TECHNISCHE
UNIVERSITÄT
WIEN
Vienna University of Technology

DIPLOMARBEIT

Integration of relative and global orientation methods using a moving ToF camera

ausgeführt am
Department für Geodäsie und Geoinformation
Forschungsgruppe Photogrammetrie

unter der Anleitung von
Univ.Prof. Dipl.-Ing. Dr.techn. Norbert Pfeifer
Dipl.-Ing. Wilfried Karel

durch

BSc. Sabine Horvath

Matr.-Nr.: 0751881
sabine.horvath@gmx.at

Wien, September 2014

Abstract

The aim of this work is the estimation and enhancement of a trajectory formed by a moving ToF camera. The camera pose has to be known for each image in order to reconstruct the object in a common reference frame, recorded from multiple views. The quality of the model certainly depends on the quality of the orientation. This work deals with the reachable precision of the camera trajectory and with its improvement potential.

The camera poses are estimated by means of the global orientation method - the bundle adjustment - on the one hand and by the relative orientation method, which bases on the range and optical flow in the image, on the other hand. The bundle adjustment produces unbiased, but noisy estimates. In contrast, the relative orientation shows a smooth trajectory with a large accumulated error. Those deficiencies of each method led to the formulation of the hypothesis, that the integration of those methods results in a smooth unbiased estimate of the trajectory.

The bundle adjustment is an easily adaptable method, which simultaneously estimates all unknown parameters. Therefore, the relative orientation is integrated in the bundle adjustment. The relative orientation result, consisting of the difference translation and the difference rotation between two subsequent frames, is expressed as a function of the parameters of the bundle adjustment. The stochastic model is extended by the covariance matrix of the relative orientation. The range measurements, performed by the ToF camera, are introduced into the model as well. The weighting of the different observation types is accomplished by the variance component estimation (VCE).

The approach is verified on two data sets. Saturation and motion blur mainly influenced the data acquisitions. The imperfections of the data recorded by a ToF camera are discussed as well.

Data set one delivers precise bundle adjustment results, but low quality relative orientations because of noise and less contrast in the texture. The integration of the relative orientation does not show much improvement because of already precise bundle adjustment results.

The bundle adjustment results of data set two are of lower precision, which is likely caused by motion blur. However, due to less noise and a higher contrast in the images, better results of the relative orientation are achieved. The integration of both methods decreases the trajectory error by a factor of 14 and the precision of the parameters improved by a factor of 3. However, the relative orientation results are not Gaussian distributed, thus the preconditions for an optimal estimate and the application of the VCE are not given. Hence the results, may not be adopted uncritically.

Although gross errors have been eliminated and systematic errors have been regarded, the distribution of the range residuals does not correspond to a normal distribution. The trajectory, derived by integrating the relative orientation and the ranges, has an offset. Due to not considered systematic range errors the observations got deformed and the camera trajectory is moved towards the object.

Kurzfassung

Das Ziel dieser Arbeit ist die Bestimmung und Verbesserung der Trajektorie einer bewegten Distanz-Kamera. Um ein Objekt mittels Bildern rekonstruieren zu können, muss die Position und die Stellung der Kamera im Raum zum Zeitpunkt der Aufnahme bekannt sein. Die Genauigkeit des geometrischen Modells hängt von der Genauigkeit der äußeren Orientierung der Kamera ab, deswegen widmet sich diese Arbeit der Schätzung und Genauigkeitssteigerung der Kamera-Trajektorie.

Einerseits wird die Kamera-Trajektorie diskret auf Basis der äußeren Orientierungen der Bündelblockausgleichung hergestellt, andererseits liefert der Range Optical Flow in den Amplituden- und Distanzbildern eine relative Orientierung zwischen aufeinanderfolgenden Bildern. Durch Aufsummierung der einzelnen relativen Orientierungen erhällt man ebenfalls die Kamera-Trajektorie. Basierend auf einem Vergleich dieser Trajektorien, wird die These formuliert, dass die Integration beider Verfahren zu einer glatten, erwartungstreuen Trajektorie führt.

Das Modell der Bündelblockausgleichung lässt sich einfach erweitern, weshalb die relative Orientierung in diese Ausgleichung integriert wird. Dafür muss ein funktionaler Zusammenhang zwischen den relativen Orientierungs-Ergebnissen und den Parametern der Bündelblockausgleichung definiert werden. Das stochastische Modell wird ebenfalls um die Kovarianzmatrix der relativen Orientierung erweitert. Weiters werden die Distanzmessungen der ToF-Kamera in das Modell inkludiert. Die Gewichtung verschiedener Beobachtungstypen wird auf Basis der Varianzkomponentenschätzung durchgeführt.

Diese These wird anhand von zwei Datensätzen evaluiert, die wesentlich von Sättigungseffekten und Bewegungsunschärfe bestimmt sind; folglich werden auch die Unzulänglichkeiten des Messsystems behandelt.

Der erste Datensatz liefert präzise äußere Orientierungen der Bündelblockausgleichung, aber im Vergleich zum zweiten Datensatz schlechtere relative Orientierungen. Die Integration der beiden Orientierungsmethoden führt zu keiner sichtlichen Verbesserung der Trajektorie, weil die Ausgangsdaten bereits ein sehr hohes Genauigkeitsniveau aufweisen.

Die Bündelblockausgleichungsergebnisse des zweiten Datensatzes sind - wahrscheinlich durch Bewegungsunschärfe verursacht - um einen Faktor zwei ungenauer als jene des ersten. Die relative Orientierung ist präziser als beim ersten Datensatz aufgrund von geringerem Rauschen und einem höheren Bildkontrast. Die Integration der relativen Orientierung verbessert den Trajektorienfehler um einen Faktor 14 und die Präzision der äußeren Orientierung um einen Faktor 3. Da Voraussetzungen der kleinsten Quadrate Ausgleichung sowie der Varianzkomponentenschätzung nicht erfüllt sind, dürfen diese Ergebnisse nicht unreflektiert übernommen werden.

Obwohl grobe Fehler entfernt und systematische Fehler in den Distanzbeobachtungen berücksichtigt werden, entspricht die Verteilung der Verbesserungen nicht einer Normalverteilung. Auch die geschätzte Kamera-Trajektorie zeigt, dass noch systematische Fehler in den Distanzen enthalten sind.

Contents

1. Introduction

In many applications, measurement devices are mounted on a moving platform to determine the location of an object requested. For that purpose, the position and the attitude of the measurement device need to be estimated. This can be done by applying direct or indirect orientation methods. This work focuses on the indirect ones, which comprise marker-based and data-driven orientation methods. Both techniques are applied and integrated in further consequence.

In case of the marker-based approach, planar targets are placed in the field of view and based on these targets a spatial section can be calculated. This global orientation method is the most accurate one. However, the placement of the targets requires additional effort.

Data-driven approaches use geometry and other properties of the acquired data to determine a transformation in-between scans (Pfeifer *et al.*, 2013). Such a relative orientation method will be introduced.

The measurement device used is a Time-of-Flight (ToF) camera. Due to the underlying Time-of-Flight principle, the camera delivers an amplitude image and additionally a range image. Frame rates about 20 fps can be achieved. Therefore, the camera perfectly fits for motion estimation. The simultaneous range measurements facilitate a real-time point cloud generation.

A possible application of a moving ToF camera is to record an indoor environment and create a 3D model of it. The quality of the model certainly depends on the quality of the orientation. Therefore, a precise orientation of a ToF camera is needed. The question that will be answered in this work is, how precise a camera trajectory can be estimated? Additionally, how much is the gain, if two different methods are integrated? Another aim is to include the range measurements to sustain the orientation. However, the problems, which are encountered with this approach, will be discussed.

This topic will become more important in geodesy, if range cameras deliver more precise distance measurements. In robotics and computer vision it is already a hot topic and developments grow by leaps and bounds.

1.1 Motivation

Ghuffar *et al.* (2013) developed a relative orientation method based on the range and optical flow. This data-driven approach is only applicable for image sequences or similar images. The complete image content is used to determine the relative orientation of a camera.

The standard method in photogrammetry is the bundle adjustment, which bases upon markers on the object recorded. For that reason, images of various perspectives can be applied.

Evaluating a recorded data set of a first experiment with those two independent orientation techniques,

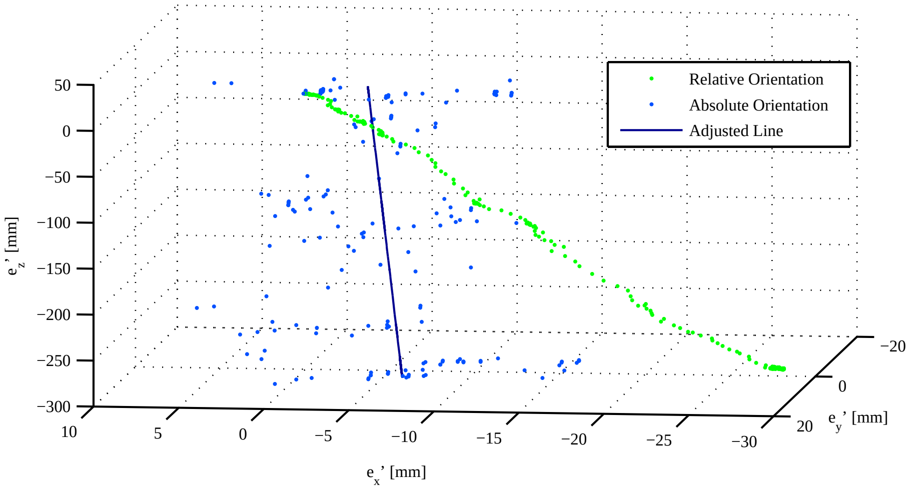

Figure 1.1: Trajectory of the moving ToF camera

leads to Figure 1.1. In that experiment, the camera was moved by 30 cm in $\boldsymbol{e}_z$-direction. The green graph shows the trajectory derived from the relative orientation and the blue dots show the camera positions estimated by the bundle adjustment. The relative orientation yields a smooth graph, but the error of each step sums up to a complete deviation of approximately 3 cm. On the contrary, the bundle adjustment supplies an unbiased result, easier recognisable by the adjusted line in dark blue. However, a low resolution of the image sensor and random errors cause the strong variation in the camera position. Consequently, the idea comes up, to combine those two methods, to receive a smooth unbiased estimate of the true trajectory.

The advantage of the relative orientation is, that no targets need to be placed in the field of view. However, this data-driven approach requires a variation in the texture in the amplitude and range image to recognise the motion of the camera and it is a relative method, i.e. it describes the motion between two consecutive images.
Due to the placed targets the results of the bundle adjustment are always in reference to the global coordinate frame. Hence, an integration of both methods will cancel out the weaknesses of each other.

1.2 Related work

This work is related to the dissertation of Ghuffar (2014). There are three intersections. Firstly, the used relative orientation has been derived and accomplished by Ghuffar *et al.* (2013) and has been made available to me. Data set one, which will be introduced in Chapter 3, has been acquired by Ghuffar (2014). The third intersection regards the integration of different methods, on which we both worked independently.
The camera calibration of Karel & Niedermayr (2010) has been used.
In Pfeifer *et al.* (2013) a good overview of the emerging range errors and the orientation methods of a ToF camera is given and Chapter 2.1.2 and 2.3 mainly base on it. The adjustment theory of Niemeier (2008) has been quoted.

1.3 Outline of the thesis

The thesis comprises various topics and the most important ones are explained in the fundamentals - starting with the range imaging technology, which determines the data generated, heading to the orientation methods and the adjustment theory, where a method, to integrate those two orientation approaches is introduced. Of importance is an unambiguous definition of the coordinate systems and the rotations, already with respect to the integration of both methods. Moreover, numerical aspects of solving large equation systems are treated. The evaluation of the trajectory error completes the fundamentals.
In Chapter 3 the two data sets recorded are presented. In the data preprocessing, gross errors are eliminated and systematic errors are considered. The utilisation of the two orientation methods on the acquired data is accomplished independently and a first comparison of the two methods is presented.
The integration of both methods is performed in the 4th chapter. The relative orientation results are included in the bundle adjustment as presented in the functional model. The stochastic model is extended by the covariance matrix of the relative orientation. Furthermore, the approximate values for the exterior orientations are determined. Questions of the implementation in OrientAl are discussed. The observation groups of the variance component estimation are specified and limitations of the method are treated.
In Chapter 5 the results are analysed and in Chapter 6 conclusions are met and an outlook is given.

The integration of both methods has been implemented in Python, an open-source programming language. The relative orientation, the data preprocessing and analysis are mainly computed in Matlab, a product of the Mathworks Inc. Equations are specially denoted by round brackets e.g. (2.1).

2. Fundamentals

The second chapter gives a theoretical introduction to the topic. It comprises the range imaging technology, the definition of coordinate systems and rotations as well as the underlying geometric model. The relative orientation method based on the range and optical flow and the global orientation method - the bundle adjustment is presented. Other aspects are the adjustment theory and the numerical equation solving methods used, which become highly relevant when solving large equation systems with 1,000 observations and more. For the comparison of different trajectories, an error measure is defined.

2.1 Range Imaging Technology

Range Imaging is a technique, which produces an image with pixel values that correspond to object distances. The principle works as follows: modulated infrared light emitted by the camera's internal lighting source is reflected by objects in the scene and travels back to the camera, where its precise time of arrival is derived simultaneously and independently at each sensor pixel.
The distance measurement is derived by the Time-of-Flight (ToF) principle. The measured quantity is the runtime of light. The two common modulation principles are the pulse modulation and the continuous wave modulation (CW). Due to cheaper and technical simpler components, the continuous wave modulation is the prevailing method. The camera used bases on the continuous wave modulation and the technique will be explained in more detail in Chapter 2.1.1.

The development of ToF cameras creates many new applications in various fields. It is an interesting tool for auto mobile navigation and safety. Targeted applications range from parking assistants and detecting obstacles on the road to recognising the attention or fatigue of drivers. Moreover, close range navigation of robotics is an important application and the entertainment industry further advances the development.
In photogrammetry, real-time 3D imaging becomes possible. The ToF camera combines the functionalities of a passive camera with those of a laser scanner, which results in simultaneous distance determination. Thus, dynamic scene analysis or motion estimation can be done through the use of a ToF camera. However, it is not adaptable for many potential applications, as many error sources are inherent to the system, as will be discussed later in Chapter 2.1.2. Such potential applications could be deformation measurements e.g. for landslide displacements. Currently, it can only be used for close-range applications because of the used modulation frequency and the limited power of the internal lighting source. The emitted power cannot be exceeded due to reasons of eye safety. Furthermore, in the presence of background light (artificial- or sunlight) a worse signal-to-noise ratio will be reached, which leads to a disturbance of the distance measurement. Thus, the utilisation is mainly restricted to indoor environments.

Figure 2.1: Swiss Ranger SR-3000

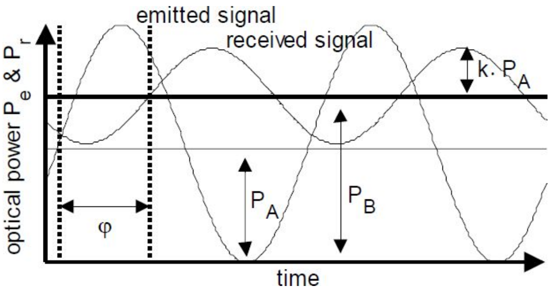

Figure 2.2: Emitted and received signal of a ToF camera (Buettgen *et al.*, 2005)

The technique providing the functionality to address each sensor pixel is called Complementary Metal-Oxide-Semiconductor (CMOS). Instead of transporting the charge across the chip and reading it at one corner of the array, as this is the case with a Charge Coupled Device (CCD) array, the CMOS technique uses several transistors at each pixel to amplify and move the charge. Therefore, it is possible to realise the Time-of-Flight measurement at each pixel. The advantages of using the CMOS technique are lower power consumption and cheaper production costs. However, the disadvantages of the CMOS array in comparison to the established CCD array are a poor signal-to-noise ratio, lower image quality and as well lower image resolution.

Mesa Imaging AG - a Swiss company - uses a combination of CCD and CMOS for their image sensors in order to benefit from the advantages of both. At about the same time the German PMD Technologies GmbH developed their products based solely on CMOS technology. Both ToF cameras base on the continuous wave modulation. The video game console Xbox bases on the Kinect, which allows to control the console by body movement. The Xbox has been sold 24 million times since 2010. The new Kinect will also be a ToF camera, which further pushes the development of this technique.

2.1.1 Functional principle of the camera used

The measurement device of choice is the SwissRanger SR-3000, a prototype of Mesa Imaging released in 2005. As shown in Figure 2.1, the internal light source is formed by an array of 55 light emitting diodes (LED). Each LED emits amplitude modulated near-infrared light with a central wavelength of 850 nm. The overall total mean power amounts to 1 Watt. The modulation frequency f_{mod} of 20 MHz corresponds to a wavelength λ_{mod} of 15 m, resulting in a maximum unambiguous range of 7.5 m. Remarkably, the SR-3000 captures data in real-time at video frame rates around 25 fps. The implemented sensor has QCIF resolution (176 x 144 pixels), where each pixel provides a well-controlled demodulation performance as described in Oggier *et al.* (2003) and Buettgen *et al.* (2005).

As mentioned before, the distance is determined by the continuous wave method, which exploits the phase delay φ between the emitted and received signal, as illustrated in Figure 2.2. The received signal is offset-shifted to a mean optical power P_B and the amplitude is reduced by a factor k. The additional offset B is mainly caused by background light and describes the total intensity. Due to optical losses the amplitude is reduced by the factor k. The incoming amplitude $A = k \cdot P_A$ relates to the signal-to-background ratio (signal-to-noise ratio). Hence, it is an accuracy measure (Buettgen *et al.*, 2005).

The demodulation is achieved by sampling the incoming optical signal four times per modulation period.

Each sample is shifted by 90 degrees. Thus, the modulation signal can be reconstructed unambiguously. Based on the four samples A0, A1, A2, A3, the three signal parameters total intensity B, incoming amplitude A and phase φ can be extracted as presented in Formulae (2.1) to (2.3).

$$B = \frac{A0 + A1 + A2 + A3}{4} \tag{2.1}$$

$$A = \frac{\sqrt{(A0 - A2)^2 + (A1 - A3)^2}}{2} \tag{2.2}$$

$$\varphi = arctan\left(\frac{A0 - A2}{A1 - A3}\right) \tag{2.3}$$

$$\rho = \frac{c}{4\pi \cdot f_{mod}} \cdot \varphi \tag{2.4}$$

The distance ρ between the image sensor and the object can be derived due to the phase delay φ. As written in Equation (2.4), the distance is a function of the four samples, the speed of light ($c = 299,792,459\frac{m}{s}$) and the modulation frequency f_{mod}.

The modulation frequency f_{mod} determines the maximum unambiguous range as well as the resolution. A signal can be resolved down to 0.1 % of its modulation wavelength. A range value beyond the ambiguity interval is mapped onto $\left[0, \frac{\lambda_{mod}}{2}\right)$ and results in a wrapped range image. An unwrapping would be possible but should be avoided (Pfeifer *et al.*, 2013).

An important acquisition parameter is the integration time. It determines, how many modulation periods are considered to compute the amplitude and the range image. In photography it is called the exposure time, which specifies, how much light arrives on the sensor. In case of the SR-3000 an integration time between 200 μs and 51.2 ms in 200 μs-steps can be chosen. Integrating the samples over a longer period, increases the signal-to-noise ratio. The better the signal-to-noise ratio (SNR) is, the more precise are the measured distances. Generally, the distance measurement precision amounts to a few centimetres.

2.1.2 Data and Errors

As mentioned before, each pixel of the image sensor detects the backscattered optical power four times during the modulation period. Based on those four samples the amplitude and the phase angle and thus the distance are derived. As a result, the amplitude map and the range map of the illuminated scene occur as presented in Figure 2.3.

Figure 2.3: Received data: Amplitude and range image

Error Sources	**Error Type**	**Modelling**
Noise	Random	-
Overexposure	Gross	-
Motion blur	Gross	
Amplitude	Systematic	Empirically
Position in sensor		Empirically
Lens distortions		Physically
Instrumental errors		Empirically
Internal scattering		Physically Empirically
Multipath in object space		Empirically
Camera warm up		Empirically
Range wrapping		Physically

Table 2.1: Error sources, types and the underlying modelling approach according to Pfeifer *et al.* (2013)

The integration time can be chosen individually for each task. With increasing integration time, generally the data contains less noise and more precise results are accomplished. However, if too much light arrives on a pixel before readout, saturation effects emerge. This means that for different tasks, the right integration time has to be found. In case of motions either due to moving objects or a moving range camera, a shorter integration time is of advantage, as it reduces motion blur.

If the object distances and the object reflectivities vary strongly within a scenery or in contiguous frames, it is difficult to determine an appropriate integration time. The dynamic range describes the contrast of distances and the contrast of reflections within the scenery. The back-scattered power from close and bright objects causes thermal noise. In the follow-up model of the SR-3000 a pixel-wise integration has been implemented, i.e. the integration of samples stops in a pixel, if a predefined threshold is exceeded (Buettgen *et al.*, 2005).

In Table 2.1 various error sources are summarised. They are listed as mentioned in the text, except the range wrapping, which has been discussed in 2.1.1. It is important to distinguish between random, gross or systematic errors, as they are treated differently. In case of random errors the measurements will be averaged, gross errors have to be eliminated and systematic errors need to be modelled, either with physical or empirical models as mentioned in Table 2.1.

Random errors mainly arise by shot noise and dark current noise. Dark current noise is caused by thermal noise of the semiconductor, which can be decreased by cooling. Background lighting causes increased photon shot noise. Finally, the circuit noise completes the system noise. Due to the fact that random errors occur to the same amount with positive and negative algebraic sign, averaging reduces this effect. The averaging can be done either in the time domain by averaging of frames or spatially by modelling the scene with geometric primitives.

Gross errors have to be eliminated. Overexposure originates from a too high integration time. Consequently, saturation effects occur in the amplitude image and the distance measurements will be distorted. Motion blur arises from a changing instantaneous field of view of a pixel during the acquisition. Thus, the backscattered energy from objects at different distances is projected onto one pixel's field of view. However, motion blur is not an error of the measurement technique; it is an error of data acquisition.

Such gross errors can be detected by robust methods of parameter estimation.

Systematic errors. Most of the errors presented in Table 2.1 belong to the systematic errors. Only when considering those errors, an accurate result can be guaranteed, otherwise they cause a deviation from the true value. Averaging is not the right correction method for systematic errors, because they do not follow any statistical distribution. Thus, in a first step the influences have to be quantified and an appropriate physical or empirical model has to be derived.

According to the radar equation for large scatterers, the amplitude is indirectly proportional to the squared distance. Thus, distant objects show a decreasing amplitude, which is why those ranges are less precise. Furthermore, the strength of the received signal depends on the object's reflectivity (brightness) and the incidence angle. Consequently, those factors have an influence on the distance determination as well. Range errors related to the amplitudes are positive (i.e. distances appear to be longer), large in case of low back-scattered energy and they decrease for higher detected amplitudes. As no physical basis for this systematic error exists, empirical models are still in development.

A physical cause has neither been found for the range errors due to the position on the image plane. A notable light fall-off from the centre to the image borders has been detected (Pfeifer *et al.*, 2013).

The radial lens distortion causes similar effects like isotropy and is equal to zero at the principal point. However, modelling by a polynomial function is possible. The decentring lens distortions are typically smaller than the radial distortions. They are caused by imperfections in the production, e.g. the centres of compound lenses and the sensor are not properly aligned and for their correction radial and tangential terms have to be considered. They can be modelled physically.

Scene-independent range errors or instrumental range errors comprise the offset between the range measurement origin and the perspective centre, periodic range errors and errors due to the serial read-out. Those effects are corrected by pixel-wise look-up tables or surface models, B-splines or polynomials and linear functions in case of the serial read-out.

"Internal scattering is caused by multiple reflections within the camera between the lens, the optical filter and the sensor" (Pfeifer *et al.*, 2013). The detected energy at a single pixel is a combination of the backscattered energy from the illuminated pixel footprint and the scattered light, reflected by other sensor pixel. Consequently, near objects with high reflectivity deteriorate the phase angle of the direct signal path. The impact on the amplitudes is small, but the phase angles and the derived distances are highly influenced. A high dynamic range increases the internal scattering error and it can amount up to 40 cm (Pfeifer *et al.*, 2013). This error has to be modelled physically as well as empirically.

Multipath effects in object space describe the same behaviour as internal scattering. With the difference that the backscattered energy in a single pixel is a mixture of direct and multipath signals in object space. The errors have been quantified in different experiments.

The camera warm-up error is a function of temperature and time. The magnitude of the drift in the SR-3000 is in the order of several centimetres and a warm-up period of half an hour should be considered before using it. After the warm-up a good long-term stability can be reached.

The coefficients of those models will be determined in an instrument calibration. A standard ToF camera calibration comprises the amplitude-dependent errors, lens distortions and the two instrumental errors - range offset and periodic range errors.

Figure 2.4: A draft of the global coordinate system C. The coordinate system sketched needs to be lowered to the main ground level.

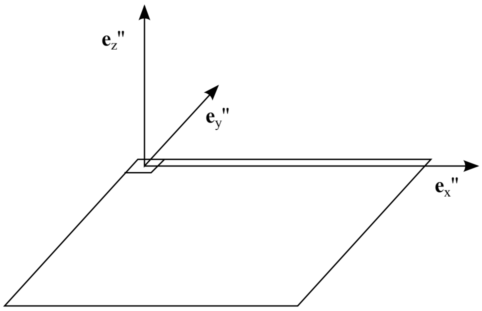

Figure 2.5: The image coordinate system C''

2.2 Coordinate Systems and Rotations

Coordinate System. In this work only cartesian, positive-oriented coordinate systems are used. Three different coordinate systems are applied: the image-, the camera- and the global coordinate system.
The global coordinate system, defined as $C = (\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z)$ is the one of the measurement laboratory of the Department of Geodesy and Geoinformation, as presented in Figure 2.4.
In order to measure image points, an image coordinate system $C'' = (\boldsymbol{e}''_x, \boldsymbol{e}''_y, \boldsymbol{e}''_z)$ needs to be defined. As shown in Figure 2.5, it is defined as a positive-oriented system, which is located in the centre of the upper- and leftmost image element. The image is represented by an image matrix and the equally spaced elements are called pixel (**Pi**cture **x el**ement). Thus, the matrix indices correspond to the pixel coordinates x'', y'' on the image plane and the z''-coordinate will be assumed to be zero, as shown in Table 2.2.
The image coordinate system C'' is not identical to the camera coordinate system $C' = (\boldsymbol{e}'_x, \boldsymbol{e}'_y, \boldsymbol{e}'_z)$. The image points need to be transformed to the camera coordinate system, which will be defined in the next section.

Rotations are linear transformations of the n-dimensional space. The two important characteristics are the speciality and the orthogonality. Speciality means, the determinant of a rotation matrix $|R| = +1$ has to be one and positive. A matrix is orthogonal, if $R^T = R^{-1}$ applies. Thus, rotations are special, orthogonal n-dimensional linear mappings $SO(n)$ according to McGlone (2004, p. 37). Moreover, they are conformal and length-preserving.
A $SO(3)$ rotation matrix contains nine elements, but merely three parameters suffice to describe a rotation since six conditions of orthonormality exist according to Kager *et al.* (2002, p. 184). Rotations can be formulated based on three angles, by axis and angle or by quaternions. Rotation by angles is the most intuitive representation. In this work rotations are described by the composition of three elementary rotations (McGlone, 2004, p. 40). However, the so called Euler angles representation causes singularities for different concatenations of the three coordinate axes for specific rotations. A universal solution might be derived, applying quaternions.

Rotations can be seen as a rotation of an object or as a rotation of a coordinate system. In this work, they are used in both ways. Generally, rotations of objects and of coordinate systems are related to each other by the transpose of a matrix.

In case of the relative orientation, it is a rotation of the tripod attached to the moving camera. Due to small motions of the camera, the rotation angles are defined as ω, φ, κ and the appropriate rotation order is specified as $\boldsymbol{e}'_x, \boldsymbol{e}'_y$ and $\boldsymbol{e}'_z$. The full rotation matrix is presented in Formula (2.5). However, the small angle approximation in (2.6) is assumed to be a sufficient parametrization of the problem. A verification can be found in Chapter 3.4.

$$\boldsymbol{R}_{\omega,\varphi,\kappa} = \begin{bmatrix} \cos\varphi \cdot \cos\kappa & -\cos\varphi \cdot \sin\kappa & \sin\varphi \\ \cos\omega \cdot \sin\kappa + \sin\omega \cdot \sin\varphi \cdot \cos\kappa & \cos\omega \cdot \cos\kappa - \sin\omega \cdot \sin\varphi \cdot \sin\kappa & -\sin\omega \cdot \cos\varphi \\ \sin\omega \cdot \sin\kappa - \cos\omega \cdot \sin\varphi \cdot \cos\kappa & \sin\omega \cdot \cos\kappa + \cos\omega \cdot \sin\varphi \cdot \sin\kappa & \cos\omega \cdot \cos\varphi \end{bmatrix} \tag{2.5}$$

$$\boldsymbol{R}_{\omega,\varphi,\kappa} = \begin{bmatrix} 1 & -\kappa & \varphi \\ \kappa & 1 & -\omega \\ -\varphi & \omega & 1 \end{bmatrix} \tag{2.6}$$

In case of the bundle adjustment a rotation of a coordinate system is performed. In that case the orientation of the coordinate systems involved are decisive. The α, ζ, κ-representation has been chosen due to a rather horizontal viewing direction of the ToF camera in reference to the global coordinate system. The rotation axis of the primary rotation α is $\boldsymbol{e}_z$, while for the secondary rotation ζ it is the (rotated) $\boldsymbol{e}_y$ and for the tertiary rotation κ it is the (rotated) $\boldsymbol{e}_z$ (Kager *et al.*, 2002, p. 224). In Formula (2.7) the complete rotation matrix is shown. Singularities will occur, if images are taken vertically.

$$\boldsymbol{R}_{\alpha,\zeta,\kappa} = \begin{bmatrix} \cos\alpha \cdot \cos\zeta \cdot \cos\kappa - \sin\alpha \cdot \sin\kappa & -\cos\alpha \cdot \cos\zeta \cdot \sin\kappa - \sin\alpha \cdot \cos\kappa & \cos\alpha \cdot \sin\zeta \\ \sin\alpha \cdot \cos\zeta \cdot \cos\kappa + \cos\alpha \cdot \sin\kappa & -\sin\alpha \cdot \cos\zeta \cdot \sin\kappa + \cos\alpha \cdot \cos\kappa & \sin\alpha \cdot \sin\zeta \\ -\sin\zeta \cdot \cos\kappa & \sin\zeta \cdot \sin\kappa & \cos\zeta \end{bmatrix} \tag{2.7}$$

$$\boldsymbol{x} = \boldsymbol{R}\boldsymbol{x}'', \quad \boldsymbol{R} = \begin{bmatrix} {}_G\boldsymbol{x}'' & {}_G\boldsymbol{y}'' & {}_G\boldsymbol{z}'' \end{bmatrix} = \begin{bmatrix} \left({}_{''}\boldsymbol{x}^G\right)^T & \left({}_{''}\boldsymbol{y}^G\right)^T & \left({}_{''}\boldsymbol{z}^G\right)^T \end{bmatrix}^T \tag{2.8}$$

Formula (2.8) specifies the rotation matrix. $\boldsymbol{R}$ performs a rotation from the image coordinate space C'' into the global space C. Actually, Formula (2.8) represents a transformation and the translation is disregarded. The column vectors of $\boldsymbol{R}$ are the basis of C'' in reference to the coordinate system C (denoted with G for global), while the rows of $\boldsymbol{R}$ are constituted by the base vectors of C represented in C''. This knowledge simplifies the estimation of approximate values.

2.3 Image Geometry

The underlying geometric model is the pinhole camera model, which describes a central perspective relation between object space i and image space j as shown in Figure 2.6. All rays coming from an object, pass straight through one single point - the perspective centre $\boldsymbol{p}_0 = [x_0 \quad y_0 \quad z_0]^T$. In that point, the camera coordinate system C' has its origin. Its $\boldsymbol{e}'_x, \boldsymbol{e}'_y$-plane is parallel to the image plane and the $\boldsymbol{e}'_z$-direction is pointing into the opposite direction as data is acquired, as represented in Figure 2.6. The orthogonal projection of the perspective centre onto the image plane defines the principal point (x''_0, y''_0) and the principal

Parameter	in C	in C'	in C''
Perspective centre	$\boldsymbol{p_0} = [x_0 \quad y_0 \quad z_0]^T$	$\boldsymbol{p_0'} = [0 \quad 0 \quad 0]^T$	$\boldsymbol{p_0''} = [x_0'' \quad y_0'' \quad c]^T$
Image point	-	$\boldsymbol{p_{ij}'} = (\boldsymbol{p_{ij}''} - \boldsymbol{p_0''})$ $\boldsymbol{p_{ij}'} = [x_{ij}'' - x_0'' \quad y_{ij}'' - y_0'' \quad -c]^T$	$\boldsymbol{p_{ij}''} = [x_{ij}'' \quad y_{ij}'' \quad 0]^T$

Table 2.2: Perspective centre and image points in the different coordinate systems.

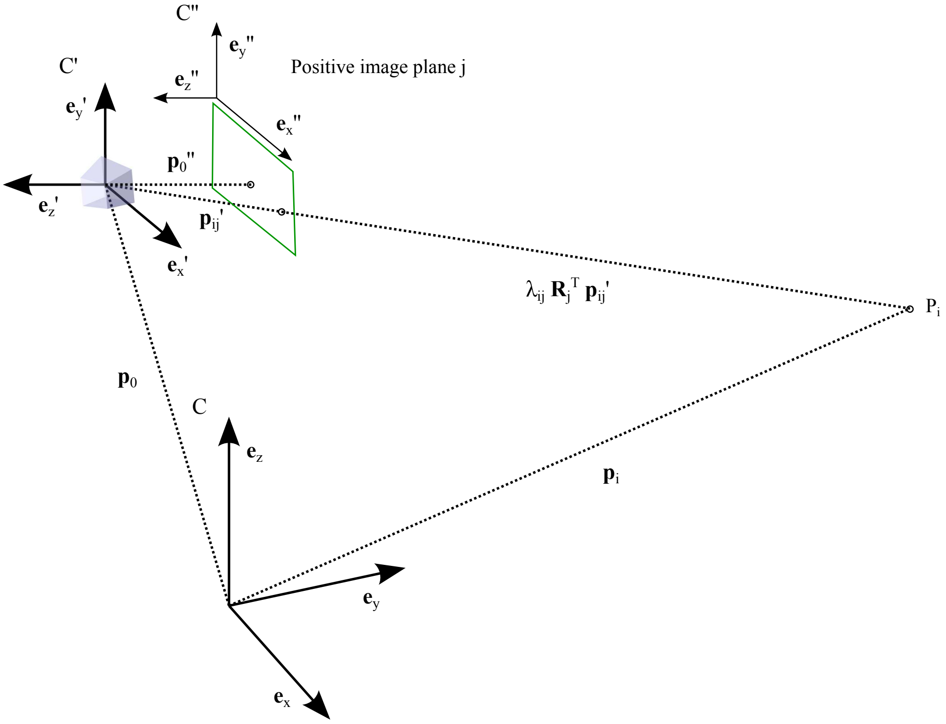

Figure 2.6: The geometric model including the three coordinate systems.

distance c, which specifies the relation of the camera coordinate system C' and the image coordinate system C''. Table 2.2 shows a comparison of the perspective centre and the image points in all three coordinate systems. Due to the aligned coordinate systems, the transformation between those coordinate systems is defined by the translation $\boldsymbol{p_{ij}'}$ only. It describes the transition from the image point p_{ij}'' to the image vector in the camera coordinate system $\boldsymbol{p_{ij}'}$. In Figure 2.6 the image plane is displayed between the perspective centre and the object. This positive camera position corresponds to the mathematical representation. Physically, the image plane lies behind the perspective centre in positive $e_z{}'$-direction. The image vector in the camera coordinate system $\boldsymbol{p_{ij}'} = (\boldsymbol{p_{ij}''} - \boldsymbol{p_0''})$ builds the projection ray from the image point to the object point.

The collinearity condition states: an object point $\boldsymbol{p_i} = [x_i \quad y_i \quad z_i]$, the corresponding image point $\boldsymbol{p_{ij}''}$ and the camera's perspective centre $\boldsymbol{p_0'}$ lie on a straight line. The relation of those three points is formulated by the collinearity equation in (2.9). The equation represents a similarity transformation.

$$\boldsymbol{p_i} = \boldsymbol{p_0} + \lambda_{ij} \boldsymbol{R} \left(\boldsymbol{p_{ij}''} - \boldsymbol{p_0''} \right) \tag{2.9}$$

An object point $\boldsymbol{p_i}$ can be determined, by adding the scaled and rotated image vector $\boldsymbol{p'_{ij}} = (\boldsymbol{p''_{ij}} - \boldsymbol{p''_0})$ to the perspective centre in the object space $\boldsymbol{p_0}$. The perspective centre $\boldsymbol{p_0}$ is a three-dimensional position vector and the rotation matrix $\boldsymbol{R}$ is parametrised by three angular arguments. They can be summarised as the exterior orientation of a camera. It describes the position and the angular orientation of the camera coordinate system C' in reference to the global one C. The interior orientation $\boldsymbol{p''_0}$ comprises the principal point (x''_0, y''_0) and the principal distance c, or the camera constant. It defines the position of the perspective centre in the image coordinate system C''. Especially, the interior orientation is a crucial factor for using a camera as a measurement device. Distortions will lead to errors in object space and hence the camera calibration is essential. A measurement error in image space will be multiplied by the image scale λ_{ij} to receive the error in object space.

In photogrammetry, where originally passive cameras were used, the scale factor λ_{ij} in (2.9) is derived by stereo-photogrammetry. The interior orientation may be assumed to be known and the exterior orientation will be determined through the application of an orientation method. This work focuses on the bundle adjustment - a global orientation method - which is explained in Chapter 2.3 in more detail.

The geometric modelling approaches of a passive camera can be easily adapted for a ToF camera. It leads to a simplification in Equation (2.9), due to the fact that no second image is needed to receive the depth scale λ_{ij}. The scale factor is now replaced by the observed range ρ_{ij} in Formula (2.10). Consequently, object points $\boldsymbol{p_i}$ can be determined from a single ToF camera view, if the exterior and interior orientation parameters are known.

$$\boldsymbol{p_i} = \boldsymbol{p_0} + \rho_{ij}\boldsymbol{R}\frac{\boldsymbol{p'_{ij}}}{||\boldsymbol{p'_{ij}}||} \tag{2.10}$$

2.4 Bundle Adjustment

As mentioned before, in order to determine the exterior orientation of a camera an orientation method has to be applied. In this section the bundle adjustment as a technique solving these unknowns is introduced.

The bundle adjustment is a method applicable to any number of images. As the name implies, the orientation is determined in the course of an adjustment, which is discussed in Chapter 2.5. The basic relation is formulated in the collinearity equation (2.9) stating, that the camera's perspective centre $\boldsymbol{p'_0}$, a measured image point $\boldsymbol{p''_{ij}}$ and the corresponding object point $\boldsymbol{p_i}$ lie on a straight line, or set up a ray. To put it in the context of adjustment calculus, the observations have to be expressed as a function of the unknowns (2.12). Therefore, the rotation matrix has been split into column vectors, as shown in Equation (2.11).

$$\boldsymbol{R} = [\boldsymbol{r}_1 \quad \boldsymbol{r}_2 \quad \boldsymbol{r}_3] \tag{2.11}$$

$$x''_{ij} = x''_0 - c \cdot \frac{\boldsymbol{r}_1(\boldsymbol{p_i} - \boldsymbol{p_0})}{\boldsymbol{r}_3(\boldsymbol{p_i} - \boldsymbol{p_0})}, \qquad y''_{ij} = y''_0 - c \cdot \frac{\boldsymbol{r}_2(\boldsymbol{p_i} - \boldsymbol{p_0})}{\boldsymbol{r}_3(\boldsymbol{p_i} - \boldsymbol{p_0})} \tag{2.12}$$

In case of many images, bundles of such spatial rays comprising the perspective centres $\boldsymbol{p_{0,j}}$ and the image points $\boldsymbol{p'_{ij}}$ can be defined. The image points either correspond to control points, which are marked and their coordinates are known or they correspond to tie points, which are well-defined points without known coordinates. Therefore, it can be denoted as a marker-based orientation method. In the course of the bundle adjustment the defined bundles will be shifted and rotated to fit into the control and tie points. Thus, the section geometry of the bundles is decisive. Typical unknowns comprise the exterior orientation, the object points, a scale or the interior orientation as well as sensor distortions.

The bundle adjustment is the most efficient and accurate image orientation technique in photogrammetry, because of a direct relation between the image and the object point. If no camera calibration is known, the parameters of the interior orientation can be treated as unknowns and are estimated in the adjustment. Thus, it is an easily adaptable method. This is also helpful, when direct georeferencing data needs to be included or systematic errors need to be considered. Another notable advantage of this orientation technique is the simultaneous estimation of all unknown parameters (Kraus, 1996).

Solving normal equations with large numbers of unknowns, dealing with and eliminating gross errors as well as acquiring approximate values are the challenges in bundle adjustment. In this work the first two points are considered to be more important and are discussed later on.

The geometric precision of the absolute orientation is mainly based on the image sensor. Thereby, the position of the sensor elements (arranged in a regular grid) as well as the sensor planarity is of importance (Luhmann, 2000, p. 168). In case of the range camera the low geometric resolution has to be considered.

Image Measurements. Another important aspect is the measurement of the control points in the images and the reached accuracy. The image coordinates of the control points are measured manually in the first image and are tracked in subsequent images. For that purpose, a 9 by 9 neighbourhood around the previous determined image point is considered. The maximum amplitude is identified and its centroid is calculated in a neighbourhood of 3 by 3 Pixels. The centroid operator is suitable for small circular targets as in the present case.

The manual measurement precision ranges from 0.3 to 0.5 pixel. The centroid can be located with a standard deviation of 0.1 pixel in case of a signalled object, rich in contrast to the neighbouring pixel. However, the reached measurement precision strongly depends on the noise in the amplitudes (Luhmann, 2000, p. 399-401).

2.5 Range and Optical flow

Optical flow is the image motion, transforming one image to the next image according to Horn & Schunck (1981). It is a technique, originally used in computer vision and can be solved by differential, region-based, energy-based or phase-based techniques.

Horn & Schunck (1981) used a differential approach to solve the optical flow. The intensity, respectively brightness at a point (x'', y'') and at a time t can be denoted by $I(x'', y'', t)$. The underlying assumption is that the brightness of a particular point in the pattern is constant. The pattern is the structure of the scene recorded, which changes relative to the camera in time - due to a moving camera or moving objects .

Based on the brightness constancy, the well-known optical flow constraint in Equation (2.13) can be obtained (Horn & Schunck, 1981). I_x, I_y, I_t are the spatial and temporal derivatives of the image brightness in C'' and $\dot{x}'', \dot{y}''$ represent the optical flow.

$$I_x \dot{x}'' + I_y \dot{y}'' + I_t = 0 \tag{2.13}$$

Range flow. Due to the fact, that ToF cameras detect the amplitude and the phase of the emitted modulated signal, a range flow can also be utilized. The assumption made by Spies *et al.* (2002) is a local time-varying surface in C viewed as a depth function $\rho(x, y, t)$. The derived range flow is presented in Formula (2.14). Whereby ρ_t describes the change at the considered image position in C'' in time and $\dot{z}$ is due to the assumption of the time-varying surface in the global space C.

$$\rho_x \dot{x}'' + \rho_y \dot{y}'' + \rho_t = \dot{z} \tag{2.14}$$

Relative orientation. The range and optical flow can also be used to determine the relative orientation of a Range Imaging Camera. Instead of estimating the flow $\dot{x}''$ and $\dot{y}''$ at each pixel, as accomplished by Horn & Schunck (1981), the unknown parameters of the camera pose will be computed. This was done by Ghuffar *et al.* (2013), who relates the relative orientation in Formula (2.15) with the flow constraints of Equation (2.13) and (2.14).

In this work the motion is generated due to a movement of the camera in a static environment. The relative orientation between two consecutive camera positions can be written as in Equation (2.15), where the relative orientation parameters are given with respect to the camera coordinate system C' at time $j+1$.

$$\begin{bmatrix} x'_{j+1} \\ y'_{j+1} \\ z'_{j+1} \end{bmatrix} = \begin{bmatrix} 1 & R_{z'} & -R_{y'} \\ -R_{z'} & 1 & R_{x'} \\ R_{y'} & -R_{x'} & 1 \end{bmatrix} \begin{bmatrix} x'_j \\ y'_j \\ z'_j \end{bmatrix} + \begin{bmatrix} t_{x'} \\ t_{y'} \\ t_{z'} \end{bmatrix}, \qquad \boldsymbol{p}'_{j+1} = \boldsymbol{R}\boldsymbol{p}'_j + \boldsymbol{t}' \tag{2.15}$$

The original unknowns $\dot{x}'', \dot{y}'', \dot{z}$ in the range and optical constraint are substituted by the parameters of the relative orientation - the three translations $t_{x'}, t_{y'}, t_{z'}$ and the three rotations $R_{x'}, R_{y'}, R_{z'}$. In Formula (2.16) the relative orientation according to Ghuffar *et al.* (2013) is presented. This equation can be written for each pixel of the images. Thus, a highly overdetermined equation system is reached, which is solved through the use of robust adjustment. The observations are the intensity and range derivatives, hence it is a data-driven orientation method.

$$\begin{gathered} \begin{bmatrix} -I_t \\ -\rho_t \end{bmatrix} = \begin{bmatrix} -I_x \frac{c}{z} & -I_y \frac{c}{z} & (I_x \frac{x''}{z} + I_y \frac{y''}{z}) \\ -\rho_x \frac{c}{z} & -\rho_y \frac{c}{z} & (\rho_x \frac{x''}{z} + \rho_y \frac{y''}{z}) \end{bmatrix} \begin{bmatrix} t_{x'} \\ t_{y'} \\ t_{z'} \end{bmatrix} + \\ \begin{bmatrix} (I_x \frac{x''y''}{c} + I_y \frac{c^2+y''^2}{c}) & -(I_y \frac{x''y''}{c} + I_x \frac{c^2+x''^2}{c}) & (I_x y'' - I_y x'') \\ (\rho_x \frac{x''y''}{c} + \rho_y \frac{c^2+y''^2}{c}) & -(\rho_y \frac{x''y''}{c} + \rho_x \frac{c^2+x''^2}{c}) & (\rho_x y'' - \rho_y x'') \end{bmatrix} \begin{bmatrix} R_{x'} \\ R_{y'} \\ R_{z'} \end{bmatrix} \end{gathered} \tag{2.16}$$

Incremental Orientation. The relative orientation describes the change of the camera pose between two contiguous frames. In order to receive a trajectory, the single motions have to be concatenated. As mentioned before, the rotations and translations of the relative orientation are related to the camera position at time $j+1$. In Figure 2.7 the transformation of Equation (2.15) is illustrated. To attain the trajectory with regard to the first camera position of all, Equations (2.17) and (2.18) have to be applied. Accordingly,

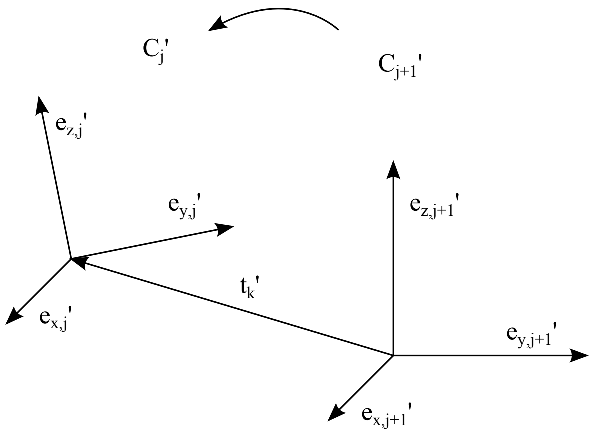

Figure 2.7: Relative Orientation in respect to the camera coordinate system C'_{j+1}

the translations $\boldsymbol{t}_k$ are transformed by $-\boldsymbol{R}_k^T \boldsymbol{t}_k$ into the foregoing camera coordinate system and the euler angles $\omega_k, \varphi_k, \kappa_k$ will be set negative.

$$\boldsymbol{R}_{incr,k} = \boldsymbol{R}_{incr,1}^T \boldsymbol{R}_{incr,2}^T \cdots \boldsymbol{R}_{incr,k}^T \tag{2.17}$$

$$\boldsymbol{t}'_{incr,k} = \boldsymbol{t}'_{incr,k-1} - \boldsymbol{R}_{incr,k} \boldsymbol{t}'_k \tag{2.18}$$

2.6 Adjustment Theory

Adjustment is a fundamental statistical method to estimate unknown parameters and their dispersion from observed quantities. Due to the law of large numbers in probability theory, as well as being not immune to measurement errors, an overdetermined equation system is a prerequisite. In this case only, the user is able to ensure the quality of the results. Firstly, the derivation of the functional model is shown and in the second part the stochastic model will be included in the system. Furthermore, the theoretical aspects of the variance component estimation will be discussed.

Functional model. n observations l_i, $(i = 1, ..., n)$ are used to estimate u unknowns x_j, $(j = 1, ..., u)$, whereby over-determination $(n > u)$ is assumed. As the term *functional model* specifies, there must be a reasonable relation between observations and unknown parameters, which can be formulated. The design matrix $\boldsymbol{A}$ will be constructed based on this functional relation. The random errors of the observations are considered in the residual vector $\boldsymbol{v}$. Hence, the problem can be defined as in Formula (2.19). Treating the adjustment as a linear problem implies the advantage of using well-known numerical linear algebra calculus. In case of n observations and n unknowns an unique solution can be reached. An overdetermined problem has no unique solution, since the residuals $\boldsymbol{v}$ in Equation (2.20) cannot be made to reach zero. However, they can be made as small as possible (Trefethen & Bau, 1997).

$$\boldsymbol{l} + \boldsymbol{v} = \boldsymbol{A}\boldsymbol{x} \tag{2.19}$$

$$\boldsymbol{v} = \boldsymbol{A}\hat{\boldsymbol{x}} - \boldsymbol{l} \tag{2.20}$$

Thus, the approach of the **Least Squares** (LS) Method is minimising the 2-norm of the residuum vector $\boldsymbol{v}$. The linear equation system transforms from (2.19) to a residual equation system in (2.20). The minimum constraint leads to the estimation of the unknowns $\hat{\boldsymbol{x}}$ as derived in (2.21) to (2.23).

$$\boldsymbol{v}^T\boldsymbol{v} = min \quad = (\boldsymbol{A}\hat{\boldsymbol{x}} - \boldsymbol{l})^T(\boldsymbol{A}\hat{\boldsymbol{x}} - \boldsymbol{l}) = \hat{\boldsymbol{x}}^T\boldsymbol{A}^T\boldsymbol{A}\hat{\boldsymbol{x}} - 2\boldsymbol{l}^T\boldsymbol{A}\hat{\boldsymbol{x}} + \boldsymbol{l}^T\boldsymbol{l} \tag{2.21}$$

$$\frac{\partial(\boldsymbol{v}^T\boldsymbol{v})}{\partial\hat{\boldsymbol{x}}} = 2\hat{\boldsymbol{x}}^T\boldsymbol{A}^T\boldsymbol{A} - 2\boldsymbol{l}^T\boldsymbol{A} = 0 \tag{2.22}$$

$$\hat{\boldsymbol{x}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{l} \tag{2.23}$$

The matrix $\boldsymbol{A}$ is defined as a Jacobi-matrix (2.24) and consists of the partial derivatives of $v_i(x_1, x_2, ..., x_u)$, whereby the partial derivatives $\frac{\partial v_i}{\partial x_j}$ describe the influence of the unknown parameters x_j on the function v_i (Niemeier, 2008, p. 64). If the relation between observations and unknowns is described as a non-linear problem, the equation will be linearised by the first-order term of the Taylor series and higher-order terms will be neglected. In this case, approximate values for $\boldsymbol{x}$ are needed and their quality determines the convergence.

$$\boldsymbol{A} = \begin{bmatrix} \frac{\partial v_1}{\partial x_1} & \cdots & \frac{\partial v_1}{\partial x_u} \\ \vdots & \ddots & \vdots \\ \frac{\partial v_n}{\partial x_1} & \cdots & \frac{\partial v_n}{\partial x_u} \end{bmatrix} \tag{2.24}$$

Stochastic model. Measurements are not completely controllable effects. Different errors as well as correlations between observations affect the measurements. However, producers test their instruments or users accomplish calibrations to specify the accuracies of observations. The precision is expressed as the standard deviation σ_i or variance σ_i^2. The covariance matrix Σ_{ll} is built up of those variances as presented in (2.25).

$$\boldsymbol{\Sigma}_{ll} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix} \tag{2.25}$$

The diagonal elements of $\boldsymbol{\Sigma}_{ll}$ are the variances of the observation and the off-diagonal elements present the covariances. Covariances have to be considered, in case observations are correlated. If no a priori knowledge of the observation precisions is available, assumptions must be made. However, the precision relation between different observations and the presence of correlations are of importance. The cofactor matrix $\boldsymbol{Q}_{ll}$ is based on these precision relations $\frac{\sigma_i^2}{\sigma_0^2}$ and the true precision level σ_0^2 will be estimated in the course of the adjustment, as shown in (2.26). σ_0^2 can be used as a preconditioning factor. The inverse of the cofactor matrix is the weight matrix $\boldsymbol{P}$ in (2.27). In case $\boldsymbol{Q}_{ll}$ is a diagonal matrix without covariances, the elements of its inverse $\boldsymbol{P}$ will be $\frac{\sigma_0^2}{\sigma_i^2}$. If $\sigma_0^2 = 1$, the elements will be $\frac{1}{\sigma_i^2}$. As can be seen, a small variance has more influence than a large one (Niemeier, 2008, p. 126).

$$\boldsymbol{\Sigma}_{ll} = \sigma_0^2 \cdot \boldsymbol{Q}_{ll} \tag{2.26}$$

$$\boldsymbol{P} = \boldsymbol{Q}_{ll}^{-1} \tag{2.27}$$

The essential steps are the definition of the functional and stochastic model. If the functional model describes a non-linear problem, approximate values are needed due to linearisation. The unknowns are solved as in (2.28) and the corresponding cofactors can be found in $\boldsymbol{Q}_{\boldsymbol{xx}}$. $\boldsymbol{A}^T\boldsymbol{A}$ is called the normal equation matrix $\boldsymbol{N}$, which is regular and therefore the inverse can be computed. The residuals are calculated by plugging in the estimated unknowns into the initial equation system in (2.29). The adjusted observation precisions are given by $\boldsymbol{Q}_{\hat{l}\hat{l}}$. In order to receive the covariance matrices, the cofactor matrices $\boldsymbol{Q}_{xx}$ and $\boldsymbol{Q}_{\hat{l}\hat{l}}$ have to be multiplied with the adjusted variance factor $\hat{\sigma}_0^2$, which is discussed later in more detail. In addition, it is possible to compute a covariance matrix of the residuals $\boldsymbol{Q}_{vv}$ as presented in (2.30). This matrix is of importance for reliability measures (Niemeier, 2008).

$$\hat{\boldsymbol{x}} = (\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{l} \quad = \boldsymbol{N}^{-1}\boldsymbol{n}, \qquad \boldsymbol{Q}_{xx} = (\boldsymbol{A}^T\boldsymbol{P}_{ll}\boldsymbol{A})^{-1} = \boldsymbol{N}^{-1} \tag{2.28}$$

$$\boldsymbol{v} = \boldsymbol{A}\hat{\boldsymbol{x}} - \boldsymbol{l}, \qquad \boldsymbol{Q}_{\hat{l}\hat{l}} = \boldsymbol{A}\boldsymbol{Q}_{xx}\boldsymbol{A}^T \tag{2.29}$$

$$\boldsymbol{Q}_{vv} = \boldsymbol{Q}_{ll} - \boldsymbol{Q}_{\hat{l}\hat{l}} \tag{2.30}$$

Variance component estimation. A main task in this work is to combine different observation types. To achieve reasonable results, the different observation types are weighted. The used method is the variance component estimation (VCE) as described in Niemeier (2008, p. 318f).

First, to avoid any misunderstanding, the variance factor σ_0^2 is explained. It defines the precision level of the observations. In the course of the global test of the adjustment model, the fundamental relation in Formula (2.31) is retrieved, showing that the residual square sum Ω is an unbiased estimator of the variance factor $\hat{\sigma}_0^2$ (Niemeier, 2008, p. 167).

$$E(\Omega) = (n-u)\sigma_0^2 \tag{2.31}$$

Thus, the variance factor a posteriori $\hat{\sigma}_0^2$ is defined in Formula (2.32) as the sum of squared weighted residuals divided by the total redundancy r, which is the difference of the number of observations n and the number of parameters u. The number of the total redundancy should lie between u and $2u$ for a reasonable estimation of the variance factor.

$$\hat{\sigma}_0^2 = \frac{\Omega}{r} = \frac{\boldsymbol{v}^T\boldsymbol{P}\boldsymbol{v}}{n-u} \tag{2.32}$$

If an adjustment problem includes different observation types with different precisions e.g. image observations of a photogrammetric utilisation, direction and distance measurements of a total station and GNSS vectors, the determination of a priori precisions of the different observations is challenging. The approach of the variance component estimation is to formulate a variance factor for each observation group, called variance component σ_{0i}^2. It can be achieved by adding error vectors $\boldsymbol{\epsilon}_i$ to the original adjustment model as done in (2.33). For each error vector $\boldsymbol{\epsilon}_i$ an own cofactor matrix $\boldsymbol{Q}_i$ has to be considered. Due to the computation of the variance factor after the adjustment, it is an iterative method. Therefore, an initial value $\alpha_{0i,0}^2$ for $\sigma_{0i,v}^2$ is needed (2.34). The initial value $\alpha_{0i,0}^2$ shall be chosen in that way, that the variance components $\hat{\sigma}_{0i,v}^2$ are approximately 1. In each iteration step $\hat{\sigma}_{0i,v}^2$ of each observation group is computed as in Equation (2.32) and then multiplied by $\alpha_{0i,v}^2$. The method converges, if $\hat{\sigma}_{0i,v}^2 = 1$. The true variance component is the improved initial value of $\alpha_{0i,v}^2$.

The cofactor matrices are connected according to the variance components. Overlapping or ranked variance components can be considered. In case of overlapping variance components e.g. two cofactor matrices fill the same position in the $\boldsymbol{Q}_{ll}$. Ranked variances means, that for each observation group one variance

component is estimated and the cofactor sub-matrices $\bar{\boldsymbol{Q}}_i$ are placed one by one, as presented in Formula (2.35). In this work, only ranked variance components will be treated.
Furthermore, no covariance components σ_{0ij} between the groups are assumed.

$$\boldsymbol{l} = \boldsymbol{A}\boldsymbol{x} + \epsilon_1 + \epsilon_2 \tag{2.33}$$

$$\boldsymbol{\Sigma}_{ll,v} = \sigma^2_{01,v}(\alpha^2_{01,v}\boldsymbol{Q_1}) + \sigma^2_{02,v}(\alpha^2_{02,v}\boldsymbol{Q_2}) \tag{2.34}$$

$$\boldsymbol{Q}_1 = \begin{bmatrix} \bar{\boldsymbol{Q}}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \qquad \boldsymbol{Q}_2 = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \bar{\boldsymbol{Q}}_2 \end{bmatrix} \tag{2.35}$$

A reliable estimation of the variance components is only guaranteed, if the total redundancies of the observation groups show a sufficient magnitude and if they are approximately of the same size. Furthermore, a correct and complete functional and stochastic model as well as no gross errors in the data are crucial.

2.7 Equation Solving

Bundle adjustment typically leads to a large sparse linear estimation problem. When solving such matrix systems, iterative methods are of advantage. In comparison to direct methods they are more efficient, due to a smaller number of operations. The effort per iteration step mainly depends on the structure of the matrix (Trefethen & Bau, 1997, p. 246), therefore it is of importance to exploit structural parameters like sparsity. A matrix is sparse, if a predominance of zero entries is recorded. The normal matrix $\boldsymbol{A}^T\boldsymbol{A}$ often exhibits more zero entries than non-zero ones, but its inverse $\boldsymbol{Q}_{xx}$ is dense.

Optimisation method. Solving an optimisation problem means, that a correction term Δx is determined for the initial value x in each iteration step. Thus, the convergence depends on the starting value x, the method used and on whether a step control δx will be applied.
The computation of a bundle adjustment can be separated in algorithms of first and second order, which refers to the used elements of the Taylor series. First order methods converge slower due to their linear behaviour, but each iteration step is much cheaper than a second order one. Examples for first order or approximate bundle algorithms are an alternation of resection and intersection, the Gauss-Seidel iteration and the gradient descent. However, poor convergence, slow propagating corrections and disregarded relations between variables lead to more historical interest in these methods. However, in the unusual case of a dense problem with strong geometry (in case of site modelling) they are still preferable. Second order bundle algorithms are efficiently organised algorithms and suitable for large sparse problems. The technique is very effective for aerial blocks due to the structured data (Triggs *et al.*, 1999, p. 312).
The software used is based on such a second order algorithm. The optimisation method is the Levenberg-Marquardt algorithm, including a step control δx. Such a method is called damped Newton and will lead to convergence. OrientAl can also be used as a robust adjustment algorithm, because robust loss functions can be defined and computed. Hence, non-linear least-squares adjustment problems can be solved by OrientAl.

Linear Solver. In adjustment theory many formulae contain matrix inverses. However, matrix inversion is almost never done. Instead, those formulae are solved using forwards and backwards substitution done by Cholesky-, LU- or QR-decomposition. "This is much faster and numerically more accurate than explicit use of the inverse" (Triggs *et al.*, 1999, p. 317). Factorization is the key computational cost. The efficiency

in this step can be improved by exploiting sparsity, symmetry and other problem structure (Triggs *et al.*, 1999).
In case of the Cholesky decomposition, the order of the different unknown parameters is crucial. As mentioned before, the structure of a matrix mainly determines the convergence speed. The correct ordering is as follows: object points, exterior orientation and interior orientation (Kraus, 1996, p. 76).

Scaling and Preconditioning. Numerical performance is influenced by variable scaling and preconditioning. It mainly has to be considered, when using linearly convergent optimisation methods. However, using second order methods with damping strategies are also scaling-dependent (Triggs *et al.*, 1999, p. 318).

Automatic differentiation. The Jacobi-matrix is computed using an automatic or algorithmic differentiation method. Numerical differentiation is based on the difference quotient and convergence behaviour mainly depends on the starting values. In contrast, automatic algorithms apply the chain rule on elementary arithmetic operations and arithmetic functions (Corliss & Griewank, 1993). Thus, convergence is no problem. Moreover, it can be done accurately to working precision.

2.8 Evaluation of the Trajectory Error

To evaluate estimated trajectories, the following two measures can be computed: the relative pose error and the absolute trajectory error (Sturm et al. 2012, cited in Ghuffar, 2014, p. 49).
In order to determine an error, a precise reference (ground truth) is needed. The unbiased camera poses of the bundle adjustment are used for that purpose. In order to evaluate the trajectory, a sequence $(j = 1, ..., n)$ of camera poses $(\boldsymbol{R}_{\alpha,j}, \boldsymbol{p}_{0,j})$ of the ground truth $\boldsymbol{G}_j \in SE(3)$, and a sequence of camera poses $(\boldsymbol{R}_{incr,j}, \boldsymbol{t}'_{incr,j})$ of the estimated incremental trajectory $\boldsymbol{F}_j \in SE(3)$ are given. Those camera poses $\boldsymbol{G}_j$ and $\boldsymbol{F}_j$ are denoted as a special euclidean group $SE(3)$ as written in Formula (2.36).

The **relative pose error** (RPE) is a local accuracy measure. Deviations of each relative camera pose are evaluated. In Equation (2.37) the relative pose error is presented. The term in the brackets depict the difference of subsequent camera poses, which corresponds to the relative orientation. Multiplying those two expressions leads to the deviation of those relative orientations. The root mean squared error of the translational component $trans(\boldsymbol{RPE}_j)$ is written in Equation (2.38).
The rotational error is parametrised by rotation axis and angle. The measures are computed as shown in McGlone (2004, p. 46). The vector $\boldsymbol{a}$ in (2.39) corresponds to the eigenvector of the rotation matrix. The rotation angle can be unambiguously determined from the trace of $\boldsymbol{R}$ and the length of the vector $\boldsymbol{a}$ (2.40). The rotation axis $\boldsymbol{d}$ will be defined as in Equation (2.41). However, these computations are only possible, if the angle α is not equal to 0 or π.

$$\boldsymbol{T}_{4\times4} = \left[\begin{array}{c|c} \boldsymbol{R} & \boldsymbol{t} \\ \hline \boldsymbol{0}_{1\times3} & 1 \end{array}\right] \qquad \boldsymbol{R} \in SO(3), \boldsymbol{t} \in \mathbb{R}^3 \tag{2.36}$$

$$\boldsymbol{RPE}_{j:n-1} = (\boldsymbol{G}_j^{-1}\boldsymbol{G}_{j+1})^{-1}(\boldsymbol{F}_j^{-1}\boldsymbol{F}_{j+1}) \tag{2.37}$$

$$RMSE(\boldsymbol{RPE}_{j:n-1}) = \left(\frac{1}{n-1} \sum_{j=1}^{n-1} ||trans(\boldsymbol{RPE}_j)||^2 \right)^{\frac{1}{2}} \tag{2.38}$$

$$\boldsymbol{a} = -\begin{bmatrix} r_{23} - r_{32} \\ r_{31} - r_{13} \\ r_{12} - r_{21} \end{bmatrix} = 2\boldsymbol{d} \sin \alpha \tag{2.39}$$

$$\alpha = atan2\,(||\boldsymbol{a}||, tr\boldsymbol{R} - 1) \tag{2.40}$$

$$\boldsymbol{d} = \frac{\boldsymbol{a}}{||\boldsymbol{a}||} \tag{2.41}$$

The **absolute trajectory error** (ATE) presents the global consistency of the estimated trajectory, which is evaluated by comparing the distances between the two trajectories $\boldsymbol{G}_j, \boldsymbol{F}_j$. To establish a comparison, the coordinate system of the bundle adjustment and the relative orientation need to be aligned. Based on the transformation matrix $\boldsymbol{G}_1$, the estimated trajectory of $\boldsymbol{F}_j$ is transformed into the ground truth trajectory $\boldsymbol{G}_j$. As discussed in Chapter 2.5, the incremental orientation transforms the trajectory into the coordinate system of the first camera. $\boldsymbol{G}_1$ describes the position and the attitude of the first camera coordinate system in reference to the global one.

According to Equation (2.42) and (2.43) the ATE and the root mean squared translational error is achieved. Again, axis and angle are computed to analyse the rotational error (see relative pose error).

$$\boldsymbol{ATE}_j = \boldsymbol{G}_j^{-1} \boldsymbol{G}_1 \boldsymbol{F}_j \tag{2.42}$$

$$RMSE(\boldsymbol{ATE}_j) = \left(\frac{1}{n} \sum_{j=1}^{n} ||trans(\boldsymbol{ATE}_j)||^2 \right)^{\frac{1}{2}} \tag{2.43}$$

3. Data and Methods

This chapter comprises the data acquisition and necessary preprocessing steps as well as the utilisation of the two orientation methods. In the last section a first comparison of the two methods is presented.

3.1 Data Acquisition

The scene recorded has to fulfil various requirements due to the integration of two different orientation methods. On the one hand enough variation in brightness and range has to be available for flow calculations, as on the other hand control points have to be included in the scene for the bundle adjustment. Therefore, the existing measuring field in the laboratory of the Department of Geodesy and Geoinformation was chosen for data acquisition. The ToF camera was mounted on a rail-movable platform, approximately 1 m above the ground and was moved diagonal with respect to the scene. Only linear motions were accomplished.

To gather reasonable data with respect to their prerequisites was a challenging task in the chosen measuring field. The control points are realised as retro-reflecting targets, which are the bright spots in the amplitude images of Figure 3.1 and 3.2. According to the figures, the scene recorded is nearly the same for both data acquisitions. However, the ToF camera was rotated around the optical axes by 100 gon in the second one. The scene shows variations in range and brightness. The dynamic range makes it difficult to find an appropriate integration time (It) for the whole image.
As written in Table 3.1, the dimensions of the images are 176 by 144 pixel and one image element has an extent of 40 μm. Standard passive cameras already feature resolutions of several Mega pixel and few μm of pixel size. In case of the SR-3000 the pixel footprint has a maximum extent of 2.5 cm ($\sim$ 5 m range). Consequently, at a distance of 5 m, only structures bigger than 2.5 cm can be resolved. The reflecting area of the targets are imaged smaller, however due to their reflection characteristics they are still well visible.

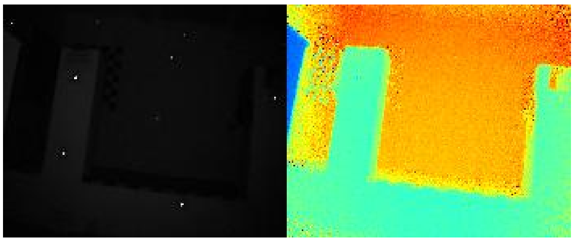

Figure 3.1: An amplitude and range image of data set one (It 2 ms)

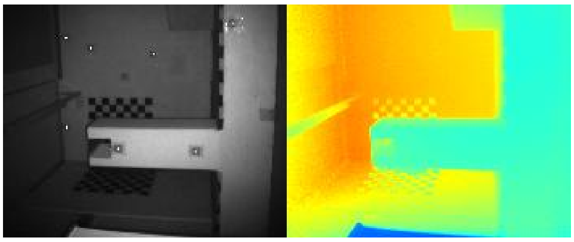

Figure 3.2: An amplitude and range image of data set two (It 30 ms)

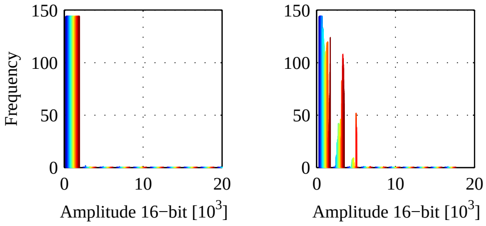

Figure 3.3: Histogram of the measured amplitudes of the first and the second data acquisition for a single frame

Parameter	**Quantity**
Image sensor	176 x 144 pixel
Pixel size	40 x 40 μm
Field of view	47.5 x 39.6 degree
Principal length c	8 mm/201.52509 pixel
Principal point x''_0, y''_0	84.43407/-54.72392 pixel
Norm. Radius	50.
PolynomRadial3	-0.385889

Table 3.1: Specifications of the SR-3000 as in SwissRanger (2008) and parameters of the interior orientation derived by Karel & Niedermayr (2010)

Figure 3.1 shows an amplitude and a range image of the first data set, which comprises very noisy images. The reason for this, is the chosen integration time, which had to be low due to the retro-reflecting targets. As those illuminated targets reflect well, they rapidly cause saturation and the range and amplitude measurements are distorted. Furthermore, an integration time of 2 ms results in an amplitude image with less contrast. Thus, a trade-off between measuring the targets and gathering images with less noise has to be made.

The second data set had been recorded in order to receive less noisy images. Therefore, the area of the retro-reflecting targets had been made narrower. In Figure 3.2 one frame of the second data set is presented. The amplitude and range image show less noise due to an integration time of 30 ms.

While recording data with the SR-3000, several error sources discussed in Chapter 2.1.1 had to be taken into account. First of all, a warm-up period of half an hour was considered. As mentioned before, noise and overexposure mainly determined the camera settings. Motion blur mainly arises for a longer integration time. Thus, the first data set should be less affected by it than the second one. However, in case of data set one the camera was moved faster. Certainly, other errors are still included in the range observations and need to be modelled. Hence, they will be discussed in the next section.

3.2 Data Preprocessing

The data preprocessing step includes checking the data on gross errors, considering systematic errors and masking observations for the relative orientation.

For the bundle adjustment the amplitude image of the control points is of importance. Gross errors such as overexposure and motion blur can distort the image of the targets. Consequently, distorted image coordinates would be observed. If the bundle adjustment is extended by the range measurements, gross and systematic range errors have to be considered to receive accurate ranges.

In case of the relative orientation, gross errors influence the spatial and temporal derivatives. Systematic errors also need to be considered, if they are a function of the position in the image. Furthermore, the image of the retro-reflecting targets is masked to reach an independent solution.

The SR-3000 quantises with 16 bits (0 - 65,535). The range values ρ have been normalised by the unambiguity range. The range calibration, described in more detail later on, is applied on the ranges

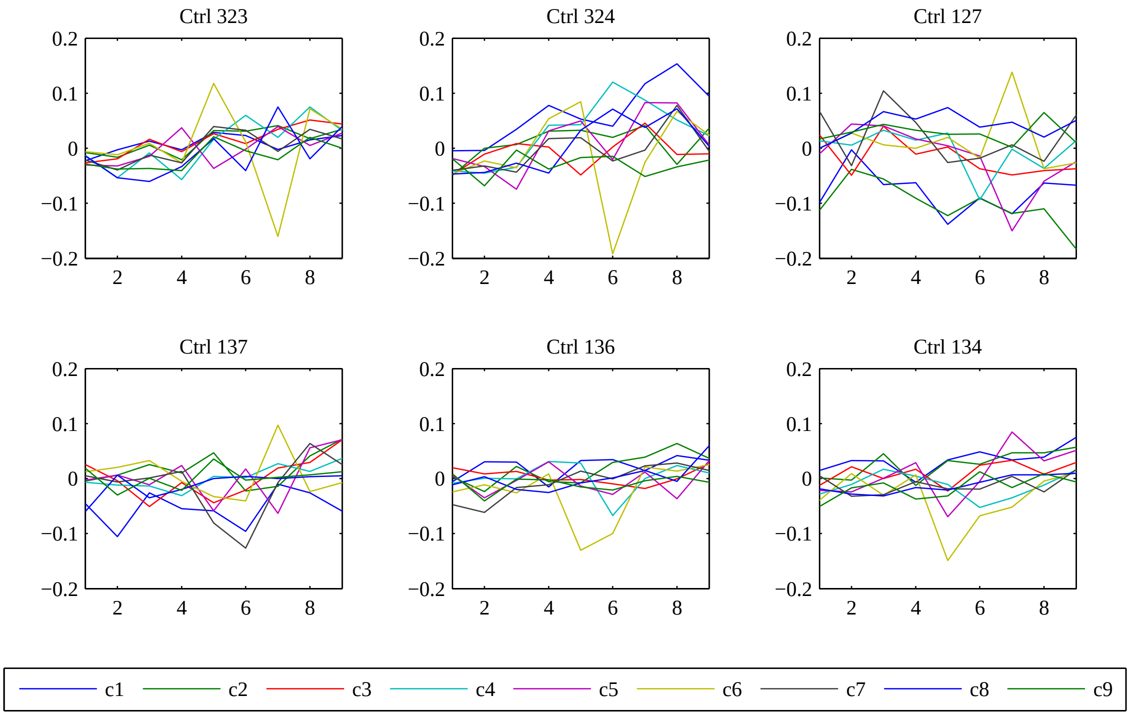

Figure 3.4: The six control points (Ctrl) of frame 24 of data set two are shown. Each plot presents the range deviations from the median in a 9 by 9 neighbourhood of the target. Each line represents one column c_i of that neighbourhood. The influence of the high reflectivity of the targets should be visible in column $c5$ at row 5.
The deviations are nearly the same for all control points and a systematic in the range deviations is not recognisable.

directly after the normalisation. Thus, in the following preprocessing steps the calibrated ranges are considered. For visualisation and measurement of the image points, the bit depth of the amplitude image has been reduced to 8 bits.

The original amplitude spectra are illustrated in the histograms in Figure 3.3. The values of the first data set mainly range from 8 to 700, its amplitude spectrum shows an underexposed image. The spectrum of the second one is more spread out and the amplitudes basically reach up to 12 bit. In both data sets few observations show a significant higher value than the mentioned upper threshold, as visible in Figure 3.3. Those values are related to the control points. It had to be verified, if those pixel are overexposed.

Gross Errors. According to the histograms in Figure 3.3, they are not. A typical overexposed image would show a similar behaviour as the underexposed (first plot in Figure 3.3), but with amplitude values concentrated on the upper side of the histogram. Generally, the retro-reflecting targets are too small in size to figure out in that histogram, if saturation arose or not. Due to maximum amplitude values of 14 bit it can be concluded, that no saturation occurred.

Even the range data could contain saturation effects and needs to be checked. If neighbouring pixel of the retro-reflecting targets are affected in range, then saturation took place. In Figure 3.4 the distributions of the range observations in a 9 by 9 vicinity of the control points are shown. The plots do not show a consistent trend of influenced neighbouring pixel. The range measurements, which belong to the targets themselves, are affected. However, the deviation is within the measurement variance, but saturation

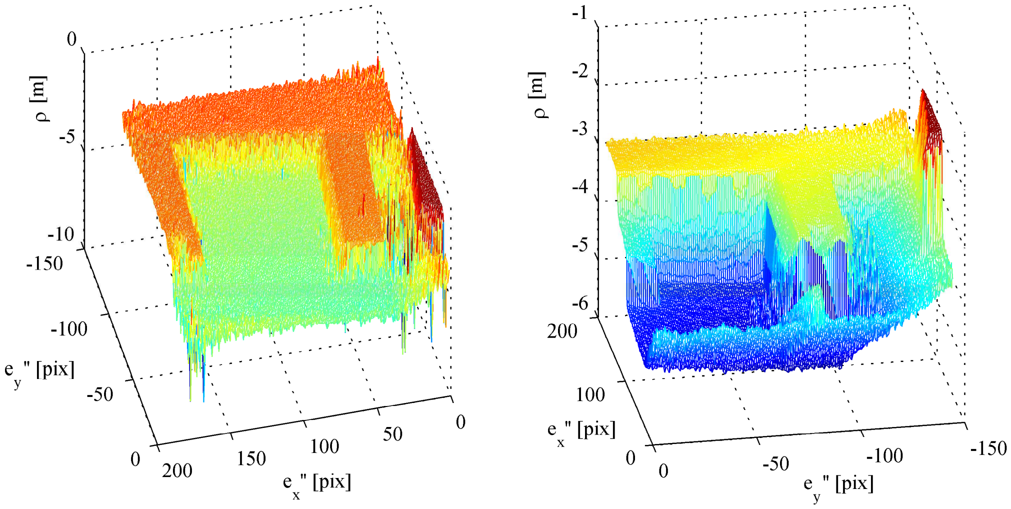

Figure 3.5: Left: Shows the range data of one frame of the first data set; Right: Presents the range data of one frame of the second acquirement

effects would be significant larger than the measurement variance. Thus, no saturation took place. Furthermore, no systematic effects are recognisable.

Figure 3.5 presents the raw range data ρ of the first and the second data acquisition. The first one contains much more noise compared to the second data set. The range distribution of the second one is smoother, clearly recognizable at the object edges. The range errors at the edges likely arise from motion blur. The chequerboard behind the left pillar in the scene is visible in both data sets. In this area the range measurements of the first data set show conspicuous deviations from their vicinity. Through the use of a median filter in both data sets, those isolated too optimistic distance measurements are removed and noise in the image is reduced. Hence, gross and random errors are eliminated.

One preprocessing step is the calibration of the measurement device. The calibration used was accomplished by Karel & Niedermayr (2010). In their self-calibration, parameters of the interior orientation and the ranging system have been estimated. The interior orientation comprises the principal length c, the principal point (x_0'', y_0'') and the lens distortions. The distortions are described by the normalisation radius and the coefficient of the radial lens distortion. The values in Table 3.1 have been achieved.

Systematic Range Errors. The errors of the ranging system have been modelled with respect to the ranges, the amplitudes, the integration time and the position in the image. The application of the calibration functions of Karel & Niedermayr (2010) leads to the following corrections: The range error dependent on the position in the image lies between +0.1 and -0.5 m. The range values are corrected by approximately 0.1 m for low amplitudes up to 200 (see Figure 3.3). The ranges themselves are improved by a maximum of -0.06 m. Furthermore, the calibration leads to a range correction of 0.017 m for an integration time of 30 ms.

However, this calibration was done years before and the interior orientation of the SR-3000 is not stable. Additionally, the materials and other camera components deform with time and hence a re-calibration

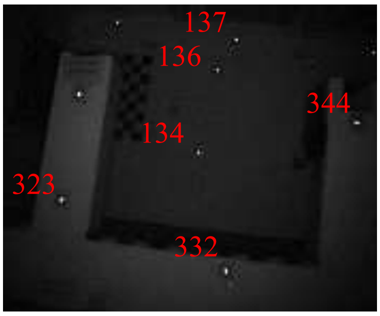

Figure 3.6: Data set 1: Distribution of the control points used

Figure 3.7: Data set 2: Distribution of the control points used

would be necessary. While analysing the results, keep in mind that systematic effects are not completely compensated and some error sources are not considered at all e.g. internal scattering and multipath.

Masking of Observations. The range and optical flow algorithm uses the complete image information to compute the relative motion. It is a complementary method to the bundle adjustment, which bases on the control points (retro-reflecting targets). To achieve a relative orientation result independent of the artificial texture, the retro-reflecting targets are masked in the amplitude image. Therefore, a median filtered amplitude image is processed as a reference. The difference of the original data and the filtered image has been computed in order to automatically detect the retro-reflecting targets. Based on an empirically determined quantile the largest deviations, which correspond to the targets, are found and filtered in a 5 by 5 window.
In the course of the median filtering of the ranges, the retro-reflecting targets have already been masked in the range images.

3.3 Bundle Adjustment

Bundle adjustment is a global orientation method, as discussed in Chapter 2.3. Images are connected and related to a global coordinate frame simultaneously. The basic constraint is that all rays belonging to one control point intersect optimally.

A general work flow for a marker-based orientation method can be described as follows:

- Preparing object information (e.g control points, geometric elements)
- Measurement of image points (e.g. centroid calculation)
- Calculating the orientation parameters (exterior and interior orientation)
- Reconstruction of the object or the camera trajectory (new points, geometric elements) (Luhmann, 2000, p. 222)

Input. Due to the data acquisition in the measurement laboratory of the department, precisely determined control points are available. The standard deviation σ_{ctrl} has been estimated with 1 mm. The spatial distribution of the control points is shown in Figure 3.6 and 3.7. Thereby, a datum needs to be defined.

Parameter	**Quantity**
# Images	91
Control points	134, 136, 137, 323, 332, 344
# Observations n	1092
# Unknowns u	546
# Redundancy r	546
Trajectory length	0.678 m
Estimated unknowns:	
$\boldsymbol{p}_{0,1}$	$[5.183, 1.814, 1.121]^T$ m
$\boldsymbol{p}_{0,91}$	$[4.587, 2.136, 1.130]^T$ m
$avg(\alpha, \zeta, \kappa)$	5.6, 111.6, 107.4 gon
Estimated accuracies:	
$\hat{\sigma}_0^2$	0.125
$\hat{\sigma}_{img}$	0.03 pixel
$median(\hat{\sigma}_{p_0})$	4.4, 5.3, 4.4 mm
$median(\hat{\sigma}_\alpha, \hat{\sigma}_\zeta, \hat{\sigma}_\kappa)$	0.0989, 0.0852, 0.0891 gon

Table 3.2: Specifications of the first data set (It 2 ms)

Parameter	**Quantity**
# Images	136
Control points	127, 134, 136, 137, 323, 324
# Observations n	1632
# Unknowns u	816
# Redundancy r	816
Trajectory length	0.628 m
Estimated unknowns:	
$\boldsymbol{p}_{0,1}$	$[4.382, 2.272, 1.015]^T$ m
$\boldsymbol{p}_{0,136}$	$[4.946, 1.996, 0.995]^T$ m
$avg(\alpha, \zeta, \kappa)$	11.1, 111.9, 0.1 gon
Estimated accuracies:	
$\hat{\sigma}_0^2$	0.160
$\hat{\sigma}_{img}$	0.04 pixel
$median(\hat{\sigma}_{p_0})$	10.2, 11.0, 7.4 mm
$median(\hat{\sigma}_\alpha, \hat{\sigma}_\zeta, \hat{\sigma}_\kappa)$	0.1551, 0.1347, 0.1574 gon

Table 3.3: Specifications of the second data set (It 30 ms)

In case of a constrained datum definition, the control points are considered as error free or constant. This datum should only be used, if the accuracy level of the control points is 5 to 10 times better than the photogrammetric point estimation (Luhmann, 2000, p. 273). In addition, if tensions in the geometric distribution of the control points are included, the bundles of rays will be deformed. A datum, which considers the standard deviations of the control points, is the weighted datum definition. Thereby, tensions are expressed in the a posteriori accuracy $\hat{\sigma}_{ctrl}$.
The measurement of the retro-reflecting targets was proceeded semi-automatically, as described in Chapter 2.3. After a manual determination of the target position in the first image, it was automatically computed for the consecutive frames. Based on a centroid calculation, a position accuracy in sub-pixel range is achieved. The image measurement accuracy σ_{img} is assumed to be 0.5 pixel.
The interior orientation of Karel & Niedermayr (2010) in Table 3.1 has been applied.

The first part of Table 3.2 and 3.3 contains general information about the two bundle blocks, e.g. the number of frames, the number of used control points. The unknown parameters are the exterior orientations of all frames, which consist of the projection centre $\boldsymbol{p}_0$ and the angles of the rotation matrix $\boldsymbol{R}_{\alpha,\zeta,\kappa}$. The control points and the interior orientation are considered to be constant. Each image shows six control points, hence 12 observation equations as in (2.12) are formulated. Subtracting the 6 unknown parameters of the exterior orientation delivers a redundancy of 6 for each image. Multiplied by the number of images results in the total redundancy r of 546 and 816.
Only translations have been accomplished. Therefore, the angular orientation of the camera remains nearly the same for all frames and average values are presented in Table 3.2 and 3.3. As mentioned earlier, for the second data acquisition the camera was rotated around its optical axis. Therefore, κ differs by approximately 100 gon between data set one and two.

Output. The accuracy levels in Table 3.2 and 3.3 show the inner accuracy of the photogrammetric utilisation. The different accuracy levels of data set one and two might be due to a different intersection geometry, a different detectability of the control points or tensions in the control points.
Figure 3.8 presents the trajectories of the two data sets. The assumption of a poor intersection geometry

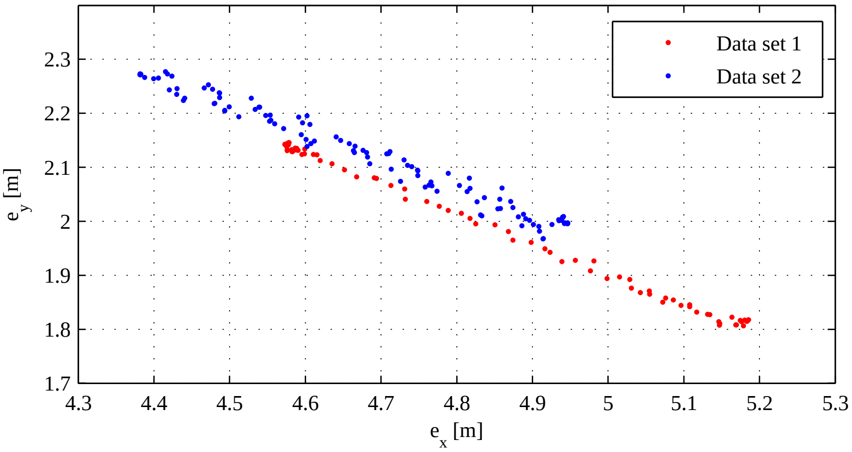

Figure 3.8: The trajectories of the first and the second data set in the global coordinate frame

in case of the second data set can not be confirmed, due to a quite similar acquisition geometry compared to the first one. The control points measured in the two data sets differ only in two points, as shown in Figure 3.6 and 3.7. The results in Table 3.2 and 3.3 base on a constrained datum definition. Computing the same adjustment with control points as observed unknowns shows, that the accuracy levels of the control points $\hat{\sigma}_{ctrl}$ and the image observations $\hat{\sigma}_{img}$ stay the same. Thus, it can be concluded, that no tensions are included in the geometric distribution of the control points. The detectability of the control points was nearly the same for both data sets. Motion blur might have caused the reduction in quality of data set two. A verification of this still needs to be done.

The bundle adjustment is sensitive to gross errors and systematic errors. If gross errors are still included in the observations and systematic errors, as summarised in Table 2.1, are not compensated, they will have a negative influence on the result.

3.4 Relative Orientation

The relative orientation is computed as discussed in Chapter 2.5. The observations of the relative orientation are the range $\rho(x'', y'', t)$ and amplitude values $I(x'', y'', t)$ of two consecutive images. One image provides 25,344 amplitude and range observations. As mentioned earlier, the relative orientation is based on a differential range and optical flow. The spatial image derivatives I_x, I_y and ρ_x, ρ_y are computed by the 5-point Simoncelli filter (Simoncelli 1994, cited in Ghuffar, 2014, p. 30), see Formula (3.1). The temporal image derivatives I_t, ρ_t are the difference of corresponding pixel in two consecutive frames.

$$f(x'') = \begin{bmatrix} 1 & -8 & 0 & 8 & -1 \end{bmatrix} \cdot \frac{1}{12} \tag{3.1}$$

To calculate the spatial derivative, the filter is applied through the convolution $I_x = I(x'', y'') * f(x'')$ in $\boldsymbol{e}''_x$-direction and $I_y = I(x'', y'') * f(x'')^T$ in $\boldsymbol{e}''_y$-direction. The two instantaneous neighbouring pixel are involved with two thirds of their original value and the two second neighbours only with one twelfth. The

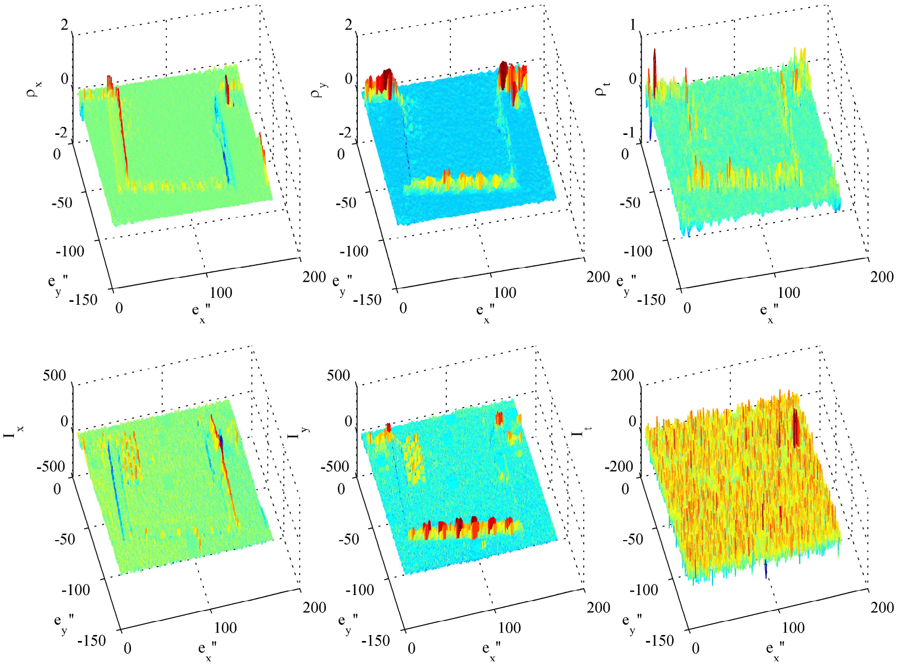

Figure 3.9: Filtered image pair of data set **one** - The first and center column show the result of a convolution with the 5-point Simoncelli derivative filter in $\boldsymbol{e}_x''$- and $\boldsymbol{e}_y''$-direction. The right column shows the change in time ρ_t and I_t of two consecutive frames.

spatial image derivatives define the areas, where texture is present and motions can be detected. In areas with no texture, the terms to the right of the equal sign in Equation (2.16) become zero. If a motion of more than 2.5 pixel is accomplished, then the areas with non-zero values of the spatial and the temporal derivatives will not match and a wrong solution will be achieved. Consequently, the detectable motion is restricted by this filter and a maximum movement of 2.5 pixel per frame can be performed.
The results in Figure 3.9 and 3.10 are obtained by applying the derivative filter on the preprocessed data. When filtering in $\boldsymbol{e}_x''$-direction, changes perpendicular to that direction can be extracted. The same is true for the $\boldsymbol{e}_y''$-direction. The temporal derivative ρ_t, I_t is the difference of corresponding pixel of two consecutive frames. It represents the same trend as in ρ_x and ρ_y. However, the difference image ρ_t shows much more noise than the filtered ones. Comparing Figure 3.9 with 3.10 shows, that data set two resolves much finer structure than data set one. Due to the filtering, the three outer most pixel of each frame are not considered. Hence, the number of observations reduces to 46,920.

As mentioned in Chapter 2.2, the rotation matrix applied is parametrised by ω, φ, κ. Due to small motions it has been assumed, that the small angle approximation is a sufficient parametrisation. The error of using Formula (2.6) instead of the full rotation matrix (2.5) is shown in Figure 3.11. The maximum angle is limited by the filter length, as discussed before. If the change is within 2.5 pixel, it will be detected. Those 2.5 pixel correspond to a maximum angle of 0.8 gon, which amounts to an error of 0.7 mgon, according to Figure 3.11. The maximum error at the object is less than 0.1 mm, thus the approximation is valid.

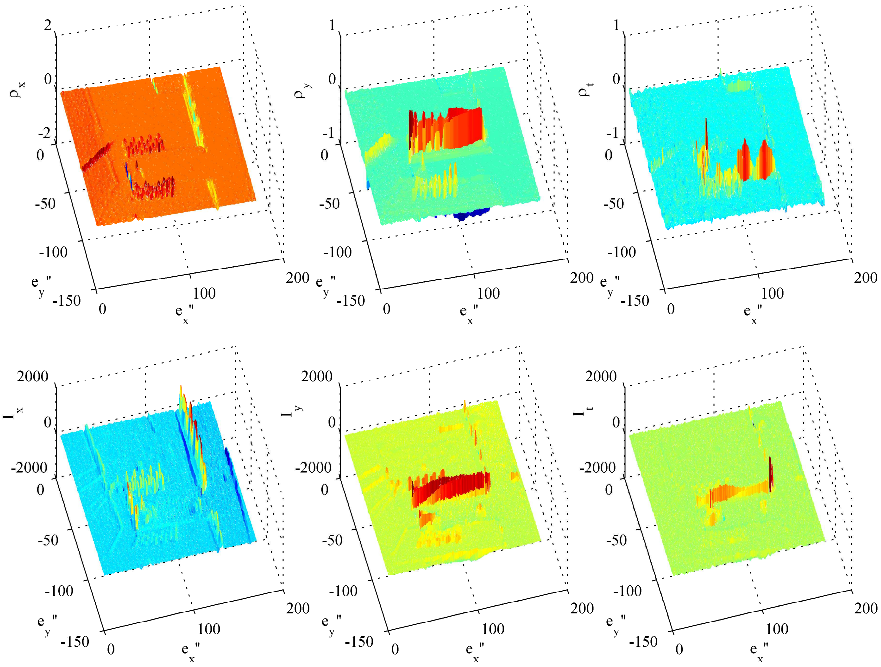

Figure 3.10: Filtered image pair of data set **two** - The first and center column show the result of a convolution with the 5-point Simoncelli derivative filter in $\boldsymbol{e}_x''$- and $\boldsymbol{e}_y''$-direction. The right column shows the change in time ρ_t and I_t of two consecutive frames.

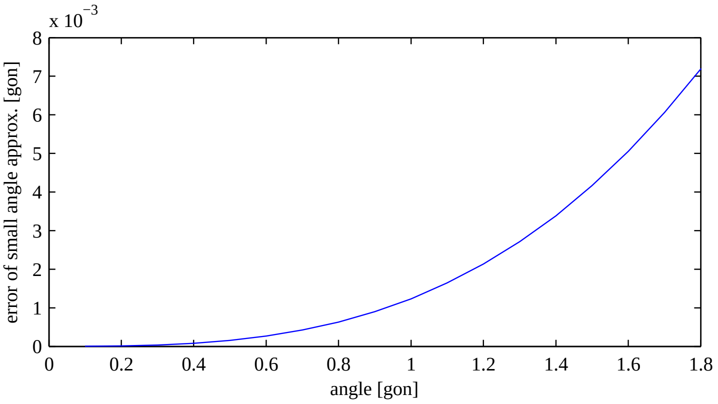

Figure 3.11: Small angle approximation error of rotation matrices

wfun	**Mean**	**Median**	**Max**	**RMSE**		**Mean**	**Median**	**Max**	**RMSE**	
ols	0.0058	0.0051	0.0145	0.0066	trans[m]	0.0034	0.0031	0.0100	0.0038	α[gon]
huber	0.0049	0.0042	0.0147	0.0059		0.0031	0.0025	0.0091	0.0035	
bisquare	0.0048	0.0042	0.0160	0.0058		0.0030	0.0026	0.0088	0.0035	
talwar	0.0049	0.0040	0.0161	0.0059		0.0030	0.0026	0.0088	0.0034	

Table 3.4: Data set 1: Relative pose error for different adjustment types

wfun	**Mean**	**Median**	**Max**	**RMSE**		**Mean**	**Median**	**Max**	**RMSE**	
ols	0.0055	0.0046	0.0223	0.0072	trans [m]	0.0033	0.0028	0.0133	0.0044	α [gon]
huber	0.0053	0.0038	0.0219	0.0071		0.0033	0.0027	0.0139	0.0043	
bisquare	0.0055	0.0044	0.0217	0.0073		0.0033	0.0027	0.0141	0.0044	
talwar	0.0055	0.0043	0.0213	0.0074		0.0033	0.0028	0.0142	0.0044	

Table 3.5: Data set 2: Relative pose error for different adjustment types

The parameter estimation of the highly over-determined relative orientation is processed by the Matlab function *robustfit*:

$$output = robustfit(\ \boldsymbol{X},\ \boldsymbol{y},\ wfun)$$

Whereby, $\boldsymbol{X}_{n\times u}$ is the predictor, which corresponds to the design matrix $\boldsymbol{A}$. u predictors are estimated by n observations. $\boldsymbol{y}_{n\times 1}$ contains the observed response. The method uses an iteratively, re-weighted least squares algorithm with various weighting functions *(wfun)* implemented (Matlab, 1984-2010). It also supports the ordinary least squares adjustment.

No independently moving objects are in the scene, consequently a robust estimation is not necessary. In case of occurring saturation effects, a robust estimation would be reasonable. In Figure 3.12 the trajectories with different weighting functions applied are presented. In Table 3.4 and 3.5 the relative pose error (RPE), in Table 3.6 the absolute trajectory error (ATE) of those trajectories are shown. As a ground truth the adjusted line of the bundle adjustment results are considered because of the low precision reached for data set two. *huber*, *bisquare* and *talwar* are M-estimators, which means that a loss function is minimised. In case of Tukey's *bisquare* and the *talwar* function, convergence behaviour can cause a problem.

No convergence arises by applying the *talwar* and *bisquare* loss function on data set two, visible in Table 3.5 and 3.6. The translational relative pose error as well as the absolute trajectory error increase for both weighting functions. Figure 3.12 (second row) shows the too short trajectories of the not end-iterated algorithm. However, it is not possible to change the iteration limit for the *robustfit* function.

Therefore, the *huber* estimator is used in case of a robust estimation. Only observations with a weight larger than 0.8 are considered in the following ordinary least squares adjustment.

The standard deviation of each orientation difference (RPE) is approximately 7 mm for the translational and less than 0.4 mgon for the rotational part, according to Table 3.4 and 3.5. The influence of the robust adjustment on an estimated trajectory has not been investigated. In case of data set one, it seems to improve the results. However, if the utilisation is justified, is an open question.

Therefore, the ordinary least-squares solution is chosen. Furthermore, the data has been checked

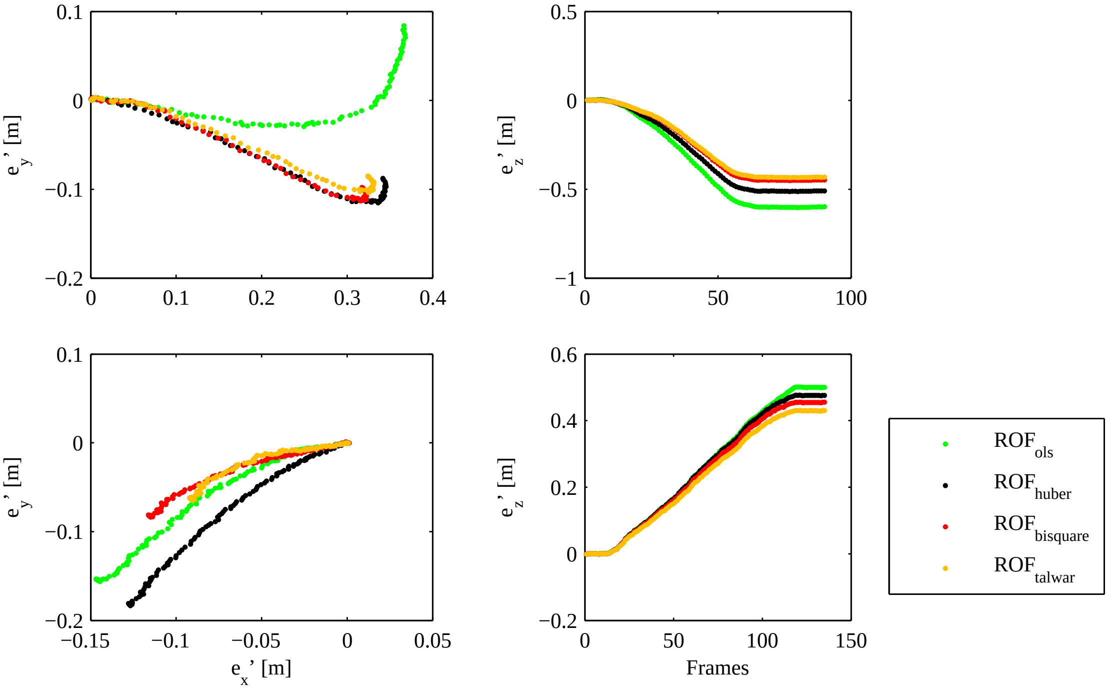

Figure 3.12: Comparison of ordinary least squares adjustment and various robust loss functions - in the first row the result of the incremental orientation of the first data set is presented and in the second row the result for the second one. The main motion has been in $\boldsymbol{e}_z'$-direction.

wfun	**Max** [m]	**RMSE** [m]	α_1 [gon]	α_n [gon]		**Max** [m]	**RMSE** [m]	α_1 [gon]	α_n [gon]	
ols	0.2234	0.1212	0.0004	0.0710	1	0.2363	0.1615	0.0002	0.0719	2
huber	0.0755	0.0474	0.0003	0.0316		0.2118	0.1417	0.0002	0.0445	
bisquare	0.1312	0.0919	0.0002	0.0348		0.3031	0.2006	0.0003	0.0542	
talwar	0.1467	0.1012	0.0004	0.0346		0.3261	0.2144	0.0002	0.0552	

Table 3.6: Absolute Trajectory Error of data set one and two

Parameter	**Quantity**
# Image pairs	90
# Avg. obs/image n	46,920
# Unknowns/image u	6
# Avg. red/image r	46,914
Trajectory length	0.707 m
Estimated unknowns:	
$\boldsymbol{t}'_{incr,90}$	$[0.366, 0.084, -0.599]^T$ m
Estimated accuracies:	
$median(\hat{\sigma}_{t'})$	$(3.9, 4.2, 2.5) \times 10^{-4}$ m
$median(\hat{\sigma}_{\omega}, \hat{\sigma}_{\varphi}, \hat{\sigma}_{\kappa})$	$(1.1, 1.0, 0.9) \times 10^{-4}$ gon

Table 3.7: Data set 1: Specifications of the relative orientation.

Parameter	**Quantity**
# Image pairs	135
# Avg. obs/image n	46,920
# Unknowns/image u	6
# Avg. red/image r	46,914
Trajectory length	0.543 m
Estimated unknowns:	
$\boldsymbol{t}'_{incr,135}$	$[-0.147, -0.154, 0.500]^T$ m
Estimated accuracies:	
$median(\hat{\sigma}_{t'})$	$(1.6, 1.6, 1.1) \times 10^{-4}$ m
$median(\hat{\sigma}_{\omega}, \hat{\sigma}_{\varphi}, \hat{\sigma}_{\kappa})$	$(4.3, 4.6, 4.1) \times 10^{-5}$ gon

Table 3.8: Data set 2: Specifications of the relative orientation.

on gross errors. In Table 3.7 and 3.8 the parameters of this adjustment are summarised. A comparison of the estimated precisions shows, that better results are achieved in case of the second data set. The median standard deviation of the first data set shows an accuracy level half as good as for data set two. Figure 3.9 and 3.10 already illustrate a difference of quality in the observations.

The standard deviations of the translations amount to 0.6 mm and 0.25 mm (data set one and two, according to Table 3.7 and 3.8). In comparison to the relative pose error, the results are too optimistic. One reason for that deviation is the artificial high redundancy of the adjustment system, as the complete image information is used. As shown in Figure 3.9 and 3.10, only areas presenting texture contribute to the parameter estimation.

It is assumed, that the amplitude and the range observations are independent and no correlations between the two observation groups exist (Ghuffar, 2014, p. 30). The amplitude terms are weighted in that way, that range and amplitude derivatives have the same mean value.

In Equation (3.2) the a posteriori correlation coefficients of $\boldsymbol{t}$ and ω, φ, κ are presented. It shows, that the $t_{x'}$-coordinate and φ as well as $t_{y'}$ and ω are highly correlated.

$$C_{correl,1} = \begin{bmatrix} 1 & -0.05 & -0.60 & -0.03 & \mathbf{0.97} & -0.04 \\ -0.05 & 1 & 0.01 & \mathbf{-0.97} & -0.08 & -0.39 \\ -0.60 & 0.01 & 1 & 0.11 & -0.53 & 0.23 \\ -0.03 & \mathbf{-0.97} & 0.11 & 1 & 0.1 & 0.42 \\ \mathbf{0.97} & -0.08 & -0.53 & 0.01 & 1 & 0.03 \\ -0.04 & -0.39 & 0.23 & 0.42 & 0.03 & 1 \end{bmatrix} \tag{3.2}$$

3.5 Comparison

In Figure 3.13 a first comparison of the two methods is presented. For that purpose, the results of the relative orientation are transformed into the global coordinate frame C. This is done by the similarity transformation as written in Equation (3.3). Based on the global orientation parameters of the first camera position, the relative orientation results are transformed. In the first row of Figure 3.13 the results of the first data set are plotted. The light blue line are the results of the bundle adjustment and in dark blue an optimal line has been adjusted. The incremental relative orientation result, represented by the green

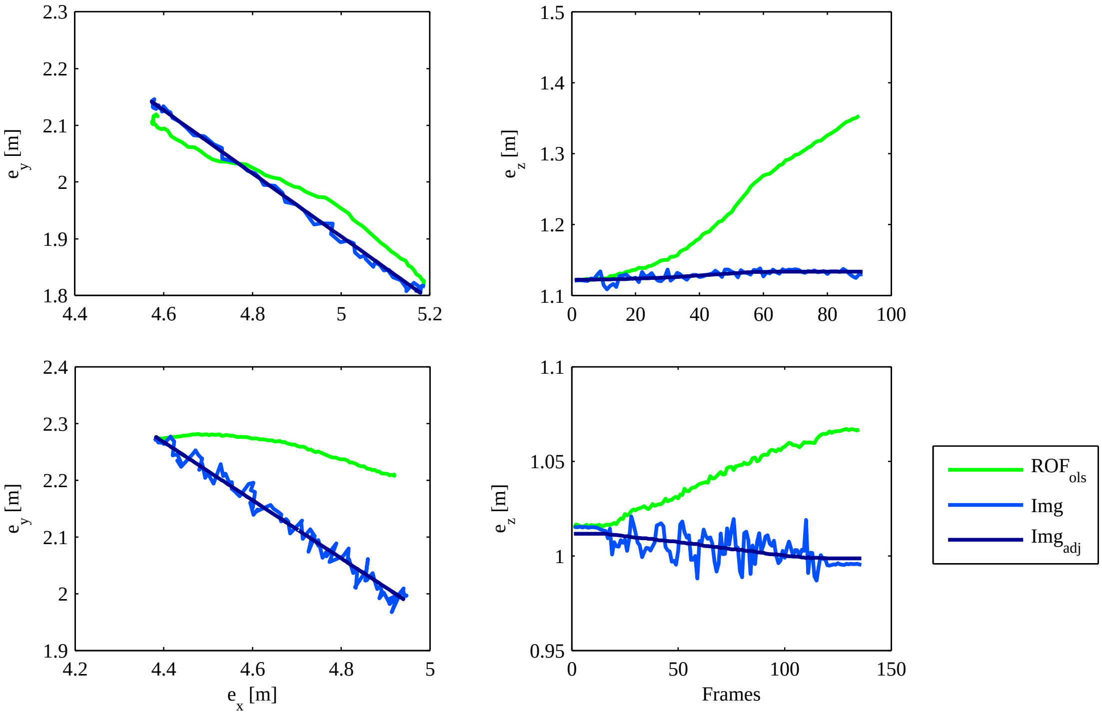

Figure 3.13: Comparison of the global (Bundle adjustment - BA) and the relative (Range and Optical flow - ROF) orientation method. In the first row the results of data set one and in the second one the results of data set two are illustrated.

graph, shows the typical accumulation of the error. In data set one the deviations are stronger in the $\boldsymbol{e}'_z$-direction, data set two in the second row of Figure 3.13 shows greater deviations in the $\boldsymbol{e}'_y$-direction. The absolute trajectory error of Table 3.6 amounts to 22.3 cm and 23.6 cm for data set one and data set two, respectively.

$$\boldsymbol{t}_{incr,i} = \boldsymbol{R}_{\alpha,1}\boldsymbol{t}'_{incr,i} + \boldsymbol{p}_{0,1} \tag{3.3}$$

The global orientation method is an unbiased estimator. However, different error sources and a low image sensor resolution cause the variation of the bundle adjustment results. In contrary, the relative orientation yields a smooth result, but the error of each orientation sums up to a large drift.

Consequently, the following assumption can be formulated: A combination of those two methods leads to a smooth, unbiased estimate of the true trajectory.

4. Implementation

As postulated in Chapter 3.5, a combination of the bundle adjustment and the relative orientation leads to a smooth, unbiased estimate of the true trajectory. In this chapter the relative orientation is integrated into the existing bundle adjustment of OrientAL. Additionally, the range measurements are included. Therefore, the functional and stochastic model are extended. Furthermore, questions of the implementation and the variance component estimation are treated.

4.1 Functional Model

A prerequisite for the integration of the relative orientation is, that a reasonable relation between observations and unknown parameters can be formulated, as mentioned in Chapter 2.6. The relative orientation is included in the established adjustment of indirect observations. Due to the additional estimated covariance matrix of the relative orientation (stochastic model), it is possible to use the relative orientation results. Those relative translations and rotations are treated as observations in further consequence. Thus, a function has to be set up, that expresses the results of the relative orientation as a function of the unknowns of the bundle adjustment. Hence, the equations of the methods are rewritten and the common ground is emphasised.

The **bundle adjustment** bases on the image observations of control points. In this chapter, Equation (2.9) is used as a similarity transformation, which describes the relation of the global coordinate system C and a camera coordinate system C'_j. Hence, instead of using the scale λ_{ij} and the image vector in the camera coordinate system $\boldsymbol{p}'_{ij}$, the vector to the point P_i with respect to the camera coordinate system of frame j is denoted as $\boldsymbol{p}'_{i,j}$.
An object point $\boldsymbol{p}_i$ can be computed by adding the rotated vector $\boldsymbol{p}'_{i,j}$ to the unknown camera position $\boldsymbol{p}_{0,j}$, as presented in Equation (4.1). The index i corresponds to the object in the global space. The index j belongs to the frame and runs from 1 to n. The bundle adjustment delivers n exterior orientation parameters $\boldsymbol{p}_{0,j}$ and $\boldsymbol{R}_{\alpha,j}$ for n images.

$$\boldsymbol{p}_i = \boldsymbol{R}_{\alpha,j}\boldsymbol{p}'_{i,j} + \boldsymbol{p}_{0,j} \tag{4.1}$$

The **relative orientation** is a method, which uses the whole image area to compute the relative motion between two consecutive frames. It is a transformation between camera coordinate systems C' of subsequent frames. As written in Equation (4.2), the relative orientation parameters $\boldsymbol{t}'_k$ and $\boldsymbol{R}_{\omega,k}, (k = 2, ..., n)$ are defined with respect to the camera coordinate system of frame $j + 1$. Therefore, the index k starts at

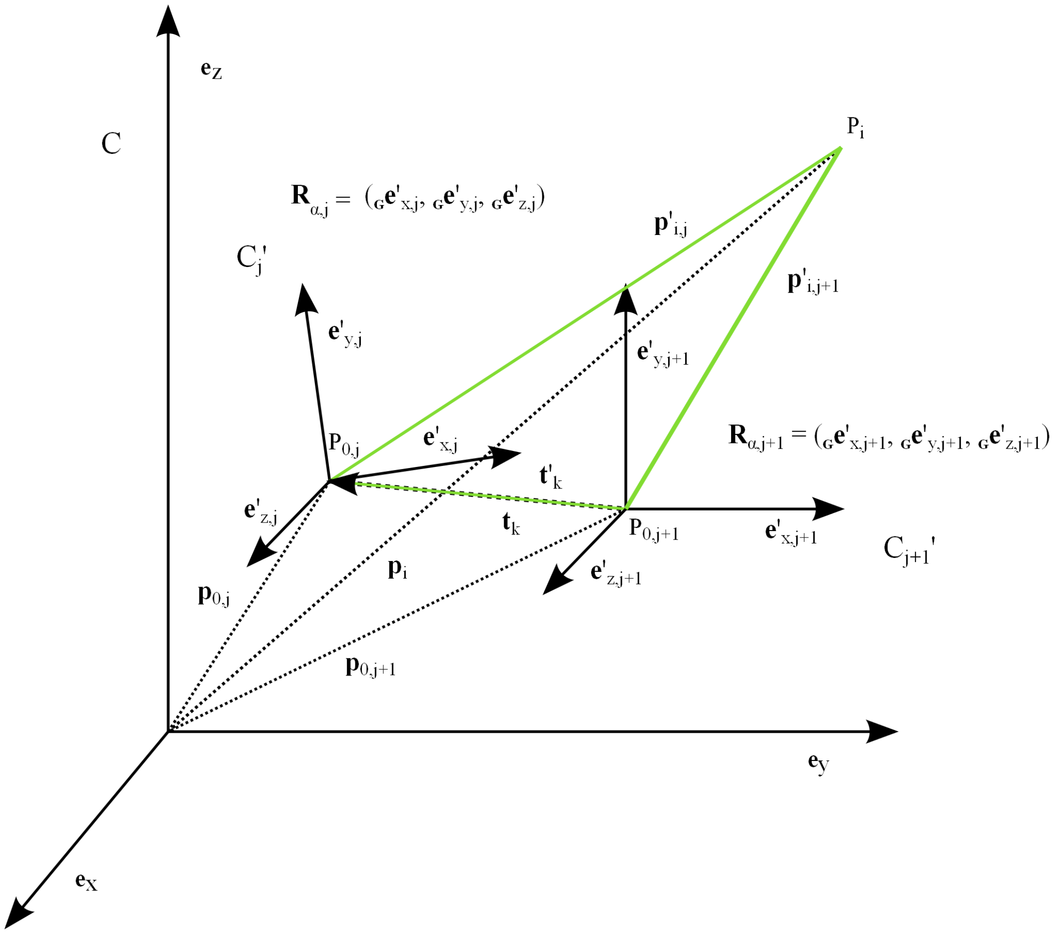

Figure 4.1: Relation of relative and global orientation.

2. The relative orientation yields $n-1$ translations and rotations.

$$\boldsymbol{p}'_{i,j+1} = \boldsymbol{R}_{\omega,k}\boldsymbol{p}'_{i,j} + \boldsymbol{t}'_k \tag{4.2}$$

Equation (4.1) describes a transformation from the camera to the global space. Whereas, the relative orientation describes the camera motion from one camera coordinate frame to the next.

In Figure 4.1 the orientation parameters of the two methods are illustrated. Crucial is the knowledge of the reference coordinate system of the relative orientation. In case the reference coordinate system is the one at frame $j+1$, the difference vector $\boldsymbol{t}_k$ leads from $\boldsymbol{p}_{0,j+1}$ to $\boldsymbol{p}_{0,j}$. To transform this directional vector $(\boldsymbol{p}_{0,j} - \boldsymbol{p}_{0,j+1})$ from the global to the camera coordinate system of frame $j+1$, the rotation matrix $\boldsymbol{R}^T_{\alpha,j+1}$ is applied. Hence, the relation of the translation is established in Equation (4.3).

The columns of the rotation matrix $\boldsymbol{R}_\alpha$ are the basis vectors of the camera coordinate system, represented in the global one (G), as presented in Figure 4.1. The columns of the difference rotation $\boldsymbol{R}_\omega$ are the basis vectors of the camera coordinate system of frame j, represented in the one of frame $j+1$.
The rotation matrix $\boldsymbol{R}_{\alpha,j+1}$ is right-multiplied by the difference rotation $\boldsymbol{R}_{\omega,k}$ and that results in the rotation

matrix $\boldsymbol{R}_{\alpha,j}$, as written in Equation (4.4). Consequently, the relative rotation matrix $\boldsymbol{R}_{\omega,k}$ is the product of $\boldsymbol{R}^T_{\alpha,j+1}\boldsymbol{R}_{\alpha,j}$.

$$\boldsymbol{t}'_k = \boldsymbol{R}^T_{\alpha,j+1}(\boldsymbol{p}_{0,j} - \boldsymbol{p}_{0,j+1}) \tag{4.3}$$

$$\boldsymbol{R}_{\alpha,j+1}\boldsymbol{R}_{\omega,k} = \boldsymbol{R}_{\alpha,j} \quad \rightarrow \quad \boldsymbol{R}_{\omega,k} = \boldsymbol{R}^T_{\alpha,j+1}\boldsymbol{R}_{\alpha,j} \tag{4.4}$$

Equation (4.3) and (4.4) show, that the relative orientation results can be equated with the unknowns of the bundle adjustment. No additional parameters need to be considered and so the integration of the relative orientation produces six additional observation- or residual equations per image and no additional unknowns.

Equation (4.3) gives three observation equations. The rotation matrix consists of nine elements. Due to the small angle approximation of the relative rotation matrix (4.5), the angles ω, φ, κ are directly present. Therefore, the three elements $R_{2,1}, R_{3,1}, R_{3,2}$ are chosen to build the other three observation equations.

$$\begin{bmatrix} 1 & -\kappa_i & \varphi_i \\ \kappa_i & 1 & -\omega_i \\ -\varphi_i & \omega_i & 1 \end{bmatrix} = \boldsymbol{R}^T_{\alpha,j+1}\boldsymbol{R}_{\alpha,j} \tag{4.5}$$

The equations (4.3) and (4.4) can be derived in closed form in two ways.

The first approach bases on the relative orientation equation (4.2). The vector $\boldsymbol{p}'_{i,j+1}$ is transformed into the global coordinate system as done by Equation (4.6). So, the vector $\boldsymbol{p}_i$ in the global coordinate system can be expressed by the vector $\boldsymbol{p}'_{i,j}$ with respect to the camera coordinate system at time j. Consequently, the product of the rotation matrices has to be equal with the rotation matrix $\boldsymbol{R}_{\alpha,j}$ and the terms summarised by the second horizontal bracket in (4.7) correspond to the projection centre $\boldsymbol{p}_{0,j}$ at time j. Equation (4.7) confirms the graphical solution described above.

$$\boldsymbol{R}^T_{\alpha,j+1}(\boldsymbol{p}_i - \boldsymbol{p}_{0,j+1}) = \boldsymbol{R}_{\omega,k}\boldsymbol{p}'_{i,j} + \boldsymbol{t}_k \tag{4.6}$$

$$\boldsymbol{p}_i = \underbrace{\boldsymbol{R}_{\alpha,j+1}\boldsymbol{R}_{\omega,k}}_{\boldsymbol{R}_{\alpha,j}}\boldsymbol{p}'_{i,j} + \underbrace{\boldsymbol{R}_{\alpha,j+1}\boldsymbol{t}'_k + \boldsymbol{p}_{0,j+1}}_{\boldsymbol{p}_{0,j}} \tag{4.7}$$

The second and more direct proof was done by Ghuffar (2014). He compared the same object point $\boldsymbol{p}_i$ in two subsequent frames using the similarity transformation of Formula (4.1). Equating the two formulae of (4.8) leads to an equation, which is similar to the relative orientation. Equation (4.9) shows the relation of a point in the camera coordinate system at frame j with respect to the same point at frame $j+1$. It can be concluded, that the product of the rotation matrices corresponds to the difference rotation $\boldsymbol{R}_{\omega,k}$ and the rotated difference vector is equal to the relative translation $\boldsymbol{t}'_k$.

$$\boldsymbol{p}_i = \boldsymbol{p}_{0,j} + \boldsymbol{R}_{\alpha,j}\boldsymbol{p}'_{i,j}, \qquad \boldsymbol{p}_i = \boldsymbol{p}_{0,j+1} + \boldsymbol{R}_{\alpha,j+1}\boldsymbol{p}'_{i,j+1} \tag{4.8}$$

$$\boldsymbol{p}'_{i,j+1} = \underbrace{\boldsymbol{R}^T_{\alpha,j+1}\boldsymbol{R}_{\alpha,j}}_{\boldsymbol{R}_{\omega,k}}\boldsymbol{p}'_{i,j} + \underbrace{\boldsymbol{R}^T_{\alpha,j+1}(\boldsymbol{p}_{0,j} - \boldsymbol{p}_{0,j+1})}_{\boldsymbol{t}'_k} \tag{4.9}$$

Both closed form approaches clearly show the relation of the global and the relative orientation. In the first one, transforming the relative orientation in the global space leads to the global transformation parameters of frame j. Whereas, the second derivation bases on two consecutive camera positions in the global space, which are equated and the relative transformation parameters are achieved.

Range Measurements. A further improvement is expected by including the range measurements to the control points. The range corresponds to the distance (euclidean norm) between the perspective centre of the camera $\boldsymbol{p}_{0,j}$ and the control point $\boldsymbol{p}_i$, as presented in Formula (4.10). In this case the index i corresponds to the control points - specific points in the global space. This equation produces one residual and due to the only unknown parameter - the perspective centre - each equation increases the total redundancy by one.

$$\rho_{i,j} = \|\boldsymbol{p}_{0,j} - \boldsymbol{p}_i\| \tag{4.10}$$

4.2 Stochastic Model

The stochastic model is extended by the covariance matrix of the relative orientation. The covariance matrix $\boldsymbol{\Sigma}_{ll,ROF}$ of the last frame in data set two is presented in (4.11). It is a full 6×6 matrix. As mentioned before, the translations and rotations are correlated and the accuracy level is too optimistic, which is caused by the artificial high over-determination of the relative orientation and by correlations in amplitude and range between neighbouring pixel, which have not been considered.

$$\boldsymbol{\Sigma}_{ll,ROF,136} = \begin{bmatrix} 3.3 \times 10^{-9} & -2.8 \times 10^{-10} & -6.9 \times 10^{-10} & 4.3 \times 10^{-11} & 8.3 \times 10^{-10} & -7.9 \times 10^{-11} \\ -2.8 \times 10^{-10} & 2.1 \times 10^{-9} & 1.1 \times 10^{-9} & -5.5 \times 10^{-10} & -4.8 \times 10^{-11} & -9.3 \times 10^{-11} \\ -6.9 \times 10^{-10} & 1.1 \times 10^{-9} & 1.8 \times 10^{-9} & -2.3 \times 10^{-10} & -1.4 \times 10^{-10} & -2.4 \times 10^{-12} \\ 4.3 \times 10^{-11} & -5.5 \times 10^{-10} & -2.3 \times 10^{-10} & 1.6 \times 10^{-10} & 6.3 \times 10^{-12} & 2.5 \times 10^{-11} \\ -2.3 \times 10^{-10} & 1.6 \times 10^{-10} & 6.3 \times 10^{-12} & 2.5 \times 10^{-11} & 8.3 \times 10^{-10} & -4.8 \times 10^{-11} \\ -1.4 \times 10^{-10} & 6.3 \times 10^{-12} & -2.4 \times 10^{-12} & 2.5 \times 10^{-11} & -2.3 \times 10^{-11} & 2.8 \times 10^{-10} \end{bmatrix} \tag{4.11}$$

If range measurements are considered in the adjustment as well, a stochastic measure is needed. As mentioned earlier, the measurement precision of the SR-3000 amounts to a few centimetres. Therefore, the ranges are assumed to present a standard deviation σ_{Range} of 0.1 m.

4.3 Approximate Values

Due to a non-linear functional problem, approximate values of the exterior orientations are needed. They can be estimated easily as many points in the measuring field of the laboratory are known. In Table 4.1 the approximations of the unknown parameters are presented. Figures 4.2 and 4.3 show the orientation of the SR-3000 for both data sets. In case of the second data acquisition the ToF camera was rotated around its optical axis by approximately 100 gon.

Parameter	**Quantity**
$\boldsymbol{p}_{0,ap1}$	$[5.0, 2.0, 1.0]^T$ m
$\alpha_{ap1}, \zeta_{ap1}, \kappa_{ap1}$	0.0, 100.0, 100.0 gon
$\boldsymbol{p}_{0,ap2}$	$[4.0, 2.0, 1.0]^T$ m
$\alpha_{ap2}, \zeta_{ap2}, \kappa_{ap2}$	0.0, 100.0, 0.0 gon

Table 4.1: Approximate values of data set one (ap1) and two (ap2)

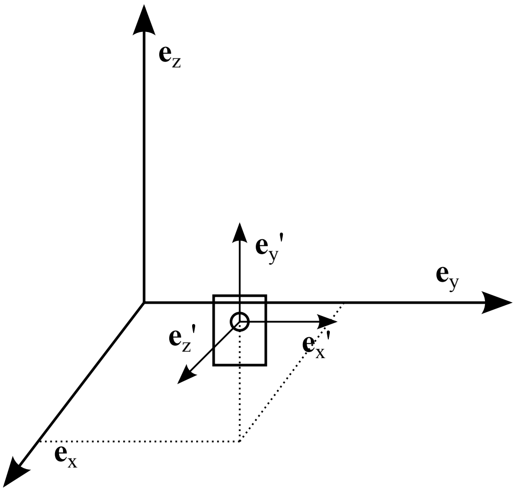

Figure 4.2: Data set 1: Approximate exterior orientation of the camera ($\boldsymbol{R}_{\alpha,ap1}, \boldsymbol{p}_{0,ap1}$)

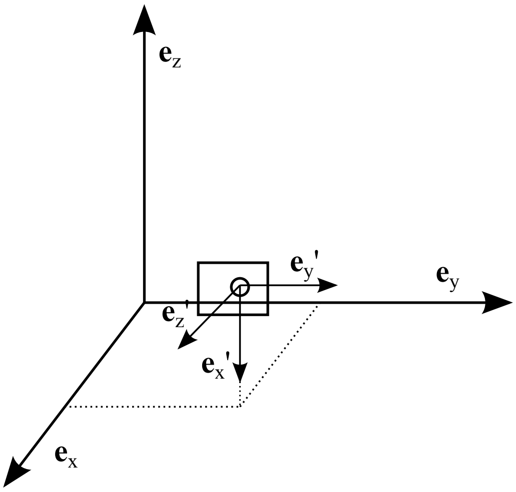

Figure 4.3: Data set 2: Approximate exterior orientation of the camera ($\boldsymbol{R}_{\alpha,ap2}, \boldsymbol{p}_{0,ap2}$)

4.4 Implementation in OrientAl

The programming language used is Python version 2.7.6. It is an open-source, high-level scripting language. Special emphasis is placed on the code readability, which expresses in the indentations used. In addition, no variable types need to be specified in the beginning. Python supports multiple programming paradigms like object-oriented, functional or procedural programming. This work has been scripted in an object-oriented way.

The library used is OrientAL. The adjustment module of OrientAL is based on the C/C++ library Ceres (Agarwal *et al.*, 2013). "Ceres provides low-level administration of observations and parameters, automatic and numeric differentiation, and a common interface to various numerical solvers" (Karel *et al.*, 2013). Due to Python-bindings, this functionality and the high-level functionality of Python is combined in OrientAl. Additionally, the Python packages NumPy and SciPy are used. NumPy arrays correspond to matrices and standard operators are available.

Modelling. A robust non-linear least-squares problem can be formulated as in Equation (4.12). The expression in the sum is denoted as the **residual block**, which consists of the loss function ρ_i and the cost function f_i. If outliers are included in the data, a loss function should be applied. In this work the "Trivial" loss function is considered, hence an ordinary least-squares adjustment is computed.

$$min \quad \frac{1}{2}\sum_i \rho_i(||f_i(x_{i,1}, ..., x_{i,k})||^2) \tag{4.12}$$

The cost function f_i expresses the elements of a residual block as functions of the unknown elements of the **parameter blocks** $(x_{i,1}, ..., x_{i,k})$. The three components of a translation vector as well as a single scalar build up a parameter block (Agarwal *et al.*, 2013). The cost function corresponds to the residual equation system (2.20). The residuals $\boldsymbol{v}$ are a function of the parameters $\boldsymbol{x}$ and the observations $\boldsymbol{l}$. The adjustment in OrientAl is organised in residual and parameter blocks.

The cost function is defined as in Listing 4.1. An important characteristic is the inheritance of attributes and methods from the base class. Inheritance is obtained by listing the base class *AutoDiff* in parenthesis after the class name *CostFunction*. Due to a non-linear problem, first order derivations are needed. Therefore, the automatic differentiation method of Ceres is applied. The attributes and methods of *AutoDiff* are inherited and the residual equations have to be put in this context.

```python
class CostFunction( oriental.adjust.cost.AutoDiff ):
    def __init__( self, obs ):
        # Initialise an instance of CostFunction
        self.obs = obs
        # Initialise the base class
        numResiduals = 1
        parameterSizes = (3,3)
        super( CostROF, self ).__init__( numResiduals, parameterSizes )

    def Evaluate( self, parameters, residuals ):
        dt = type(parameters[0][0])
        par1 = parameters[0]
        par2 = parameters[1]
        predict = f(par1,par2)
        residuals[0] = predict - dt(self.obs)
        return True
```

Listing 4.1: Cost function scheme

The method *__init__* is the initialisation method, which is called, when a new instance of the class is created. *Self* is always the first argument of a python class method. Further arguments correspond to the observations, which are accessed with the dot notation *(self.obs)*.
With *super* the basis class *AutoDiff* is accessed. The arguments are the number of residuals and the sizes of the parameters.
In the method *Evaluate* the functional model is defined. The essential measures are the parameters and the residuals. Based on the observations and the approximate values of the unknowns, either the residuals or the Jacobi matrix is computed.
Therefore, the data type *Jet* is of importance, denoted as *dt* in Listing 4.1. It supports the automatic differentiation. It allows to address the derivative elements of the Jacobi matrix. It is composed of one scalar and four accessible derivative elements. If more than four derivatives are needed, *Evaluate* is called multiple times.

The definition of the least squares problem in OrientAl does not consider a priori weights. However, the weight matrix $\boldsymbol{P}$ has to be included in the cost function f_i, otherwise a variance component estimation can not be established. The idea is to substitute $\sqrt{\boldsymbol{P}}\boldsymbol{A}$ and $\sqrt{\boldsymbol{P}}\boldsymbol{l}$ for the design matrix $\boldsymbol{A}$ and the observation vector $\boldsymbol{l}$, to achieve Formula (2.28) in further consequence.
Thus, the complete functional model needs to be extended by the square root of the weight matrix $\sqrt{\boldsymbol{P}}$. The residual equation changes according to (4.13). Minimising the 2-norm of the residuum vector leads to the estimation of the parameters $\hat{\boldsymbol{x}}$ (4.14). Due to the fact that $\boldsymbol{P}$ is a symmetric matrix it is equal to its transpose.

$$\boldsymbol{P}^{\frac{1}{2}}\boldsymbol{v} = \boldsymbol{P}^{\frac{1}{2}}\boldsymbol{A}\boldsymbol{x} - \boldsymbol{P}^{\frac{1}{2}}\boldsymbol{l} \tag{4.13}$$

$$\boldsymbol{v}^T\boldsymbol{P}^{\frac{1}{2}}\boldsymbol{P}^{\frac{1}{2}}\boldsymbol{v} = min \quad \rightarrow \quad \hat{\boldsymbol{x}} = \left(\boldsymbol{A}^T\boldsymbol{P}^{\frac{1}{2}}\boldsymbol{P}^{\frac{1}{2}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{P}^{\frac{1}{2}}\boldsymbol{P}^{\frac{1}{2}}\boldsymbol{l} \tag{4.14}$$

To compute the square root of a matrix, the spectral decomposition is used. The eigenvalues are the diagonal elements of $\boldsymbol{D}$ and the orthogonal matrix $\boldsymbol{O}$ is the eigenbasis of $\boldsymbol{P}$. The decomposition is presented in Formula (4.15). The square root of the diagonal matrix $\boldsymbol{D}$ yields the square root of the original matrix $\boldsymbol{P}$.

$$\boldsymbol{P}^{\frac{1}{2}} = \boldsymbol{O}\boldsymbol{D}^{\frac{1}{2}}\boldsymbol{O}^T \tag{4.15}$$

In Listing 4.2 the essential part of the implementation of the weight matrix $\boldsymbol{P}^{\frac{1}{2}}$ is shown. Therefore, the covariance matrix of the relative orientation (4.11) is considered. Computing the inverse and the square root of it leads to $\sqrt{\boldsymbol{P}_{ROF}}$. In Listing 4.2 it is denoted as *sqrtPBlock*. It is accessed by the dot notation, thus it is an object property, but must not cause a change in the number of residuals.

The already defined residuals are replaced by the *residuals_weighted*. In the course of the two *for*-loops, the variance and the covariances are assigned to the residual and the appropriate derivatives. The *sqrtPBlock* is a 6 × 6 matrix, therefore the index *iRow* and *iCol* run from 0 to 5.

As mentioned before, by means of the data type *Jet* the derivative elements of the Jacobi matrix are accessible. The variable *residuals_weighted* is defined as *Jet* in the second row of Listing 4.2. Without this functionality an inclusion of the weight matrix $\boldsymbol{P}$ in OrientAl would not be possible.

```
# Assign the right element of sqrt(P) to the appropriate derivation
residuals_weighted = np.zeros( residuals.shape, dtype=dt )
for iRow in xrange(6):
    for iCol in xrange(6):
        residuals_weighted[iRow] += residuals[iCol] * self.sqrtPBlock[iRow,iCol]

residuals[:] = residuals_weighted
return True
```

Listing 4.2: Considering the weight matrix $\boldsymbol{P}$

The *CostFunction* is applied as presented in Listing 4.3.

First, the *Problem* needs to be defined. To create a least squares problem, the method *AddResidualBlock* is used. Thus, a residual block is added to the problem. A residual block consists of a cost function, an optional loss function and the parameters. The observations and the square root of the weight matrix are the constructor arguments of the *CostFunction* (Listing 4.3). The optional loss function and the approximate values of the parameters (*par1*, *par2*) are further arguments to *AddResidualBlock*. As mentioned earlier, an ordinary least-squares adjustment is computed, hence the *Trivial* loss is used, as defined in the second row of Listing 4.3.

Any parameter block can be set constant using *SetParameterBlockConstant*. In Listing 4.3 the parameter *par2* is set constant. In this case the corresponding entries in the Jacobi matrix become zero. Those zero entries cause singularities and therefore, this parameter block will not be evaluated.

The Problem is solved based on the following arguments: the defined problem, the solver options and a summary. The class *Options* controls the behaviour of the solver and the default settings can be changed e.g. the linear solver, number of iterations.

```
problem = adjust.Problem()
loss = adjust.loss.Trivial()

problem.AddResidualBlock(CostFunction.CostFunction ( obs1, sqrtPBlock ),
                                            loss,
                                            par1,
                                            par2)
problem.SetParameterBlockConstant( par2 )

options = adjust.Solver.Options()
summary = adjust.Solver.Summary()
adjust.Solve( options, problem, summary )

jacobi = problem.Evaluate( evalOpts, residuals=False, jacobians=True )
```

Listing 4.3: The utilisation scheme of the CostFunction

An instance of the class *Problem* is able to manage a number of residual blocks. Those residual blocks depend on the cost function, the loss function and the parameters. Each residual block produces one or more residuals (Karel *et al.*, 2013).
In this thesis one type of residual blocks is based on the cost function **PhotoTorlegard**. This cost function computes the collinearity equations (2.12). It is a standard cost function of OrientAl.
The second type of residual blocks uses the cost function **CostROF**, which has been implemented on basis of the functional model in Chapter 4.1.
The third type of residual blocks bases on the cost function **CostRange** and computes Equation (4.10).
Furthermore, the residual block **CostCtrlPts** is determined for estimating the control points, i.e. a weighted datum shall be established.

As the terms residual- and parameter blocks imply, the adjustment is built on blocks. One residual block is computed for each image point. Therefore, the cost functions are added in loops, to proceed it for all frames. The last command in Listing 4.3 affects the evaluation of the residuals $\boldsymbol{v}$ or the design matrix $\boldsymbol{A}$, depending on which argument is set True. Through the *EvaluateOptions (evalOpts)* the user is able to define, which parameter blocks shall be estimated. If a parameter block is set constant, then it will not be considered in this list. The *Evaluate* method executes the computation of the residuals or the design matrix.

The cofactor matrix of the parameters $\boldsymbol{Q}_{xx}$ can be determined through the use of the OrientAl method *Covariance*. First, the covariances of each block are computed and then the matrix is built. Therefore, the blocks have to be arranged in the same order as the Jacobi matrix. Thus, knowledge of the structure of the design and the covariance matrix is a prerequisite. The cofactor matrix of the parameters $\boldsymbol{Q}_{xx}$ shows the precision of the parameters and is needed to estimate the cofactor matrix of the residuals $\boldsymbol{Q}_{vv}$. This is of importance for the following variance component estimation.

For further considerations, the residuals $\boldsymbol{v}$ and the design matrix $\boldsymbol{A}$ have to be corrected. In the step of including the weight matrix into the problem, the complete functional model has been extended by $\sqrt{\boldsymbol{P}}$, as presented in (4.13). Therefore, the estimated measures are right-multiplied by the square root of the cofactor matrix $\sqrt{\boldsymbol{Q}_{ll}}$ because of the relation $\boldsymbol{P}^{-\frac{1}{2}} = \boldsymbol{Q}_{ll}^{\frac{1}{2}}$.

$$\sqrt{\boldsymbol{Q}_{ll}}\boldsymbol{v}, \quad \sqrt{\boldsymbol{Q}_{ll}}\boldsymbol{A} \tag{4.16}$$

This step implicated some errors. First, only the design matrix has been corrected and the still extended residuals have been used for further computations.

4.5 Variance Component Estimation

The Variance Component Estimation (VCE) is the method chosen, to integrate observations of different precisions. Therefore, observation groups are built and variance components are estimated, as discussed in Chapter 2.6. Four different types of observations are considered. First of all the image observations yield the camera pose based on the collinearity equations. The second group arises because of the integration of the relative orientation results. Group three is built on the range measurements. In case the control points are treated as observations and unknowns (weighted datum), they are an own observation group.

An essential step is the computation of the variance factor (2.32) for each group, called the variance component $\sigma^2_{0g,v}$. The variables needed are the residuals, the weight matrix and the redundancy number, as shown in Equation (4.17). In each iteration step the variance component is computed and multiplied to the initial value $\alpha^2_{0g,0}$ (4.18).

$$\hat{\sigma}^2_{0g,v} = \frac{\boldsymbol{v}^T_{g,v}\boldsymbol{P}_{g,v}\boldsymbol{v}_{g,v}}{r_{g,v}} \tag{4.17}$$

$$\alpha^2_{0g,v} = \alpha^2_{0g,0} \cdot \prod_{v=1}^{n} \hat{\sigma}^2_{0g,v} \tag{4.18}$$

The redundancy number corresponds to the trace of the cofactor matrix of the residuals $\boldsymbol{Q}_{vv}$ (2.30). Therefore, the cofactor matrices of the unknowns $\boldsymbol{Q}_{xx}$ (2.28) and the one of the adjusted observations $\boldsymbol{Q}_{\hat{l}\hat{l}}$ (2.29) are required. All these mentioned quantities are split up into groups and therefore assigned with the index g in Formulae (4.17) and (4.18).

The implementation scheme of the VCE is presented in Listing 4.4. Thereby, two groups are considered: BA stands for the bundle adjustment and ROF are the initials of range and optical flow and signifies the relative orientation. The VCE is an iterative method and therefore computed in a loop. Outside of the loop, the initial values $\alpha^2_{0g,v}$ and the variance factors $\hat{\sigma}^2_{0g,v}$ are defined. The initial values have been chosen to be one.

As defined in Formula (2.34), the product of all estimated variance factors is multiplied with the cofactor matrix. In case of the bundle adjustment an a priori standard deviation σ_{img} of 0.5 pixel has been assumed. Thus, the standard deviation is multiplied by the square root of the initial value to include it in the stochastic system.

The covariance matrix of Formula (4.11) constitutes the stochastic model of the relative orientation. The *ROF_sll_orig* is multiplied with the initial value *alpha2_ROF_n*. Thereof, the square root of the inverse matrix is computed and is integrated in the cost function.

After computation and multiplication of the new variance components, the procedure starts again with the improved initial value $\alpha^2_{0g,v}$ until the termination criterion is reached.

```
alpha2_BA_n   = 1
alpha2_ROF_n = 1
sig2_apost_ROF = 0
sig2_apost_BA = 0
while True:
    sig2_ROF = sig2_apost_ROF
    sig2_BA   = sig2_apost_BA
    # Problem

    stdDevs = np.array([0.5,0.5]) * np.sqrt(alpha2_BA_n)
    sqrtPBlock = np.diag(1./stdDevs)
    # AddResidualBlock( PhotoTorlegard )

    ROF_sll = ROF_sll_orig * alpha2_ROF_n
    for i in xrange(0,noROF,6):
        sqrtP_ROF[i:i+6] = sp.linalg.sqrtm(np.linalg.inv(ROF_sll[i:i+6]))
    # AddResidualBlock( CostROF )

    # Solver
    sig2_apost_BA   = res_apost_BA.T.dot(p_BA).dot(res_apost_BA)/sum(redAn_BA)
    sig2_apost_ROF = res_apost_ROF.T.dot(p_ROF).dot(res_apost_ROF)/sum(redAn_ROF)
    alpha2_BA_n    *= sig2_apost_BA
    alpha2_ROF_n   *= sig2_apost_ROF

# Termination criterion
    if abs(sig2_apost_ROF-sig2_ROF) and abs(sig2_apost_BA-sig2_BA) <= 10**-10:
        # Data output
```

Listing 4.4: Implementation scheme of the VCE

Termination criterion. The method converges if $\hat{\sigma}^2_{0g,v}$ of all groups g becomes one. The termination criterion has been defined as a difference measure. The threshold is defined based on the working precision, which amounts to 10^{-16}. The variance factor shall be estimated up to two thirds of the working precision. The absolute difference of two subsequent variance factors has to be equal or smaller than the threshold of 10^{-10}, for the algorithm to terminate, as presented in Equation (4.19).

$$\sum_{g=1}^{n} |\hat{\sigma}^2_{0g,v} - \hat{\sigma}^2_{0g,v-1}| \leq 10^{-10} \tag{4.19}$$

4.6 Numerical Aspects

As discussed in Chapter 2.7, bundle adjustment typically is a large sparse linear estimation problem. The computation of the problem can be split up in an overall optimisation method and the key-computational

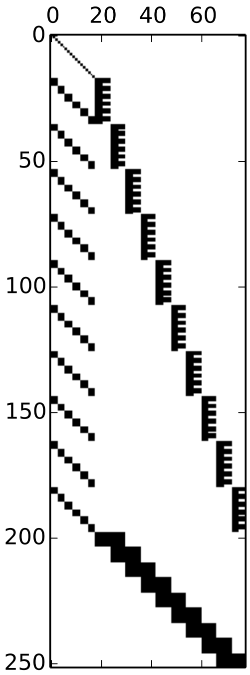

Figure 4.4: Structure of the Jacobi matrix $\boldsymbol{A}$ (10 frames: 252 observations, 78 unknowns)

cost, the solver. OrientAl, which bases on Ceres, uses the second order algorithm Levenberg-Marquardt for optimisation.

Linear Solver. OrientAl provides different linear solver to compute the cofactor matrix of the unknowns $\boldsymbol{Q}_{xx}$. Which one shall be chosen depends on the structure of the problem treated. As discussed in Chapter 2.7, structural parameters like sparsity and symmetry shall be exploited. This will lead to an increasing efficiency of the solver and hence a faster convergence.

In Figure 4.4 the sparsity of the Jacobi matrix $\boldsymbol{A}$ is illustrated. The design matrix is only built out of ten frames, in order to see the structure more clearly. The matrix elements coloured in black correspond to the non-zero values. The percentage of non-zero elements amounts to 1 %. Thus, it is a highly sparse matrix.
Another important aspect is the ordering of the different parameters. Mainly the Cholesky factorisation benefits from the right order of the parameters. First the object points shall be considered, followed by the exterior and the interior orientation.

As Figure 4.4 shows, the first 18 columns correspond to the six control points $\boldsymbol{p}_i$. The remaining 60 columns are the estimated exterior orientation parameters $\boldsymbol{R}_{\alpha,j}, \boldsymbol{p}_{0,j}$ (10 times 6).
The first block describes the control points as parameters and observations (weighted datum). The following 180 rows present the image (10 times 6×2) and the range (10 times 6) residual equations, which are engaged with one another. Each image observation bases on the control points $\boldsymbol{p}_i$, the perspective

Time [s]	Sparse Schur	Sparse Cholesky	Iterative Schur
Preprocessor	0.015	0.013	0.013
Residual	7.177	6.690	4.056
Jacobian	25.796	23.033	16.438
Linear solver	0.472	0.343	0.369
Minimiser	**33.504**	**30.106**	**20.898**
Total	33.521	30.122	20.912

Table 4.2: Comparison of the run-time performance of the three mentioned linear solvers.

centre $\boldsymbol{p}_{0,j}$ and the angles of the rotation matrix $\boldsymbol{R}_{\alpha,j}$. The interior orientation is set constant.
The difference between perspective centre and the object point builds the range measurement. Thus, only the perspective centres $\boldsymbol{p}_{0,j}$ and the control points $\boldsymbol{p}_i$ are considered in this case, as visible in Figure 4.4. The last 54 rows belong to the relative orientation (9 times 6). Two subsequent exterior orientations correspond to the relative orientation result. Therefore, only nine blocks of relative orientation residual equations are present in Figure 4.4.

OrientAl automatically chose the *Sparse Schur* solver. It corresponds to the Schur decomposition, which solves special structured problems more efficient than the Cholesky factorisation.
For a general sparse problem the *Sparse Normal Cholesky* is recommended. In case of large problems the *Iterative Schur* with the preconditioner *Jacobi Schur* offers an excellent balance of speed and accuracy (Agarwal *et al.*, 2013). This has been proven based on the run-time performance, which is presented in the full report of the adjustment. In Table 4.2 a comparison of the run-time performance of those three linear solvers is shown. The *Iterative Schur* computes the problem with 3,276 observations and 834 parameters in 21 s, that is about one third faster than the sparse Schur and the Cholesky decomposition. Comparing the Schur and the Cholesky factorisation, the faster computation was accomplished by the *Sparse Cholesky*.
To choose the right solver is a decisive step. For equation systems with more than 10,000 observations it becomes even more important.

Preconditioning and Scaling. The values of the covariance matrices of the relative orientation $\boldsymbol{\Sigma}_{ll,ROF}$ lead to an ill-conditioned numerical problem. Therefore, the *ROF_sll_orig* in Listing 4.4 has been multiplied with 10^5, to result in a better conditioned one.

5. Analysis

The analysis presents the results of the integration of additional observations into the bundle adjustment. To achieve reasonable results, a variance component for each observation type is estimated and presented. The influence of those observations on the trajectory is analysed and compared with the result of Chapter 3.5. Furthermore, the limitations of the variance component estimation are treated.

As demonstrated in Equation (5.1), an error term ϵ_g is included in the adjustment model for each observation group. The stochastic model is extended similarly, as shown in Equation (5.2). The product of all variance factors of a group $\hat{\sigma}^2_{0g}$ results in $\alpha^2_{0g,v}$, which presents the final variance component of the group g.

$$\boldsymbol{l} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\epsilon}_{Img} + \boldsymbol{\epsilon}_{ROF} + \boldsymbol{\epsilon}_{R} + \boldsymbol{\epsilon}_{Ctrl} \tag{5.1}$$

$$\boldsymbol{\Sigma}_{ll,v} = \alpha^2_{0Img,v}\boldsymbol{Q}_{Img} + \alpha^2_{0ROF,v}\boldsymbol{Q}_{ROF} + \alpha^2_{0R,v}\boldsymbol{Q}_{R} + \alpha^2_{0Ctrl,v}\boldsymbol{Q}_{Ctrl} \tag{5.2}$$

First, the image observations of the bundle adjustment *Img* and the relative orientation results *ROF* are the only two observation groups considered. Step by step the range measurements of the control points *R* and the control positions (weighted datum) *Ctrl* are introduced as observation group.
The last section deals with the change of the trajectory due to the reduction of the number of control points.

The variance component estimation (VCE) completely relies on a correct functional model and on data, free of gross errors. Therefore, the residuals of the adjustment are tested on their reliability.
The redundancy of an observation r_i describes to which extent the observation contributes to the parameter estimation and outliers might be detected in the residuals. If the diagonal elements of $\boldsymbol{Q}_{vv}\boldsymbol{P}_{ll}$, which correspond to the redundancy of observations r_i, are large, then outliers can be found reliably.

$$w_i = \frac{v_i}{\hat{\sigma}_{v_i}} = \frac{v_i}{\alpha_{0g,v}\sqrt{(q_{vv})_{ii}}} \tag{5.3}$$

The standardised residuals w_i are calculated as presented in Equation (5.3). The expected value is 0 and the standard deviation is ±1 - they represent a simple test procedure. If the standardised residuals are smaller than a specific quantile of the underlying distribution - Gaussian distribution, no outlier is included in the data (Kraus, 1996, p. 225).

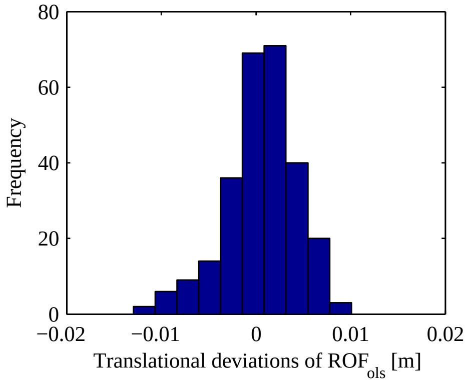

Figure 5.1: Data set 1 - Histogram of the translations of the RPE corresponding to Chapter 3.4

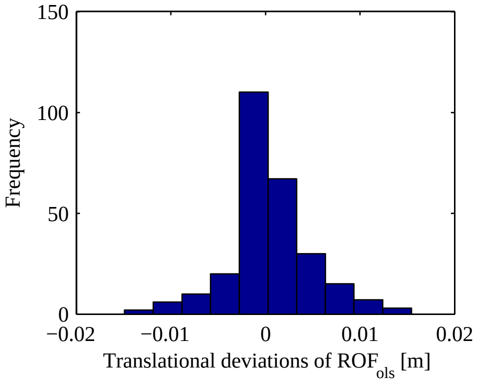

Figure 5.2: Data set 2 - Histogram of the translations of the RPE corresponding to Chapter 3.4

5.1 Integration of the Relative Orientation

In this section, the relative orientation results are integrated in the bundle adjustment. Two observation groups are built: the image observations of the bundle adjustment *Img* and the relative orientation results *ROF*. The adjustment bases upon a constrained datum definition.

The least squares adjustment delivers the best estimations for Gaussian or nearly Gaussian distributed observations, due to the reason that the least squares adjustment model only considers random errors, which are Gaussian distributed. It is assumed, that the distribution of the residuals follows the distribution of the observations. If the observations only include random errors, the distribution of the residuals should be normal distributed as well. A deviation from a normal distribution is mainly caused by gross errors. In this case, robust adjustment methods lead to better results (Niemeier, 2008, p. 208).

In order to use the relative orientation results in the least squares adjustment, the distribution of the relative orientation results need to be verified. The Figures 5.1 and 5.2 base on the translational part of the relative pose error (RPE) of Chapter 3.4. The histograms show the deviation of the relative orientation from the adjusted line of the bundle adjustment results. The data of the first acquisition shows a distinct left tail. In comparison, data set two in Figure 5.2 presents a symmetric distribution, but with noticeable tails as well. For both applies an average error of 7 mm. Thus, the translations vary in between $\pm 2\sigma$. However, even 98 % (data set one) or 94 % (data set two) of the relative orientations are within the standard deviation of 7 mm.

The hybrid adjustment of image and relative orientation observations produces the following residual distributions before the variance component estimation.
In the first row of Figure 5.3 the distribution of the standardised residuals of data set one are plotted. The red line illustrates the Standard-Gaussian distribution, which affords a better comparison. The image residuals are normal distributed. The residuals of the relative orientation results follow the distribution of Figure 5.1, which is the RPE of the relative orientation result of data set one. In comparison to the Gaussian distribution it shows a divergent behaviour. The scatter plots in the second row present the

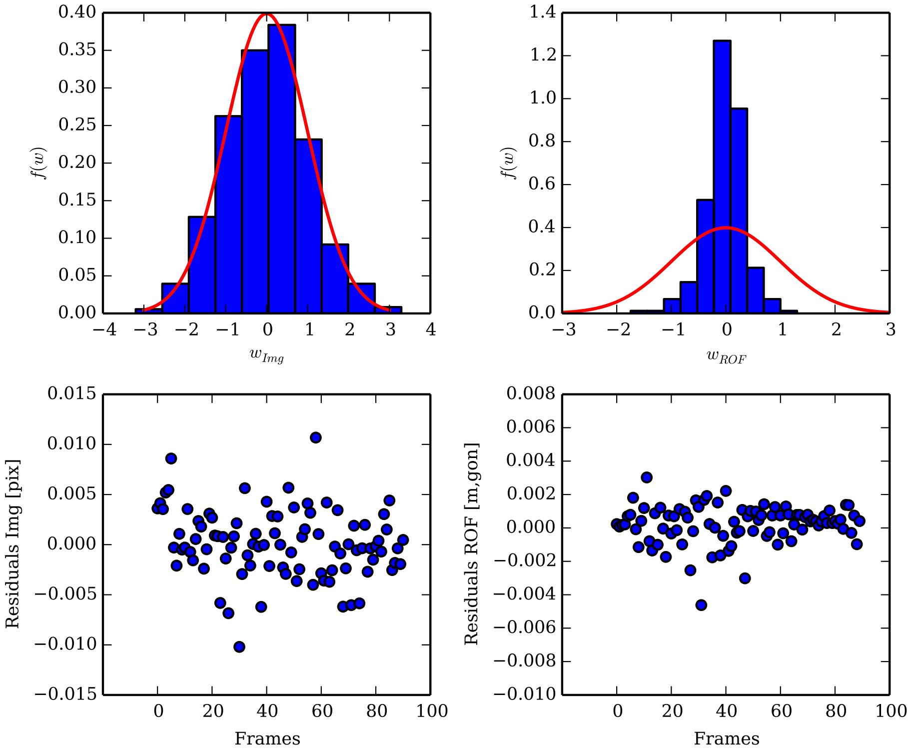

Figure 5.3: Data set 1 before VCE (Img+ROF) - First row: Normalised histograms of the standardised residuals; Second row: Scatter plots of the residuals per frame

residuals per image. For all observations, belonging to one frame the mean value was computed in order to recognise the influence of outliers. The plots show the chronology of the deviations.

The relative orientation comprises the translational and the rotational part. They are different observations with inherently different precisions. To be proper, they should build two groups. However, due to high correlations between those groups, covariance components must be estimated, which demands further effort and would have gone beyond the scope of this work. Thus, the relative rotations and translations have not been considered separately.

In Figure 5.4 the distribution of the standardised residuals of the second data acquisition is presented. Similar to data set one the image residuals before the VCE correspond to a Gaussian distribution. The normalised histogram of the w_{ROF} matches to the one in Figure 5.2, characterised by a distinct peak and long tails. However, it is not conform with a Gaussian distribution.

The scatter plots present less variation in the first and the final frames, which comes along with no or less motion accomplished.

The Figures 5.5 and 5.6 show the distributions of the residuals after the VCE. They illustrate the influence of the VCE on not normal distributed observations included in a least squares adjustment. The two standardised spectra (image and relative orientation) are aligned i.e. both peaks of the distribution have the same percentage. Interestingly, the range of the image residuals increase by a factor of two and the relative orientation *ROF* residuals decline by a factor of two, as presented in the scatter plots.

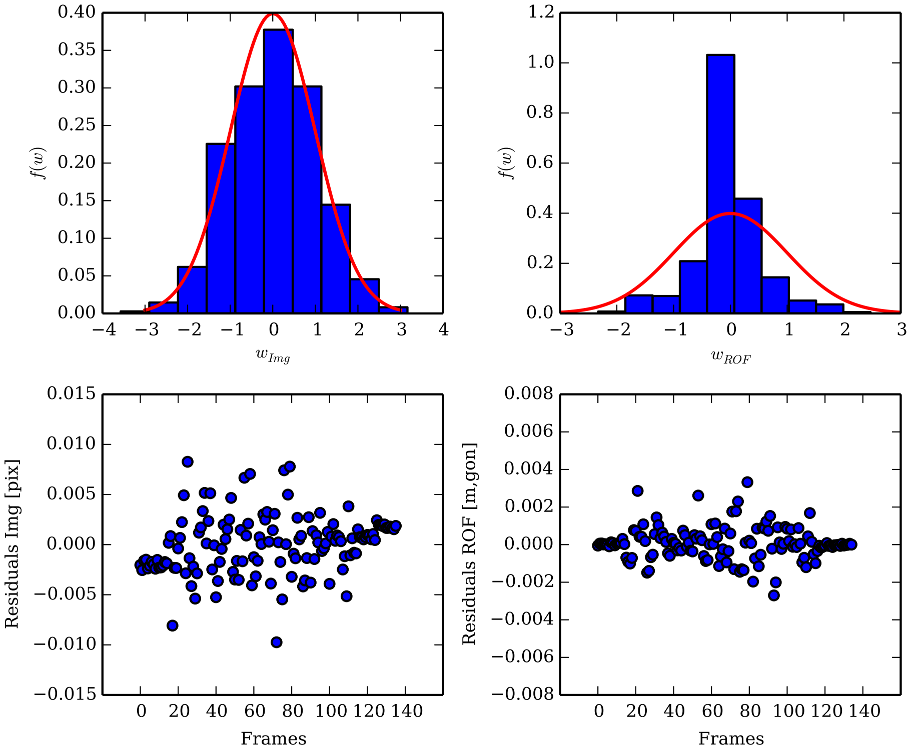

Figure 5.4: Data set 2 before VCE (Img+ROF) - First row: Normalised histograms of the standardised residuals; Second row: Scatter plots of the residuals per frame

As mentioned before, the least squares adjustment requires normal or nearly normal distributed observations. That does not apply for the relative orientation observations. Consequently, the least squares adjustment does not result in the best unbiased estimate of the parameters.

Prerequisites for the variance component estimation are a complete and correct functional model and no gross errors included in the data. As discussed in Chapter 3.2, the data recorded by the SR-3000 still includes systematic range errors and possibly gross errors as well. The histograms of Figure 5.1 and 5.2 were also computed for the robust relative orientation results. For that purpose, the *huber* weighting function was applied, as mentioned in Chapter 3.4. The RPE of the robust adjustment results show a similar distribution as the ordinary least squares adjustment. Thus, the relative orientation results are mainly influenced by systematic range errors. Those not considered systematic errors lead to deformations of all results in the bundle adjustment. Therefore, the precondition of a complete and a correct functional model for the estimation of the relative orientations is not fulfilled.

In addition, the implemented range optical flow shows limitations as well. It is a relative orientation method, presenting a high drift. The optical flow is a 2D method. The range flow assumes a local time-varying surface $\dot{z}$ in the 3D space, but no 3D models of the objects recorded are considered. Furthermore, the perspective of the camera does not enter the model either.

All those mentioned restrictions - belonging to the data acquired or to the relative orientation based upon the range optical flow - influence the variance components estimated.

The adjustment of the relative orientation, discussed in Chapter 3.4, induces a precision of the esti-

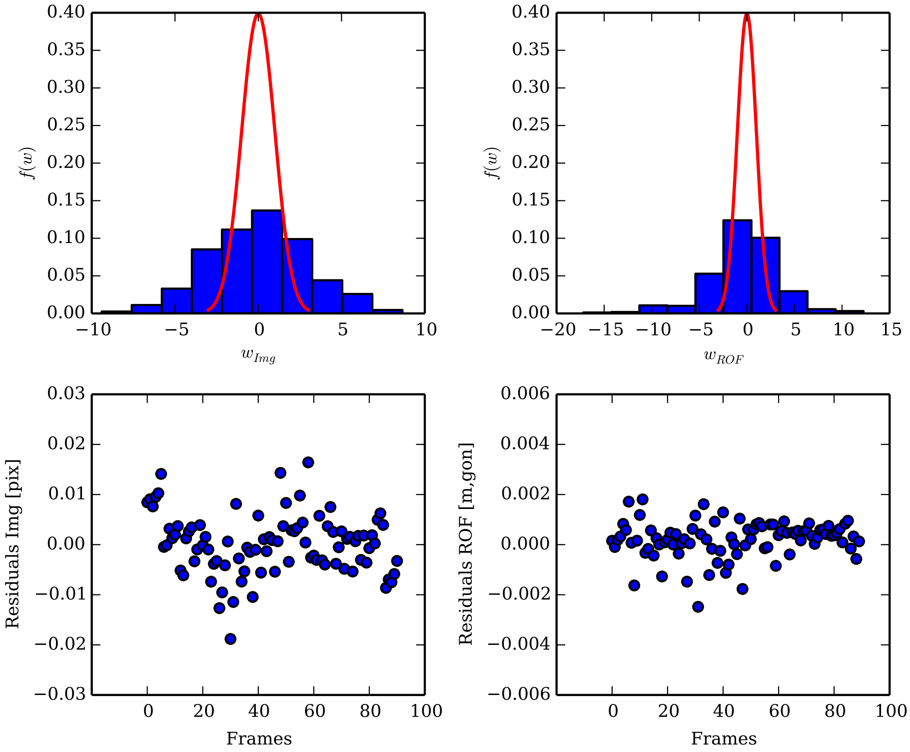

Figure 5.5: Data set 1 after VCE (Img+ROF) - First row: Normalised histograms of the standardised residuals; Second row: Scatter plots of the residuals per frame

mated parameters of approximately 0.4 mm (Table 3.7), while the RPE leads to a standard deviation of 7 mm. The precision level of the relative orientation in the hybrid adjustment is 3 mm (Table 5.1 and 5.2). The values are summarised in Formula (5.4).

$$\hat{\sigma}_{ROF,trans} \approx 0.4mm, \qquad \sigma_{RPE,trans} \approx 7mm, \qquad \hat{\sigma}_{Hyb,trans} \approx 3mm \tag{5.4}$$

The reason for the optimistic estimation of the relative orientation precision is the artificially high redundancy of the adjustment system. The complete image content is used for the estimation of the parameters. However, only areas with texture, as shown in Figure 3.9 and 3.10, contribute to the parameter estimation. Additionally, correlations in amplitude and range between neighbouring pixel that have not been considered, also cause a too optimistic result.

One aim of this work is to estimate a smooth unbiased camera trajectory. The plotted trajectories in Figure 5.7 present a smooth, unbiased trajectory, which fits well to the bundle adjustment result of Chapter 3.3. The second data set (second row) corresponds even better to the accomplished horizontal moving direction.

The image observations are predominant in this composition of observations. As shown in Table 5.1 and 5.2, the redundancy of the image observations amounts to nearly 70 or 80 % of the total redundancy. The total redundancy is approximately two times the number of unknowns u, which allows a reliable variance component estimation. However, the observation groups should be more balanced. The variance compo-

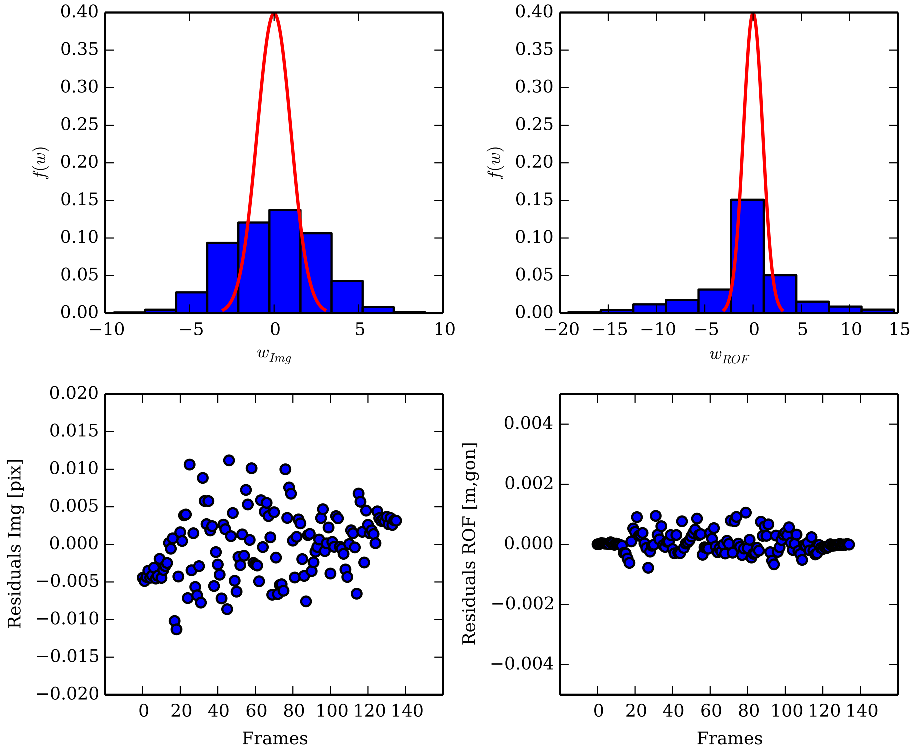

Figure 5.6: Data set 2 after VCE (Img+ROF) - First row: Normalised histograms of the standardised residuals; Second row: scatter plots of the residuals per frame

nent estimation terminates after 18 iterations for both data sets.

The variance component of the image observations amounts to 0.12 and to 0.14 pixel in case of data set two. The assumed variance factor was 0.25 pixel.

The variance component of the relative orientation amounts to 9 and 12 mm. For numerical reasons the covariance matrix of the relative orientations is preconditioned - multiplied by a factor of 10^5. Consequently, the true variance component is 940 or 1180 in case of data set two. The precisions of the relative orientation are too optimistic by a factor of approximately 30.

$$\alpha^2_{0Img,v} \approx 0.13, \qquad \alpha^2_{0ROF,v} \approx 0.01 * 10^5 = 10^3 \tag{5.5}$$

The essential measures of the adjustment are presented in Table 5.1 and 5.2. The estimated unknowns are similar to those in Chapter 3.3. The precision of the unknown parameters is nearly the same for both data sets. In comparison to the results, reached in Chapter 3.3 the standard deviations of the unknowns in data set two are about a factor of 3 better. Whereas the precision of the exterior orientation of data set one improved by approximately 25 %.

In Table 5.3 the relative pose error based on the integrated result *Img+ROF* is presented. For a better comparison the results of the relative orientation ROF_{ols} are also listed. The root mean square error of the translational part improved by 1 mm for the second data set recorded. The precision of the rotational part highly increased for data set one. Its RMSE is not half as much of the relative orientation result.

The absolute trajectory error (ATE) in Table 5.4 enhanced from 23 to 1.5 cm for a trajectory length of

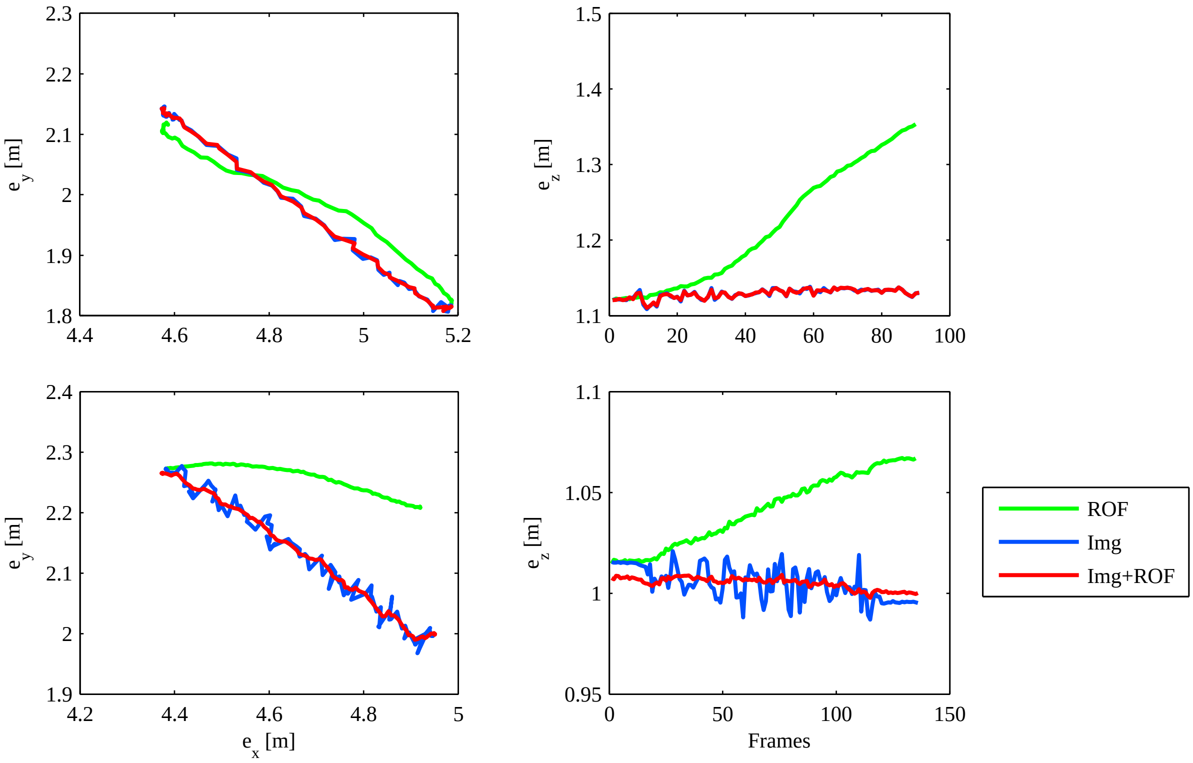

Figure 5.7: The trajectory estimation, based upon the integration of the relative orientation into the bundle adjustment (red line) is presented in comparison to the trajectories, derived separately.

Parameter	**Quantity**
# Images	91
Control points	134,136,137,323,332,344
# Observations n	1632
# Unknowns u	546
# Redundancy r	1086
$r_{Img,v}$	754.6 (69.5 %)
$r_{ROF,v}$	331.4 (30.5 %)
Trajectory length	0.682 m
Estimated unknowns:	
$\boldsymbol{p}_{0,1}$	$[5.184, 1.804, 1.120]^T$ m
$\boldsymbol{p}_{0,91}$	$[4.588, 2.135, 1.133]^T$ m
$avg(\alpha, \zeta, \kappa)$	5.5, 111.7, 107.5 gon
Estimated accuracies:	
$\alpha^2_{0Img,v}$	0.117 pixel
$\hat{\sigma}_{0Img,v}$	0.03 pixel
$\alpha^2_{0ROF,v}$	0.0094 m
$\hat{\sigma}_{0ROF,trans}$	1.4, 3.7, 4.1 mm
$\hat{\sigma}_{0ROF,rot}$	1.2, 0.4, 1.0 mgon
$median(\hat{\sigma}_{p_0})$	3.4, 1.9, 3.2 mm
$median(\hat{\sigma}_\alpha, \hat{\sigma}_\zeta, \hat{\sigma}_\kappa)$	0.0350, 0.0553, 0.0604 gon

Table 5.1: Specifications of the first data set (It 2 ms)

Parameter	**Quantity**
# Images	136
Control points	127,134,136,137,323,324
# Observations n	2442
# Unknowns u	816
# Total Redundancy r	1626
$r_{Img,v}$	1284.6 (79.0 %)
$r_{ROF,v}$	341.4 (21.0 %)
Trajectory length	0.6358 m
Estimated unknowns:	
$\boldsymbol{p}_{0,1}$	$[4.374, 2.266, 1.008]^T$ m
$\boldsymbol{p}_{0,136}$	$[4.951, 1.999, 1.000]^T$ m
$avg(\alpha, \zeta, \kappa)$	11.1, 111.9, 0.0 gon
Estimated accuracies:	
$\alpha^2_{0Img,v}$	0.144 pixel
$\hat{\sigma}_{0Img,v}$	0.03 pixel
$\alpha^2_{0ROF,v}$	0.0118 m
$\hat{\sigma}_{0ROF,trans}$	1.1, 3.4, 2.9 mm
$\hat{\sigma}_{0ROF,rot}$	1.0, 0.3, 1.0 mgon
$median(\hat{\sigma}_{p_0})$	3.5, 4.0, 1.7 mm
$median(\hat{\sigma}_\alpha, \hat{\sigma}_\zeta, \hat{\sigma}_\kappa)$	0.0668, 0.0279, 0.0549 gon

Table 5.2: Specifications of the second data set (It 30 ms)

	trans [m]				α [gon]			
	Mean	**Median**	**Max**	**RMSE**	**Mean**	**Median**	**Max**	**RMSE**
Data set 1:								
ROF_{ols}	0.0058	0.0051	0.0145	0.0066	0.0034	0.0031	0.0100	0.0038
Img+ROF	0.0053	0.0047	0.0162	0.0063	0.0013	0.0011	0.0039	0.0015
Data set 2:								
ROF_{ols}	0.0055	0.0046	0.0223	0.0072	0.0033	0.0028	0.0133	0.0044
Img+ROF	0.0048	0.0038	0.0210	0.0062	0.0031	0.0028	0.0102	0.0039

Table 5.3: Relative pose error for the different adjustment types

approximately 60 cm. The RMSE of the ATE is approximately the half of the trajectory error (Max).

As shown in the first row of Figure 5.7, the trajectory of the bundle adjustment and the trajectory of the integrated solution fit very well. The outcome of the bundle adjustment of data set one already implies a high precision so as not to much improvement is visible.

The second data set (second row) represents the expected smooth unbiased trajectory despite all deficiencies.

5.2 Integration of Range Measurements

Now the model of Chapter 5.1 is extended by the range observations of the control points R. The range measurements, performed by a ToF camera are subject to many errors.

As discussed in Chapter 3.2, gross errors were eliminated, as far as detected, and a camera calibration was considered. Therein systematic range errors with respect to the ranges, the amplitudes, the integration time and the position in the image are modelled, but some systematic errors are not considered at all. The

	Max [m]	**RMSE** [m]	α_1 [gon]	α_n [gon]
Data set 1:				
ROF$_{ols}$	0.2234	0.1212	0.0004	0.0710
Img+ROF	0.0144	0.0062	0.0004	0.0010
Data set 2:				
ROF$_{ols}$	0.2363	0.1615	0.0002	0.0719
Img+ROF	0.0160	0.0099	0.0036	0.0017

Table 5.4: Absolute Trajectory Error of data set one and two

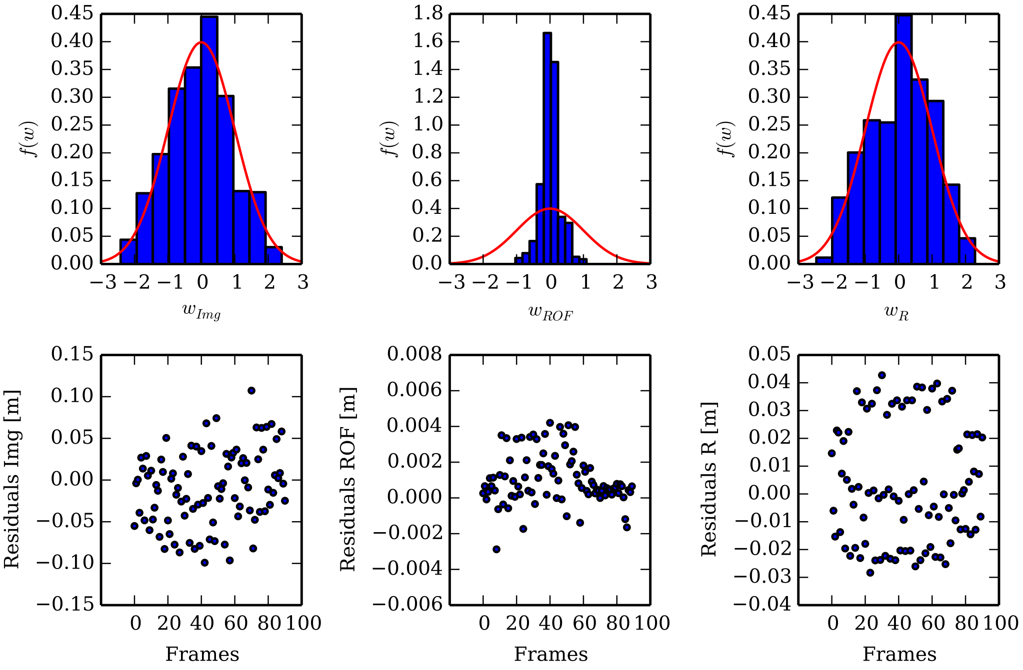

Figure 5.8: Data set 1 before VCE (Img+ROF+R) - First row: Normalised histograms of the standardised residuals; Second row: Scatter plots of the residuals per frame

integration of the range measurements shows the influence of those errors, if the ranges are affected.

The residual spectra of the three observation groups are presented in Figure 5.8 and 5.9. The distributions of the image and relative orientation residuals of data set one are similar to the ones in Figure 5.3, but the peak of the relative orientation spectrum is even more distinct. The distribution of the standardised residuals of the ranges (rightmost) fits quite well to the Gaussian distribution. The scatter plot of the range residuals shows three areas with similar residuals (-0.02, 0 and 0.035).
An even more distinct tendency presents data set two, which shows the same two areas with deviations around 0 and -0.02 m. However, the third area scatters around a residual value of 8 cm, which is two times as high as for data set one. The deviations of the image residuals are split up in two areas as well. Errors in the relative orientation and in range lead to a falsification of the image observations; blurring effects arise.

Figure 5.10 presents the trajectory of those three integrated observations. Data set one (first row) already implies a high precision and the range measurements do not affect the result. However, data

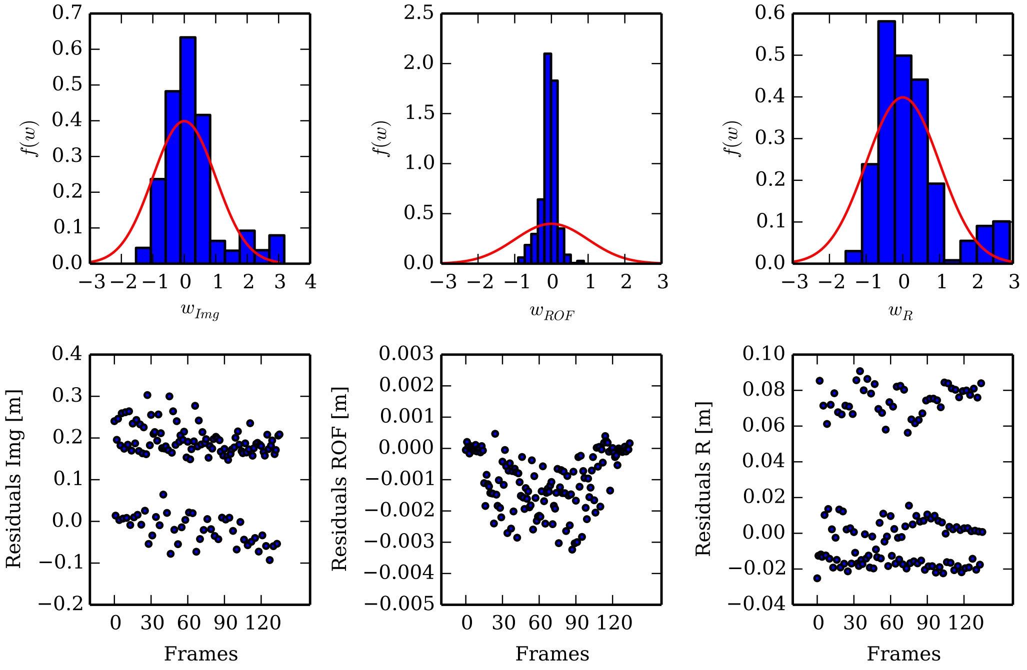

Figure 5.9: Data set 2 before VCE (Img+ROF+R) - First row: Normalised histograms of the standardised residuals; Second row: Scatter plots of the residuals per frame

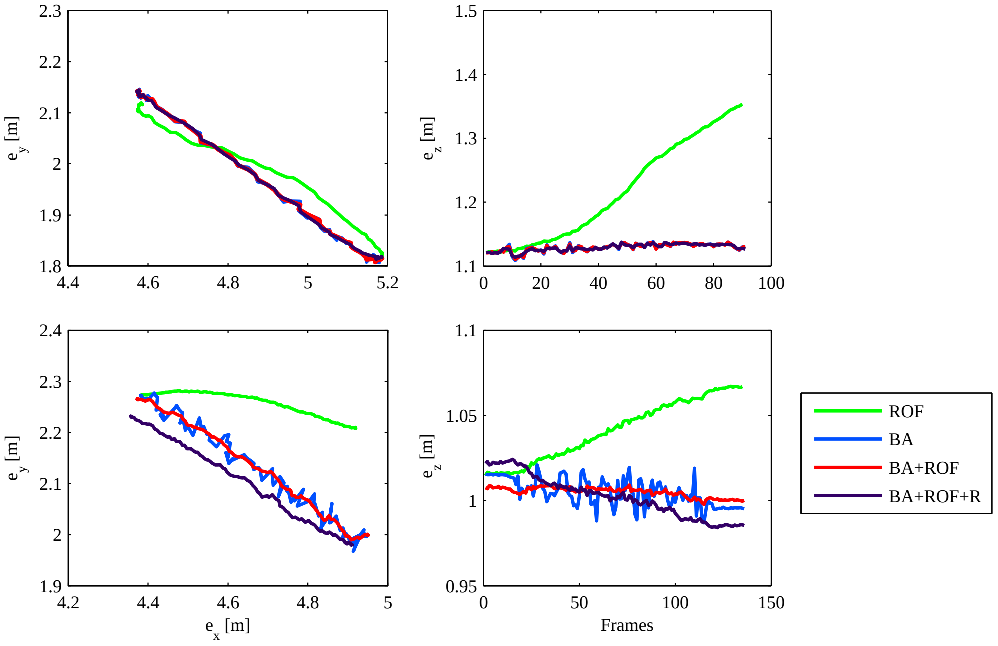

Figure 5.10: The purple line of data set two (second row) shows the influence of systematic errors in the distances. The camera trajectory is moved towards the object ($y = 0$).

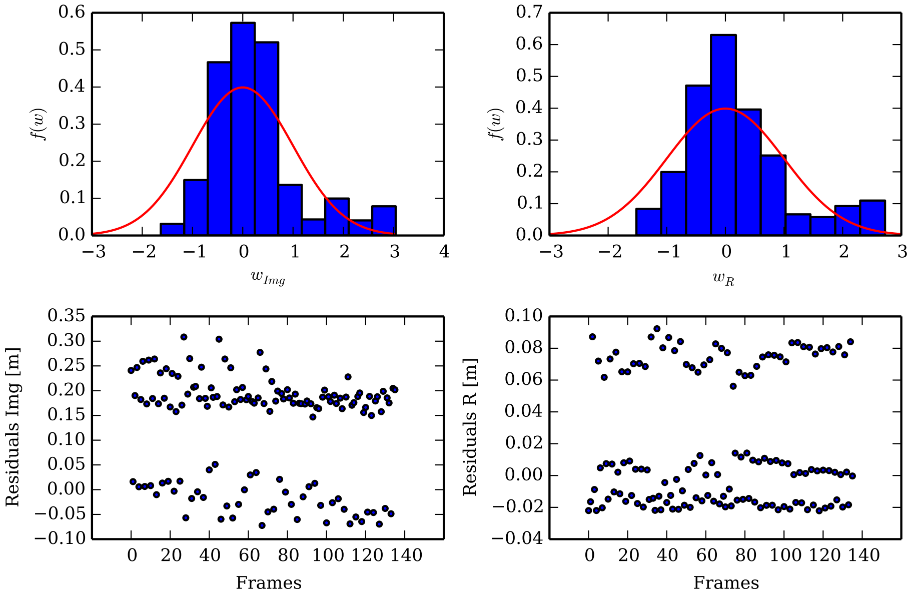

Figure 5.11: Data set 2 before VCE (Img+R) - First row: Normalised histograms of the standardised residuals; Second row: Scatter plots of the residuals per frame

set two (second row) is strongly influenced by range errors. The purple line is drawn to the object recorded (located in $y = 0$). The redundancy of the range observations amounts to 25 %. For the relative orientation it is nearly the same. The image observation group still takes up 50 % of the total redundancy.

As demonstrated in Chapter 5.1, the relative orientations should not be used in this context due to their deficiencies. Thus, the relative orientations are disregarded and the focus is directed on to the image and the range observations. In Figure 5.11 the image and range residuals of data set two are shown. The normalised distribution of the standardised residuals do not correspond to a Gaussian distribution. The right tails of the distributions might refer to gross errors. The histograms present the same tendency as in Figure 5.9. The blurring of the image observations is mainly caused by the range measurements.

To verify, if the range observations include systematic or gross errors, they need to be compared to a reference. The bundle adjustment results of Chapter 3.3 are suitable to calculate reference distances as defined in (4.10). Figure 5.12 presents the deviations of the observed ranges with respect to the calculated ones. The deviations are plotted in dependence of the ranges. No gross errors are visible. The range measurements of the control points at a distance of 5 m vary between ±2 cm. The standard deviation of the ranges belonging to the control points 323 and 324 amounts to ±4 cm, which originates from the range dependency.

The plot clearly shows, that systematic effects are still included in the observations. The deviations around 3 and 5 m respectively do not follow the same rules. The higher deviations correspond to the border areas of the image, as shown in Figure 3.6 and 3.7. Possibly, still errors due to the position on the sensor are present.

The high reflectivity of the close pillar in the scene might cause internal scattering and multipath effects

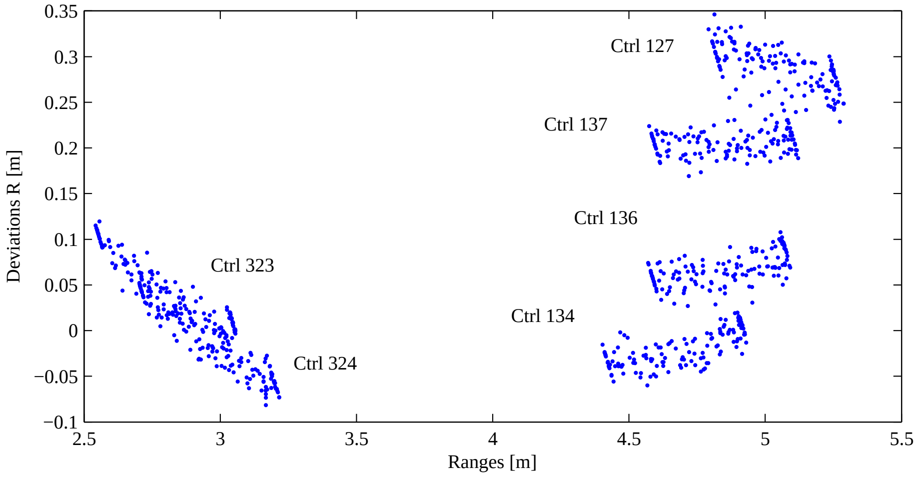

Figure 5.12: Data set 2: The range deviations as a function of the ranges - The range deviations are calculated based on the bundle adjustment results of Chapter 3.3.

have also not been considered, which mainly occur on object edges.

All errors need to be considered, before using the range data in the least squares adjustment. The integration of the range observations is not meaningful as well, otherwise the results get deformed, as shown in Figure 5.10.

5.3 Integration of the Control Points

When considering the control points as observed unknowns, tensions caused by the control points are assigned to the control points as well and are not only transferred to the image observations. If a weighted datum is used, those observations represent another observation type. Primarily, they are determined by a total station. Therefore, its precision shall be estimated in the course of the VCE as well.
However, one requirement of the VCE is that the observation groups shall show a similar number of observations or redundancy. In Table 5.5 the redundancies of the different observation groups after one iteration are shown. The redundancy of the control points does not even amount to 1 % of the total redundancy, as only six control points are considered for all frames. In comparison to the other groups its redundancy is so small, that it converges to zero after few iterations. So the question arises, which ratio of redundancies still leads to reasonable results.
The estimation of the control point precision is not feasible in the course of the VCE in this adjustment system.

5.4 Reduction of Image Observation Redundancy

The image observations strongly influence the trajectory as shown in various figures. In this section the effect of reducing the redundancy of the image observations is presented. This is verified on the integrated adjustment of image and relative orientation observations.

Observation Group	# Observations	Redundancy	Redundancy in %
Img	1092	783.4	48.5
ROF	540	436.9	27.1
R	546	390.6	24.2
Ctrl	18	3.2	0.2
Total	2196	1614.1	100.0

Table 5.5: Data set 1: The redundancy number of the different observation groups are shown.

Observation Group	# Observations	Redundancy	Redundancy in %
Data set 1:			
Img	546	284.4	46.0
ROF	540	291.6	54.0
Total	1086	540.0	100.0
Data set 2:			
Img	810	548.7	67.7
ROF	816	261.3	32.3
Total	1626	810.0	100.0

Table 5.6: The redundancy number of the different observation groups are shown.

In case of considering only three instead of the six control points, the number of image observations is nearly aligned with the relative orientations, as written in Table 5.6. The number of redundancy does not change in the same amount.

Figure 5.13 shows the trajectories calculated on basis of image observations of three control points per image and the relative orientations. The yellow line illustrates the influence of the reduction of redundancy. The large accumulated error of the relative orientation method appears in the trajectory, but it corresponds much better to the unbiased estimation (Img).

In Table 5.7 the absolute trajectory error of the different variants are shown. For data set one the consideration of only three control points produces a similar ATE (Max) as derived for six control points (Img+ROF$_6$), but its precision changed about 30 % for the worse. For data set two the ATE amounts to 10 cm and accords with the trajectory computed (Figure 5.13) and its root mean square error raised by a factor of 5.

Obviously, a reduction of the image observations by a factor of two, leads to an increased deviation of the trajectory (factor 6 in case of data set two).

	Max [m]	**RMSE** [m]	α_1 [gon]	α_n [gon]
Data set 1:				
ROF$_{ols}$	0.2234	0.1212	0.0004	0.0710
Img+ROF$_6$	0.0144	0.0062	0.0004	0.0010
Img+ROF$_3$	0.0172	0.0085	0.0063	-0.4996
Data set 2:				
ROF$_{ols}$	0.2363	0.1615	0.0002	0.0719
Img+ROF$_6$	0.0160	0.0099	0.0036	0.0017
Img+ROF$_3$	0.0972	0.0539	0.0190	0.0099

Table 5.7: Absolute Trajectory Error of data set one and two

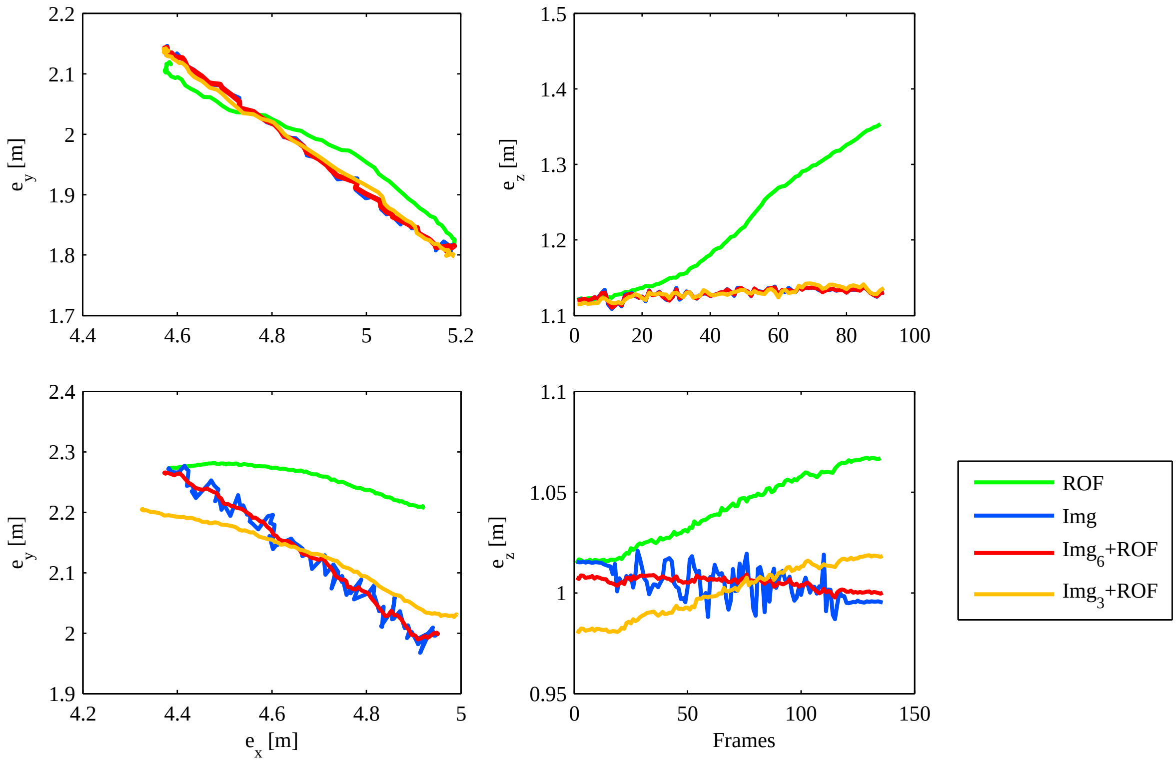

Figure 5.13: Comparison of different trajectories - The influence of a reduced number of control points on the trajectory is shown.

6. Conclusion

The aim of this work is to estimate the trajectory of a moving ToF camera. Two complementary methods are chosen. The bundle adjustment is the standard method in photogrammetry. It computes the camera pose in reference to the global coordinate system. In contrast, the relative orientation method based on the range optical flow of Ghuffar *et al.* (2013) delivers relative orientation parameters between two subsequent images. Both orientation methods show deficiencies. The thesis formulated cancels out the weaknesses of each method and shall result in a more precise estimation of the camera trajectory.

The first step was the acquisition of the data. Therefore, the existing measuring field in the laboratory of the Department of Geodesy and Geoinformation was chosen. To record reasonable data, knowledge of the functional principle of the ToF camera, as well as of the occurring errors is crucial. The control points, realised by the retro-reflecting targets are not the best choice of markers for data acquisition with a ToF camera. Those targets are so bright, that they heavily influence the data acquisition. The set-up parameters needed to be adapted. Two data sets are used, which have been recorded with different integration times (It).
The camera calibration of (Karel & Niedermayr, 2010) was applied, which presents a calibration of the ranging system. However, not all known systematic errors are modelled. Gross errors are eliminated.

Each method is applied on both data sets and the results are presented in Chapter 3.3 and 3.4. The parameters of data set one are estimated more precisely in case of the bundle adjustment, but the relative orientation results are inferior to that of data set two.
Data set two (It 30 ms) shows a better signal to noise ratio and the contrast in amplitude and range is higher than for data set one. For that reason, data set two presents more precise results for the relative orientation. The standard deviations of the estimated parameters in the bundle adjustment are worse by a factor of 2; this is possibly caused by motion blur. The absolute trajectory error (ATE) of both trajectories amounts to approximately 23 cm.

The trajectory of the relative orientation (incremental orientation) is smooth and shows a large accumulated error. In contrast, the exterior orientations of the bundle adjustment are unbiased. However, each image is oriented independently and the chronology of the orientations show a high variation. No information of subsequent images enters the bundle adjustment.
The thesis formulated is to integrate the relative orientation results into the bundle adjustment, to achieve a smooth unbiased trajectory. To set up the functional relation, it is necessary to know, if the relative orientation parameters are in reference to the first or the second camera coordinate system. The functional relation is derived geometrically and in closed form in Chapter 4.1.

A further extension of the adjustment model is applicable due to range measurements accomplished by the ToF camera. The range observations are treated critically due to the knowledge, that not all errors have been eliminated or modelled, which confirms later on.

Those different observations - image, relative orientations and range observations - present different precisions. In the course of a variance component estimation (VCE) the precision level for each observation type can be computed. It is an iterative procedure due to the estimation of the variance factor after the adjustment. To reasonably estimate those variance components, the total redundancy should lie between u and two times the number of unknowns u. Furthermore, a complete and correct functional model and data, free of gross errors, are prerequisites of the VCE.

During the analysis it turned out, that some preconditions are not fulfilled, more precisely the relative orientation results do not follow a Gaussian distribution. Thus, the least squares adjustment does not lead to the best unbiased estimation. The VCE should not be applied due to an incomplete functional model. Systematic errors in and possibly limitations of the relative orientation cause the deviation from the normal distribution.

Despite not fulfilled preconditions, the VCE of the integrated relative orientations converges after 18 iterations and produces a trajectory as expected. The outcome of the bundle adjustment of data set one already implies a high precision so as not to much improvement is visible. The absolute trajectory error amounts to approximately 1.5 cm for both data sets. The precision of the estimated parameters increases by 25 % for data set one and by a factor of 3 for data set two.

The variance component of the relative orientation amounts to 10^3, which means that the precisions of the relative orientation adjustment are estimated too optimistic by a factor of approximately 30. The reasons are the artificially high redundancy of the adjustment system and correlations in amplitude and range between neighbouring pixel, that have not been considered.

As expected, the range measurements are still affected by systematic errors, so the range observations deteriorate the results.

In this work, a simple approach of integrating different observation types was chosen. However, the variance component estimation can be extended by covariance components and by parameters of an error function. Both extensions are feasible on this adjustment problem.

The relative orientation comprises the translational and the rotational part. They are different observations with inherently different precisions and should build separate groups. Due to high correlations between those groups, covariance components would need to be estimated as well. Förstner (1979) shows a way of considering covariance components.

The systematic errors included in the range measurements, can be handled as overlapping variance components. The range variances are split up in a constant and a distance-dependent parameter and estimated in the course of the VCE.

A. Mathematical Notation

Vectors are bold small letters. If not explicitly noted, vectors are column vectors. Matrices are represented by bold capital letters and scalars are denoted as small letters.

The coordinate systems are defined by their unit vectors $\boldsymbol{e}_i$ in the following way: $C = (\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z)$. The origin and the orientation are described in the text. The coordinates are represented by x, y, z.

The nomenclature for adjustment calculus is used according to Niemeier (2008), e.g. variables after the adjustment (a posteriori) are assigned with the $\hat{}$-operator.

Generally, the index i is used as an enumeration index, but also corresponds to the object in reference to the global coordinate system C in specific. The index j belongs to the camera and the frame respectively. k is the index for the relative orientation.

In some cases a uniquely utilisation of variables or indices was not possible, but a distinction is given in the context of its usage.

The various symbols defining special variables are presented in Table 1.1.

Entity	**Example**	**Description**
Object point	P_i	capital letter with index i
Vector in C (C', C'')	$\boldsymbol{p}$ ($\boldsymbol{p}'$, $\boldsymbol{p}''$)	bold small letter
Perspective center in C	$\boldsymbol{p}_0$	bold small letter with index 0
Image vector in C'	$\boldsymbol{p}'_{ij}$	bold small letter with index ij
Vector in C' pointing to P_i	$\boldsymbol{p}'_{i,j}$	bold small letter with index i,j
Translation vector	$\boldsymbol{t}$	bold small letter
Rotation matrix	$\boldsymbol{R}$	bold capital letter
Column vector of $\boldsymbol{R}$	$\boldsymbol{r}$	bold small letter
Element of $\boldsymbol{R}$	$R_{1,1}$	capital letter with index
Range resp. Range Image	ρ	small letter
Amplitude (Brightness) resp. Amplitude Image	I	capital letter
Spatial amplitude- and range derivation	I_x, ρ_x	capital letter with index x or y
Temporal amplitude- and range derivation	I_t, ρ_t	capital letter with index t
Zero matrix	$\boldsymbol{0}$	bold zero

Table A.1: Notation of used variables

B. Source Code

B.1 Cost function - Relative Orientation

```
import oriental
import numpy as np
from oriental import adjust
from oriental.adjust import parameters
from oriental.adjust.cost import cos, sin

class CostROF( oriental.adjust.cost.AutoDiff ):

    def __init__( self, dT1, dT2, dT3, dR1, dR2, dR3, sqrtP ):
        self.dT1 = dT1
        self.dT2 = dT2
        self.dT3 = dT3
        self.dR1 = dR1
        self.dR2 = dR2
        self.dR3 = dR3
        self.sqrtPBlock = sqrtP

        numResiduals = 6
        parameterSizes = (3,3,3,3)
        super( CostROF, self ).__init__( numResiduals, parameterSizes )

    def Evaluate( self, parameters, residuals ):
        dt = type(parameters[0][0])
        rad = np.pi/200
        T1 = parameters[0]
        T2 = parameters[1]
        R1 = parameters[2]
        R2 = parameters[3]

    # ROTATION
        al = R1[0]*rad
        ze = R1[1]*rad
        ka = R1[2]*rad
        rot1 = self.alzeka( al, ze, ka )
```

```
        al = R2[0]*rad
        ze = R2[1]*rad
        ka = R2[2]*rad
        rot2 = self.alzeka( al, ze, ka)

        R = rot2.T.dot( rot1 )

        om = dt(self.dR1)
        ph = dt(self.dR2)
        ka = dt(self.dR3)
        diff_R = [[   1, -ka,  ph],
                  [ ka,    1, -om],
                  [-ph,  om,    1]]

        residuals[3] = R[1][0]-diff_R[1][0]
        residuals[4] = R[2][0]-diff_R[2][0]
        residuals[5] = R[2][1]-diff_R[2][1]

    # TRANSLATION
        d = T1 - T2
        R = rot2.T
        prediction_X = R.dot(d)

        residuals[0] = prediction_X[0] - dt(self.dT1)
        residuals[1] = prediction_X[1] - dt(self.dT2)
        residuals[2] = prediction_X[2] - dt(self.dT3)

        residuals_weighted = np.zeros( residuals.shape, dtype=dt )
        for iRow in xrange(6):
            for iCol in xrange(6):
                residuals_weighted[iRow] += residuals[iCol] * self.sqrtPBlock[iRow,iCol]
        residuals[:] = residuals_weighted

        return True


    def alzeka( self, al, ze, ka ):
        "returns the rotation matrix defined by (alpha, zeta, kappe) in [radians]"
        assert type(al)==type(ze) and type(al)==type(ka)
        dt = type(al)
        cAl = cos(al)
        sAl = sin(al)
        cZe = cos(ze)
        sZe = sin(ze)
        cKa = cos(ka)
        sKa = sin(ka)
```

```
        rot = np.empty( (3,3), dtype=dt )
        rot[0][0] =  cAl*cZe*cKa - sAl*sKa;
        rot[0][1] = -cAl*cZe*sKa - sAl*cKa;
        rot[0][2] =  cAl*sZe;
        rot[1][0] =  sAl*cZe*cKa + cAl*sKa;
        rot[1][1] = -sAl*cZe*sKa + cAl*cKa;
        rot[1][2] =  sAl*sZe;
        rot[2][0] = -sZe*cKa;
        rot[2][1] =  sKa*sZe;
        rot[2][2] =  cZe;

        return rot
```

Listing B.1: Cost function of relative orientations

B.2 Cost function - Range

```
import oriental
import numpy as np
from oriental import adjust
from oriental.adjust.cost import sqrt

class CostRange( oriental.adjust.cost.AutoDiff ):

    def __init__( self, Range , sqrtP ):
        self.Range = Range
        self.sqrtP = sqrtP

        numResiduals = 1
        parameterSizes = (3,3)
        super( CostRange, self ).__init__( numResiduals, parameterSizes )

    def Evaluate( self, parameters, residuals ):
        dt     = type(parameters[0][0]) #Jet
        prjCtr = parameters[0]
        X      = parameters[1]

        range_calc = sqrt((prjCtr[0]-X[0])*(prjCtr[0]-X[0]) +
                          (prjCtr[1]-X[1])*(prjCtr[1]-X[1]) +
                          (prjCtr[2]-X[2])*(prjCtr[2]-X[2]))
        residuals[0] = (range_calc - dt(self.Range))  * self.sqrtP

        return True
```

Listing B.2: Cost function of the range

B.3 Cost function - Control Point

```
import oriental
import numpy as np
from oriental import adjust

class CostCtrlPts( oriental.adjust.cost.AutoDiff ):

    def __init__( self, pointsx, pointsy, pointsz, sqrtP ):
        self.pointsx = pointsx
        self.pointsy = pointsy
        self.pointsz = pointsz
        self.sqrtPBlock = sqrtP

        numResiduals = 3
        parameterSizes = (3,)
        super( CostCtrlPts, self ).__init__( numResiduals, parameterSizes )

    def Evaluate( self, parameters, residuals ):

        dt = type(parameters[0][0]) # Jet
        X = parameters[0]

        residuals[0] = (X[0] - dt(self.pointsx))*self.sqrtPBlock[0,0]
        residuals[1] = (X[1] - dt(self.pointsy))*self.sqrtPBlock[1,1]
        residuals[2] = (X[2] - dt(self.pointsz))*self.sqrtPBlock[2,2]

        return True
```

Listing B.3: Cost function of the control point

Acknowledgement

I want to thank Prof. Pfeifer for the topic, I was able to work on, for his contribution of ideas and incitations. Willi, thank you for your competent and mostly fast assistance - working with you was enriching. Sajid, I want to thank you for the knowledge you provided to me.

Thanks to my colleagues Helene, Marlies and Martin for listening to me and changing my way of looking at a problem. Thank you Markus for the discussions about choosing right indices to create a unique formulation of a problem. Furthermore, I want to thank my family for always supporting me.

References

Agarwal, S., Mierle, K. and Others (2013). Ceres solver. `https://code.google.com/p/ceres-solver/`.

Buettgen, B., Oggier, T., Lehmann, M., Kaufmann, R. and Lustenberger, F. (2005). Ccd/cmos lock-in pixel for range imaging: challenges, limitations and state-of-the-art. In *In Proceedings of 1st Range Imaging Research Day*, p. 21–32.

Corliss, G.F. and Griewank, A. (1993). Operator overloading as an enabling technology for automatic differentiation. Tech. rep., Mathematics and Computer Science Division, Argonne National Laboratory.

Förstner, W. (1979). Ein verfahren zur schätzung von varianz- und kovarianzkomponenten. *Allgmeine Vermessungs Nachrichten (AVN)*, 11-12, p. 446–452.

Ghuffar, S. (2014). *Motion Estimation from Integration of Range and Intensity Data*. Ph.D. thesis, Vienna University of Technology.

Ghuffar, S., Ressl, C. and Pfeifer, N. (2013). Relative orientation of videos from range imaging cameras. In *Proceedings of the SPIE 8791, Videometrics, Range Imaging, and Applications XII; and Automated Visual Inspection*, 879114.

Horn, B. and Schunck, B. (1981). Determining optical flow. *Artificial Intelligence*, 17, p. 185–203.

Kager, H., Rottensteiner, F., Kerschner, M. and Stadler, P. (2002). *Orpheus user manual 3.2.1*. Institute of Photogrammetry and Remote Sensing, Vienna University of Technology.

Karel, W. and Niedermayr, S. (2010). Photogrammetrie + laserscanning = range imaging. *VGI 98*, 2, 1–8.

Karel, W., Doneus, M., Verhoeven, G., Briese, C., Ressl, C. and Pfeifer, N. (2013). Oriental - automatic geo-referencing and ortho-rectification of archaelogical aerial photographs. *ISPRS Annals of the Photogrammetry,*, Volume II-5/W1.

Kraus, K. (1996). *Photogrammetrie*, Vol. Band 2. Dümmler/Bonn, 3rd edn.

Luhmann, T. (2000). *Nahbereichsphotogrammetrie*. H. Wichmann Verlag.

Matlab (1984-2010). *MatLab R2010b Documentation*. The MathWorks, Inc.

McGlone, J.C., ed. (2004). *Manual of Photogrammetry*. American Society for Photogrammetry and Remote Sensing, 5th edn.

NIEMEIER, W. (2008). *Ausgleichungsrechnung*. de Gruyter.

OGGIER, T., LEHMANN, M., KAUFMANN, R., SCHWEIZER, M., RICHTER, M., P. METZLER, G.L., LUSTENBERGER, F. AND BLANC, N. (2003). An all-solid-state optical range camera for 3d real-time imaging with sub-centimeter depth resolution (swissrangertm). *SPIE*, 5249, p. 634–545.

PFEIFER, N., LICHTI, D., BÖHM, J. AND KAREL, W. (2013). *TOF Range-Imaging Cameras*, chap. 3D Cameras: Errors, Calibration and Orientation, p. 117–138. Springer.

SPIES, H., JÄHNE, B. AND BARRON, J. (2002). Range flow estimation. *Computer Vision and Image Understanding*, 85, p. 209–231.

SWISSRANGER (2008). *SR-3000 Data Sheet*. MESA Imaging AG.

TREFETHEN, L. AND BAU, D. (1997). *Numerical linear algebra*. siam - Society for Industrial and Applied Mathematics.

TRIGGS, B., ZISSERMAN, A. AND SZELISKI, R. (1999). Bundle adjustment - a modern synthesis. In *Vision Algorithms: Theory and Practice*, p. 298–373, Springer-Verlag.

List of Figures

List of Tables