



KUNGL
TEKNISKA
HÖGSKOLAN

CVAP
*Computational Vision &
Active Perception
Laboratory*

Department of Numerical Analysis and Computing Science
TRITA-NA-9503 • ISSN 1101-2250 • ISRN KTH/NA/P-9503-SE • CVAP172

Some Aspects of Zoom-Lens Camera Calibration

Mengxiang Li and Jean-Marc Lavest

Some Aspects of Zoom-Lens Camera Calibration

Mengxiang Li^{*} and Jean-Marc Lavest[†]

^{*}Computational Vision and Active Perception Laboratory(CVAP)
Department of Numerical Analysis and Computing Science
Royal Institute of Technology (KTH), S-100 44 Stockholm, Sweden
mxli@bion.kth.se

[†]Laboratoire des Sciences et Matériaux pour l'Electronique, et d'Automatique (LASMEA)
URA 1793 of the CNRS, Blaise-Pascal University of Clermont-Ferrand
63177 Aubiere cedex. France
lavest@le-eva.univ-bpclermont.fr

February, 1995

^{*}The presented work was performed under the ESPRIT-BRA 7108 *Vision as Process* (VAP-II) project. The support from the Swedish National Board for Industrial and Technical Development, NUTEK, is gratefully acknowledged.

[†]This work was realized during a scientific exchange between CVAP, KTH-Sweden and LASMEA, UBP-France. The financial support from the french *Conseil Régional d'Auvergne* is gratefully acknowledged.

Abstract

Camera calibration has been a studying subject in both photogrammetry and computer vision community for a long time. Most of the previous work has been concerned with algorithms and techniques for geometric properties of CCD camera. Very little attention has been paid on the radiometric properties. Calibration results does not only depend on the mathematical tools used to compute the *intrinsic* and *extrinsic* parameters. A lot of cares have to be taken in the electronic stability of camera and frame grabber, the construction of a 3D reference object and the way calibration points are measured and detected in images. In this paper we address some practical aspects of camera calibration, in particular, of a *zoom lens* system. Through a systematic approach we try to describe all the keys points that have to be checked in order to obtain accurate calibration results. Experimental results on calibrating the KTH head-eye system, which is an active vision system consisting a pair of cameras with motor controlled zooming, focusing and aperture, are presented and analyzed.

1 Introduction

Camera calibration is considered important and useful, if not necessary, in many computer vision problems, e.g., 3D reconstruction from stereo, vergence, fixation, etc.. Basically, there are two problems associated with camera calibration: (i) calibration of the internal parameters of a camera, the so-called *intrinsic* parameters, including optical and mechanical (geometrical) properties of a camera, such as focal length, lens distortion parameters, the intersection point of the optical axis with the image plane – *principal point*, aspect ratio of the pixel array, etc. Sometimes, manufacturers provide partial of these parameters, but they are not accurate enough. Also some of these parameters may vary from time to time, while some of them may be calibrated once for all, depending on the stability of the mechanic and optic construction of the camera; (ii) pose estimation problem of a camera (system) relative to a 3D world reference system, including rotation and translation between these two systems. These are the so-called *extrinsic* parameters. These parameters are not directly related to the camera itself, but the set-up of a camera, which means that they have to be calibrated for each set-up.

Camera calibration has been a studying subject in photogrammetry for more than thirty years, see e.g., [4, 5, 17] and references there in. This problem has also attracted great attention in the computer vision community for the last two decades. There has been a lot research on this topic. It is not the intention to give literature review here, reader is referred to [19, 20] for a more complete review. Most of the previous work on camera calibration has been concerned with algorithms and techniques for calibrating geometric properties of a camera. For CCD cameras, very little attention has been paid on their radiometric properties, except in [6, 3]. Also very little has been presented in literature on *practical aspects of camera calibration*, which is also important if one wants to achieve the best results with a given algorithm.

Zoom-lens system becomes more and more important in vision systems, e.g., for depth reconstruction [8], for active vision [12, 13], due to its flexibility and controllability. In an active vision system, optical parameters, i.e., zoom, focus and aperture, are usually (computer) motor controlled. Zoom is used to acquire images at different resolution, e.g., simulating foveation, focus is used to (automatically) refocus on objects at different distance, or to compute depth [7], while aperture is used to automatically adjust the iris according to lighting condition. Calibration of such a system becomes important, as the pinhole model used for single-lens system can not be applied directly in a zoom-lens system [8]. Previous work on zoom-lens camera calibration can be found in [16]. This paper is not concerned with theoretic or algorithm of camera calibration, rather with some practical and experimental aspects of camera calibration, in particular, for motor controlled zoom-lens system. We first discuss some electronic or radiometric properties of a CCD camera and the problems associated with them in image acquisition. We also propose some measures to deal with these problems. Secondly, we discuss some practical aspects of camera calibration, such as calibration reference object, geometric configuration of the camera set-up, target point detection and precise localization, etc., We also propose some solutions on how to achieve the best results for a given calibration object and an algorithm. Then finally we present some experimental results on the calibration of a head-eye system with motor controlled zoom-lens camera system. The (intrinsic parameter) calibration of such a system implies determining the intrinsic parameters of the system at all the settings of the optical parameters (motors) and the combinations of them. As the resolution of the motor is very high (50,000 steps for each motor), It is not possible to calibrate at each possible setting (and the combination of them). We propose an indexing approach: calibrate each optical parameter at some limited intervals, construct an indexing function and compute the intrinsic parameters of any setting by interpolation. The interpolation function is surface with zoom setting as one dimension and focus setting as another dimension.

In Section 2, some electronic properties of CCD cameras and frame grabber are discussed. Section 3 describes the mathematical models and algorithm used in the experiments for calibration. In Section 4 and 5 we discuss some practical aspects of calibration: calibration reference

object and how to detect image points accurately. Experimental results on calibrating the KTH head-eye system are presented and analyzed in Section 6.

2 Electronic Stability

Even if the geometric and radiometric characteristics of a solid state sensor make it ideally for accurate purpose, electronic and mechanical design of a CCD camera are not usually designed for metrology. Thus, electronic stability of both camera and frame grabber has to be checked before doing calibration experiments. Two possible sub-problems can be distinguished:

- Adjustment of parameters: gain and offset signals of an analogic camera, the Low-Pass-Filter effects, the automatic iris control, to name only a few, can introduce measurement errors if they are not well set and several of them can lead to systematic defaults.
- The synchronization problem between the frame-grabber-clocks and camera-clocks when composite-video signal is sampled to be stored in memory (so-called *line-jitter effect*).

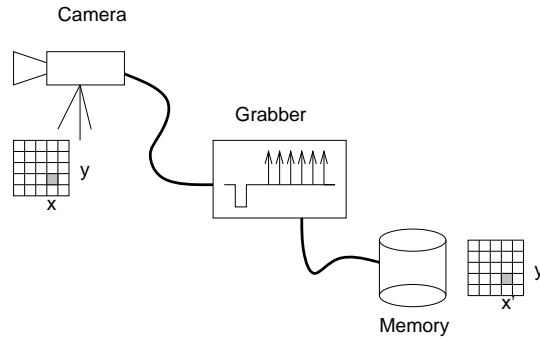


Figure 1: *Image Grabbing Scheme.*

Let us briefly recall the basic scheme of an image acquisition system using an analogic camera (Fig. 1). First of all, the CCD matrix content is read (line by line) and transformed into a video composite signal. The signal includes the top-frame (beginning of an image), top-line (beginning of each line), the black reference and the video information. This analogic signal is sent (via a coaxial cable, several meters long) to the frame grabber. Then inside the frame grabber the analogic signal is sampled, converted to a numerical data and stored in memory. All the transformation scheme pose the relationship problem between the data stored at coordinate (x', y') in memory and the real photonic energy received by the CDD matrix at the same point. In [3], these phenomena have been observed and solutions are proposed to reduce or avoid their effects.

2.1 Electronic Parameters

In the following, we briefly discuss some of the electronics parts which are sensitive for image acquisition.

- Switching of the automatic control of camera gain and iris. These options which are usually included in cameras for visual purpose (surveillance) can introduce a large change in the grey level values of an object according to the background. Furthermore it is really interesting to command the iris position directly from the computer.

- **Black level reference.** This electronic problem is also called *DC-restoration*. Its aim is to maintain the black reference of the composite-video-signal to 0 volt. This device must be able to react quickly, it allows to obtain a perfect grey level correspondence between the video signal and the CCD information but also plays a major role in the frame grabber synchronization.
- **Low Pass Filter (LPF).** A low pass filter is commonly used in the grabbing card, and take place just before the Analogic to Digital Converter (CAN). It is supposed to suppress all high frequency disturbances introduced during data transfer from the camera to the grabbing system. But this filter, usually of high order to react quickly, induces undesirable effects for people who want to make accurate measurement: *a non symmetric behavior*. For example, step edges generate frequencies close to the cut-off of the filter and lead to well known shapes of high order filters close to instability. As shown in Fig. 2, the LPF introduces ringing effects for each edge or pulse detected in the image and induces measurement errors such as edge shifting. The LPF effects can be easily avoid by removing the filter inside the grabbing card [3, 2, 6].

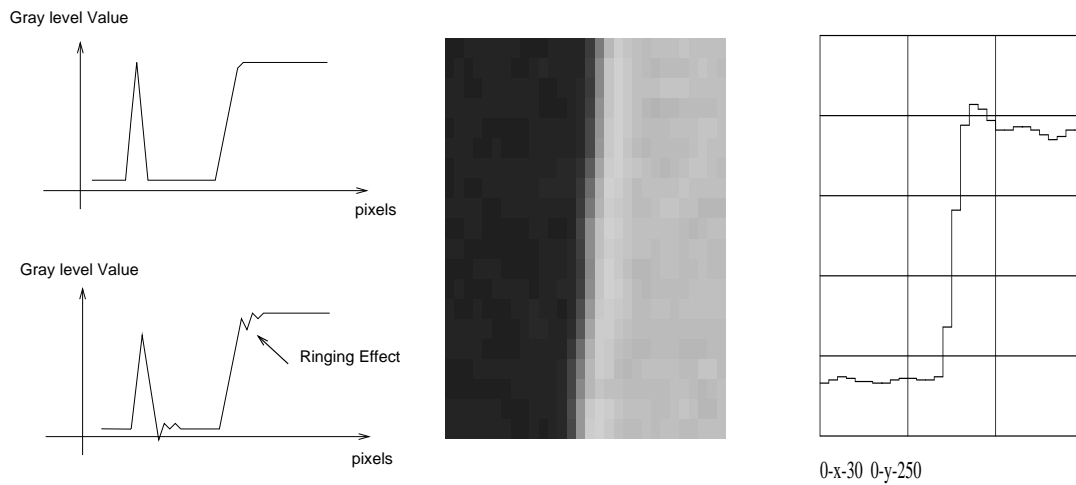


Figure 2: *Ringling effects of the Low-Pass-Filter (LPF)*

- **Warming up effects.** The thermal changes introduced in electronic components are stabilized after a warming up time from 30 to 60 minutes. Thus it is necessary to turn on both camera and frame grabber approximatively one hour before starting calibration experiments.

2.2 Frame Grabber Synchronization

This is surely one of the most tricky problems of image acquisition device. When the composite video signal is sent to the frame grabber, grabber clocks have to be reconstructed (from video top frame and top line) in order to sample the video information and store it as numerical data. The basic idea is to perform this task as to obtain the most accurate matching between the original sampled signal (on CCD matrix) and pixels written in memory.

In order to avoid all divergence phenomenon between the input signal and those generated by the frame grabber (see Fig. 3), a PLL (Phase Locked Loop) is used. Any imprecision in the phase comparison or in the PLL loop control will result in a random displacement of the sampling position. This effect is called *line-jitter*. The line-jitter shifting takes place at the beginning of each line and its random range leads to errors and instability measurements up to 1.0 pixel.

How to avoid or to minimize the line-jitter effects

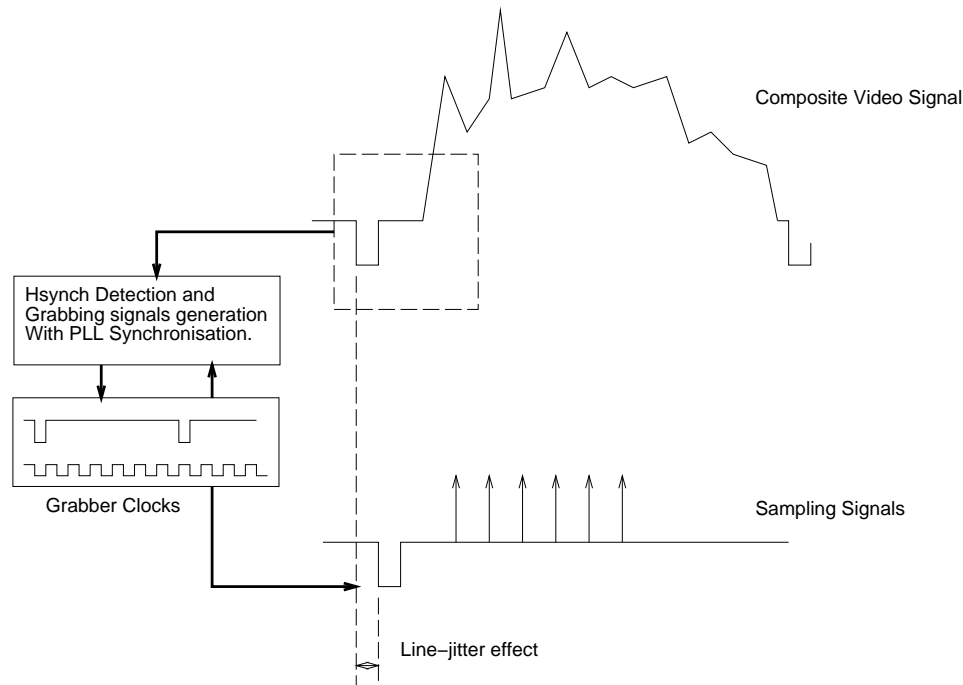


Figure 3: *Generation of sampling signals.*

- Averaging pictures. This is the simplest way to minimize line-jitter effects. It is always possible to send commands to the frame-grabber in order to average several consecutive images. By this way the random noise is reduced but not totally removed. It also supposes that both camera and scene remains stable during grabbing time.
- Using camera-Hsync-clock in the PLL loop. Instead of reconstructing all the grabber-clocks-signals from the video signal, the Hsync (Horizontal synchronization) is sent from the camera. Then the PLL stability is increased and the jitter effects are reduced. But it supposes that first, two different signals are sent from the camera (video and Hsync) and also, that the frame grabber allows to take into account an external clock.
- Using a reference Master-clock. Instead of using a PLL to generate and synchronize grabber-clock signals, a fixed master clock placed in the grabber is used. Clock signals are created by dividing this reference. When a camera-Hsync-top is detected, it resets the process and starts the generation of sampling signals. By this way the jitter effects are reduced to $\frac{1}{8}$ of a pixel.
- Using camera-Hsync and camera-Pixel clocks. This is the most accurate way for analogic camera to remove jitter-effects. The camera sends to the grabbing system both video-signal, Hsync and Pixel clocks. Then a boolean combination between these two signals starts the sampling of video information.
- Using a digital CCD-camera. Inside a digital CCD camera, only digital data are sent from the camera to the frame grabber (instead of the video composite signal). It means that the sampling process is realized directly inside the camera. Thus it is the same clock signals that are used to read the CCD matrix and to sample the analogic signal. By this way all jitter effects are definitively suppressed.
- Using a digital CID-camera. Several solid-state technologies exist to create an image. The most usual one is the CCD one (Charge-Couple-Device). But there also exists CID technology (Charge-Injection-Device) more dedicated to measurement. This sensor offers

the possibility to address each pixel separately (instead of reading line by line for CCD matrix). Then each point is directly stored in memory without any analogic transfer.

Readers interested in different solutions to reduce line-jitter effects can find more analysis in [3].

2.3 Experiments

In order to estimate line-jitter effects the following experiments have been performed. Let us consider an accurate grid with a black background and white lines mounted on a rigid support in front of the camera. Each line intersection (called cross) coordinates detected in the grid are compared between two consecutive images grabbed from the KTH head-eye system [12], see Fig. 4. The specific algorithm implemented to accurately detect the cross position is detailed in section 5.2 and a local approximation of window size 31×31 pixel² around each cross is used to perform a sub-pixel location.

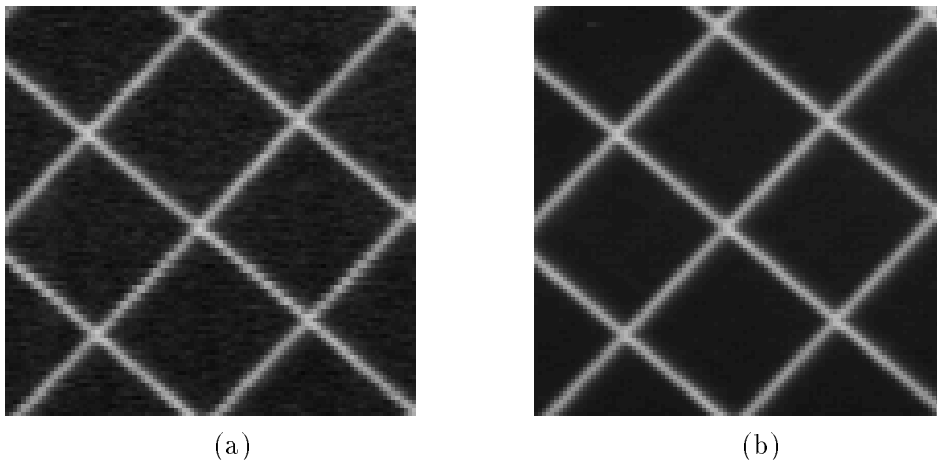


Figure 4: *Details of the grid images used for experiments. (a) single-frame image; (b) an image of averaged 32 frames. Line-jitter effects can be noticed on the left image.*

Table 1: *Line-Jitter effect. Difference of cross coordinates between two consecutive images*

	single frame		32-frame averaged		digital CID camera	
	x	y	x	y	x	y
<i>mean</i>	0.0205	0.0168	0.0044	-0.0027	0.0002	-0.0003
<i>sigma</i>	0.2024	0.0653	0.0522	0.0102	0.0042	0.0038
<i>max</i>	0.5600	0.4230	0.1320	0.0550	0.0104	0.0097

In Tab. 1, a statistical study of the differences of cross localizations between two consecutive images has been performed. It shows the line-jitter effects on the x-coordinate. The mean (*mean*), standard deviation (*sigma*) and maximum (*max*) of differences between two points of two consecutive images have been computed with about 300 crosses. This experiment shows localization errors (between two consecutive images without moving the camera and the grid) of 0.2 pixel on x-coordinate. It is non-ignorable value compared with the window size used for the local cross detection. Of course an error along x-coordinate induces a lower error measurement in y-coordinate.

In the next experiment 32 consecutive frames have been averaged to compose an image, in this way it is expected to reduced the line-jitter effects. As it can be noticed in Tab. 1 error range is reduced by a factor of four, but still remains visible in x-coordinate. It also points out that the

maximum accuracy given by this grabbing system is limited to 0.05 pixel. Noise introduced by line-jitter effect can be seen in Fig. 4 (a). One can also notice the *visual* improvement realized by averaging 32 frames.

The last experiments, shows the results obtained with the same reference grid and a digital CID camera at UBP-France. Here, jitter effects don't exist and the accuracy achieved between two consecutive pictures is much better than in the previous cases. This last experiment points the fact that in this case, major calibration errors are not introduced by the grabbing device.

Conclusion: to conclude, one can say that electronics instability introduces systematic errors up to 1.0 pixel, due to in general jitter effect. Therefore, using sub-pixel detectors and introducing distortion parameters in calibration algorithm don't make sense if these parameters are not taken into account.

3 Mathematical Tools for Simultaneously Calibration by Least Squares Technique

In our study, we use a simplified camera model, i.e., the *pin-hole* model, as depicted in Figure 5. The pin-hole model does not hold for zoom lens [16], [8], but for each step of the zoom, the lens system can be abstracted as a pin-hole model. Through this section, the following notations are used.

- $W-XYZ$ is a right-handed 3-D coordinate system as the world reference coordinate system.
- $o-xy$ is the 2-D image pixel system with the origin at the low-left corner of the image.
- $\bar{o}-\bar{x}\bar{y}$ is a 2-D image coordinate system with x and y parallel to the ones of $o-xy$ and with origin at the principal point \bar{o} .
- $c-\bar{x}\bar{y}\bar{z}$ is the 3-D camera coordinate system with origin at the optical center c and z -axis coincides with the optical axis and x, y parallel to the ones of $o-xy$.

The intrinsic parameters we need to calibrate are: the principal point $\bar{o}(x_0, y_0)$, the focal length f , the pixel sizes of the CCD array or the aspect ratio of them, and the distortion parameters of the optical system. The extrinsic parameters are the rotation matrix \mathbf{R} as well as the translation vector \mathbf{T} between $W-XYZ$ and $c-\bar{x}\bar{y}\bar{z}$.

The methods presented in this section is strictly a least squares technique in the sense that we minimize the errors of the measurements. Notations and reference systems are those conventionally used in photogrammetry.

3.1 The Mathematical Model

We assume a perspective projection between 2D image and 3D object (under the pin-hole camera model). The relationship between a 3D point and its 2D image projection is described by the following equations:

$$\begin{pmatrix} \bar{x}_i \\ \bar{y}_i \\ \bar{z}_i \end{pmatrix} = \lambda_i \mathbf{R} \begin{pmatrix} X_i - T_x \\ Y_i - T_y \\ Z_i - T_z \end{pmatrix} \quad (1)$$

where $(\bar{x}_i, \bar{y}_i, \bar{z}_i)$ is the image point in the camera system as defined in Figure 5 and $-\bar{z}_i \equiv f$, i.e., the focal length of the camera, λ_i is a scale factor which maps a point in object space to the image plane, (X_i, Y_i, Z_i) is the object point in the world-coordinate system $W-XYZ$, (T_x, T_y, T_z)

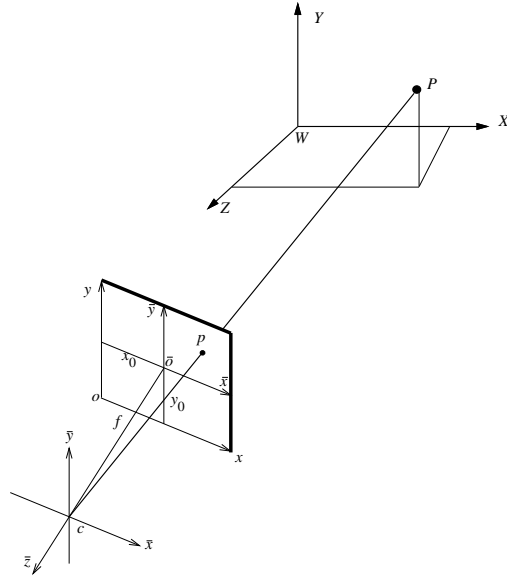


Figure 5: *The pin-hole camera model, image geometry and coordinate systems.*

is the translation vector, and \mathbf{R} is the rotation matrix, which is parameterized by three rotation angles: α rotating around X -axis, β around Y -axis, and γ around Z -axis:

$$\mathbf{R} = \begin{pmatrix} \cos \gamma \cos \beta & \sin \gamma \cos \alpha + \cos \gamma \sin \beta \sin \alpha & \sin \gamma \sin \alpha - \cos \gamma \sin \beta \cos \alpha \\ -\sin \gamma \cos \beta & \cos \gamma \cos \alpha - \sin \gamma \sin \beta \sin \alpha & \cos \gamma \sin \alpha + \sin \gamma \sin \beta \cos \alpha \\ \sin \beta & -\cos \beta \sin \alpha & \cos \beta \cos \alpha \end{pmatrix} \quad (2)$$

Eliminating λ_i in (1) and omit the subscript i , we have the following so called *collinearity equations* in photogrammetry:

$$\left. \begin{aligned} \bar{x} &= -f \frac{r_{11}(X-T_x) + r_{12}(Y-T_y) + r_{13}(Z-T_z)}{r_{31}(X-T_x) + r_{32}(Y-T_y) + r_{33}(Z-T_z)} \\ \bar{y} &= -f \frac{r_{21}(X-T_x) + r_{22}(Y-T_y) + r_{23}(Z-T_z)}{r_{31}(X-T_x) + r_{32}(Y-T_y) + r_{33}(Z-T_z)} \end{aligned} \right\} \quad (3)$$

If we transform (\bar{x}, \bar{y}) into the pixel coordinate system (x, y) , that is

$$\left. \begin{aligned} \bar{x} &= (x + v_x - x_0 - dx)s_x \\ \bar{y} &= (y + v_y - y_0 - dy)s_y \end{aligned} \right\} \quad (4)$$

here dx, dy are lens distortion components, which consist of two parts: *radial* and *tangential*, i.e., $dx = dx_r + dx_t$ and $dy = dy_r + dy_t$. We use two models which are often used in photogrammetry [1]:

$$\left. \begin{aligned} dx_r &= (x - x_0)(a_1 r^2 + a_2 r^4 + a_3 r^6) \\ dy_r &= (y - y_0)(a_1 r^2 + a_2 r^4 + a_3 r^6) \end{aligned} \right\} \quad (5)$$

$$\left. \begin{aligned} dx_t &= \left[p_1[r^2 + 2(x - x_0)^2] + 2p_2(x - x_0)(y - y_0) \right] (1 + p_3 r^2) \\ dy_t &= \left[p_2[r^2 + 2(y - y_0)^2] + 2p_1(x - x_0)(y - y_0) \right] (1 + p_3 r^2) \end{aligned} \right\} \quad (6)$$

where in (4), (5), and (6), (x, y) are image coordinates in the pixel system, (v_x, v_y) are measurement error of (x, y) , (x_0, y_0) are the principal point in the pixel system, a_1, a_2, a_3 are the radial lens distortion parameters, p_1, p_2, p_3 are the tangential distortion parameters, and s_x, s_y are the scale factors of a pixel in x and y directions respectively, and $r = \sqrt{(x - x_0)^2 + (y - y_0)^2}$, is the radial distance from the principal point. As r can be as large as $\frac{\sqrt{2}}{2}L$ (L is the image size), r^4

and r^6 tend to be very large, (5) will result in numerical instability. One way to overcome this problem is by rewriting (5) in the following way:

$$\left. \begin{aligned} dx_r &= (x - x_0)[a'_1((\frac{r}{r_0})^2 - 1) + a'_2((\frac{r}{r_0})^4 - 1) + a'_3((\frac{r}{r_0})^6 - 1)] \\ dy_r &= (y - y_0)[a'_1((\frac{r}{r_0})^2 - 1) + a'_2((\frac{r}{r_0})^4 - 1) + a'_3((\frac{r}{r_0})^6 - 1)] \end{aligned} \right\} \quad (7)$$

where $a'_i = a_i r_0^{2i}$, ($i = 1, 2, 3$). (7) implies that we set the distortion to 0 at radial distance r_0 . This results in a focal length change of $\Delta f = \frac{f}{r_0} dr_0$. Substituting (4), (5) and (6) into (3), and let $f_x = \frac{f}{s_x}$ and $f_y = \frac{f}{s_y}$, we have the following:

$$\left. \begin{aligned} x + v_x &= x_0 + dx_r + dx_t - f_x \frac{r_{11}(X-T_x) + r_{12}(Y-T_y) + r_{13}(Z-T_z)}{r_{31}(X-T_x) + r_{32}(Y-T_y) + r_{33}(Z-T_z)} = P(\Phi) \\ y + v_y &= y_0 + dy_r + dy_t - f_y \frac{r_{21}(X-T_x) + r_{22}(Y-T_y) + r_{23}(Z-T_z)}{r_{31}(X-T_x) + r_{32}(Y-T_y) + r_{33}(Z-T_z)} = Q(\Phi) \end{aligned} \right\} \quad (8)$$

where Φ is the parameter vector, i.e.,

$$\Phi = [x_0, y_0, a_1, a_2, a_3, p_1, p_2, p_3, f_x, f_y, T_x, T_y, T_z, \alpha, \beta, \gamma]^T$$

3.2 Solving the Problem

The problem is now to solve for Φ by minimizing $\sum_{i=1}^n (v_{x_i}^2 + v_{y_i}^2)$. In (8) $P(\Phi)$ and $Q(\Phi)$ are non-linear functions of Φ , the minimization is a non-linear optimization problem. One way of solving the problem is to linearize (8) with some initial value Φ_0 and solve for $\Delta\Phi$. Then by adding $\Delta\Phi$ to Φ_0 as new initial value and repeat the process until certain convergency criterion is satisfied. Given n 3D points and their corresponding 2D image points, we can write the $2 \times n$ linearized measurement or error equations in matrix form: (for detail derivation of the linearization see Appendix A):

$$V = A \Delta\Phi - L \quad (9)$$

Let the weight matrix of the measurements¹ be W , the least squares solution to (9) is a minimization problem of

$$\min_{\Delta\Phi \in \mathbb{R}^{15}} (V^T W V) \quad (10)$$

The solution of (10) can be obtained as (for details, see Appendix B):

$$\Delta\Phi = (A^T W A)^{-1} (A^T W L) \quad (11)$$

Remark: In general, the solution of the minimization problem (10) is not guaranteed and the linearization requires some initial values for the parameters. In our case, e.g. for zoom-lens purpose, we get our initial values from the calibration using vanishing points (for the focal length and the extrinsic parameters [10]) and from zooming properties (for the principal point [8]). These values are always very close to the optimal ones, so the solution is obtainable in practice without any problem. We have tested different methods for solving (10), i.e., singular value decomposition (SVD), QR-factorization, and normal equation. They all give more or less the same results due the factor that we have very close initial values, but the normal equation is the fastest.

¹In normal case, W is a $2n \times 2n$ diagonal matrix. If all the measurements are made with the same accuracy and there is no correlation between measurements, then W is identical matrix multiplied by a constant, i.e., $W = cI$.

3.3 Multi-image Calibration

One major source of errors in the calibration is the measurement errors. One way to improve this is to combine more than one images taken by the same camera from different or the same positions. In such a case, the intrinsic parameters are the same for all images, the calibration is to compute a parameter vector of

$$\Phi_{10+6m} = [x_0, y_0, a_1, a_2, a_3, p_1, p_2, p_3, f_x, f_y, T_x^1, T_y^1, T_z^1, \alpha^1, \beta^1, \gamma^1, \dots, T_y^m, T_z^m, \alpha^m, \beta^m, \gamma^m]^T$$

And the coefficient matrix \mathbf{A} as in (9) will look like this:

$$\mathbf{A}_{2mn \times (10+6m)} = \left[\begin{array}{c|ccc} & \mathbf{A}_{2n \times 6}^{11} & & 0 \\ & & \ddots & \\ \mathbf{A}_{2mn \times (10+6m)}^1 & & & \mathbf{A}_{2n \times 6}^{ii} \\ & & & \ddots \\ & 0 & & \mathbf{A}_{2n \times 6}^{mm} \end{array} \right]$$

where m is the number of images and n is the number of points on each image. The total number of equations will be $(2mn)$ and the total number of parameter is $(10 + 6m)$. The redundancy number r is $r = 2mn - 10 - 6m$, which is much larger than the one in the single-image calibration case. As we will show in the next subsection that the larger the redundancy the higher the reliability.

We can also use multi-image approach to solve for partial of intrinsic parameters, e.g., the principal point, from more than one images taken from different position and with different zoom and/or focus settings under the assumption that zooming and focusing do not alter the optical axis. Or we can use this technique to compute the same extrinsic parameters or partial of them from multiple images taken from different optical parameter settings, e.g, by zooming or focusing and keep the camera still under the assumption that zooming and focusing do not alter the optical axis and the optical center. The multi-image technique has also been used by Tsai [19].

3.4 Accuracy Assessment

From the least squares estimation of (10), we can compute the estimate of the residual vector \mathbf{V}

$$\hat{\mathbf{V}} = [\mathbf{A}(\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W} - \mathbf{I}] \mathbf{L} \quad (12)$$

and the estimate of the so called *standard error of unit weight*, which is the posteriori estimation of the noises σ_0 of the image coordinates:

$$\hat{\sigma}_0^2 = \frac{\hat{\mathbf{V}}^T \mathbf{W} \hat{\mathbf{V}}}{N - u} \quad (13)$$

and the estimate of the *co-factor* matrix of the parameters Φ :

$$\mathbf{C}_\Phi = (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \quad (14)$$

And for each individual parameter ϕ_i , we can then compute the estimate of its precision (standard deviation):

$$\hat{\sigma}_{\phi_i}^2 = \hat{\sigma}_0^2 c_{ii} \quad (15)$$

One way of measuring reliability (quality) of a least squares adjustment system is the relative redundancy [18], i.e.,

$$q = \text{tr}[\mathbf{I} - \mathbf{A}(\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T] = \frac{r}{N} \quad (16)$$

where in (13) – (15), r is the redundancy number, N is the total number of measurement equations, u is the total number of unknown parameters, and c_{ii} is the i 'th diagonal element of the co-factor matrix \mathbf{C}_Φ . As we can see that the relative redundancy number in the multi-image case is much larger than the single-image case, so we have a higher reliability in the multi-image case, $q_{multi} = \frac{r}{2mn} = 1 - \frac{10+6m}{2mn}$, comparing to the single-image case, $q_{single} = 1 - \frac{13}{2n}$. If $m > 1$, then $q_{multi} > q_{single}$.

There are many factors which effect the accuracy of the computed parameters, thus the reconstructed 3D data using these parameters. In the previous paragraph we have shown some error sources. Let us summarize the major factors:

- The number and accuracy of the measurements. How well we can measure 2D points in the image, and the accuracy of the 3D object points are the main factor which effect the accuracy of the calibration. Usually, we assume the 3D points are free of error, which is not quite true. Assuming both errors are random errors (with a certain distribution), this type of error will result in a random residual error in the image after the least squares adjustment.
- Geometrical configuration of the set up, i.e., the relation between the camera and the calibration object. This factor determines the internal strength of the least squares solution, i.e., the design matrix \mathbf{A} of the system. Improper geometrical configuration will result in poor accuracy and reliability.
- The mathematical models used for the calibration. We model the camera lens as a *pin-hole* mode, this is just an approximation, and so is the lens distortion modelled as radial and tangential distortion. The lens system is a very complicated optical and mechanical construction, there is no exact model for it. The imperfect model will result in a systematic error in the image after the calibration.

In order to see the effects of various factors, we use simulation to test them. The simulation is done as follows: given a set of 3-D object points (e.g., the calibration object) and a set of camera parameters Φ , we compute the back projected 2-D image points of the 3-D points using (8) (only radial distortion is used here), and then add some pseudo-noise v with Gaussian distribution $g(0, \sigma_0)$ to the image coordinates. And then using these 2-D image points to calibrate camera parameters $\hat{\Phi}$. We have performed a serial of experiments using different σ_0 and different number of images for multi-image calibration. Figure 6 through Figure 7 show the results.

Figure 6 shows the estimates of standard deviations of all parameters as a function of noises σ_0 of the image coordinates. As we can see clearly from the figure that the standard deviations of the computed parameters increase linearly with the increasing of the noise σ_0 . This is expected from (15) as the geometry and the model is the same, i.e., c_{ii} remains the same. This suggests that *we should reduce our measurement error* as much as possible in order to get high accuracy for the calibration.

Figure 7 shows the results of the accuracy of the calibrated parameters as a function of the number of images used for the multi-image calibration. While from Figure 7, we can see that increasing the number of images increases the accuracy of the computed parameters dramatically.

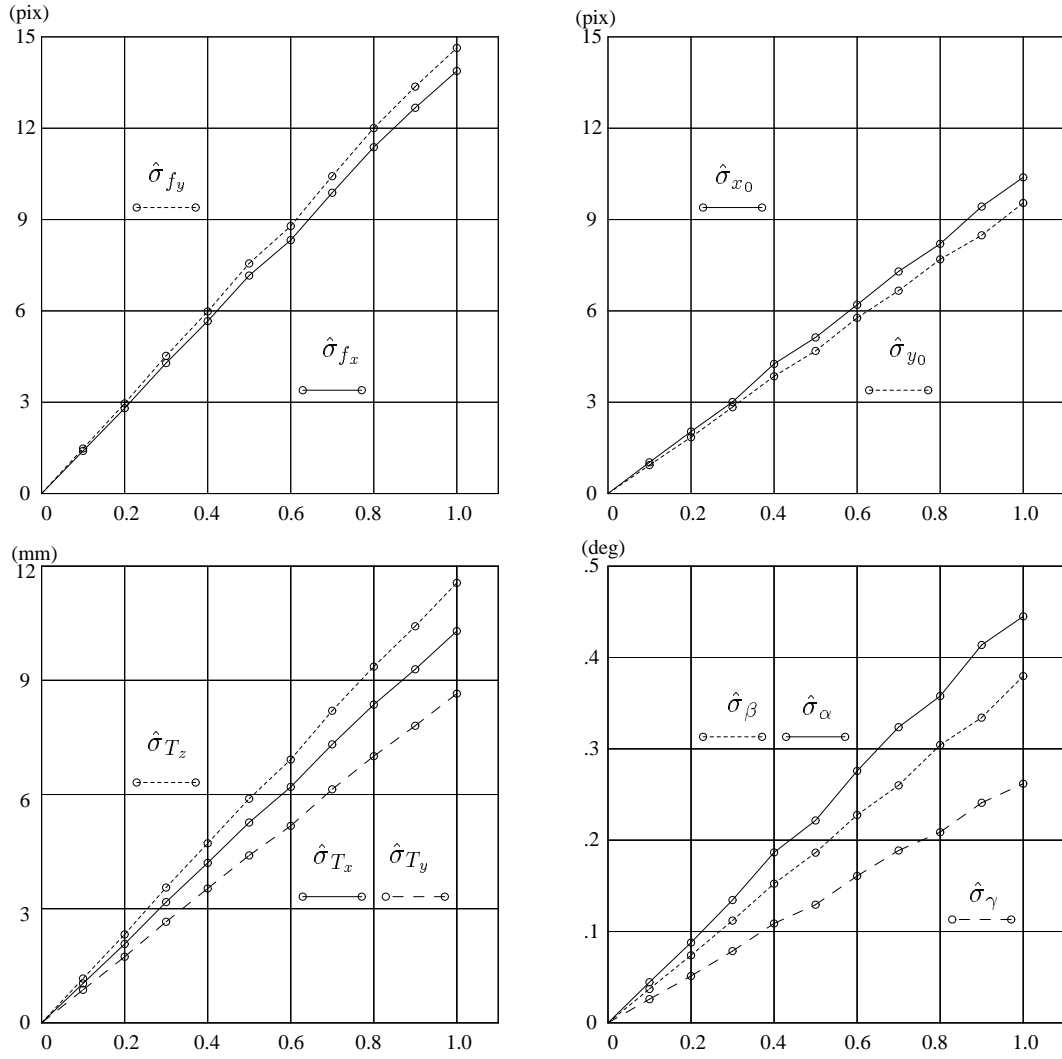


Figure 6: Estimated standard deviation $\hat{\sigma}_{\phi_i}$ of calibrated parameters as function of noises of image coordinates σ_0 (the horizontal axes in pixels). **Up:** for intrinsic parameters f_x, f_y, x_0 and y_0 , **Low:** for translation T_x, T_y, T_z , and rotation angle α, β, γ .

4 Consideration on Calibration Reference Object

In order to achieve accurate measurements and to achieve good calibration results, particular care has to be taken to the way the 3D reference object is constructed and also to the way the coordinates of calibration points are measured. Such an achievement is not a straightforward task, and the loss of accuracy during this calibration step leads to erroneous results and introduces a major instability in the resolution of the non linear system involved by calibration. An accurate reference object requires time, money and sometimes *several of them* are needed in case of zoom-lens calibration.

In the following sections, we first quickly describe different solutions to realize the calibration object. Then, we show that the geometric configuration between the camera and the reference object also effects to the calibration results. Finally, the KTH reference object used during experiments is presented.

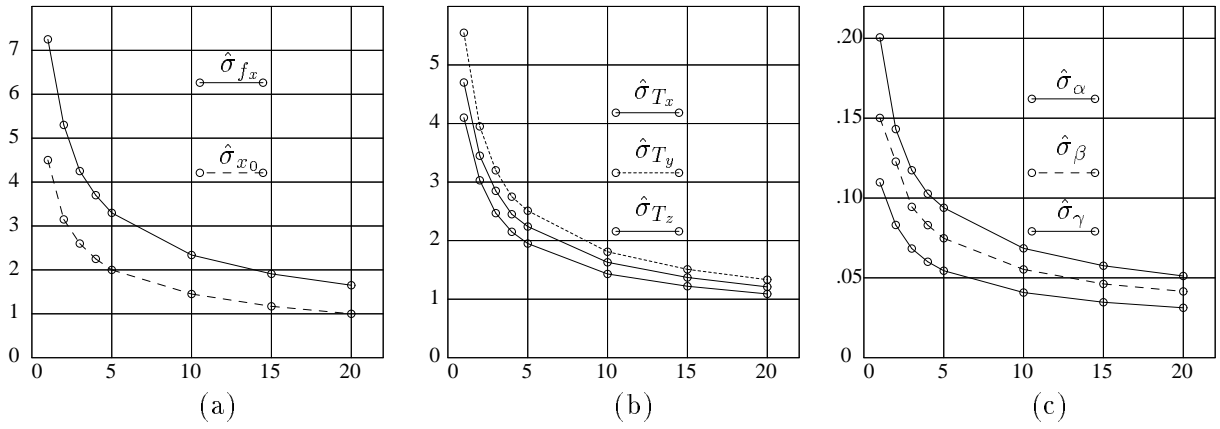


Figure 7: *Estimated standard deviation $\hat{\sigma}_{\phi}$ of calibrated parameters as function of the number of images (horizontal axes) used for multi-image calibration, i.e., all the images have the same intrinsic parameters but different extrinsic parameters. The examples shown here are simulated with $\sigma_0 = 0.5$. (a) for intrinsic parameters $f_x(f_y)$ and $x_0(y_0)$, (b) for translation T_x, T_y, T_z , and (c) for rotation angles α, β, γ .*

4.1 3D Reference Object Construction

At least three approaches can be used to construct a 3D calibration object.

- Translation of a flat calibration pattern [20]. It is not really difficult to find an accurate flat pattern and a lot of calibration experiments are done using regular grids. In [20] a 16×16 -in optical glass plate is used, it is done by a high-precision photographic process. The pattern is placed horizontally inside a custom-made stand that can move it vertically. The camera has also to be placed vertically at the top of the device.

Several considerations have to be taken into account with such a solution:

1. The flat pattern needs to be accurately *placed and moved* perpendicular to the translation axis (i.e., a perpendicular uncertainty of 0.2 deg in orientation, combines with a pattern translation of 300mm leads to 3D measurement errors up to $1mm^3$ for calibration points).
 2. The pattern can only be used for a few number of focal lengths and the camera has to be placed in a special way during experiments which could be a major handicap when cameras are fixed in an active-head for example.
- 3D cube [10, 8]. Most of calibration experiments are done with a 3D cube. Even if each plane can be accurately processed, it is not a straightforward task to guarantee that faces are perpendicular. For close-range applications [8] which requires a small target, a full cube can be accurately manufactured and the uncertainties in 3D coordinates points are lower than $0.5mm^3$. Larger targets [10] require to put together three flat patterns and the perpendicular problem becomes the major source of uncertainty. One advantage of using cube is to us vanishing points technique for calibration [10].
 - Construction of a test-field [3]. This may be the best solution to realize a 3D reference object for calibration purpose. It allows you to make experiments for both close and remote conditions. A test-field usually consists of targets (black circle on white background) placed on a wall and on a structure in front of the wall. The 3D measurements of the test-field is done with theodolites. This calibration reference object is time consuming but finally [3] achieves uncertainty lower than $0.2mm^3$ for a test-field size $(4.2 \times 2.0 \times 1.1m^3)$.

4.2 Geometric Configuration between Camera and Calibration Object

The geometric configuration between the camera and the calibration object also plays an important rule in calibration. The same object, but different configurations will result in different correlation among parameters and thus different accuracies. In this section, we will show this effect through simulation. Given camera parameters and a 3D calibration object, which is plane pattern shifted three times in the normal direction to the plane as to constitute a 3D system, we project the 3D points back to the image plane according to the model presented in Section 3. Then add some noises to the image points (with Gaussian noise $\sigma_o \simeq 0.2\text{pixel}$). Finally, using these noise contaminated points to compute the camera parameters.

Table 2: *Accuracy of different configurations*

<i>parameter</i>	case I: $\hat{\sigma}_0 = 0.1922$			case II: $\hat{\sigma}_0 = 0.1920$		
	<i>true</i>	<i>computed</i>	$\hat{\sigma}_x$	<i>true</i>	<i>computed</i>	$\hat{\sigma}_x$
f_x (pix)	1000.0	999.363	0.3068	1000.0	1000.705	0.3646
f_y (pix)	1000.0	999.314	0.3068	1000.0	1000.710	0.3645
x_o (pix)	350.0	350.012	0.2771	350.0	349.504	0.2994
y_o (pix)	350.0	350.479	0.2771	350.0	349.378	0.2964
T_x (mm)	0.0	0.101	0.1331	800.0	800.586	0.2804
T_y (mm)	0.0	0.032	0.1331	800.0	800.374	0.2812
T_z (mm)	2300.0	2298.866	0.5190	2000.0	2001.184	0.5608
α (dgr)	0.0	0.026	0.0158	29.0	-29.033	0.0184
β (dgr)	0.0	0.002	0.0158	26.5	26.524	0.0167
γ (dgr)	0.0	-0.003	0.0028	5.0	5.011	0.0084

Table 3: *Correlation between parameters. Case I: 3-plane object parallel to image plane.*

	f_x	f_y	x_o	y_o	T_x	T_y	T_z	α	β	γ
f_x	1.00	0.95	0.00	0.00	0.00	0.00	0.97	0.00	0.00	-0.00
f_y	0.95	1.00	0.00	0.00	0.00	0.00	0.97	-0.00	-0.00	0.00
x_o	0.00	0.00	1.00	-0.00	0.01	-0.00	0.00	-0.00	-0.97	0.00
y_o	0.00	0.00	-0.00	1.00	0.00	0.01	0.00	0.97	-0.00	-0.00
T_x	0.00	0.00	0.01	0.00	1.00	0.00	0.00	0.00	0.24	-0.00
T_y	0.00	0.00	-0.00	0.01	0.00	1.00	0.00	-0.24	-0.00	0.00
T_z	0.97	0.97	0.00	0.00	0.00	0.00	1.00	0.00	-0.00	-0.00
α	0.00	-0.00	-0.00	0.97	0.00	-0.24	0.00	1.00	0.00	-0.00
β	0.00	-0.00	-0.97	-0.00	0.24	-0.00	-0.00	0.00	1.00	-0.00
γ	-0.00	0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.00	-0.00	1.00

Tab. 2 shows two cases in which the same object, but different camera configurations is used. In the first case patterns are parallel to the image plane, in the second one patterns are rotated to the image plane ($\simeq 30\text{deg}$). As we can see that in both cases they have the same $\hat{\sigma}_0$ (the estimation of the noises, or the *standard error of unit weight*, see Section 3.4), but different standard derivations for each computed parameters. Tab. 3 – 4 show the normalized correlation among the computed parameters for these two cases. The numbers in the tables correspond to the correlation between two parameters. The larger (absolute value), the higher correlation. If two parameters are highly correlated, they are mixed up in the computation (not separable) and compensate for each other and . It is clear from the tables that in case I parameters are less correlated than case II. Tab. 5 shows compensation effect of two correlated parameters. In the table, in one case we use $x_o = 355$, which is 5 pixels away from the true value, as fixed value in the calibration, and in other case use $y_o = 355$ as fixed value. As we can see from Tab. 3 that x_o is highly correlated with β , while y_o is highly correlated with α . If we introduce some error in x_o and fix it in the calibration, all the parameters, except β , will be computed correctly, while β

will have *incorrect* value, because it has to compensate the error introduced by x_o . Similar thing happens between y_o and α as they are highly correlated. This can be noticed from Tab. 5.

What these simulations tell us is that correlation between calibration parameters and therefore calibration accuracy depends on the choice of reference object and the way it is placed in front of the camera.

Table 4: *Correlation between parameters. Case II: 3-plane object rotated to image plane.*

	f_x	f_y	x_o	y_o	T_x	T_y	T_z	α	β	γ
f_x	1.00	0.96	-0.01	-0.00	0.89	0.88	0.97	0.08	-0.06	-0.07
f_y	0.96	1.00	-0.02	-0.00	0.88	0.89	0.97	0.08	-0.06	-0.08
x_o	-0.01	-0.02	1.00	-0.03	-0.02	-0.06	-0.08	0.05	-0.96	-0.09
y_o	-0.00	-0.00	-0.03	1.00	-0.03	0.01	-0.06	0.96	0.12	-0.90
T_x	0.89	0.88	-0.02	-0.03	1.00	0.77	0.85	0.05	0.05	-0.04
T_y	0.88	0.89	-0.06	0.01	0.77	1.00	0.85	-0.02	-0.04	0.01
T_z	0.97	0.97	-0.08	-0.06	0.85	0.85	1.00	0.05	-0.02	-0.04
α	0.08	0.08	0.05	0.96	0.05	-0.02	0.05	1.00	0.04	-0.94
β	-0.06	-0.06	-0.96	0.12	0.05	-0.04	-0.02	0.04	1.00	0.01
γ	-0.07	-0.08	-0.09	-0.90	-0.04	0.01	-0.04	-0.94	0.01	1.00

Table 5: *Compensation effect of two correlated parameters (case I)*

para- meter	true value	full calib		fixed $x_o = 355.0$		fixed $y_o = 355.0$	
		$\hat{\sigma}_0 = 0.1922$		$\hat{\sigma}_0 = 0.2488$		$\hat{\sigma}_0 = 0.2397$	
		computed	$\hat{\sigma}_x$	computed	$\hat{\sigma}_x$	computed	$\hat{\sigma}_x$
f_x	1000.0	999.363	0.3068	999.402	0.3971	999.408	0.3826
f_y	1000.0	999.314	0.3068	999.364	0.3971	999.348	0.3826
x_o	350.0	350.012	0.2771	355.000	0.0001	350.016	0.3455
y_o	350.0	350.479	0.2771	350.483	0.3586	355.000	0.0001
T_x	0.0	0.101	0.1331	0.133	0.1723	0.101	0.1660
T_y	0.0	0.032	0.1331	0.032	0.1723	0.061	0.1660
T_z	2300.0	2298.866	0.5190	2298.975	0.6718	2298.966	0.6472
α	0.0	0.026	0.0158	0.027	0.0205	0.276	0.0051
β	0.0	0.002	0.0158	-0.274	0.0053	0.001	0.0198
γ	0.0	-0.003	0.0028	-0.003	0.0036	-0.003	0.0035

4.3 Target used in Experiments at KTH

Two cubes have been used during experiments of zoom-lens calibration. Each of them is composed by three flat metallic anodized patterns. As it can be noticed on Fig. 8.(a), we use a black background and with lines. It appeared to be more convenient to realize good lighting conditions with such a target. Each face has been fitted together using accurate mechanical device.

The two cubes are respectively $(0.4 \times 0.4 \times 0.4m^3)$ and $(0.15 \times 0.15 \times 0.15m^3)$ in sizes. As we can note, zoom-lens calibration requires several reference objects if the different parameters have to be estimated for a large scale of zooming set.

We choose the cube object as calibration target is because we use the vanishing points technique to compute the approximate values of the camera parameters *automatically* [10].

5 Calibration Points Detection

As has been discussed in Section 3.4, measurement error of image point is one of the major error source in camera calibration. Accurate detection of image points is an important step in

a calibration process. This section describes some practical tools used with the KTH active head-eye system, to match automatically image calibration points to their corresponding 3-D points.

The three faces of the KTH calibration cube (Fig. 8.(a)), are orthogonal to each other. All lines are grouped into three groups which are orthogonal to each other. The three edges of the cube also serve as 3D coordinate system as shown in the Fig. 8 (a), and the intersection points of all perpendicular lines on each face are measured in this 3D coordinate system as control points.

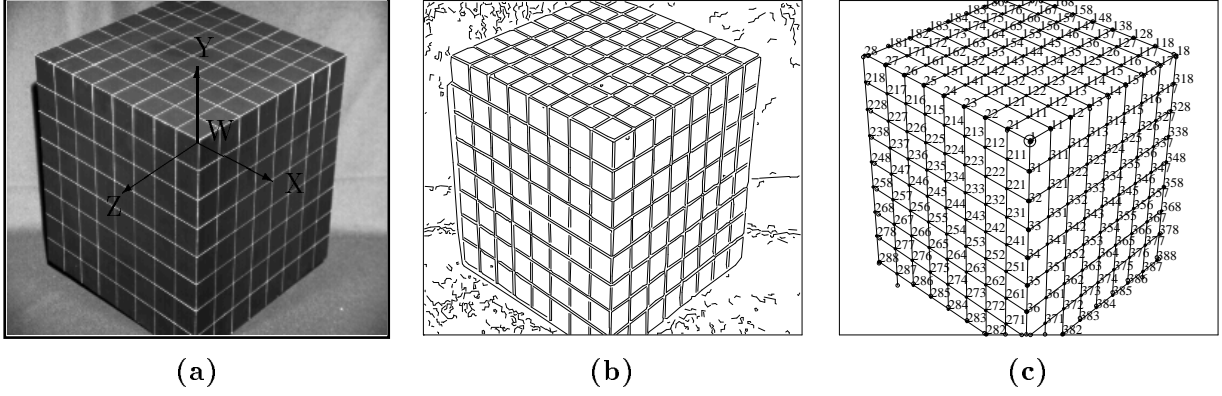


Figure 8: (a) The calibration cube with multi-orthogonal straight lines on each face. The three edges define a right-hand 3D coordinate system as the world reference system. The intersection points on each face, are used as 3D control points for the calibration. (b) Results of edge detection, (c) The identified line segments and points.

5.1 Automatic localization of calibration points using edge detection

The automatic procedure for calibration points detection and matching is done as follows:

- Take a image and run a edge detection and chaining (note on Fig. 8 (b) that edge detector introduces a double response for each straight line).
- perform some segmentation and grouping operations on the edge segments to extract line segments. Try also to connect together short collinear segments (b).
- compute the middle line position of each double lines and the intersection points of the line segments and identify them. Identification is done by searching the central point (connected to three branches) near the image center. Then the three edges are searched and identified, finally all the other points are found and matched using invariant properties of points on a straight line and points on a plane (c).

After applying this automatic process, all image calibration points are detected and matched into the 3D world coordinate system.

5.2 Accurate computation of calibration point coordinates

In this subsection we describe a special algorithm used to re-compute the coordinates of each intersection points (called cross position). Instead of using noisy information obtained by edge detection, the basic idea is to use the grey level data directly to determine the cross coordinates. The software used to compute the sub-pixel localization was introduced by [14], in which the author has found the expression of an analytical surface whose shape is really close to the photonic response detected around a cross position (see Fig. 9). Let us consider the function $F(x, y)$:

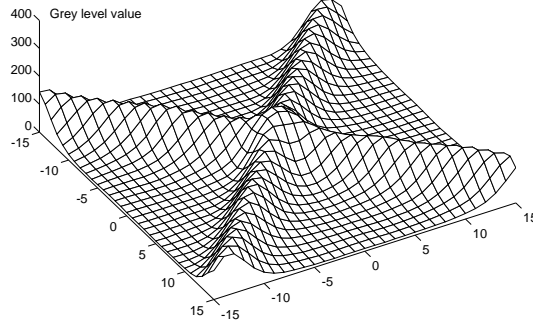


Figure 9: Real grey level signal around a cross position

$$F(x, y) = a + bx + cy + d_1^2 \exp^{-e_1^2(x \cos \theta_1 + y \sin \theta_1 + l_1)^2} + d_2^2 \exp^{-e_2^2(x \cos \theta_2 + y \sin \theta_2 + l_2)^2} \quad (17)$$

where:

- coefficient a represents the mean grey level value of the background,
- coefficients b, c take into account a regular slope according to the background grey level values (useful when cross surface is not parallel to the image plane),
- coefficients d_1, d_2 represent the two branch heights respectively,
- coefficients e_1, e_2 represent the branch widths,
- coefficients θ_1, θ_2 represent the branch orientations,
- coefficients l_1, l_2 represent the orthogonal distance from the branches to the local origin point.

Let us suppose that the cross position is approximatively found (using edge detection previously presented), then the analytical surface $F(x, y)$ can be fitted inside the grey level data by adjusting the function parameters. The problem is to find

$$\mathbf{V} = (a, b, c, d_1, e_1, l_1, \theta_1, d_2, e_2, l_2, \theta_2)^T$$

such that

$$\sum_{x=-n}^n \sum_{y=-n}^n (I(x, y) - F(x, y))^2$$

is minimized. Where $I(x, y)$ is the real grey level value at point (x, y) and n is the half window size of the local approximation.

Levenberg-Marquardt algorithm [15] has been used to solve this non-linear minimization. Even if the number of parameters to be fitted is important, straightforward tools can be used to start the process with an approximate value V' close to the solution. Usually, less than ten iterations per cross are needed to reach a stable solution.

In Fig. 10 we have drawn the analytical solution (a) found from real cross data (b). Note that the central part of the theoretical cross doesn't fit perfectly with real data (higher pick). So, for practical purpose these middle points are not taken into account in the minimization. The standard deviation of residual errors between the two surfaces is lower than 5 grey level value.

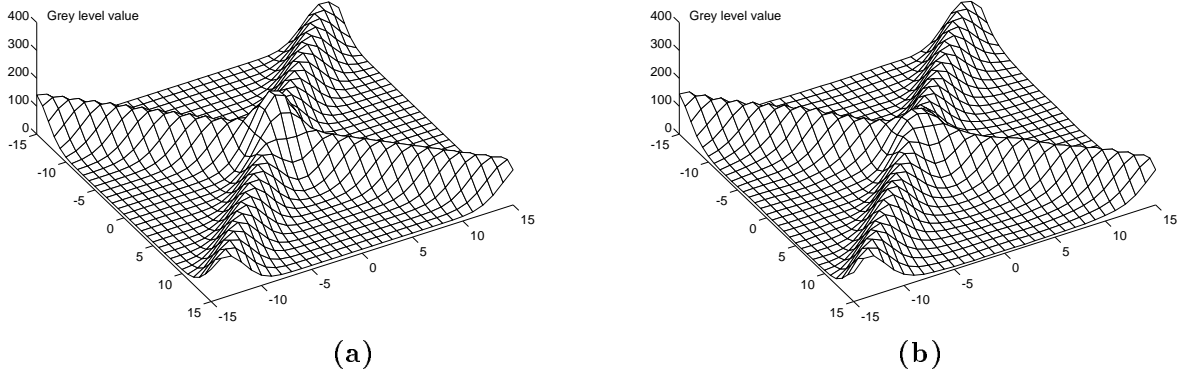


Figure 10: Convergence result of cross detection according to a flat grid image as in Fig. 4

In Fig. 11 the same experiment has been applied on a cube calibration point (Fig. 8 a). The noise is there more obvious, but it fits correctly.

After the fitting, the sub-pixel location of a cross point can be computed straightforward. According to $F(x, y)$ formula, the cross center location (x_0, y_0) is computed by solving the following linear system:

$$\begin{cases} x_0 * \cos \theta_1 + y_0 * \sin \theta_1 + l_1 = 0 \\ x_0 * \cos \theta_2 + y_0 * \sin \theta_2 + l_2 = 0 \end{cases}$$

5.3 Conclusion

Most of calibration articles never deal with the detection of the calibration targets. But experiments show that using classical edge detectors don't make it possible to obtain reliable results. For our purpose, a cross detection is used to perform a sub-pixel location of reference points in the CCD matrix. Solution using point detection (spot) can also be found in [3].

6 Experiments and Results on Zoom-lens Calibration

In this section we present some experimental results. The experiments were performed on the KTH head-eye system (Fig. 12) which consists of **a pair of cameras with zoom-lens** (eyes) mounted on a head and the head is mounted on a shoulder (platform). The system has 13 degrees of freedom (DOFs), which are: each eye has 5 degrees of freedom – tilt and pan for rotation, focus, zoom and iris for optical parameters; the neck has 3 degrees of freedom – tilt and pan for rotation and one for changing the baseline (this is an experimental feature of the system). All the rotations have a resolution of 0.0072° or 0.000126 rad, and a top speed of $180^\circ/\text{second}$. For detailed description of the system see [12] and [13]. The calibration of the system includes two parts: the first part is the camera-zoom-lens calibration problem, i.e., to calibrate intrinsic and extrinsic parameters of a camera at a particular set-up, and for a serial of such setups, we can build up a *look-up-table* for the intrinsic parameters. The second part is the *kinematic calibration*, which is addressed in [11].

6.1 Determining Principal Point by Zooming

6.1.1 Zooming Properties

Using the pin-hole model, the *principal point* is defined as the intersection point of the optical axis and the image plane. Assuming that the optical axis of the camera is perpendicular to the

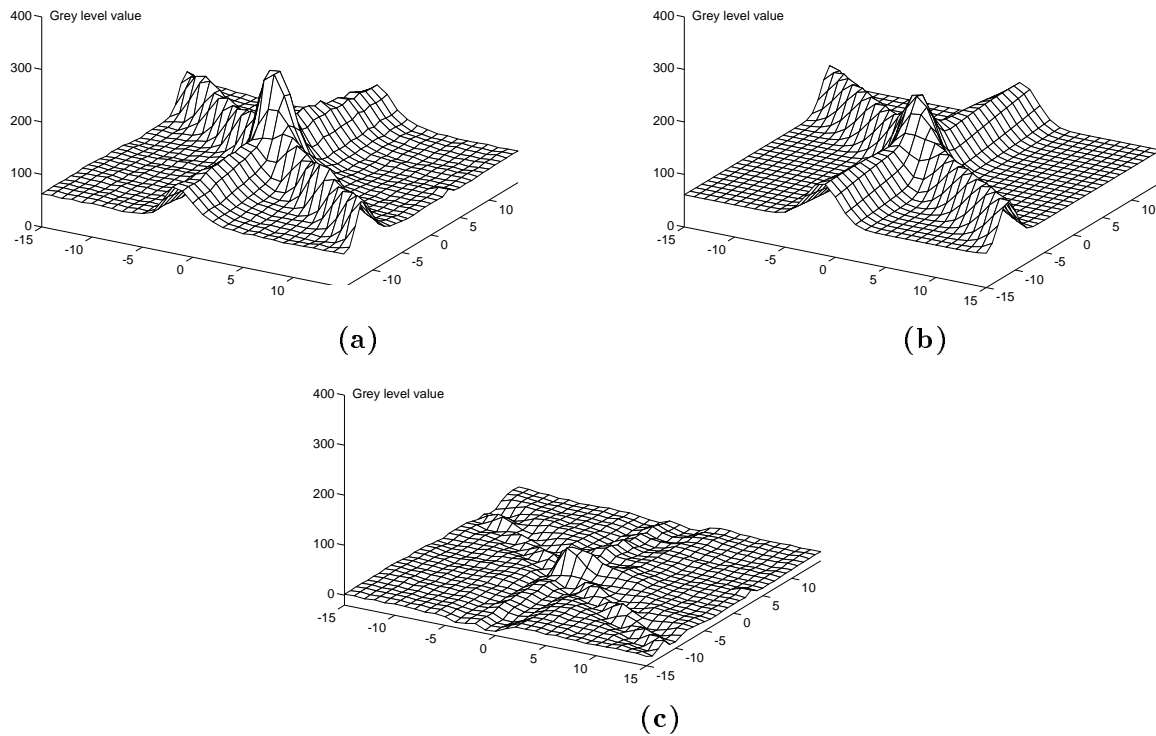


Figure 11: Convergence result of cross detection according to a cube image. (a) real noisy data corresponding to a cube calibration point, (b) fitted surface, (c) difference between the two surfaces

image plane and the optical system of the camera is stable during focusing and zooming, i.e., the optical axis does not alter, then the principal point is fixed in the image plane during focusing and zooming. In a pin-hole camera model, focusing or zooming is equivalent to changing the distance between optical center and image plane, see Fig. 13. When focusing or zooming, each image point will move radiately along a line passing through the principal point. If we take a sequence of images by setting different focus and/or zooming, find the corresponding image points on each image and overlay them together (see Fig. 13.(b)), the lines determined by the corresponding image points at different images will intersect at a common point, which is the principal point. Due to various errors, the lines will not intersect precisely at a common point, a least squares estimation of the principal point is needed by minimizing the sum of the squared distances from the point to all the lines. This technique has been proposed by Tsai [19] and used by Lavest et. al. [8].

6.1.2 Results of Principal-Point Calibration by Zooming

The principal point is calibrated first separately using the zooming technique previously described. A regular and accurate pattern is used for the calibration. The dot in the grid is used for ease recognition of the pattern.

The three motor steps of zoom, focus and iris are mapped to $[0, 1.0]$ with “continuous” stepping. By keeping a calibration object still and taking image by stepping 0.05 steps, we get a sequence of 22 images of the same object **for a given focus**. See Fig. 14 for a few examples. Then using the cross detection algorithm all the crosses coordinates are computed and identified (referred to the dot in the center). By tracing the same points (the cross) through the sequence (some of the points will fall out of the image when zooming out), a number of line segments are obtained, see Fig. 15 (a). By fitting line model to line segments, line parameters are computed for

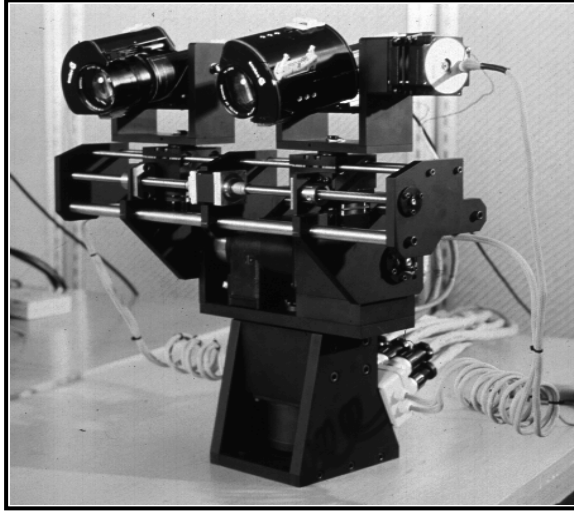


Figure 12: *The KTH-head-eye system with 13 degrees of freedom.*

all the line segments. And finally, the common intersection point of all these lines are computed by least squares method, which minimizes the distances of the intersection point to lines see Fig. 15 (b).

In order to check the stability of the principal point when zooming, we have chosen different combinations of images to determine the common intersection point. Tab. 6 and 7 show a few examples of the results and Fig. 16 shows the trace of movement of principal point under zooming and zooming back sequence. As it can be seen from the figures, the principal point moves about two pixels. But the value computed with all pictures (see the large cross in Fig. 16) remains stable and its movement is lower than 0.5 pixel. The figures also show the stability over different distances of objects, which is equivalent to different focus setups, and time as those images were taken at a time interval of 3 weeks.

Table 6: *Results of calibrating principal-point by zooming. First experiment, Seven views*

<i>focus step</i>	<i>zoom combination</i>	<i>zoom</i>		<i>zoom back</i>	
		x_0	y_0	x_0	y_0
0.23	1 - 3	368.58	283.00	370.25	283.54
	2 - 4	369.98	283.36	369.69	283.12
	3 - 5	369.89	283.54	369.65	283.42
	4 - 6	369.89	283.77	370.19	284.11
	5 - 7	370.07	284.04	370.38	284.43
	1 - 7	369.84	283.72	369.95	283.70

The conclusions we can draw from these experiments are that the stability of principal point during focusing, zooming and other mechanical and optical movements of the camera depends on the camera lens itself. In [8] a high quality zoom lens leads to a more stable result of of (x_0, y_0) coordinates. But even in this case the displacement remains small.

6.1.3 Principal Point Influence

In order to check the influence of (x_0, y_0) coordinates in the calibration results, we have done the following experiments: Let us consider the same simulation as in Section 4.2 with two different

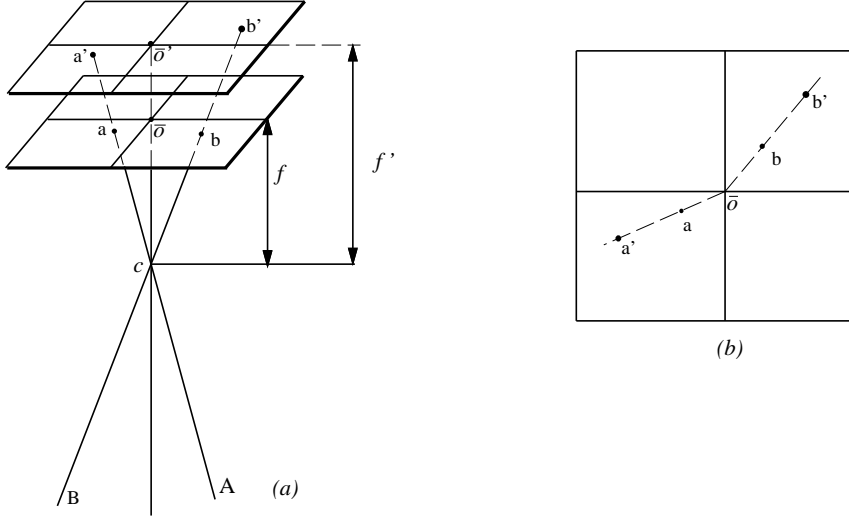


Figure 13: *Determine principal point by focusing and/or zooming.*

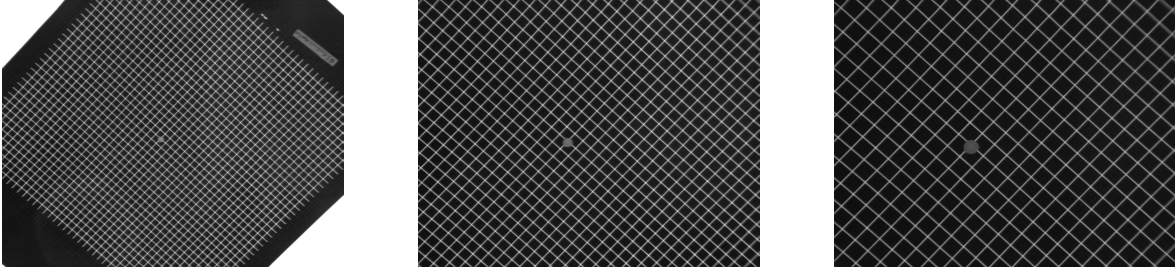


Figure 14: *A few images out of a sequences used for determining principal point by zooming.*

objects and known 3D points (Tab. 8 and 9). A Gaussian distributed noise ($\sigma = 0.2pix$) has been added on image data. The first object is the 3D cube used in calibration experiment. The second one corresponds to a flat pattern moving along a straight line. In Tab. 8 and 9, the first column corresponds to the true parameters used to construct synthetic data. The second column shows the parameters computed by the algorithm without any constraint on (x_0, y_0) coordinates. In the last three columns a systematic error has been added to (x_0, y_0) coordinates and used as fixed (known) parameters in the calibration to compute the rest of parameters. The last row $\hat{\sigma}_0$ represents the estimated standard error of unit weight.

We can note from both experiments, the weak influence of (x_0, y_0) coordinates in calibration parameters results. Its effect, on the residual errors σ_0 , is less than 0.05 pixel for the worst condition (more than 5 pixel from the true solution).

6.1.4 Conclusion

Experience in calibration algorithms shows that the principal point is highly correlated to the other calibration parameters. Its computation often leads to unstable results. So, given the relative stability obtained for (x_0, y_0) coordinates in section 6.1.2 and their weak influence on calibration results (for a position not so far from the true one, of course !), we think that using zooming properties is an efficient way to compute *independently* the principal point coordinates in case of zoom-lens. This method was successfully used in [8, 9] and we will assume in the following calibration experiments that the principal point coordinates are already computed.

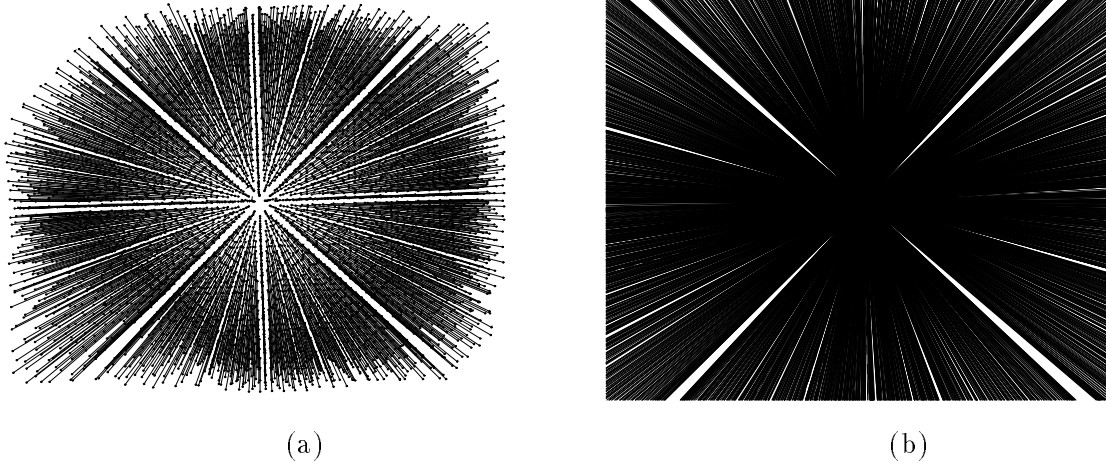


Figure 15: *Determining principal point by zooming. (a) Line segments obtained by tracing corresponding points through the sequence of images; (b) common intersection point of line segments.*

Table 7: *Results of calibrating principal-point by zooming. Second experiment, 11 views taken after a time interval of 3 weeks.*

focus step	zoom combination	zoom		zoom back	
		x_0	y_0	x_0	y_0
0.43	1-3	369.54	284.31	369.60	284.56
	2-4	369.33	284.21	368.60	284.55
	3-5	369.46	284.28	369.82	284.74
	4-6	369.79	284.09	370.42	284.62
	5-7	369.91	283.89	370.28	284.51
	6-8	370.10	283.75	369.84	283.77
	7-9	370.14	283.53	369.44	282.88
	8-10	369.95	283.60	369.53	282.62
	9-11	369.67	283.46	369.17	282.28
	1-11	369.59	283.98	369.75	284.21

6.2 Least Squares Calibration

In the following we present results of various experiments realized in zoom-lens calibration. All results shown here are done using fixed principal point calibrated by zooming. As lenses are moving when zoom or focus setting is modified, calibration is performed for each combination of these two parameters. A special paragraph is devoted to iris influence on calibration parameters. For technical purpose, the motor controllers allow to use range command values between (1-32) for each setting.

- The focus step 1 corresponds to the closest distance between the object and the camera (0.8m). The step 32 corresponds to the infinity.
- Zoom steps 1 and 32 are respectively the minimum and maximum magnification.
- Iris steps 1 and 32 are respectively the closed and open position.

Note that the residual value σ_0 (difference between measurements and data estimated by calibration parameters) achieved by the calibration algorithm is lower than 0.16pixel . 137 calibration points were used in this experiment.

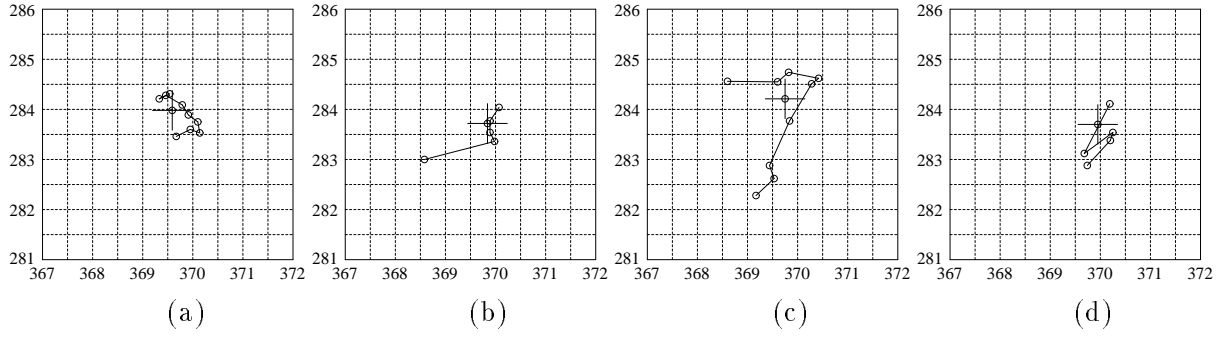


Figure 16: *Shifting of principal point under zooming (a-c) and zooming back (b-d). Two experiments done at different focus setups and times.*

Table 8: *Principal-point influence. (3D cube standard)*

	<i>true value</i>	<i>full calibration</i>	$d = (1\sqrt{2})$	$d = (2\sqrt{2})$	$d = (4\sqrt{2})$
x_0 (pix)	370.00	368.73	371.00	372.00	374.00
y_0 (pix)	370.00	370.53	371.00	372.00	374.00
f_x (pix)	2000.00	1999.04	1999.19	1999.34	1999.67
f_y (pix)	2000.00	1998.84	1999.04	1999.25	1999.69
T_x (mm)	600.00	599.53	599.66	599.79	600.07
T_y (mm)	700.00	699.85	700.03	700.22	700.61
T_z (mm)	600.00	599.47	599.52	599.57	599.68
R_α (deg)	-45.00	-45.02	-44.97	-44.93	-44.85
R_β (deg)	30.00	29.99	29.98	29.97	29.96
R_γ (deg)	30.00	30.01	29.99	29.96	29.92
σ_0 (pix)	0.190	0.190	0.191	0.193	0.200

6.2.1 Focal Lengths (f_x, f_y) and Pixel Ratio ($\frac{f_y}{f_x}$)

The experiment is done as follows: keeping iris fixed, calibration has been performed with different zoom and focus settings. As it can be noticed on Figure 17 all focus values were not usable in experiments. The reference object was too small to be detected and used for calibration purpose, after range 22 of the focus setting. Actually, the two 3D cubes have been placed in a range from $0.8m$ from the camera up to $4.5m$. These distances correspond to the action area where the KTH-head is supposed to work.

Figure 18, 19 and 20 show the calibrated focal length f_x and f_y as a function of zoom and focus motors steps. From the figures, we can noticed the follows.

- First, the focal-length changes as a continuous function of both zoom and focus parameters.
- The calibration results for f_x and f_y are almost the same and the pixel ratio ($\frac{f_y}{f_x}$) is equal to 0.9999 . This value, close to 1.0, is due to a special way the video signal is sampled. The sampling frequency is set in the frame grabber in order to obtain a square pixel.
- Calibration results change when the focus is modified. Focus effect is not so important for short focal-length but increases greatly for long ones. (See Figure 19 and 20 left part).

Construction of the look-up table:

Table 9: *Principal-point influence. (3D moving pattern)*

	<i>true value</i>	<i>full calibration</i>	$d = (1\sqrt{2})$	$d = (2\sqrt{2})$	$d = (4\sqrt{2})$
x_0 (pix)	350.00	350.19	351.00	352.00	354.00
y_0 (pix)	350.00	349.43	351.00	352.00	354.00
f_x (pix)	1000.00	999.75	999.76	999.77	999.81
f_y (pix)	1000.00	999.72	999.73	999.74	999.78
T_x (mm)	0.00	0.12	0.12	0.13	0.13
T_y (mm)	0.00	-0.08	-0.08	-0.08	-0.07
T_z (mm)	2500.00	2499.48	2499.49	2499.52	2499.64
R_α (deg)	0.00	-0.03	0.06	0.11	0.23
R_β (deg)	0.00	-0.01	-0.06	-0.11	-0.22
R_γ (deg)	0.00	0.00	0.00	0.00	0.00
σ_0 (pix)	0.190	0.191	0.196	0.207	0.243

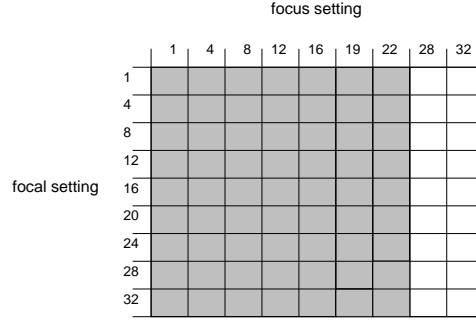


Figure 17: *Zoom-focus combination used for KTH head-eye zoom calibration*

From Figure 18 several methods can be used to compute the look-up-function and index focal length at any zoom-focus step. The simplest one is to make a local approximation (linear or bi-cubic) or to try to fit data (Table 10) by a spline interpolated surface.

6.2.2 Distortion

Figure 21 shows the results of the calibrated radial distortion with different zooming steps. The distortion is calibrated by setting the distortion to zero at $r_0 = 200$ pixels, i.e., the radial distance from the principal point. Notice by varying r_0 , we will get different distortion, eventually, different focal length.

We should point out that at lower zoom step (< 4), the calibration object projected on the image covers only a small part at the center, we cannot determine the distortion outside this region, rather than extrapolate them.

As can be seen from the figure, that camera's distortion varies as zoom-in/out, but for *certain applications*, it can be ignored within the center part of image (400×400 pixels), as it is less than 0.5 pixels.

Figure 22 shows the calibrated tangential distortions at different zooming steps. The distortion varies as zooming in or out. Notice here too, at the small zoom steps, the object covers only the center part of the image, so at the border the distortion is only an extrapolation. As previously said, a large test-field would probably lead to more accurate distortion results.

Table 10: Calibration results of focal length f_x (pix)

zoom step	focus step						
	1	4	8	12	16	19	22
1	1311.63	1283.38	1265.47	1245.11	1225.18	1226.65	1206.83
4	1487.70	1455.83	1444.05	1407.11	1386.49	1370.34	1351.43
8	1777.25	1733.63	1724.83	1680.90	1656.53	1629.45	1611.45
12	2165.44	2102.11	2082.24	2058.55	1947.01	1977.93	1946.80
16	2682.09	2624.70	2552.57	2530.42	2479.55	2416.01	2339.36
20	3639.99	3520.89	3263.93	3180.72	3127.93	3075.01	2999.92
24	4805.30	4664.28	4480.05	4123.18	4117.78	4026.74	3883.77
28	6730.21	6613.74	6332.97	6076.27	5736.30	5553.32	5351.06
32	11301.14	10517.06	9382.15	9147.95	8243.51	8002.65	7801.27

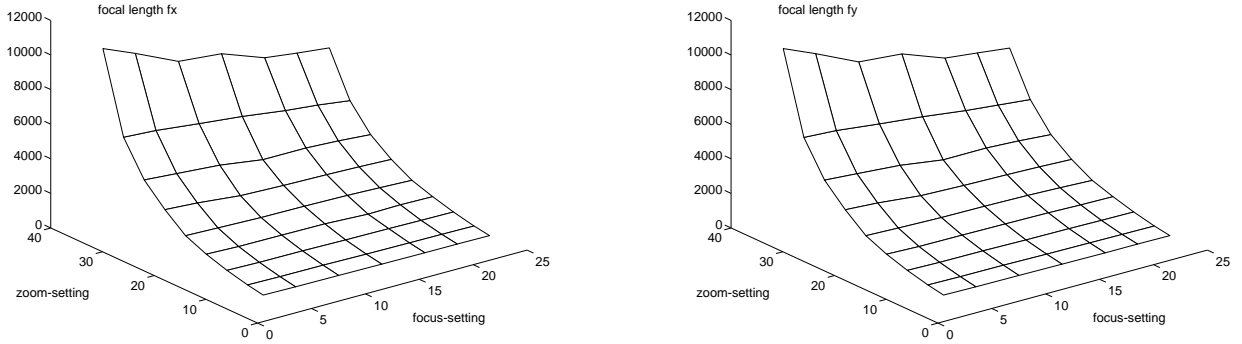


Figure 18: Calibration results of focal length (f_x, f_y) for the zoom-lens mounted on KTH active-head. The horizontal axes are zoom-focus motor steps and the vertical axe is the focal length in x and y directions expressed in pixel.

6.2.3 Does Iris change the Focal Length ?

We know that both zooming and focusing change the focal length under the *pin-hole* model. The question is does iris also change focal length? Table 11 shows the experimental results of the calibrated focal length under different iris openings. The table shows that iris does not changed the focal length systematically. The differences may simply come from measurement errors. Iris opening changes the brightness of the image. It tends to amplify images of bright objects with dark background when increasing (open) iris and vice verse. But the amplification is a local scaling referred to the center point of the object, so it does not shift nor scale the image globally.

Table 11: Results of calibrated local length under different iris openings

iris	$f_x(\text{pix})$	$\sigma_{f_x}(\text{pix})$
close	2008.574	6.039
	2009.312	6.124
	2005.799	6.187
	2005.865	6.218
↓	2007.321	6.379
	2007.529	6.397
	2009.249	6.677
	2008.732	7.739
open	2007.367	7.751
	2009.648	8.003

Table 12: Results of calibrated rotation angles (in degree) at different steps.

zoom step	angles in degree		
	α	β	γ
1	-37.976	35.069	25.481
4	-37.928	35.085	25.477
8	-37.974	35.124	25.507
12	-38.009	35.111	25.524
16	-37.956	35.051	25.487
20	-37.939	35.096	25.461
24	-37.919	35.168	25.455
28	-37.939	35.116	25.486
32	-37.952	35.144	25.458

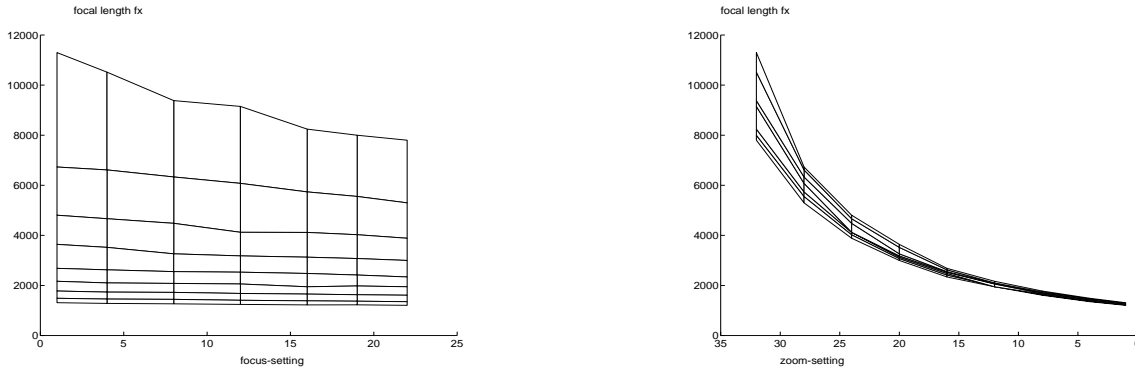


Figure 19: Details of the focal length surface f_x .

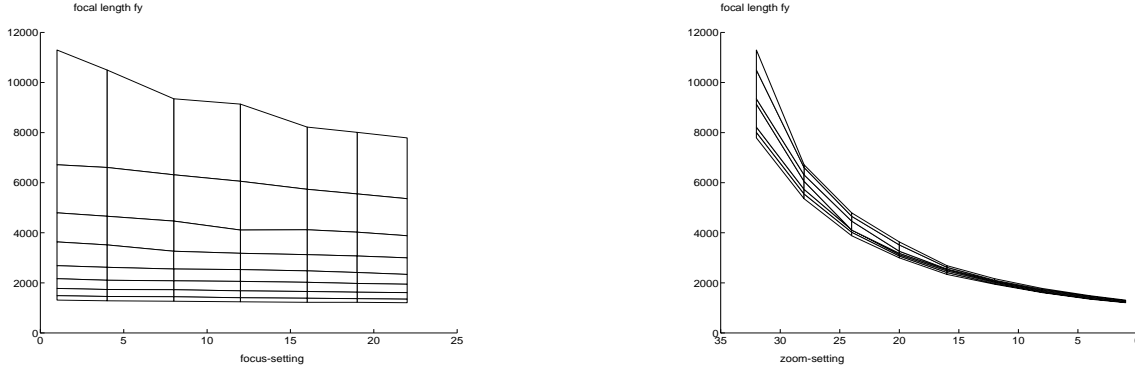


Figure 20: Details of the focal length surface f_y .

6.2.4 Optical Center Translation in W-Frame-Reference

Figure 23 shows the movement of optical center in space when zooming. We can see the optical center travels (approximately) in a straight line in space (*the optical axis*), but the range ($354.1mm$) does not corresponding to the range of the focal length.

This is due to the pin-hole model used for the zoom lens, which does not hold. In case of zoom lens the distance between the object and the principal plane (CCD matrix) can not be consider as a fixed parameter. A complete description of this phenomenon is given in [8], [9] and shows that a thick optical model has to be consider to understand it.

Table 12 shows the results of the computed rotation angles α , β , and γ . We can see that they are very close; it proves that the virtual displacement of the object in the camera-frame-reference is almost a translation.

7 Conclusion

Through this article, we have tried to describe as accurately as possible all the keys points that needed to be checked in order to achieve good calibration results.

Calibration experiments does not only depend on the mathematical tools used to compute intrinsic and extrinsic parameters. Considerable attention has to be given to the camera and frame grabber stability, to the construction and accuracy of the 3-D reference object and finally to the way the calibration points are measured and detected in images.

The calibration algorithm described in the paper is not a new one and has been already used in photogrammetry. A least squares technique is applied to minimize measurement errors in images and compute simultaneously intrinsic, extrinsic as well as radial and tangential distortion parameters. Statistical analysis can also be performed using residual and covariance matrix.

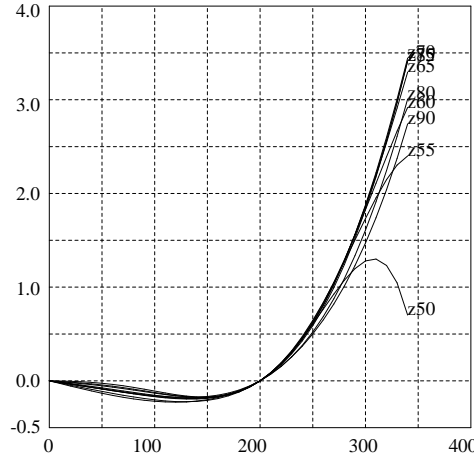


Figure 21: *Recovered radial distortions: The horizontal axe is the distances from the principal point (in pixel) and vertical axe is the distortion (in pixel).*

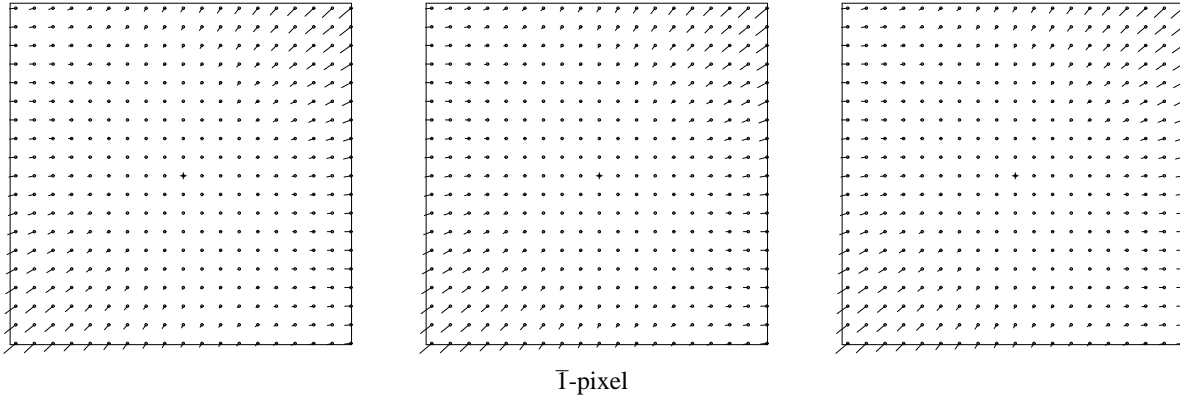


Figure 22: *Recovered tangential distortions at different zooming steps. The plotting covers a area of 480×480 pixels around the center.*

We think that calibration presented in the article could be slightly improved first using a numerical camera (we have tried to improve existing materials but noise introduced by camera and grabber is still considerable) and secondly constructing a test-field as calibration object reference. This last point, in case of zoom lens application, would allow us to use a larger working calibration space and also to get a more accurate distortion estimation.

Finally, a lot of calibration algorithms exist in the literature, each of them has advantages and drawbacks. Making a choice greatly depends on the vision task to be performed! We hope that the reader will find through our experiments some interesting ideas to improve its own calibration results.

References

- [1] American Society for Photogrammetry. *Manual of Photogrammetry*, 4th edition, 1984.
- [2] E.P. Baltsavias, H.A. Beyer, D. Fritsch, and R.K. Lenz. Fundamentals of real-time photogrammetry. *Tutorial notes of ISPRS Commission V Symposium, Zurich, Switzerland*, September 1990.

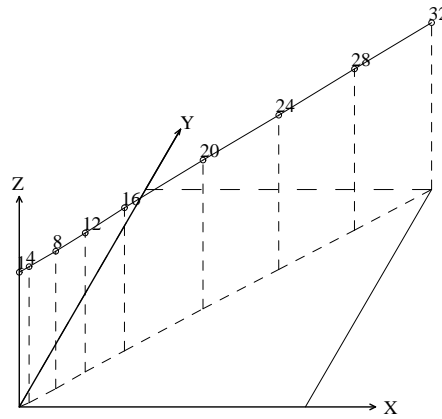


Figure 23: *Movement of optical center in space under zooming (W-frame-reference (mm)): moving range (861.1, 709.7, 959.9) \rightarrow (1065.4, 889.1, 1186.3), $D_l = 354.1\text{mm}$*

- [3] H.A. Beyer. *Geometric and Radiometric Analysis of a CCD-Camera Based Photogrammetric Close-Range System*. PhD thesis, Institut für Geodäsie und Photogrammetrie, Nr 51, ETH, Zurich, May 1992.
- [4] D.C. Brown. Decentering distortion of lenses. *Photogrammetric Engineering*, 32(3), 1966.
- [5] D.C. Brown. Close-range camera calibration. *Photogrammetric Engineering*, 32(8):855–866, 1971.
- [6] J. Dahler. Problems in digital image acquisition with CCD camera. In *Proc. of ISPRS Intercommision Conference*, pages 48–59, Interlaken, Switzerland, 1987.
- [7] A. Horii. Depth from defocusing. Technical Report ISRN KTH/NA/P--92/16--SE, Dept. of Numerical Analysis and Computing Science, KTH (Royal Institute of Technology), June 1992.
- [8] J. M. Lavest, G. Rives, and M. Dhome. 3D reconstruction by zooming. *IEEE Robotics and Automation*, 9(2):196–207, April 1993.
- [9] J.M. Lavest, G. Rives, and M. Dhome. Modelling a solid of revolution by zooming. *IEEE Transactions on Robotics and Automation*, 11(2):267–271, April 1995.
- [10] M. Li. Camera calibration of a head-eye system for active vision. In J.O. Eklundh, editor, *Proceedings of the 3'rd European Conference on Computer Vision*, pages I:543–554, Stockholm, May 1994.
- [11] M. Li, D. Betsis, and J.-M. Lavest. Kinematic calibration of the KTH head-eye system. Technical Report ISRN KTH/NA/P--94/34--SE, Dept. of Numerical Analysis and Computing Science, KTH (Royal Institute of Technology), November 1994.
- [12] K. Pahlavan and J.O. Eklundh. A head-eye system – analysis and design. *CVGIP: Image Understanding: Special Issue on Purposive, Qualitative and Active Vision*, J. Y. Aloimonos Ed., 56(1):41–56, July 1992.
- [13] K. Pahlavan, T. Uhlin, and J.O. Eklundh. Active vision as a methodology. In J. Y. Aloimonos, editor, *Active Perception*, Advances in Computers. Lawrence-Erlbaum, 1993.
- [14] B. Peuchot. Camera virtual equivalent model 0.01 pixel detector. *Computerized Medical Imaging and Graphics*, 17(4/5):289–294, July 1993.

- [15] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C, second edition*. Cambridge University Press, 1992.
- [16] K. Tarabanis, R. Y. Tsai, and D. S. Goodman. Modeling of a computer-controlled zoom lens. In *IEEE International Conference on Robotics and Automation*, pages 1545–1551, Nice, France, May 1992.
- [17] K. Torlegård. *On the Determination of Interior Orientation of Close-up Cameras under Operational Conditions using 3D test Objects*. PhD thesis, Dept. of Photogrammetry, Royal Institute of Technology, Stockholm, 1967.
- [18] K. Torlegård. Accuracy improvement in close range photogrammetry. *Schriftenreihe, Wissenschaftlicher Studiengang Vermessungswesen, Hochschule der Bundeswehr München*, Heft 5, September 1981.
- [19] R. Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, RA-3(4):323–331, August 1987.
- [20] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 14(10):965–980, October 1992.

A Linearization of Collinearity Equations

The linearization of the collinear equations (8) with respect to Φ at $\Phi = \Phi_0$ lead to the following:

$$\left. \begin{aligned} x + v_x &= P(\Phi_0) + \sum_{i=1}^t \frac{\partial P}{\partial \phi_i} \delta \phi_i \\ y + v_y &= Q(\Phi_0) + \sum_{i=1}^t \frac{\partial Q}{\partial \phi_i} \delta \phi_i \end{aligned} \right\} \quad (18)$$

where $\frac{\partial P}{\partial \phi_i}$ is the partial derivatives to ϕ_i at $\phi_i = \tilde{\phi}_i$ and the $\tilde{\cdot}$ denotes the initial value for individual parameter, and t is the number of parameters, i.e., $t = 15$. Let

$$\begin{aligned} \tilde{x} &= x - \tilde{x}_0; \quad \tilde{y} = y - \tilde{y}_0 \\ \tilde{X} &= X - \tilde{T}_x; \quad \tilde{Y} = Y - \tilde{T}_y; \quad \tilde{Z} = Z - \tilde{T}_z \\ N_x &= \tilde{r}_{11}(X - \tilde{T}_x) + \tilde{r}_{12}(Y - \tilde{T}_y) + \tilde{r}_{13}(Z - \tilde{T}_z) \\ N_y &= \tilde{r}_{21}(X - \tilde{T}_x) + \tilde{r}_{22}(Y - \tilde{T}_y) + \tilde{r}_{23}(Z - \tilde{T}_z) \\ N_z &= \tilde{r}_{31}(X - \tilde{T}_x) + \tilde{r}_{32}(Y - \tilde{T}_y) + \tilde{r}_{33}(Z - \tilde{T}_z) \end{aligned}$$

The partial derivatives are computed as following:

$$\left. \begin{aligned}
\frac{\partial P}{\partial \phi_1} &= \frac{\partial P}{\partial x_0} = 1 - \tilde{a}_1[r^2 + 2\tilde{x}^2] - \tilde{a}_2 r^2[r^2 + 4\tilde{x}^2] - \tilde{a}_3 r^4[r^2 + 6\tilde{x}^2] - 2(3p_1\tilde{x} + p_2\tilde{y}) \\
\frac{\partial P}{\partial \phi_2} &= \frac{\partial P}{\partial y_0} = -2[\tilde{x}\tilde{y}(\tilde{a}_1 + 2\tilde{a}_2 r^2 + 3\tilde{a}_3 r^4) + p_1\tilde{y} + p_2\tilde{x}] \\
\frac{\partial P}{\partial \phi_3} &= \frac{\partial P}{\partial a_1} = \tilde{x}r^2 \\
\frac{\partial P}{\partial \phi_4} &= \frac{\partial P}{\partial a_2} = \tilde{x}r^4 \\
\frac{\partial P}{\partial \phi_5} &= \frac{\partial P}{\partial a_3} = \tilde{x}r^6 \\
\frac{\partial P}{\partial \phi_6} &= \frac{\partial P}{\partial p_1} = r^2 + 2\tilde{x}^2 \\
\frac{\partial P}{\partial \phi_7} &= \frac{\partial P}{\partial p_2} = 2\tilde{x}\tilde{y} \\
\frac{\partial P}{\partial \phi_8} &= \frac{\partial P}{\partial f_x} = -\frac{N_x}{N_z} \\
\frac{\partial P}{\partial \phi_9} &= \frac{\partial P}{\partial f_y} = 0 \\
\frac{\partial P}{\partial \phi_{10}} &= \frac{\partial P}{\partial T_x} = \frac{\tilde{f}_x}{N_z^2}(N_z\tilde{r}_{11} - N_x\tilde{r}_{31}) \\
\frac{\partial P}{\partial \phi_{11}} &= \frac{\partial P}{\partial T_y} = \frac{\tilde{f}_x}{N_z^2}(N_z\tilde{r}_{12} - N_x\tilde{r}_{32}) \\
\frac{\partial P}{\partial \phi_{12}} &= \frac{\partial P}{\partial T_z} = \frac{\tilde{f}_x}{N_z^2}(N_z\tilde{r}_{13} - N_x\tilde{r}_{33}) \\
\frac{\partial P}{\partial \phi_{13}} &= \frac{\partial P}{\partial \alpha} = \frac{\tilde{f}_x}{N_z^2}\{N_z[\tilde{r}_{13}\tilde{Y} - \tilde{r}_{12}\tilde{Z}] - N_x[\tilde{r}_{33}\tilde{Y} - \tilde{r}_{32}\tilde{Z}]\} \\
\frac{\partial P}{\partial \phi_{14}} &= \frac{\partial P}{\partial \beta} = \frac{\tilde{f}_x}{N_z^2}\{N_z^2 \cos \tilde{\gamma} + N_x[\cos \tilde{\beta}\tilde{X} - \tan \tilde{\beta}(\tilde{r}_{32}\tilde{Y} + \tilde{r}_{33}\tilde{Z})]\} \\
\frac{\partial P}{\partial \phi_{15}} &= \frac{\partial P}{\partial \gamma} = -\frac{\tilde{f}_x}{N_z}N_y
\end{aligned} \right\}$$

$$\left. \begin{aligned}
\frac{\partial Q}{\partial \phi_1} &= \frac{\partial Q}{\partial x_0} = -2[\tilde{x}\tilde{y}(\tilde{a}_1 + 2\tilde{a}_2 r^2 + 3\tilde{a}_3 r^4) + p_1\tilde{y} + p_2\tilde{x}] \\
\frac{\partial Q}{\partial \phi_2} &= \frac{\partial Q}{\partial y_0} = 1 - \tilde{a}_1[r^2 + 2\tilde{y}^2] - \tilde{a}_2 r^2[r^2 + 4\tilde{y}^2] - \tilde{a}_3 r^4[r^2 + 6\tilde{y}^2] - 2(p_1\tilde{x} + 3p_2\tilde{y}) \\
\frac{\partial Q}{\partial \phi_3} &= \frac{\partial Q}{\partial a_1} = \tilde{y}r^2 \\
\frac{\partial Q}{\partial \phi_4} &= \frac{\partial Q}{\partial a_2} = \tilde{y}r^4 \\
\frac{\partial Q}{\partial \phi_5} &= \frac{\partial Q}{\partial a_3} = \tilde{y}r^6 \\
\frac{\partial P}{\partial \phi_6} &= \frac{\partial P}{\partial p_1} = 2\tilde{x}\tilde{y} \\
\frac{\partial P}{\partial \phi_7} &= \frac{\partial P}{\partial p_2} = r^2 + 2\tilde{y}^2 \\
\frac{\partial Q}{\partial \phi_8} &= \frac{\partial Q}{\partial f_x} = 0 \\
\frac{\partial Q}{\partial \phi_9} &= \frac{\partial Q}{\partial f_y} = -\frac{N_y}{N_z} \\
\frac{\partial Q}{\partial \phi_{10}} &= \frac{\partial Q}{\partial T_x} = \frac{\tilde{f}_y}{N_z^2}(N_z\tilde{r}_{21} - N_y\tilde{r}_{31}) \\
\frac{\partial Q}{\partial \phi_{11}} &= \frac{\partial Q}{\partial T_y} = \frac{\tilde{f}_y}{N_z^2}(N_z\tilde{r}_{22} - N_y\tilde{r}_{32}) \\
\frac{\partial Q}{\partial \phi_{12}} &= \frac{\partial Q}{\partial T_z} = \frac{\tilde{f}_y}{N_z^2}(N_z\tilde{r}_{23} - N_y\tilde{r}_{33}) \\
\frac{\partial Q}{\partial \phi_{13}} &= \frac{\partial Q}{\partial \alpha} = \frac{\tilde{f}_y}{N_z^2}\{N_z[\tilde{r}_{23}\tilde{Y} - \tilde{r}_{22}\tilde{Z}] - N_y[\tilde{r}_{33}\tilde{Y} - \tilde{r}_{32}\tilde{Z}]\} \\
\frac{\partial Q}{\partial \phi_{14}} &= \frac{\partial Q}{\partial \beta} = -\frac{\tilde{f}_y}{N_z^2}\{N_z^2 \sin \tilde{\gamma} - N_y[\cos \tilde{\beta}\tilde{X} - \tan \tilde{\beta}(\tilde{r}_{32}\tilde{Y} + \tilde{r}_{33}\tilde{Z})]\} \\
\frac{\partial Q}{\partial \phi_{15}} &= \frac{\partial Q}{\partial \gamma} = \frac{\tilde{f}_y}{N_z}N_x
\end{aligned} \right\}$$

The contents of the matrices in (9) for n given points are:

$$\mathbf{A} = \begin{pmatrix} \frac{\partial P^1}{\partial \phi_1} & \frac{\partial P^1}{\partial \phi_2} & \frac{\partial P^1}{\partial \phi_3} & \frac{\partial P^1}{\partial \phi_4} & \frac{\partial P^1}{\partial \phi_5} & \frac{\partial P^1}{\partial \phi_6} & \frac{\partial P^1}{\partial \phi_7} & \frac{\partial P^1}{\partial \phi_8} & \frac{\partial P^1}{\partial \phi_9} & \frac{\partial P^1}{\partial \phi_{10}} & \frac{\partial P^1}{\partial \phi_{11}} & \frac{\partial P^1}{\partial \phi_{12}} & \frac{\partial P^1}{\partial \phi_{13}} & \frac{\partial P^1}{\partial \phi_{14}} & \frac{\partial P^1}{\partial \phi_{15}} \\ \frac{\partial Q^1}{\partial \phi_1} & \frac{\partial Q^1}{\partial \phi_2} & \frac{\partial Q^1}{\partial \phi_3} & \frac{\partial Q^1}{\partial \phi_4} & \frac{\partial Q^1}{\partial \phi_5} & \frac{\partial Q^1}{\partial \phi_6} & \frac{\partial Q^1}{\partial \phi_7} & \frac{\partial Q^1}{\partial \phi_8} & \frac{\partial Q^1}{\partial \phi_9} & \frac{\partial Q^1}{\partial \phi_{10}} & \frac{\partial Q^1}{\partial \phi_{11}} & \frac{\partial Q^1}{\partial \phi_{12}} & \frac{\partial Q^1}{\partial \phi_{13}} & \frac{\partial Q^1}{\partial \phi_{14}} & \frac{\partial Q^1}{\partial \phi_{15}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial P^n}{\partial \phi_1} & \frac{\partial P^n}{\partial \phi_2} & \frac{\partial P^n}{\partial \phi_3} & \frac{\partial P^n}{\partial \phi_4} & \frac{\partial P^n}{\partial \phi_5} & \frac{\partial P^n}{\partial \phi_6} & \frac{\partial P^n}{\partial \phi_7} & \frac{\partial P^n}{\partial \phi_8} & \frac{\partial P^n}{\partial \phi_9} & \frac{\partial P^n}{\partial \phi_{10}} & \frac{\partial P^n}{\partial \phi_{11}} & \frac{\partial P^n}{\partial \phi_{12}} & \frac{\partial P^n}{\partial \phi_{13}} & \frac{\partial P^n}{\partial \phi_{14}} & \frac{\partial P^n}{\partial \phi_{15}} \\ \frac{\partial Q^n}{\partial \phi_1} & \frac{\partial Q^n}{\partial \phi_2} & \frac{\partial Q^n}{\partial \phi_3} & \frac{\partial Q^n}{\partial \phi_4} & \frac{\partial Q^n}{\partial \phi_5} & \frac{\partial Q^n}{\partial \phi_6} & \frac{\partial Q^n}{\partial \phi_7} & \frac{\partial Q^n}{\partial \phi_8} & \frac{\partial Q^n}{\partial \phi_9} & \frac{\partial Q^n}{\partial \phi_{10}} & \frac{\partial Q^n}{\partial \phi_{11}} & \frac{\partial Q^n}{\partial \phi_{12}} & \frac{\partial Q^n}{\partial \phi_{13}} & \frac{\partial Q^n}{\partial \phi_{14}} & \frac{\partial Q^n}{\partial \phi_{15}} \end{pmatrix}$$

$$\mathbf{L}_{2n \times 1} = \begin{pmatrix} x_1 - P_1(\boldsymbol{\Phi}_0) \\ y_1 - Q_1(\boldsymbol{\Phi}_0) \\ \vdots \\ x_n - P_n(\boldsymbol{\Phi}_0) \\ y_n - Q_n(\boldsymbol{\Phi}_0) \end{pmatrix}, \quad \mathbf{V}_{2n \times 1} = \begin{pmatrix} v_{x_1} \\ v_{y_1} \\ \vdots \\ v_{x_n} \\ v_{y_n} \end{pmatrix}$$

B Least Squares Solution of System: $\mathbf{V} = \mathbf{A}\boldsymbol{\Delta\Phi} - \mathbf{L}$

The solution to (9) through the minimization of $\mathbf{V}^T \mathbf{W} \mathbf{V}$ is a classic linear least squares problem. Let $\boldsymbol{\Omega} = \mathbf{V}^T \mathbf{W} \mathbf{V}$. At the minimal of $\boldsymbol{\Omega}$, we have $\frac{\partial \boldsymbol{\Omega}}{\partial \boldsymbol{\Phi}} = 0$, i.e.,

$$\frac{\partial \boldsymbol{\Omega}}{\partial \boldsymbol{\Phi}} = 2\mathbf{V}^T \mathbf{W} \frac{\partial \mathbf{V}}{\partial \boldsymbol{\Phi}} = 2\mathbf{V}^T \mathbf{W} \mathbf{A} = 0, \quad \implies \quad \mathbf{A}^T \mathbf{W} \mathbf{V} = 0$$

Replacing \mathbf{V} from (9), the above equations become:

$$\mathbf{A}^T \mathbf{W} (\mathbf{A} \boldsymbol{\Delta\Phi} - \mathbf{L}) = \mathbf{A}^T \mathbf{W} \mathbf{A} \boldsymbol{\Delta\Phi} - \mathbf{A}^T \mathbf{W} \mathbf{L} = 0$$

which leads to the solution of $\boldsymbol{\Delta\Phi}$:

$$\boldsymbol{\Delta\Phi} = (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{W} \mathbf{L})$$