand, one can quite easily provide some prior information, which simplifies the task. Our approach belongs to such a group of techniques.

1.2 Motivation

All considered cases of self-calibration, which allow magnification parameter variation throughout distinct views, suffer from a significant number of critical camera configurations [17]. It is therefore much "safer" not to change the camera settings.

Let us combine one fully calibrated image (*the reference image*) with an uncalibrated one, taken from a different viewpoint. Then, one has only one magnification parameter to estimate. What about the other intrinsic parameters? The complex relationship between calibration and camera lens control settings [19] does not allow straight-forward simplifications.

To summarize, we are interested in the following issues:

- Are there any conditions that enable the use of a priori knowledge of the intrinsic parameters?
- Can one allow for zooming/focusing, while still maintain a small family of critical situations?
- What can be done with stereo pairs, if one camera/view is fully calibrated?

1.3 Contribution

We combine off-line and on-line methods in order to calibrate a digital camera with a zoom lens and auto-focus.

We introduce a novel outlook on the self-calibration problem, by reducing to one the number of intrinsic parameters to be estimated. We provide a close-form solution for the method. Also, one has to account for only a single family of critical camera configurations [17].

By studying the behaviour of the camera intrinsic parameters as a function of variable zoom and focus, we derive approximate values for the aspect ratio, the principal point, and the magnification parameter for the boundary zoom positions. We verify the reliability of any of the approximations with statistical tests. We identify a small influence of focus upon calibration, which becomes negligible for settings larger than 2.5m. We conclude, that once a camera is calibrated for a known zoom setting, one can re-use those values any time that zoom is set. Therefore, we recommend employing minimally or maximally zoomed-in images as the reference ones, since those zoom settings can be reliably reproduced.

Furthermore, we simplify the computation of the epipolar geometry for stereo images of different resolutions, omitting a direct matching between them. The problem of matching two images of different zoom and viewpoint is therefore decomposed into two simpler matching problems: a wide baseline matching with the same zoom [2], and matching images with different zoom, shot from the same viewpoint [5].

The proposed method of "combined calibration" estimates the intrinsic parameters with even 2%-accuracy, from real images, leading to a reliable Euclidean reconstruction.

1.4 Paper organization

Section 2 defines the camera model and discusses properties of off-the-shelf digital cameras. A study on dependencies of the camera lens control settings on the calibration issue is presented in **Section 3**. **Section 4** encloses our self-calibration algorithm, along with a study of its critical cases. Further, in **Section 5** the problem of matching images with different spatial resolutions is addressed. Experiments on self-calibration of real images are presented in **Section 6**, followed by examples of a successful reconstruction of a real scene (**Section 7**), and the final conclusion (**Section 8**).

1.5 Notations

We distinguish different elements of vector space with different fonts. Also, we try to keep denoting exact space-objects with the same letters, throughout the whole paper.

P	camera projection matrix
F	fundamental matrix
K	calibration matrix
A	affine transformation matrix
R	camera rotation matrix
$\mathbf{t}$	camera translation vector
$\mathbf{M}$	world point
$\mathbf{m}, \mathbf{m}'$	points on images
$\mathbf{e}, \mathbf{e}'$	epipoles
Ω	absolute conic
ω	image of the absolute conic (3×3 matrix)
$\mathbf{e}_i$	the i-th vector from the canonical basis
$[\mathbf{v}]_\times$	antisymmetric matrix associated with vector $\mathbf{v}$
$\sim$	equality up to scale
B^T	transpose of a matrix B
B^{-1}	inverse of a matrix B
$\mathsf{B}^{-\mathsf{T}}$	transposed inverse of a matrix B

2 Camera modeling

2.1 The model

We assume the perspective camera model with the projection matrix of the form:

$$\mathsf{P} = \mathsf{K}\begin{pmatrix}\mathsf{R} & \mathbf{t}\end{pmatrix} \tag{1}$$

where R and $\mathbf{t}$ represent the orientation and the position of the camera with respect to the world coordinate system, and K is the calibration matrix:

$$\mathsf{K} = \begin{pmatrix} k\alpha & 0 & u_0 \\ 0 & \alpha & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

with the principal point (u_0, v_0), the magnification parameter α and the aspect ratio k. We assume a zero-skew.

A scene point $\mathbf{M}$ is projected onto the image onto a point $\mathbf{m}$ via $\mathbf{m} = \mathsf{P}\mathbf{M}$.

2.2 Off-the-shelf digital camera

Most often one is provided with digital cameras, which allow mechanical setting of both zoom and focus. One can specify the area of interest (and thus, its depth on the image) and focus on chosen features within the area.

We have worked with the Olympus Camedia C-2500L digital camera. It provides both auto-focus and manual-focus with discretized values from 0.3m until 15m and ∞ to be set. The zoom, on the contrary, has a continuous range and a manual drive, which makes the reproducibility of different settings difficult (with notable exceptions for the minimal and the maximal zooms).

Each $(zoom, focus)$ setting corresponds to a physical configuration of lenses, inside the camera. Since their functional dependencies are complex, we cannot specify the exact camera state, which makes the estimation of camera internal parameters difficult. When using auto-focus, the only camera settings that we are able to reproduce (and to expect the same calibration results, for an arbitrary image, taken with the same settings) are: $(zoom, focus) = (zoom_{min}, \infty)$ and $(zoom, focus) = (zoom_{max}, \infty)$.

The question is how do the entries of the calibration matrix K change with variations of zoom and focus. Experiments described in the following section suggest conditions, under which the internal camera calibration can be assumed invariant, for different $(zoom, focus)$ settings.

3 Off-line stability study over calibration

We study the stability of the camera internal parameters, under change in the camera mechanical settings, zoom and focus. We point out the parameters that do not vary much,

and can be assumed invariant. We find a small influence of focus on calibration, if the camera is far enough from the scene. Finally, we provide calibration knowledge for particular zoom settings, which is to be used a priori, in self-calibration.

3.1 A way to calibrate

We extract the calibration matrix K from the projection matrix P, estimated from correspondences between non-coplanar 3D points and their 2D images.

The form of $\mathsf{P} = (\bar{\mathsf{P}} \quad \mathbf{p})$ and (1) imply: $\bar{\mathsf{P}} = \mathsf{KR}$. Since

$$\bar{\mathsf{P}}\bar{\mathsf{P}}^{\mathsf{T}} = \mathsf{K}\underbrace{\mathsf{R}\mathsf{R}^{\mathsf{T}}}_{\mathsf{I}}\mathsf{K}^{\mathsf{T}} = \mathsf{K}\mathsf{K}^{\mathsf{T}}$$

we can simply obtain K from the Cholesky decomposition of $\bar{\mathsf{P}}\bar{\mathsf{P}}^{\mathsf{T}}$.

In order to estimate P, we run a non-linear algorithm, which minimizes the reprojection error

$$C = \sum_{i=1}^{n}(u_i - u_{mi})^2 + (v_i - v_{mi})^2 \tag{2}$$

of n image points (u_i, v_i) and reprojections (u_{mi}, v_{mi}) of the corresponding 3D points $\mathbf{M}_i$.

3.2 Optical distortion

Since imperfect camera lenses give rise to non-perspective image distortion, it is often necessary to optimize (2) using additional distortion parameters. In some cases, this extended projection model causes over-parameterization, resulting in instabilities in the estimation of all intrinsic parameters.

Based on the observation, that the bigger the zoom used, the less distortion is present in the image, we can point out experimentally a "critical" zoom, for which the *estimated* distortion coefficient does not decrease with the increase of zoom. Therefore, we omit the distortion parameters in the optimization, if a zoom is bigger than the "critical" one.

We only estimate the first term D_r of the radial distortion, which proved sufficient to provide reliable results. Overall, the employed calibration method is described in [4].

3.3 Experiments

We stepped the lens through the full range of its control parameters - zoom and focus - and performed a full camera calibration, at each step (images of a calibration grid were considered). To ensure the stability of calibration, we considered only images with a sufficiently large number of control points clearly visible on an image of the grid. (The focused-in object can be neither too close nor too far away from the camera.)

We used manual focusing. For each focus value, the zoom was examined in different positions within its range of continuity. For each $(zoom, focus)$ setting, we made several

Focus[m]	k[1]	α[pix]	u_0[pix]	v_0[pix]	D_r[1]
1	0.9993	700	321	268	-0.2393
	0.9991	698	321	267	-0.2423
	0.9999	700	317	267	-0.2363
1.2	0.9992	695	314	277	-0.2598
	0.9996	702	320	269	-0.2405
	0.9997	728	294	238	-0.1468
1.5	0.9998	731	316	232	-0.1574
	1.0007	710	325	269	-0.2469
	0.9998	723	318	234	-0.1601
2	1.0007	736	295	269	-0.1523
	1.0002	699	319	274	-0.2970
2.5	1.0001	722	318	268	-0.2207

Table 1: Calibration results: the *minimal* zoom and varying focus.

Focus[m]	k[1]	α[pix]	u_0[pix]	v_0[pix]	D_r[1]
1	0.9996	921	316	268	-0.0976
	0.9994	920	318	268	-0.0945
	0.9992	918	319	269	-0.1037
1.2	1.0010	1133	320	266	0.0445
	1.0009	1122	317	274	0.0075
	1.0008	1128	318	271	0.0218
1.5	1.0015	1384	310	297	0
	1.0005	1386	291	296	0.0294
	1.0013	1391	320	292	0
2	1.0020	1749	312	313	0
	1.0020	1740	311	310	0
	1.0013	1745	289	303	0
2.5	1.0030	1969	301	324	0.1470
	1.0008	1944	255	316	0.0306
3.5	1.0024	1959	314	356	0.0295
	1.0012	1965	290	336	0
5	1.0016	1999	301	344	0.0182

Table 2: Calibration results: the *maximal* zoom and varying focus.

images with slightly different orientations of the calibration grid. The distance camera-grid was kept identical to the value of the set focus.

We considered focus values between 1m and 5m. The images were of size 640×512 pixels. Some of the obtained estimations[1] of the internal camera parameters are listed - separately

[1] An additional numerous data used in those experiments is described in Section 3.5.1.

Focus [m]	Zoom	k[1]	α[pix]	u_0[pix]	v_0[pix]	D_r[1]
1.2	$z_1 > min$	0.9997	823	314	267	-0.1615
		0.9997	817	319	270	-0.1765
		0.9992	816	317	271	-0.1907
1.5	$z_2 > min$	1.0006	928	318	272	-0.1259
		1.0008	959	312	251	-0.0288
		1.0010	931	317	274	-0.1311
1.5	$z_3 > z_2$	1.0011	1078	313	268	0.0014
2	$z_4 > min$	1.0003	1044	318	273	-0.0678
		1.0007	1047	317	267	-0.0303
		1.0006	1042	316	270	-0.0629
2.5	$z_5 > min$	1.0001	838	312	270	-0.1661
		1.0005	833	323	278	-0.1982
2.5	$z_6 > z_5$	1.0004	1034	317	269	-0.0090
2.5	$z_7 > z_6$	1.0001	1233	316	276	0.0105
		1.0002	1238	316	274	0.0641
3.5	$z_8 > min$	1.0005	997	314	270	-0.0319
5	$z_9 > min$	1.0011	1307	316	300	-0.0533
		1.0013	1353	309	302	0.0765

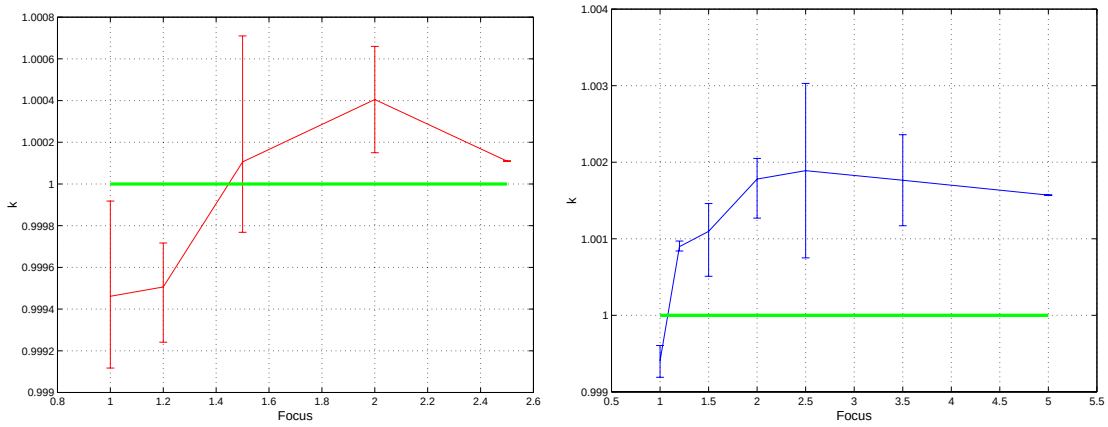
Table 3: Calibration results: varying zoom and focus.

- for the minimal zoom (Table 1), for the maximal zoom (Table 2), for varying zoom (Table 3).

3.4 Dependencies

What information can be extracted from Tables 1, 2 and 3 ?

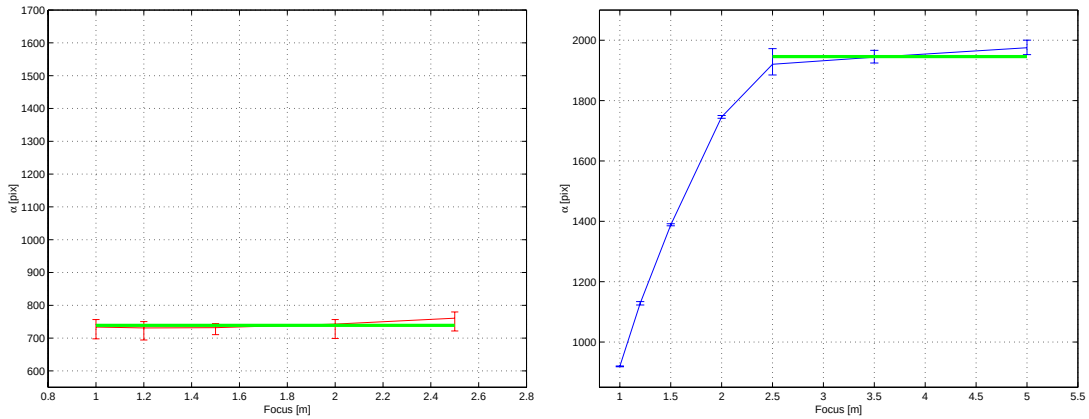
- **Aspect ratio (k).** It is close to unity. The equality $k = 1$ is valid for any $(zoom, focus)$ setting, with a relative error 0.3% (see Figure 1).
- **Magnification parameter (α).** For the minimal zoom, α stays constant relative to focusing. For the maximal zoom, the same is observed as soon as the distance camera-object is bigger than 2.5m (see Figure 2). Hence, for a chosen zoom, it is possible to represent the relevant α with a single value (e.g. its mean): $\alpha_{min} = 739$ (with 4%-relative error) for $(zoom, focus) = (zoom_{min}, focus \geq 0.3\text{m})$, and $\alpha_{max} = 1946$ (with 2%-relative error) for $(zoom, focus) = (zoom_{max}, focus \geq 2.5\text{m})$.
- **Principal point** (u_0, v_0)**.** Overall, it concentrates near the image centre (see Figures 3(c) and 4(c)). Since in general, the exact position of the principal point does not have big impact upon the quality of reconstruction, it is possible to employ approximate values: $u_0 = 311$ (with 7%-relative error), $v_0 = 280$ (with 15%-relative error).



(a) MIN zoom

(b) MAX zoom

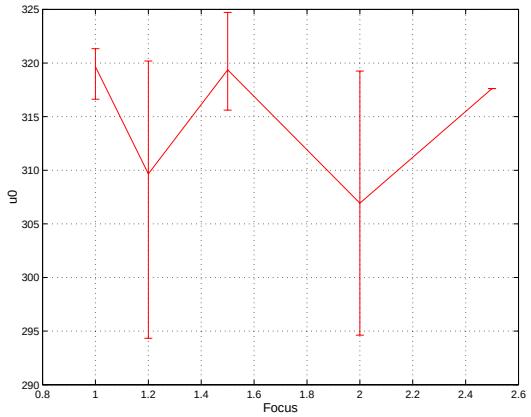
Figure 1: The aspect ratio k vs. focus: k keeps close to 1.



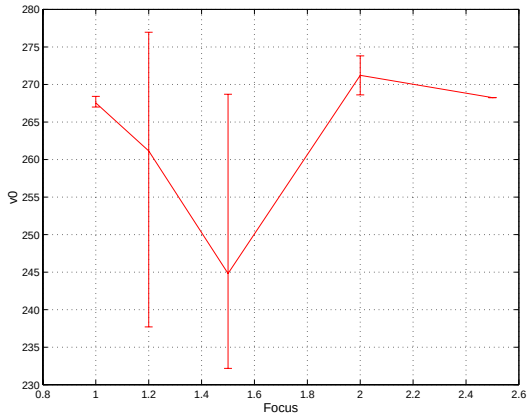
(a) MIN zoom; $\alpha_{min} = 739$.

(b) MAX zoom; approximation by $\alpha_{max} = 1946$ only if $focus \geq 2.5$m.

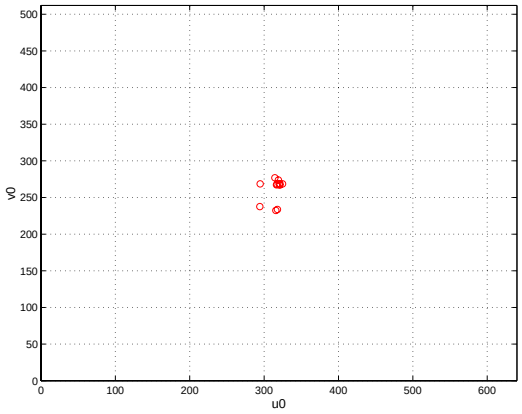
Figure 2: The magnification parameter α vs. focus: α can be approximated by its mean $(\alpha_{min}/\alpha_{max})$.



(a) Variations of u_0



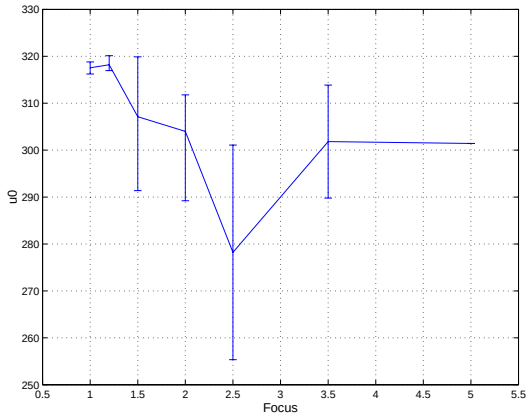
(b) Variations of v_0



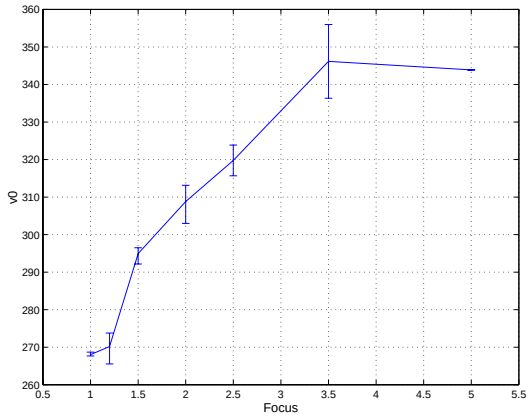
(c) (u_0, v_0)

Figure 3: The principal point (u_0, v_0) vs. focus - MIN zoom.

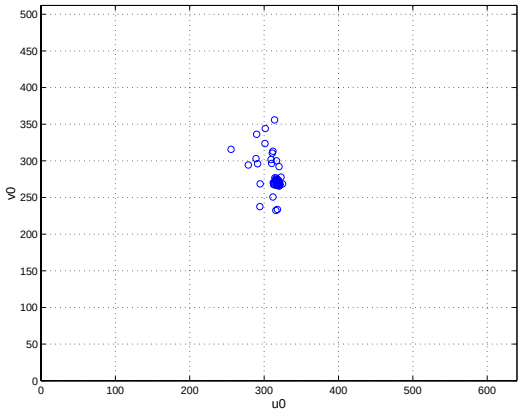
- **Auto-focusing.** For a fixed zoom, the setting of focus does not influence calibration significantly. We can use auto-focusing, and still be capable to employ calibration results for the examined zooms. We only have to keep in mind the requirement concerning the maximal zoom: the distance camera-scene has to be larger than 2.5m.



(a) Variations of u_0



(b) Variations of v_0



(c) (u_0, v_0)

Figure 4: The principal point (u_0, v_0) vs. focus - MAX zoom.

3.5 Reliability statistics

We employed a statistical method to find out the risk of assuming invariance of some of the intrinsic camera parameters - $\alpha_{min}, \alpha_{max}, u_0, v_0$.

For a variable p (meaning anyone out of the intrinsic parameters), we aim to be able to represent it by a single value of its mean μ. An estimator $\hat{p}$ of the mean μ has always some

uncertainty Δ. Therefore, our goal is to write:

$$p = \hat{p} \pm \Delta \quad \text{with a factor of risk } \gamma \tag{3}$$

which is equivalent to the likelihood of μ to belong to the area of uncertainty:

$$P(\mu \in [\hat{p} - \Delta, \hat{p} + \Delta]) = 1 - \gamma$$

In order to estimate $\hat{p}$ and Δ, we use a method which requires knowledge of p distribution being Gaussian. As our data -each one of samples of $\alpha_{min}, \alpha_{max}, u_0, v_0$- fulfills the constraint (we verify it later, in Section 3.5.2), we are able to apply the method, description of which follows. For more detailed studies over statistical properties and tests further employed, refer to [16] and [15].

Let us consider a sample of the variable p of elements $p^{(i)}$, $i = 1..N$, and let us assume that p follows the Gaussian distribution law with the mean μ and the standard deviation σ.

The experimental average is the best estimator of p. We use it as an approximation of μ, and denote it with $\hat{p}$:

$$\hat{p} = \frac{1}{N} \sum_{i=1}^{N} p^{(i)} \tag{4}$$

But, how close is μ to $\hat{p}$?

We use the Student's t-statistic to measure the difference between the two means:

$$t = \frac{\hat{p} - \mu}{\hat{\sigma}}$$

where $\hat{\sigma}$ -assumed to be equal to σ[2]- is the experimental standard deviation:

$$\hat{\sigma} = \frac{1}{N\,(N-1)} \sum_{i=1}^{N} (p^{(i)} - \hat{p})^2$$

The fact that t follows the Student's distribution law with $N-1$ degrees of freedom enables us to estimate the uncertainty Δ of the approximation of μ by $\hat{p}$. It is obtained by scaling $\hat{\sigma}$ with the Student's t-coefficient $t_{N,\gamma}$, defined for certain N and γ (γ being a factor of risk, also called *the significance level*):

$$\Delta = \hat{\sigma} \cdot t_{N,\gamma} \tag{5}$$

Altogether, in order to define a variable p in the way (3), one needs to employ equations (4) and (5).

[2] According to a sort of the Student's t-test, which measures the significance of a difference of means, when the two distributions are thought to have the same variance.

3.5.1 Experimental data

We took 48 images with the minimal zoom, and 48 images with the maximal zoom. We used auto-focusing in both cases. Focus was chosen arbitrary up to the constraint of being larger than 2.5 m, when dealing with the maximal zoom. We used the Olympus Camedia C-2500L digital camera.

We calibrated all of the 96 images, obtaining α, k, u_0, v_0 for each one of them.

Our tests concerned α, u_0 and v_0 (since k keeps very close to 1, thus can be assumed constant).

Tests for the magnification parameter α had to be made separately for different zooms, thus α_{min} and α_{max}. The principal point (u_0, v_0), however, was treated altogether for any zoom chosen. Hence, we had 48-element samples to represent separately α_{min} and α_{max}, and 96-element-samples of u_0 and v_0.

3.5.2 Is the distribution of α, u_0, v_0 Gaussian?

For a variable p (i.e. any of $\alpha_{min}, \alpha_{max}, u_0, v_0$), we verify the null hypothesis: "p follows the Gaussian distribution law with the mean μ (being the experimental mean of p) and the standard deviation σ (being the experimental standard deviation of p)".

Two kinds of tests were invoked:

- **Henry's line test** - makes use of a linear approximation of the inverse cumulative density function *icdf* of the Gaussian distribution $N(\mu, \sigma)$ at point $p^{(i)}$:

$$W_i := icdf(\text{experimental probability } P(p \leq p^{(i)})) \quad \approx \quad \frac{p^{(i)} - \mu}{\sigma}$$

  If p is of the Gaussian distribution, pairs $(p^{(i)}, W_i)$ are roughly aligned.

- χ^2 **test** - the hypothesis is accepted if: the difference between the experimental data distribution and the supposed distribution (represented by χ^2 decision statistic) is smaller than *level*, which is the inverse cumulative density function of the χ^2-distribution of $N_B - 1$ degrees of freedom, with a risk γ. (N_B is the number of "bins", that data is sorted into; N_i is the number of events observed in the ith bin; n_i is the number expected according to the supposed distribution.)

$$\chi^2 = \sum_{i=1}^{N_B} \frac{(N_i - n_i)^2}{n_i} \quad <(?) \quad level$$

  Also, it is possible to visually verify the hypothesis, in an intuitive way, by the comparison between shapes of: the data histogram (based on the defined "bin"-division of the data range) and the hypothesized distribution's density function.

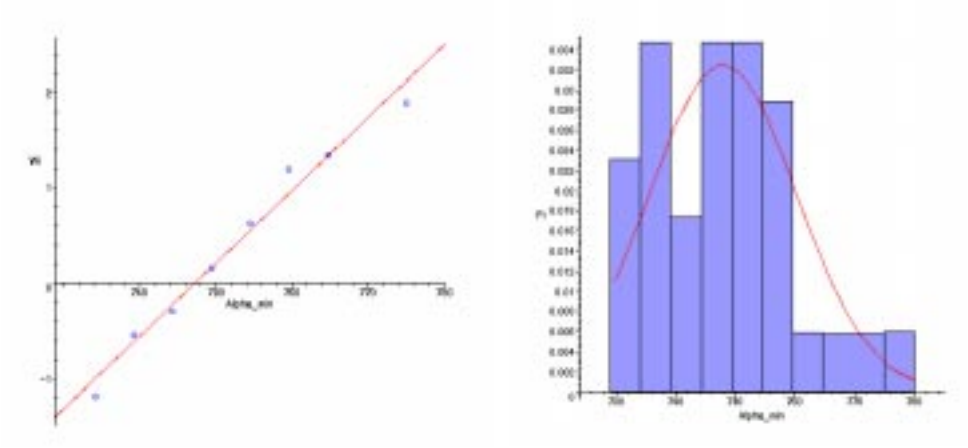
Decision statistic verification. For each parameter, its data sample was divided into such intervals ("bins"), that none of them were empty, and $N_i \geq 3$.

We supposed an equal factor of risk $\gamma = 0.05$ for every examined variable. The decision statistic χ^2 was then compared with *level*, which is different for different number of bins. Results are presented in Table 4.

param	N_B	*level*	χ^2	$\chi^2 < level$?
α_{min}	9	12.59	6.43	YES
α_{max}	13	18.31	11.13	YES
u_0	9	12.59	11.20	YES
v_0	10	14.07	6.90	YES

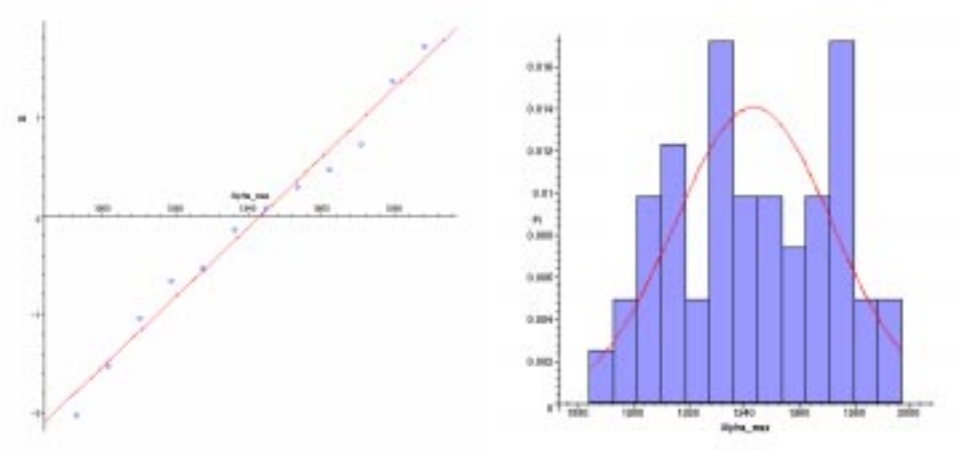
Table 4: Verification of χ^2 related to *level*.

Graphical verification. Obtained plots are presented on Figures 5, 6, 7 and 8. Henry's test is illustrated on the part (a), and χ^2 test on the part (b) of each figure. Each of the parts (b) enables a visual comparison between the data histogram and some Gaussian distribution density function, with the parameters (μ, σ) depending on the data sample $p^{(i)}$: $(\mu, \sigma) = (\hat{p}, \hat{\sigma})$. Bins are defined identically as for the χ^2 decision variable test.



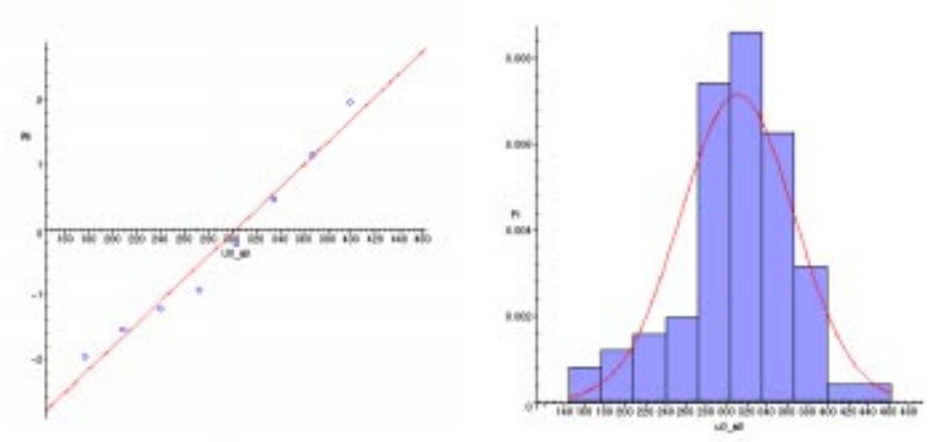
(a) Henry's line (b) χ^2 histogram

Figure 5: α_{min} - Gaussian distribution test



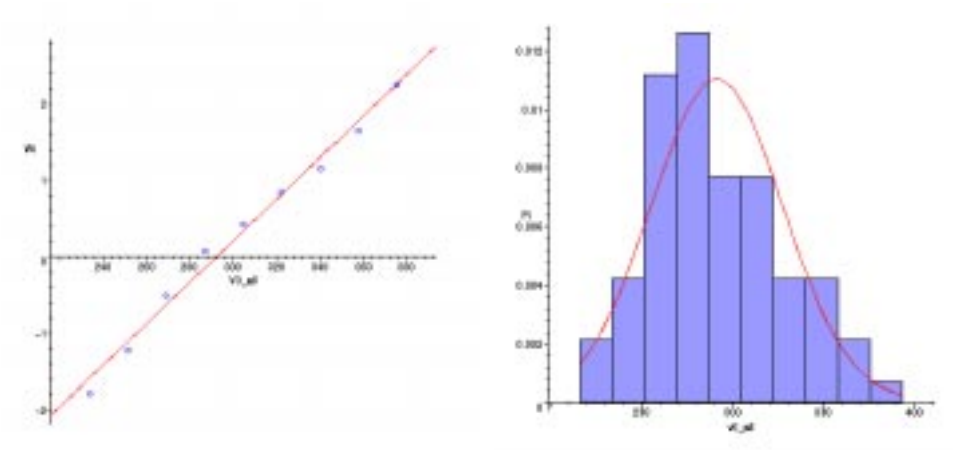
(a) Henry's line (b) χ^2 histogram

Figure 6: α_{max} - Gaussian distribution test



(a) Henry's line (b) χ^2 histogram

Figure 7: u_0 - Gaussian distribution test



(a) Henry's line (b) χ^2 histogram

Figure 8: v_0 - Gaussian distribution test

Interpretation. Parts (a) of Figures 5, 6, 7 and 8 present well aligned points, therefore -according to Henry's test- they vote for adequate internal parameters to follow the Gaussian distribution law. Similarly, parts (b) of the figures show histograms which shapes resemble quite well shapes of adequate Gaussian density functions.

More confirmation is obtained by examination of the decision variable χ^2. Since in all cases it preserves the imposed constraint (see Table 4), one gets ensured that the distribution of each one of the examined internal parameters can be assumed Gaussian, with a possible 5%-error.

3.5.3 Statistically estimated variations

We have now the *right* to refer to (3) to describe statistically variations of $\alpha_{min}, \alpha_{max}, u_0, v_0$. When allowing the factor of risk $\gamma = 0.1$, we have:

$$\alpha_{min} = 739 \pm 34 \ (4.6\%)$$
$$\alpha_{max} = 1946 \pm 47 \ (2.4\%)$$
$$u_0 = 311 \pm 74 \ (23.9\%)$$
$$v_0 = 280 \pm 55 \ (19.0\%)$$

with relative variations given in parentheses.

The obtained results enable us to assume invariance of the camera intrinsic parameters, in specific cases. One can take advantage of that by using computed empirical mean values as a priori known estimations of adequate entries of the calibration matrix. Obviously, one cannot neglect the existence of errors. On the other hand, considered samples with numerous representative data encourage us to rely on the simplified estimations.

3.6 Final results to be used in self-calibration

A view taken with the minimal/maximal zooming. We are provided with calibration matrices of reference: K_{min} for the minimal zoom case (for any focus value), and K_{max} for the maximal zoom case (for focus $\geq$ 2.5m).

A view taken with an arbitrary (unknown) zooming. One is provided with the values of k and (u_0, v_0). Hence, α remains the only calibration parameter to determine.

A summary is given in Table 5.

Zoom	Focus[m]	k[1]	α[pix]	u_0[pix]	v_0[pix]
min	≥ 0.3	1	739	311	280
max	≥ 2.5	1	1946	311	280
?	**?**	1	**?**	311	280

Table 5: Results of off-line calibration (the Olympus Camedia C-2500L digital camera, images of the size 640×512 pixels).

Images of any size. For an image of the size $dim_u \times dim_v$ pixels, the calibration parameters have to be scaled by a factor λ, which is different for different parameters (see Table 6).

param	k	α	u_0	v_0
λ	$\frac{dim_u}{640} \cdot \frac{512}{dim_v}$	$\frac{dim_v}{512}$	$\frac{dim_u}{640}$	$\frac{dim_v}{512}$

Table 6: Scale factor λ.

Then, each of the internal parameters -associated with an image of a *New Image Size*- can be obtained:

$$param_{NewImageSize} := \lambda \cdot param_{640 \times 512}$$

3.7 What about another digital camera?

We ran the off-line calibration process on another digital camera, Olympus Camedia C-1400L. The obtained estimations of the camera internal parameters (see Table 7) are according to the Student's reliability test, with a factor of risk 0.1 (refer to 3.5).

Camera type	α[pix]		k[1]	u_0[pix]	v_0[pix]
	min zoom	max zoom			
Olympus Camedia C-1400L	878 ± 31	1896 ± 38	1 ± 0.003	265 ± 96	246 ± 37
Olympus Camedia C-2500L	739 ± 34	1946 ± 47	1 ± 0.003	311 ± 74	280 ± 55

Table 7: Off-line calibration of different cameras.

One can calibrate off-line any digital camera, obtaining values with divers reliability levels. Nevertheless, the process itself is correct, and results in estimations, which can be easily employed in further processes.

4 Self-calibration

We consider a stereo pair of images: a calibrated reference image and an image taken with an unknown zoom. (In practice, we obtain the calibration for the reference image simply by taking it using the minimal or the maximal zoom, and adopting the according intrinsic parameters, obtained by the off-line calibration.) We are thus provided with calibration matrices: K_{ref}, fully known, for the reference image, and K, defined up to unknown α, for the other image. Due to Kruppa's equations [10], we derive a close-form solution for α. Also, we reveal stereo configurations, for which our self-calibration algorithm fails.

4.1 Kruppa's equations

The idea is to take into account any internal calibration knowledge provided, does not matter whether it is full or partial.

One can separate already known entries of the calibration matrix K (k, u_0, v_0) from the unknown one (α):

$$\mathsf{K} = \begin{pmatrix} k & 0 & u_0 \\ 0 & 1 & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathsf{K}_0 \mathsf{K}_\alpha$$

and make K_0 denote the partial calibration knowledge.

On employing K_{ref} and K_0, the epipolar constraint of the image pair, can be now put into a "semi-calibrated" space, obtaining $\underline{\mathsf{E}}$:

$$\underline{\mathsf{E}} = \mathsf{K}_0{}^{\mathsf{T}} \mathsf{F} \mathsf{K}_{ref} \tag{6}$$

F being the fundamental matrix of the stereo.

Presented dependences support deriving the searched solution for the unknown internal parameter α from a link between the epipolar geometry and the camera internal calibration.

Let us examine that in details.

Solving the camera internal calibration problem (estimating matrix K) is equivalent to finding the image ω of the absolute conic [7], taken by that camera. Since $\omega^{-1} \sim \mathsf{K}\mathsf{K}^\mathsf{T}$, let us denote $\mathsf{C} = \mathsf{K}\mathsf{K}^\mathsf{T}$ for the camera state to be calibrated, and $\mathsf{C}_{ref} = \mathsf{K}_{ref}{\mathsf{K}_{ref}}^\mathsf{T}$ for the camera in the reference position.

The link between the fundamental matrix F and images of the absolute conic can be expressed as follows[3]:

$$\mathsf{F}\mathsf{C}_{ref}\mathsf{F}^\mathsf{T} \sim [\mathbf{e}]_\times \mathsf{C} {[\mathbf{e}]_\times}^\mathsf{T} \tag{7}$$

where $\mathbf{e}$ is the epipole on the image, taken by the uncalibrated camera. The situation is presented at Figure 9.

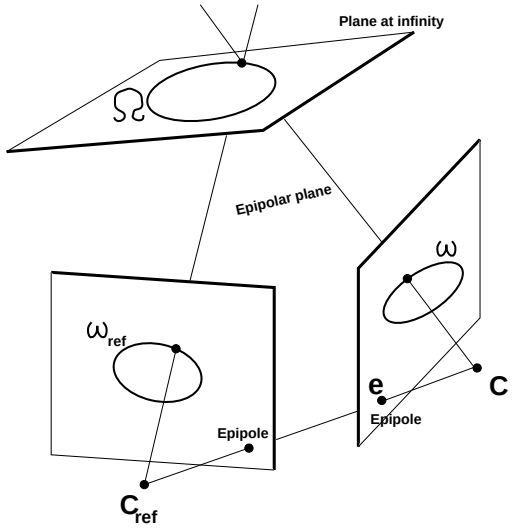


Figure 9: The epipolar geometry of the absolute conic and its images. A reference image and an uncalibrated one considered.

Multiplying (7) from the left by ${\mathsf{K}_0}^\mathsf{T}$, and from the right by K_0 gives:

$${\mathsf{K}_0}^\mathsf{T}\mathsf{F}\mathsf{K}_{ref}{\mathsf{K}_{ref}}^\mathsf{T}\mathsf{F}^\mathsf{T}\mathsf{K}_0 \sim {\mathsf{K}_0}^\mathsf{T}[\mathbf{e}]_\times \mathsf{K}_0 \underbrace{\mathsf{K}_\alpha {\mathsf{K}_\alpha}^\mathsf{T}}_{\mathsf{C}_\alpha} {\mathsf{K}_0}^\mathsf{T} {[\mathbf{e}]_\times}^\mathsf{T} \mathsf{K}_0 \tag{8}$$

[3]The derivation is given in Appendix A

with C_α of the following form:

$$\mathsf{C}_\alpha = \begin{pmatrix} \alpha^2 & 0 & 0 \\ 0 & \alpha^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{9}$$

Applying a property (valid for any matrix A and vector $\mathbf{v}$)

$$\mathsf{A}^{-\mathsf{T}}[\mathbf{v}]_\times \sim [\mathsf{A}\mathbf{v}]_\times \mathsf{A} \tag{10}$$

on a component of (8):

$$\mathsf{K_0}^{\mathsf{T}}[\mathbf{e}]_\times \mathsf{K_0} \sim [\mathsf{K_0}^{-1}\mathbf{e}]_\times \underbrace{\mathsf{K_0}^{-1}\mathsf{K_0}}_{\mathsf{I}}$$

and reminding (6), allows to write (8) as:

$$\underline{\mathsf{E}}\underline{\mathsf{E}}^{\mathsf{T}} \sim [\mathsf{K_0}^{-1}\mathbf{e}]_\times \mathsf{C}_\alpha [\mathsf{K_0}^{-1}\mathbf{e}]_\times{}^{\mathsf{T}} \tag{11}$$

Let us use the Singular Value Decomposition of $\underline{\mathsf{E}}$:

$$\underline{\mathsf{E}} = \mathsf{U}\ \mathrm{diag}(r, s, 0)\ \mathsf{V}^{\mathsf{T}} \tag{12}$$

Introducing (12) into (11), and moving U and U^{T} to the opposite side of the formula, result in:

$$\mathrm{diag}(r, s, 0)\underbrace{\mathsf{V}^{\mathsf{T}}\mathsf{V}}_{\mathsf{I}}\mathrm{diag}(r, s, 0) \sim \mathsf{U}^{\mathsf{T}}[\mathsf{K_0}^{-1}\mathbf{e}]_\times \mathsf{C}_\alpha [\mathsf{K_0}^{-1}\mathbf{e}]_\times \mathsf{U} \tag{13}$$

(remind a property: $[\mathbf{v}]_\times{}^{\mathsf{T}} = -[\mathbf{v}]_\times$, for any vector $\mathbf{v}$).

Due to (10), we can write (13) in the form:

$$\mathrm{diag}(r^2, s^2, 0) \sim [\mathsf{U}^{\mathsf{T}}\mathsf{K_0}^{-1}\mathbf{e}]_\times \mathsf{U}^{\mathsf{T}}\mathsf{C}_\alpha \mathsf{U}[\mathsf{U}^{\mathsf{T}}\mathsf{K_0}^{-1}\mathbf{e}]_\times \tag{14}$$

Let us notice, that

$$[\mathsf{U}^{\mathsf{T}}\mathsf{K_0}^{-1}\mathbf{e}]_\times = \left[\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right]_\times = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

If we denote with $(\mathbf{u}_1\ \ \mathbf{u}_2\ \ \mathbf{u}_3)$ columns of matrix U, (14) writes as follows:

$$\begin{pmatrix} r^2 & 0 \\ 0 & s^2 \end{pmatrix} \sim \begin{pmatrix} \mathbf{u}_2{}^{\mathsf{T}}\mathsf{C}_\alpha \mathbf{u}_2 & -\mathbf{u}_1{}^{\mathsf{T}}\mathsf{C}_\alpha \mathbf{u}_2 \\ -\mathbf{u}_1{}^{\mathsf{T}}\mathsf{C}_\alpha \mathbf{u}_2 & \mathbf{u}_1{}^{\mathsf{T}}\mathsf{C}_\alpha \mathbf{u}_1 \end{pmatrix} \tag{15}$$

Equalities between ratios of coefficients of the matrices in (15) form Kruppa's equations. However, only the following equality can contribute positively to the solution:

$$\frac{r^2}{s^2} = \frac{\mathbf{u}_2{}^\top \mathsf{C}_\alpha \mathbf{u}_2}{\mathbf{u}_1{}^\top \mathsf{C}_\alpha \mathbf{u}_1} \tag{16}$$

The other possible equation ($\mathbf{u}_1{}^\top \mathsf{C}_\alpha \mathbf{u}_2 = 0$) leads always to a solution $\mathsf{C}_\alpha = \mathtt{I}$, and thus $\alpha = 1$.

Remembering the form of C_α (9), one can retrieve from (16) the unknown α, by solving a quadratic equation (since the numerator and denominator of (16) are linear expressions in entries of matrix C_α):

$$\alpha = \sqrt{\frac{s^2 u_{32}^2 - r^2 u_{31}^2}{r^2(u_{11}^2 + u_{21}^2) - s^2(u_{12}^2 + u_{22}^2)}} \tag{17}$$

where u_{ij} are entries of matrix U, and r, s - the singular values, given in (12).

4.2 Outline of the algorithm

Step 0: Perform off-line calibration of the camera, obtaining $\alpha_{min}, \alpha_{max}, k, (u_0, v_0)$ - thus full calibration matrices: K_{min} and K_{max} for reference images, and a calibration matrix K (associated with another image) defined up to α.

Then, for a stereo pair (a reference image and an image of an unknown zooming), given the matching:

Step 1: Compute the fundamental matrix F.

Step 2: Move F to a "semi-calibrated" space, obtaining a new matrix $\underline{\mathsf{E}}$ - see (6).

Step 3: Apply the SVD on $\underline{\mathsf{E}}$ - see (12).

Step 4: Use entries of the matrices obtained in Step 3 to compute the unknown internal parameter α of matrix K - see (17).

4.3 Critical motions

As it has been fully studied in [17], a solution for the unknown magnification parameter is not always uniquely defined. In our case, since we consider to know all intrinsic parameters of one camera, there exist only one family of camera configurations that is critical, which is significantly less than with more general cases of self-calibration.

Let us consider a stereo pair of cameras: $\mathbf{C}_{ref}$ (fully calibrated) and $\mathbf{C}$ (with an unknown α). The algorithm is singular if the centre of camera $\mathbf{C}_{ref}$ lies on the optical axis of camera

C (Figure 10). This kind of configuration is connected with a camera movement (starting from the reference position), that consists of any rotation, followed by a translation along the optical axis of the camera. There is no constraint on orientation of camera $\mathbf{C}_{ref}$.

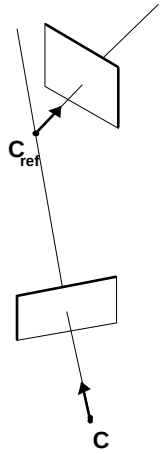


Figure 10: Critical configuration of cameras

Why does it happen?

Let us look for the exact form of the projection of the absolute conic Ω on the image formed by the camera **C**. We can find Ω as the intersection of the two cones (see Figure 11). If there exists more than one possible solution of the image ω of Ω, the problem remains unsolved.

Since the transformed calibration matrix for **C** has the form (9), ω represents a circle with the centre at the origin and its radius $i/\alpha, (i = \sqrt{-1})$. From the knowledge of ω_{ref}, we can form an isotropic cone, which contains the absolute conic. Let us take a plane, that is perpendicular to the principal axis of **C**, and let us place it between the two cameras. This plane intersects the cone in a circle centered at the optical axis of **C**, which is further projected by **C** to its image. However, if to choose another position of the plane, one gets another circle of the intersection, and thus - another image.

In practice, any camera configurations that are close to the critical ones, can cause problems in self-calibration, giving rise to inaccurate results.

5 Matching

We are interested in running our self-calibration algorithm on pairs of images of different spatial resolutions (different magnifications). Being aware of problems concerning matching such views, we propose a way to avoid it, by introducing an additional view, that allows to match the two original ones.

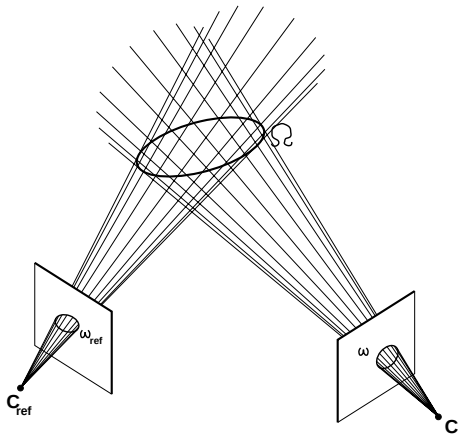


Figure 11: Determination of the absolute conic from its two images

5.1 Difficulties

Existing direct techniques for automatic matching of two images taken from different viewpoints and with different resolutions do not give satisfactory results. Since a big area on the zoomed-in image is to be correlated with a small area on the zoomed-out image, accuracy of computed epipolar lines is weak.

Even when dealing with images very similarly zoomed-in, very few algorithms cope with matching them, if the camera movement between the two views is not small. On the other hand, once one decreases the baseline between cameras (so that it would be appropriate for correlation techniques), the scene reconstruction becomes less reliable.

5.2 Our method

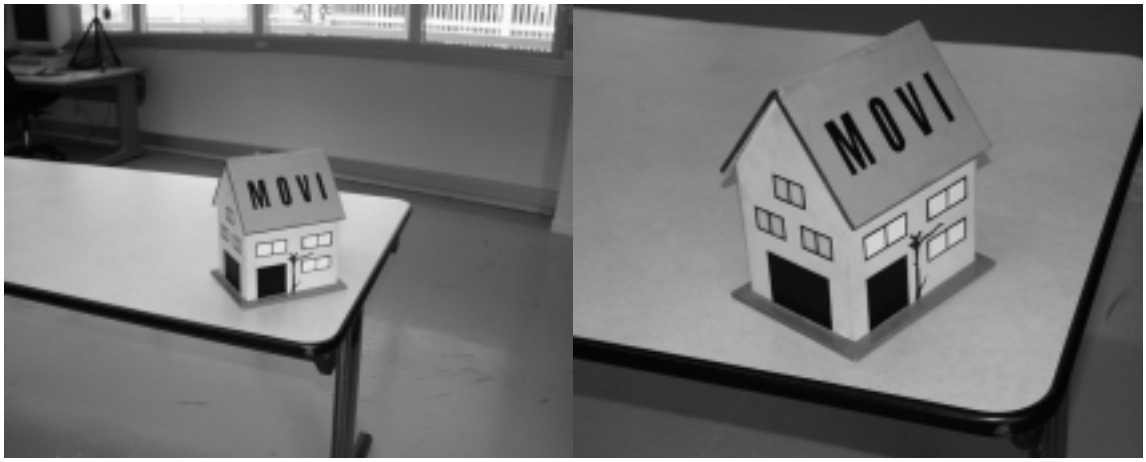
To avoid a manual specification of corresponding points, we combine two techniques:

- matching two images of *the same resolution*, taken from *different viewpoints*
- matching two images of *different resolutions*, taken from *the same viewpoint*

Let us consider an example.

We aim to estimate the epipolar geometry of a stereo pair of images: the reference image I_{ref} and a zoomed-in image I_{zoom} (see Figure 12).

We assume being provided with an additional view I_a, of the same resolution as the reference one, but taken from the same camera position as the zoomed-in one (see Figure 13).



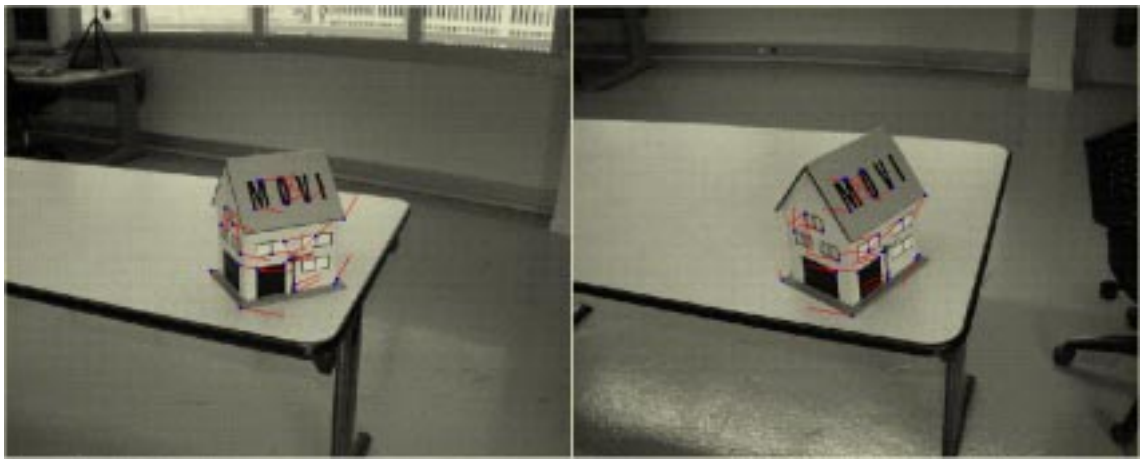
(a) I_{ref} (b) I_{zoom}

Figure 12: Stereo pair of images to be matched.

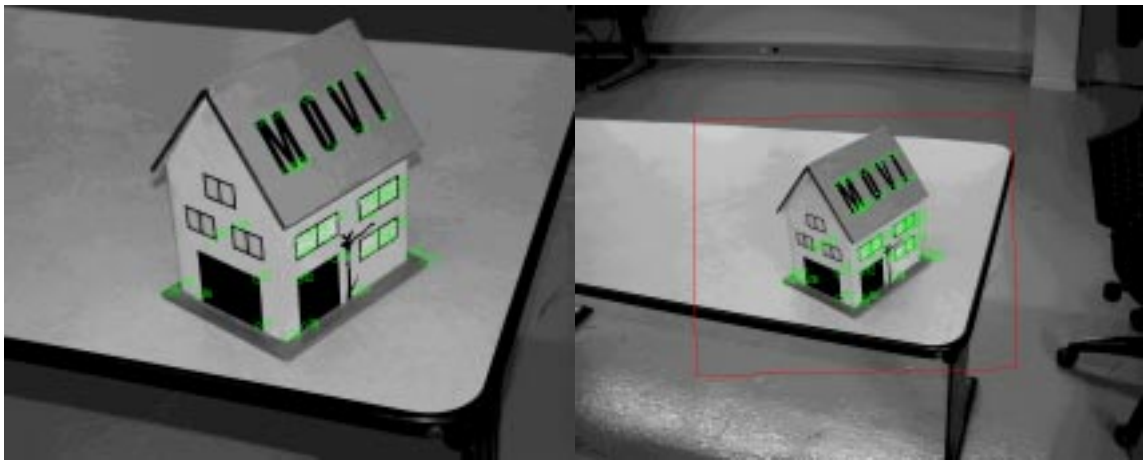
Figure 13: An additional view I_a.

Having performed an automatic matching [2] between I_{ref} and I_a (see Figure 14(a)), we compute the fundamental matrix F_a of that stereo pair (refer to Figure 15). Thus for any image points $\mathbf{m}_{ref}, \mathbf{m}_a$ (related to I_{ref} and I_a respectively):

$$\mathbf{m}_a{}^{\mathsf{T}} \mathsf{F}_a \mathbf{m}_{ref} = 0 \tag{18}$$



(a) Matching I_{ref} (on the left) and I_a (on the right) - corresponding epipolar lines are marked. As the results, the fundamental matrix F_a is computed.



(b) I_{zoom} view. It is matched with an additional view I_a, which results in an affine transformation A between them.

(c) I_a view. Matches with I_{zoom} are shown, along with a rectangle corresponding to the zoomed-in view I_{zoom}.

Figure 14: In order to perform an automatic matching between left-side views - I_{ref} and I_{zoom} - it is necessary to pass through the view on the right (I_a).

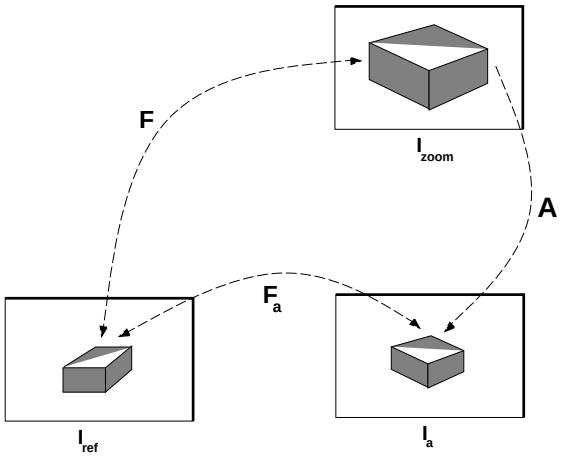


Figure 15: Connections between images: fundamental matrix F_a between I_{ref} and I_a, affine transformation A between I_a and I_{zoom}, fundamental matrix F between I_{ref} and I_{zoom}.

On employing method [5], we match I_a with I_{zoom} (see Figure 14(b-c)), and estimate an affine transformation A between them[4] - for any image points $\mathbf{m}_a, \mathbf{m}_{zoom}$ (related to I_a and I_{zoom} respectively):

$$\mathbf{m}_a = \mathsf{A}\mathbf{m}_{zoom} \tag{19}$$

Now, (18) and (19) let us find out the fundamental matrix F of the stereo pair of interest - I_{ref} and I_{zoom}:

$$(\mathsf{A}\mathbf{m}_{zoom})^{\mathsf{T}}\mathsf{F}_a\mathbf{m}_{ref} = 0$$

$$\mathbf{m}_{zoom}{}^{\mathsf{T}}\underbrace{\mathsf{A}^{\mathsf{T}}\mathsf{F}_a}_{\mathsf{F}}\mathbf{m}_{ref} = 0$$

Thus

$$\mathsf{F} = \mathsf{A}^{\mathsf{T}}\mathsf{F}_a \tag{20}$$

Equation (20) enables us to compute the epipolar geometry between images of different resolutions (I_{ref} and I_{zoom}), without being given matches between them. It is sufficient to specify correspondences between each of those images and a special additional one (I_a), and result in connections written as functions of F_a and A.

Notice, that we do not result in direct matches between I_{ref} and I_{zoom}.

[4]The ability of such an estimation is proven in Appendix B.

The algorithm does not introduce any additional constraints. However, if it is to be put into work with some other tasks (e.g. self-calibration), the positions of the camera have to be appropriately chosen (see 4.3). In practice, the critical case can only happen if the observed object is put very close (ideally: exactly in the same place) to the prior camera position.

5.3 Outline of the algorithm

Step 0: When going to take images, which are to be of different spatial resolutions (I_{ref} and I_{zoom}), perform:

1. Take image I_{ref} from a reference position, with a "reference zoom setting".
2. Move the camera to another position, from which you plan to take the I_{zoom} view (respecting the critical motion sequence of the self-calibration case); BUT: do not change the zoom.
3. With the set camera state (new position, but the "old" zoom), take an additional image I_a.
4. Without moving the camera, change zoom as desired, and take the I_{zoom} view.

Hence, we have three images: I_{ref}, I_a and I_{zoom}.

Step 1: Match I_{ref} with I_a using one of existing methods for wide baseline stereo, e.g. [2].

Step 2: Estimate the fundamental matrix F_a of the stereo $I_{ref} + I_a$ (see (18)).

Step 3: Match I_a with I_{zoom} using method [5].

Step 4: Compute an affine transformation A between views I_a and I_{zoom} (see (19)).

Step 5: Find the epipolar geometry of the stereo $I_{ref} + I_{zoom}$ (see (20)).

6 Experiments

6.1 Image pairs

We took images of an arbitrary object (a toy-sized house), with the minimal and the maximal zoom settings, from a reference viewpoint. Then, from another camera position, we shot a number of images, of variable zooming. We also took images of a calibration grid, every time a view of the house was registered. Hence, we had two sets with corresponding images (of different features), taken with identical camera settings (Figure 16(c) and (f)).

Separately for each photographed object, we combined our images in stereo pairs of a reference image and an image taken with an unknown zoom. Each image of an unknown zoom was put into 2 stereo pairs: with a minimal and with a maximal zoom reference image. We matched the views using the algorithm from Section 5.3 (an exception was done for images of the calibration grid, for which we were provided - by a calibration software - with links 3D-2D between a point on the grid and its image; given two such links associated to the same point 3D, we matched its two images 2D). Having employed results from off-line calibration (the constraint: distance camera-scene $\geq$ 2.5m had been fulfilled), we ran the self-calibration algorithm for each stereo pair, obtaining estimations for α, related to every considered zooming (see Table 8).

Zoom	Calibration	Self-calibration			
		Calibration grid		Arbitrary object	
		ref. Min	ref. Max	ref. Min	ref. Max
$zoom_0$	708	701 (1%)	763 (8%)	731 (3%)	691 (2%)
$zoom_1$	847	814 (4%)	887 (5%)	881 (4%)	829 (2%)
$zoom_2$	1018	960 (6%)	1044 (3%)	1040 (2%)	1012 (1%)
$zoom_3$	1250	1171 (6%)	1273 (2%)	1298 (4%)	1242 (1%)
$zoom_4$	1486	1387 (7%)	1508 (1%)	1592 (7%)	1442 (3%)
$zoom_5$	1729	1621 (6%)	1751 (1%)	1886 (9%)	1779 (3%)
$zoom_6$	1905	1772 (7%)	1933 (1%)	2072 (9%)	1862 (2%)

Table 8: Results of self-calibration performed on stereo pairs of the calibration grid and an arbitrary object, employing both: the minimal and maximal zoom reference images. Relative errors (with respect to calibration) are given to each estimated value of α.

6.2 Sequence of images

We took a sequence of 5 views (*Im1..Im5*) of a building, each one from a different camera position, and with different zooms (Figure 17). *Im1* was taken with the minimal zoom, magnifications of the other images were "spontaneous", and thus - unknown.

We disposed of matching of the entire sequence.

We self-calibrated images *Im2..Im5* with *Im1* as the reference. Then, assuming the obtained values as being given a priori, we made several runs of the algorithm, each time with a different image as the reference. Results are presented in Table 9.

6.3 Discussion

The algorithm recovers the unknown magnification parameter with a high accuracy, provided good quality (precise and numerous) matching of a stereo.

(a) Minimal zoom reference image (camera distance: 1.5m).

(b) Maximal zoom reference image (camera distance: 3.5m).

(c) Image taken with $zoom_3$. From self-calibration: $\alpha = 1171$ (if minimal zoom reference), $\alpha = 1273$ (if maximal zoom reference).

(d) Minimal zoom reference image (camera distance: 1.5m).

(e) Maximal zoom reference image (camera distance: 3.5m).

(f) Image taken with $zoom_3$. From self-calibration: $\alpha = 1298$ (if minimal zoom reference), $\alpha = 1242$ (if maximal zoom reference).

Figure 16: Self-calibration on different stereo pairs - (a)+(c),(b)+(c),(d)+(f),(e)+(f) - estimates α related to the same zoom (used to take images (c) and (f)). Calibration of (c) gives $\alpha = 1250$.

However, there exist cases, where the relative error grows unexpectedly, even over 10%. They show up for stereo pairs, which combine images of significantly different resolutions (e.g. a minimal zoom reference view with a strongly zoomed-in one; and vice-versa).

The reason could be related to the fact, that the considered self-calibration step does not take into account any distortion model, and thus, its results are not always consistent with off-line calibration (see Section 3.2). In particular: a distortion model, considered

(a) *Im1*; minimal zoom reference image (known $\alpha = 706$).

(b) *Im2*; retrieved $\alpha = 1148$.

(c) *Im3*; retrieved $\alpha = 689$.

(d) *Im4*; retrieved $\alpha = 739$.

(e) *Im5*; retrieved $\alpha = 805$.

Figure 17: Sequence of a building. Images (b)-(e) were taken with unknown zooms.

for the minimal zoom reference image, is "forwarded" by self-calibration to α, estimated for the other image of the stereo pair. If that image has been taken with a relatively big zoom setting, the no-distortion model has to be considered then, in order to avoid over-parameterization. For the opposite case: not taking distortion into account for the maximal zoom reference case, implies the no-distortion model for the other image, as well, which is not always correct (zoomed-out images).

A way to cope with the described inconsistence would be to employ a non-linear optimization. The self-calibration step, along with a structure from motion linear method, provide an initial guess for camera parameters (internal and external ones). Then, it would be sufficient to use an extended projection model (including distortion) in a bundle-adjustment setting.

Overall, the experiments validate that our self-calibration method is reliable, for any stereo pair. The unknown magnification parameter can be recovered with even 2%-accuracy,

Reference image	Image to calibrate				
	Im1	*Im2*	*Im3*	*Im4*	*Im5*
Im1	-	1148	689	739	805
Im2	703 (0.4%)	-	678 (2%)	660 (11%)	1108 (38%)
Im3	722 (2%)	1176 (2%)	-	689 (7%)	?
Im4	725 (3%)	1339 (17%)	729 (6%)	-	?
Im5	702 (1%)	847 (26%)	?	?	-

Table 9: Self-calibration of the sequence with the use of different images as the reference. The row with the *Im1*-reference acts as the most reliable one. Relative errors (related to the *Im1*-row) are in parentheses. Question mark (?) signifies cases, where an image pair cannot be matched sufficiently well (in order to estimate the epipolar geometry). Notice: *Im2* (of a significant magnification) is neither a good reference view to self-calibrate zoomed-out ones, nor is well self-calibrated itself, when employing zoomed-out views for the reference.

provided that the stereo pair is composed of images of similar resolutions. Therefore, it is more convenient to use a minimal zoom reference view to self-calibrate zoomed-out images, and a maximal zoom reference one, for more zoomed-in images.

7 3D reconstruction

We applied the described technique on a stereo pair of images of a chimney (Figure 18). The only knowledge we had, was that both images were taken with our camera, and that one of them had the minimal possible resolution. From off-line calibration of the camera, we knew partially the internal camera parameters: $k = 1, \alpha_{min} = 739, u_0 = 311, v_0 = 280$. Self-calibration provided us with an estimation for the unknown magnification parameter for the second setting of the camera: $\alpha = 980$.

Reliability of the obtained reconstruction of the chimney (Figure 19), with correctly retrieved depth and angles between specified planes (Table 10), certifies a high quality of the performed calibration, and thus, capability to recover the Euclidean structure.

8 Conclusion

We have presented a method to simplify the self-calibration process of a zooming camera, based only on information of a boundary (minimal or maximal) zoom, used for taking one of the images. Due to the off-line calibration preprocessing, the on-line self-calibration step has only one parameter to estimate, and thus, only one family of critical motion sequences for cameras to deal with (a situation that is not valid for more complex cases of self-calibration). We provide a close-form solution for the problem and present experiments on real images that validate the stability and reliability of our method.

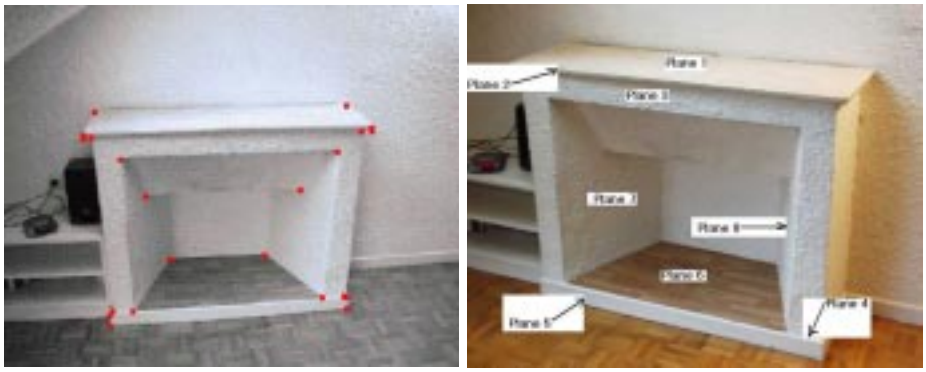


Figure 18: Stereo pair of a chimney (the reference image on the left).

Figure 19: Reconstructed chimney: a front view and a side view.

The proposed combined calibration technique can be easily used in various applications, as quite often one is provided with at least one reference image. The complex problem of dealing with wide, differently zoomed views of a scene, is decomposed into several simpler tasks, which is an important advantage of the presented approach.

Acknowledgements

We would like to thank Frederik Schaffalitzky from the Visual Geometry Group in Oxford, for making accessible a matching software. Also, we thank Adrien Bartoli for some experimental data (image sequence of a building).

Plane	*2*	*3*	*4*	*5*	*6*	*7*	*8*
1	69 (90)	88 (90)	22 (0)	88 (90)	2 (0)	89 (90)	89 (90)
2	-	20 (0)	47 (90)	19 (0)	71 (90)	74 (70)	114 (115)
3	-	-	67 (90)	2 (0)	91 (90)	71 (70)	114 (115)
4	-	-	-	66 (90)	24 (0)	97 (90)	81 (90)
5	-	-	-	-	90 (90)	70 (70)	113 (115)
6	-	-	-	-	-	88 (90)	90 (90)
7	-	-	-	-	-	-	43 (45)

Table 10: Angles (in [deg]) between planes of the chimney: retrieved values, and the real ones (in parentheses). Notice less accurate retrieval of angles at plane *2* or plane *4*; this is caused by difficulties in a precise specification of points defining the planes.

References

[1] M. Armstrong, A. Zisserman, and R. Hartley. Self-calibration from image triplets. In B. Buxton and R. Cipolla, editors, *Proceedings of the 4th European Conference on Computer Vision, Cambridge, England*, volume 1064 of *Lecture Notes in Computer Science*, pages 3–16. Springer-Verlag, April 1996.

[2] A. Baumberg. Reliable feature matching across widely separated views. In *Proceedings of the Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina, USA*, pages 774–781, 2000.

[3] S. Bougnoux. From projective to euclidean space under any practical situation, a criticism of self-calibration. In *Proceedings of the 6th International Conference on Computer Vision, Bombay, India*, pages 790–796, 1998.

[4] E. Boyer. *Reconstruction de surfaces d'objets courbes en vision par ordinateur.* PhD thesis, Institut National Polytechnique de Lorraine, December 1996.

[5] Y. Dufournaud, C. Schmid, and R. Horaud. Matching images with different resolutions. In *Proceedings of the Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina, USA*, June 2000.

[6] O. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In G. Sandini, editor, *Proceedings of the 2nd European Conference on Computer Vision, Santa Margherita Ligure, Italy*, pages 563–578. Springer-Verlag, May 1992.

[7] O. Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint.* Artificial Intelligence. MITP, 1993.

[8] R.I. Hartley. Estimation of relative camera positions for uncalibrated cameras. In G. Sandini, editor, *Proceedings of the 2nd European Conference on Computer Vision, Santa Margherita Ligure, Italy*, pages 579–587. Springer-Verlag, 1992.

[9] R.I. Hartley. Self-calibration from multiple views with a rotating camera. In *Proceedings of the 3rd European Conference on Computer Vision, Stockholm, Sweden*, pages 471–478. Springer-Verlag, May 1994.

[10] R.I. Hartley. Kruppa's equations derived from the fundamental matrix. IEEE *Transactions on Pattern Analysis and Machine Intelligence*, 19(2):133–135, February 1997.

[11] R.I. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. In *Proceedings of the Conference on Computer Vision and Pattern Recognition, Urbana-Champaign, Illinois, USA*, pages 761–764, 1992.

[12] A. Heyden and K. Åström. Minimal conditions on intrinsic parameters for euclidean recons truction. In *Proceedings of the Third Asian Conference on Computer Vision, Hong Kong*, volume II, pages 169–176, January 1998.

[13] E. Malis and R. Cipolla. Multi-view constraints between collineations: Application to self-calibration from unknown planar structures. In D. Vernon, editor, *Proceedings of the 6th European Conference on Computer Vision, Dublin, Ireland*, volume 1843 of *Lecture Notes in Computer Science*, pages 610–624. Springer-Verlag, June 2000.

[14] M. Pollefeys, R. Koch, and L. Van Gool. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. In *Proceedings of the 6th International Conference on Computer Vision, Bombay, India*, pages 90–95, January 1998.

[15] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C - The Art of Scientific Computing*. Cambridge University Press, 2nd edition, 1992.

[16] G. Saporta. *Probabilités, Analyse des Données et Statistique*. Eds Technip, Paris, 1990.

[17] P. Sturm. Critical motion sequences for the self-calibration of cameras and stereo systems with variable focal length. In T. Pridmore and D. Elliman, editors, *Proceedings of the tenth British Machine Vision Conference, Nottingham, England*, pages 63–72. British Machine Vision Association, September 1999.

[18] B. Triggs. Autocalibration from planar scenes. In *Proceedings of the 5th European Conference on Computer Vision, Freiburg, Germany*, 1998.

[19] R.G. Willson and S.A. Shafer. Modeling and calibration of zoom lenses. In A. Gruen and T.S. Huang, editors, *Camera Calibration and Orientation Determination*. Springer-Verlag, 1993.

[20] C. Zeller and O. Faugeras. Camera self-calibration from video sequences: the kruppa equations revisited. Rapport de recherche 2793, INRIA, February 1996.

A Link between images of the absolute conic and the fundamental matrix

Let us consider two views ω and ω' of the absolute conic Ω (Figure 20).

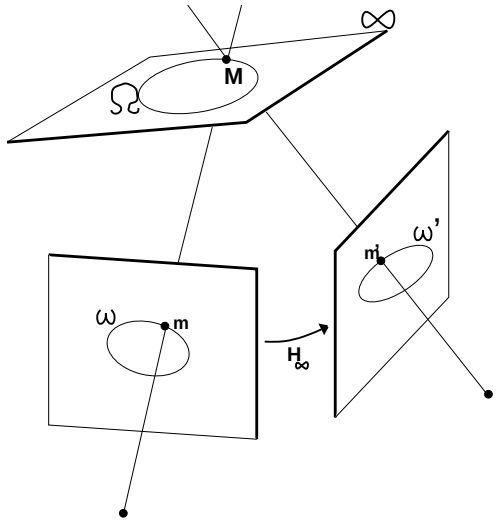


Figure 20: The absolute conic and the collineation between its images.

Images $\mathbf{m}$ and $\mathbf{m}'$ of a point $\mathbf{M}$ of the conic are related by:

$$\mathbf{m}' \sim \mathsf{H}_\infty \mathbf{m} \tag{21}$$

Applying the conic equation to both images

$$\mathbf{m}^\mathsf{T} \omega \mathbf{m} = 0 \qquad\qquad \mathbf{m'}^\mathsf{T} \omega' \mathbf{m}' = 0$$

and connecting them with (21), we get

$$\mathbf{m}^\mathsf{T} (\mathsf{H}_\infty{}^\mathsf{T} \omega' \mathsf{H}_\infty) \mathbf{m} = 0$$

and hence

$$\omega \sim \mathsf{H}_\infty{}^\mathsf{T} \omega' \mathsf{H}_\infty \tag{22}$$

what is equivalent to

$$\mathsf{H}_\infty{}^{-\mathsf{T}} \omega \mathsf{H}_\infty{}^{-1} \sim \omega' \tag{23}$$

For the reason of simplicity of derivations, we move (23) to the dual space:

$$\mathsf{H}_\infty \omega^{-1} {\mathsf{H}_\infty}^{\mathsf{T}} \sim \omega'^{-1} \tag{24}$$

Multiplying both sides of (24) left-hand-side with $[\mathbf{e}']_\times$ and right-hand-side with ${[\mathbf{e}']_\times}^{\mathsf{T}}$ gives

$$([\mathbf{e}']_\times \mathsf{H}_\infty)\omega^{-1}(\mathsf{H}_\infty{}^{\mathsf{T}}[\mathbf{e}']_\times{}^{\mathsf{T}}) \sim [\mathbf{e}']_\times \omega'^{-1} {[\mathbf{e}']_\times}^{\mathsf{T}}$$

Since $\mathsf{F} \sim [\mathbf{e}']_\times \mathsf{H}_\pi$ for an arbitrary plane π, we obtain

$$\mathsf{F}\omega^{-1}\mathsf{F}^{\mathsf{T}} \sim [\mathbf{e}']_\times \omega'^{-1} {[\mathbf{e}']_\times}^{\mathsf{T}} \tag{25}$$

where F is the fundamental matrix that characterizes the two images.

Refer to [20] for a wider study.

B Affine transformation between two views taken from the same camera position but with different resolutions

Let us consider a space point $\mathbf{M}$ projected onto two images - I_1 and I_2 - in points $\mathbf{m}_1$ and $\mathbf{m}_2$ respectively:

$$\mathbf{m}_1 = \mathsf{P}_1\mathbf{M} \qquad\qquad \mathbf{m}_2 = \mathsf{P}_2\mathbf{M}$$

If the images have been taken from the same camera position (Figure 21), the projection matrices - P_1 and P_2 - are:

$$\mathsf{P}_1 = \mathsf{K}_1 \begin{pmatrix} \mathsf{R} & \mathbf{t} \end{pmatrix} = \mathsf{K}_1\mathsf{B}$$

$$\mathsf{P}_2 = \mathsf{K}_2 \begin{pmatrix} \mathsf{R} & \mathbf{t} \end{pmatrix} = \mathsf{K}_2\mathsf{B}$$

where $\mathsf{K}_1, \mathsf{K}_2$ denote calibration matrices, and B is associated to the camera position.

Having combined the above equations, we have $\mathsf{B}\mathbf{M} = {\mathsf{K}_2}^{-1}\mathbf{m}_2$, and thus:

$$\mathbf{m}_1 = \mathsf{K}_1\mathsf{B}\mathbf{M} = \mathsf{K}_1{\mathsf{K}_2}^{-1}\mathbf{m}_2$$

Since an upper-triangular matrix -in particular- represents an affine transformation, both of the matrices K_1 and ${\mathsf{K}_2}^{-1}$ do. Affine transformations form a group, therefore

$$\mathsf{A} = \mathsf{K}_1{\mathsf{K}_2}^{-1}$$

is an affine transformation linking image points of I_1 with the ones of I_2:

$$\mathbf{m}_1 = \mathsf{A}\mathbf{m}_2$$

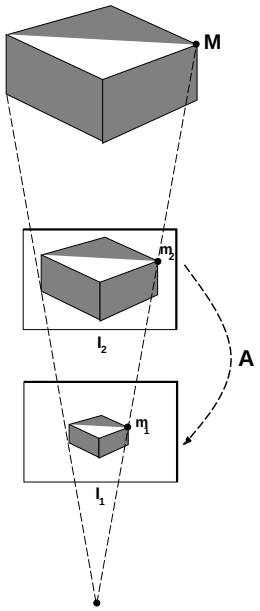


Figure 21: Geometry of two views taken from the same camera position, but with different resolutions. Affine transformation A between the images.

Moreover, if we assume the calibration matrices of forms:

$$\mathsf{K}_i = \begin{pmatrix} k\alpha_i & 0 & u_0 \\ 0 & \alpha_i & v_0 \\ 0 & 0 & 1 \end{pmatrix} \qquad\qquad i = 1, 2$$

the affine transformation A has a special form:

$$\mathsf{A} = \begin{pmatrix} s & 0 & u_0(1-s) \\ 0 & s & v_0(1-s) \\ 0 & 0 & 1 \end{pmatrix}$$

where $s = \frac{\alpha_1}{\alpha_2}$.

Unité de recherche INRIA Rhône-Alpes
655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399