Affine Multi-View Modelling for Close Range Object Measurement

A thesis submitted to the University College London for the degree of Doctor of Philosophy (Ph.D.)

Margarita Rova

Department of Civil, Environmental and Geomatic Engineering

London, 2010

Declaration of authorship

I, Margarita Rova, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed   …………………………………
Date   …………………………………

Acknowledgements

Abstract

In photogrammetry, sensor modelling with 3D point estimation is a fundamental topic of research. Perspective frame cameras offer the mathematical basis for close range modelling approaches. The norm is to employ robust bundle adjustments for simultaneous parameter estimation and 3D object measurement. In 2D to 3D modelling strategies image resolution, scale, sampling and geometric distortion are prior factors. Non-conventional image geometries that implement uncalibrated cameras are established in computer vision approaches; these aim for fast solutions at the expense of precision. The projective camera is defined in homogeneous terms and linear algorithms are employed. An attractive sensor model disembodied from projective distortions is the affine. Affine modelling has been studied in the contexts of geometry recovery, feature detection and texturing in vision, however multi-view approaches for precise object measurement are not yet widely available.

This project investigates affine multi-view modelling from a photogrammetric standpoint. A new affine bundle adjustment system has been developed for point-based data observed in close range image networks. The system allows calibration, orientation and 3D point estimation. It is processed as a least squares solution with high redundancy providing statistical analysis. Starting values are recovered from a combination of implicit perspective and explicit affine approaches. System development focuses on retrieval of orientation parameters, 3D point coordinates and internal calibration with definition of system datum, sensor scale and radial lens distortion. Algorithm development is supported with method description by simulation. Initialization and implementation are evaluated with the statistical indicators, algorithm convergence and correlation of parameters. Object space is assessed with evaluation of the 3D point correlation coefficients and error ellipsoids. Sensor scale is checked with comparison of camera systems utilizing quality and accuracy metrics. For independent method evaluation, testing is implemented over a perspective bundle adjustment tool with similar indicators. Test datasets are initialized from precise reference image networks. Real affine image networks are acquired with an optical system (~1M pixel CCD cameras with 0.16x telecentric lens). Analysis of tests ascertains that the affine method results in an RMS image misclosure at a sub-pixel level and precisions of a few tenths of microns in object space.

Περίληψη

Στη φωτογραμμετρία ο προσανατολισμός της κάμερας και ο προσδιορισμός τρισδιάστατων σημείων αποτελεί θεμελιώδες ερευνητικό θέμα. Προοπτικές κάμερες προσφέρουν τη μαθηματική βάση των αλγόριθμων που εφαρμόζονται σε επίγειες εφαρμογές. Νόρμα αποτελεί η εφαρμογή συστημάτων συνόρθωσης της δέσμης για τον ταυτόχρονο προσδιορισμό των παραμέτρων προσανατολισμού και τρισδιάστατης μέτρησης σημείων. Σε στρατηγικές που βασίζονται στη διδιάστατη προς τη τρισδιάστατη αντιστοιχία η ανάλυση της εικόνας, η κλίμακα του σένσορα, η δειγματοληψία επιφανείας καθώς και η γεωμετρική διαστροφή αποτελούν βασικούς παράγοντες. Μη συμβατικές εικονοληπτικές γεωμετρίες που εφαρμόζονται σε μη-βαθμονομημένες κάμερες είναι εδραιωμένες σε εφαρμογές της όρασης υπολογιστών; αυτές αποσκοπούν σε επιλύσεις υψηλής ταχύτητας επιβαρύνοντας τον παράγοντα της ακρίβειας. Η προβολική κάμερα προσδιορίζεται σε συστήματα ομογενών συντεταγμένων όπου εφαρμόζονται γραμμικοί αλγόριθμοι. Ένας σένσορας ιδιαίτερου ενδιαφέροντος είναι ο αφινικός, ο οποίος είναι απαλλαγμένος από προβολικές διαστροφές. Αφινικές μέθοδοι έχουν μελετηθεί στο πλαίσιο της ανακατασκευής γεωμετρίας, εξαγωγής σημείων και της δημιουργίας υφών στην όραση υπολογιστών; ωστόσο πολυ-εικονικές μέθοδοι για μέτρηση αντικειμένων με υψηλή ακρίβεια δεν είναι ακόμα ευρέως διαθέσιμες.

Αυτή η μελέτη ερευνά το αφινικό πολυ-εικονικό πρόβλημα από την φωτογραμμετρική σκοπιά. Ένα νέο αφινικό σύστημα συνόρθωσης της δέσμης έχει αναπτυχθεί για δεδομένα μέτρησης σημείων που παρατηρούνται σε επίγεια δίκτυα εικόνων. Το σύστημα επιτρέπει βαθμονόμηση, προσανατολισμό και υπολογισμό σημείων στο χώρο του αντικειμένου. Επεξεργάζεται ως ελαχιστοτετραγωνική επίλυση με υψηλή περίσσεια παρέχοντας στατιστική ανάλυση. Αρχικές τιμές ανακτώνται μέσω συνδυασμού έμμεσων προοπτικών και άμεσων αφινικών ρουτινών. Η ανάπτυξη του συστήματος εστιάζει στον υπολογισμό των παραμέτρων προσανατολισμού, συντεταγμένων σημείων αντικειμένου και εσωτερικής βαθμονόμησης με τη δυνατότητα να προσδιορίζεται το σύστημα αναφοράς (εξωτερικές ή εσωτερικές δεσμεύσεις), της κλίμακας του σένσορα και της ακτινικής διαστροφής. Η ανάπτυξη του αλγορίθμου υποστηρίζεται με περιγραφή της μεθόδου με δεδομένα προσομοίωσης. Υπολογισμός αρχικών τιμών και εφαρμογή του

αλγόριθμου εκτιμώνται με στατιστικούς δείκτες, σύγκλιση του αλγόριθμου και συσχέτιση των παραμέτρων. Ο χώρος του αντικειμένου ελέγχεται με αξιολόγηση των συντελεστών συσχέτισης για τα σημεία του χώρου και των ελλειψοειδών σφάλματος. Η κλίμακα του σένσορα ελέγχεται με σύγκριση συστημάτων κάμερας με μέτρα εσωτερικής και εξωτερικής ακρίβειας. Για την ανεξάρτητη αξιολόγηση της μεθόδου, έλεγχος εφαρμόζεται με συμβατικό σύστημα μεθόδου της δέσμης με παρόμοιους δείκτες. Ο υπολογισμός των αρχικών τιμών των πειραματικών δεδομένων έχει προέλθει από ακριβή δίκτυα εικόνας αναφοράς. Αληθή αφινικά δίκτυα εικόνων λαμβάνονται με ένα οπτικό σύστημα (~1M pixel CCD κάμερες οπλιζόμενες με 0.16x τηλεκεντρικό φακό). Ανάλυση των ελέγχων επιβεβαιώνει ότι το αφινικό μοντέλο αποδίδει αποτελέσματα με RMS σφάλμα εικόνας υποψηφιδικής τάξης και ακρίβειας της τάξης των μερικών δεκάδων μικρών στο χώρο του αντικειμένου.

Contents

List of figures

List of tables

Abbreviations

ADP	ADditional Parameter
BA	Bundle Adjustment
CAL	CALibration data structure
CCD	Charge Coupled Device
CFA	Colour Filter Array
CMOS	Complementary Metal Oxide Semiconductor
CP	Control Point
CPU	Central Processing Unit
CTF	Contrast Transfer Function
CV	Computer Vision
DEM	Digital Elevation Model
DLT	Direct Linear Transformation
DOF	Depth Of Field
DPW	Digital Photogrammetric Workstation
DSLR	Digital Single Lens Reflex
EO	Exterior Orientation
FOD	First Order Design
FOV	Field Of View
ICP	Iterative Closest Point
ID	IDentifier
IO	Interior Orientation
IS	Image Space
ISPRS	International Society for Photogrammetry and Remote Sensing
LED	Light Emitting Diode
LSE	Least Squares Estimation
LSM	Least Squares Matching
MAD	Mean of Absolute Differences
MOS	Metal Oxide Semiconductor
MTF	Modulation Transfer Function
NSMP	Normalized Span of Model Parameters
OBS	OBServation data structure
ORIENT	affine ORIENTation data structure

OS	Object Space
PHOTO	PHOTO orientation data structure
PSF	Point Spread Function
RAM	Random Access Memory
RANSAC	RANdom SAmple Consensus algorithm
RGB	Red Green Blue
ROI	Region Of Interest
SLR	Single Lens Reflex
SNR	Signal to Noise Ratio
SOD	Second Order Design
SPIE	The Society of Photo - optical Instrumentation Engineers
SVD	Singular Value Decomposition
TARGET	TARGET data structure
TOD	Third Order Design
TP	Tie Point
UCL	University College London
VMS	Vision Measurement System
ZOD	Zero Order Design
ZW	Zeng-Wang resection
1D	One Dimensional
2D	Two Dimensional
3D	Three Dimensional

1.   Introduction

Photogrammetry is primarily focused on camera calibration, orientation and object reconstruction. Since its origin, perspective-based imaging sensors have been widely utilized particularly in industrial, archaeological and medical applications. In principle, the main aim of photogrammetric processing is accuracy. The norm is to apply robust bundle adjustment tools for both parameter estimation and object reconstruction. A similar sensor-based discipline is computer vision which is generally open in variant imaging geometries focusing on the implementation of uncalibrated cameras and automation. The projective camera is defined in homogeneous terms and linear algorithms are employed. Current state of the art is system automation and often fusion of intensity and range imaging sensors. Thus, integration of photogrammetry and computer vision approaches is a fact, particularly when considering their significant overlap in close range applications.

This research reports on the investigation and development of the affine sensor model adopted from the computer vision community. It offers a processing framework in the context of multi-view modelling from affine images. Specifically a new bundle adjustment system has been developed and applied in close range images arranged in strong convergent network configurations. The developed system allows calibration, orientation and three dimensional (3D) point estimation in a photogrammetric approach. This implies that the system is processed as a least squares solution with high redundancy and that it provides statistical analysis of the achieved quality. To introduce the research covered within this thesis, the following sections are outlined to provide the context (see section 1.1.), motivation (see section 1.2.), problem statement (see section 1.3.), research objectives (see section 1.4.) and tools (see section 1.5.) as well as a summary of the thesis structure (see section 1.6.).

1.1. Context

Photogrammetric modelling and measurement approaches are highly dependent on correct camera calibration and orientation. Frame perspective-based sensors are typical in established close range systems (Maas, 2008). In industrial applications such systems are calibrated based on routine self-calibrated bundle adjustment strategies with quality assessment (Brown 1974; Granshaw, 1980; Clarke and Fryer

1998; Triggs et. al., 2000; Gruen and Beyer, 2001; Fraser, 2001). In fact processing usually implies the establishment of geometrically strong (wide baselines, angle of convergence and intersection rays) imaging networks. Planar or volumetric arrays, artificial high contrast or natural signalized point features, controlled illumination and stable calibration conditions are some of the physical requirements. It is natural that these are aspects concerning the algorithmic framework within which calibration and orientation are implemented.

Besides the establishment of camera calibration and orientation, applications in architecture, cultural heritage and medicine, the main areas of close range photogrammetry, generally require the production of complete (without voids) textured 3D models. Image based approaches applied in such instances are based on stereo or multi-image matching strategies (Remondino and Zhang, 2006). At larger mapping scales alternative techniques combine images with laser scan range data. The accuracy achieved is dependent on the registration and texture mapping methods (El Hakim et. al., 1998). Where data registration is concerned this can relate to a geometric transform between different sensors, views or temporal variations (Zitova & Flusser, 2003). Approaches that are based on intensity and feature correspondences are typical; yet they result in seamless pairwise problems that are propagated within the final model and reduce the quality as a result (El-Hakim et al., 2004). Subsequent texturing may rely on a projective transform between for example the triangle plain (of a model) and the texture (of an image). However, significant distortions can be visible at triangle edges particularly regarding radiometric differences, even if correct calibration and orientation are considered (Grammatikopoulos et al., 2005). Object geometry and texture, as well as metric requirements, dictate the approach followed at different instances.

In close range imaging the perspective sensor can present strong scale variation, non-consistent sampling as well as inner geometric distortions. In fact applications that focus on fine object detail measurement of objects that occupy volumes of a few centimetres in object space, scale recovery becomes critical. An alternative imaging situation is the affine projection which connects image and object spaces through parallel lines of sight. It is characterized by an invariant scale factor and given the

parallelism of observation lines no perspective projection is involved. As a result, the image plane can be positioned anywhere along the optical axis while oriented orthogonally to the imaging rays. The affine is a generic sensor model which is closer to the Euclidean reconstruction, it was firstly introduced by the computer vision community and it can thus be found in associated literature.

In structure from motion problems the affine sensor has been implemented with geometric approaches based on local coordinate frame methods (Koenderink and Van Doorn, 1991). Tomasi and Kanade (1992) proposed a non-local coordinate frame method applied in the total scene points, but problems regarding rank considerations have been observed. Definition of a coordinate datum is based upon the centroid of targets cluster. Shapiro (1995) follows an extended multi-view approach based on a singular value decomposition (SVD)[1] solution. In image analysis Mikolajczyk & Schmid (2004) outline a series of interest detection methods which are invariant under scale and affine transforms with main objective the performance evaluation of these methods. From the photogrammetric standpoint the affine sensor has been employed in mathematical problems that for example explore initialization of orientation procedures (Kyle, 2004) or perform long distance measurements (Ono et. al., 2004). El-Hakim et al. (2004) apply image-based registration methods based on an affine or projective model[2] for the purpose of image mosaicing. In fact, in mapping applications orthoimage generation is the result of aircraft or satellite imagery in the aerial processing domain or architectural mapping in the close range. This can be regarded as a special case of image resampling where the spatial resolution of the source image in combination with the resolution of the digital elevation model (DEM) in digital photogrammetric workstations (DPW) identify the final quality of the product (Agouris et al., 2004). Where texturing is concerned, Weinhaus & Devich (1999) have demonstrated a hybrid projection model based on a unified perspective and affine projection model that can be optionally adapted for the purpose of mapping textures onto planar polygons.

---

[1] Singular Value Decomposition (SVD): If A is a mxn real matrix with m>n, then matrix A can be written based on the SVD form: $A=UDV^{T}$ (Wolfram, 2009b).
[2] An affine transform involves six parameters and it is generally composed of translation, rotation, scale and orthogonality (scaling direction) parameters. The affine transform can be upgraded to projective by incrementing the transformation parameters by two. The resultant eight parameter projective transform will generally map parallel lines to convergent.

1.2. Motivation

Imaging projection is the process that connects a two dimensional (2D) image with 3D object space to recover geometry. From the close range viewpoint, affine images have not been widely utilized in practice. A primary reason for this is the fact that such real images are acquired with dedicated machine vision systems which are not general purpose cameras. In fact these sensors are characterized by a limited imaging footprint and range with a simultaneous increase in image scale. Nevertheless, minimal geometric distortion, as well as consistent image sampling, can significantly enhance the quality of modelling (for instance in a multi-view framework) from pure affine images. In the context of establishing a close range convergent network of affine images the modelling task becomes the intersection of the 2D to 3D lines of sight in order to calibrate, orientate the employed sensor and coordinate points in 3D. Photogrammetric processing of such data requires the ability to include full error propagation within the system for statistical analysis.

This thesis presents a new multi-view modelling algorithm for the processing of point based data structures measured on affine images in the close range. The method is appropriate for close-range convergent image networks acquired with an affine machine vision system. System initialization is performed from a set of artificially high contrast signalized geometric structures. The algorithm is processed in the form of a bundle adjustment system supported with statistical analysis. Both stages of initialization and bundle adjustment processing are evaluated in a methodological approach starting from a simulation project and subsequent testing with real world datasets. Assessment is extracted at each stage with statistical indicators, correlation analysis as well as independent checks according to the demands of testing. It is noted that the employed datasets are purpose built geometric structures of varying geometry and that they are implemented in order to evaluate practical aspects of the method.

This work focuses on the investigation and implementation of an affine multi-view modelling framework which can be seen as an initial research work towards complete 3D modelling from pure affine sensor imagery. In this context, and as far as affine multi-view modelling is concerned, the method can be applied in any application

where non-contact fine object detail measurements are required such as metrology, archaeology and medicine (see section 2.2.).

1.3. Problem statement

This research work seeks the answers to the following fundamental questions:

- Is it realistic to generate, measure and process real affine multi-view images within a modelling framework in the context of deriving precise close range object measurements?

- In the context of such a framework, how do sensor geometry (parallel projection rays, invariant scale factor and calibration) as well as local coordinate frame (datum) influence method precision and 3D point estimation (object geometry)?

- What is the quality of affine multi-image modelling in comparison to established photogrammetric solutions?

1.4. Research objectives

Built upon the problem statement the main objectives of this thesis are formulated as follows:

1. To investigate the affine sensor model for the multi-view imaging case from the photogrammetric standpoint. On this basis the main aim is to develop and propose a framework that offers the potential to accommodate sensor calibration, orientation and 3D object measurement. In addition the method needs to be capable of catering for full covariance matrices and therefore to provide measurable outcomes with regards to the method's quality evaluation. It is stated that affine sensor development involves both study of theoretical aspects (e.g. starting value derivation, algorithm design and method development with simulation data) as well as treatment of practical aspects (e.g. method application for real world test data acquired with an affine system).

2. To assess the method intrinsically in the aspects of (a) model parameters estimation (b) algorithm convergence behaviour and (c) consistency of parameters correlation.

   2.1. Model behaviour refers to method initialization (reference measurements and affine starting value estimation) as well as overall bundle adjustment results. To achieve this, evaluation measures include implicit statistical indicators (e.g. number of iterations, redundancy, a posteriori precision factor and RMS image misclosure), 3D point precisions, residual vectors and histograms visualizations, radial lens distortion profiles as well as 3D check measurements.

   2.2. Convergence behaviour assesses the aspects of algorithm convergence as well as quality of convergence. Estimation and visualization of the change in model parameters between successive pairs of iterations is one derived measure for this purpose. Additional measures are model parameters precisions and their visualization as well as the mean of absolute differences between successive iteration pairs.

   2.3. Consistency of parameter correlations are evaluated with inspection of the correlation coefficient matrix patterns for different bundle adjustment runs (e.g. external or inner constraints datum, control and tie point data implementation and inclusion or exclusion of radial distortion parameter). In addition, individual solutions are selected to illustrate correlations of parameters where these are significantly large (the term large here denotes correlation coefficients that are greater than 0.75 and 0.90 in magnitude).

3. To assess the method extrinsically in the aspects of (a) object space evaluation, as well as (b) system scale invariance.

   3.1. Object space evaluation refers to evaluation of bundle adjustment results for datasets of different image sensor and quality, object geometry and point data contribution (network geometry and point visibility). Measures utilized to assess this aspect can include the method's statistical indicators, absolute 3D correlations with point proximity as well as 3D error ellipsoids for the estimated point data.

   3.2. System scale assessment relates to method evaluation for test datasets acquired with different sensors. Prior measures applied for this purpose

involve the method's statistical indicators. Further absolute differences can be employed over external independent length measurements whereas absolute differences between estimated and reference point data can be considered as highly useful.

4. To assess the method quality aspect in relation to corresponding established photogrammetric approaches. This aims on the evaluation of the developed affine multi-view solution with regards to a well tested and well understood perspective-bundle adjustment. Besides bundle adjustment statistical results, 3D point error ellipsoids and 3D point (control and tie point data) can be utilized for the purpose of independent evaluation.

1.5. Research tools

To develop and implement the method that is demonstrated here a set of tools have been utilized. The mathematical model was written and implemented in Microsoft's Visual Studio 2005 in C/C++ (Press et al., 2005) and was subsequently upgraded in Visual Studio 2008. In support of implementation and analysis of the developed multi-view framework additional tools were utilized. Particular mathematical model testing and partial graphical output in some instances was performed in the Mathworks Matlab environment (Mathworks, 2009). Initialization, image measurement and reference data processing were implemented in the in-house photogrammetric tool VMS 8.0. Additional educational and open source tools were utilized to underpin experimental analysis and methodological testing.

1.6. Structure summary

This thesis is composed of seven main chapters supported by the references and the associated appendices (see Figure 1.1). Chapters 2 - 4 form the core background to this research work, Chapter 5 refers to the proposed modelling methodology developed and implemented for the purposes of this research work, Chapter 6 analyzes and outputs the results and finally Chapter 7 summarizes the main conclusions and discusses directions for future research.

Context, statement & goals	Background & technical fundamentals	Modelling, implementation, testing & results	Assessment of conclusions	Additional pointers
1. Introduction	2. Non-contact object measurement in the close range 3. Digital close range image formation 4. Modelling from multiple views	5. Affine multi–view modelling 6. Results and analysis	7. Conclusions and future research	References Appendix A: Essential elements for affine model formation Appendix B: Processing framework Appendix C: Processing data examples

Figure 1.1: Thesis outline.

In more detail the thesis contents are outlined as follows:

Chapter 2: Non-contact object measurement in the close range

This chapter provides the thesis' context. The main concepts of photogrammetry and computer vision are covered with reference to basic calibration and orientation procedures whilst covering examples of close range measurement applications. Prior to the analysis of any multi-view problem a first concern is to understand the digital image formation process. This is the central topic of the subsequent chapter.

Chapter 3: Digital close range image formation

This chapter emphasizes the fundamental issues governing digital images presented as a literature review. Aspects relating to close range image acquisition systems, digital image characteristics in relation to quality and geometry, as well as measurement methods are covered. The chapter closes with the fundamental camera calibration models.

Chapter 4: Modelling from multiple views

This chapter is initiated from the starting point of a bundle adjustment overview to place the current state of the art in modelling of frame cameras. Subsequently it reviews the least squares technique and fundamental modelling of perspective cameras, covering the aspects of self calibration and starting value estimation. It introduces the affine sensor linked with a description of approaches found in the literature. The concepts of datum constraints, network geometry and quality control are additionally given.

<u>Chapter 5: Affine multi-view modelling</u>

This chapter presents the proposed method, starting from the point of model initialization to the full implementation of the affine multi-view framework. The description of the method is given in analytical terms regarding its subsequent implemented stages. In support of algorithm development the method is presented from the standpoint of simulation. In this regard a synthetic test object is treated for a subset of two, three and seven view geometry cases to describe the method.

<u>Chapter 6: Results and analysis</u>

This chapter covers the research results and provides analysis of the developed method. Starting from a descriptive viewpoint of the designed test objects and image-sets, this chapter addresses the developed method through a series of extensive test cases where ad hoc aspects are evaluated. Particularly, the demonstrated approach is investigated in relation to model behaviour, object space analysis, sensor scale analysis as well as independent assessment. The test-data are initialized and premeasured from precise reference measurements. Statistical quality and accuracy measures are given for each of the test cases.

<u>Chapter 7: Conclusions and future research</u>

This chapter summarizes the findings of the research illustrating the central conclusions and proposes future research work directions regarding algorithm improvements as well as practical extensions of the developed approach.

The thesis is completed with the enclosed references and appendices that support the methods and data processing where this is critical.

2. Non-contact object measurement in the close range

This chapter is organized as a background chapter starting from the main areas of interest which are relevant to this thesis (see section 2.1.). Subsequently some application examples are given linked with an outline of some typical problems in the close range (see section 2.2.). These topics are covered in the aspect of non-contact object measurement in the close range in order to provide the thesis context.

2.1. Areas of interest

The central application areas of close range photogrammetry are industrial metrology, cultural heritage and medicine. Nowadays a range of different approaches can be applied dependent on the requirements, specifications and accuracy levels of the end product, as well as employed sensors, tools and algorithms capabilities to name a few factors. Figure 2.1 comprises an attempt to provide a diagram of the approaches that find wide applications in studies of close range object measurements. Following the structure of the diagram, these disciplines are referred to as sensor-based, subdivided as photogrammetry and computer vision. In the antipode, geometry-based infer computer graphics methods. The processes behind these areas present differences related to accuracy, processing and cost as key factors but they share overlapping interest in their suitability for registration and surface reconstruction tasks.
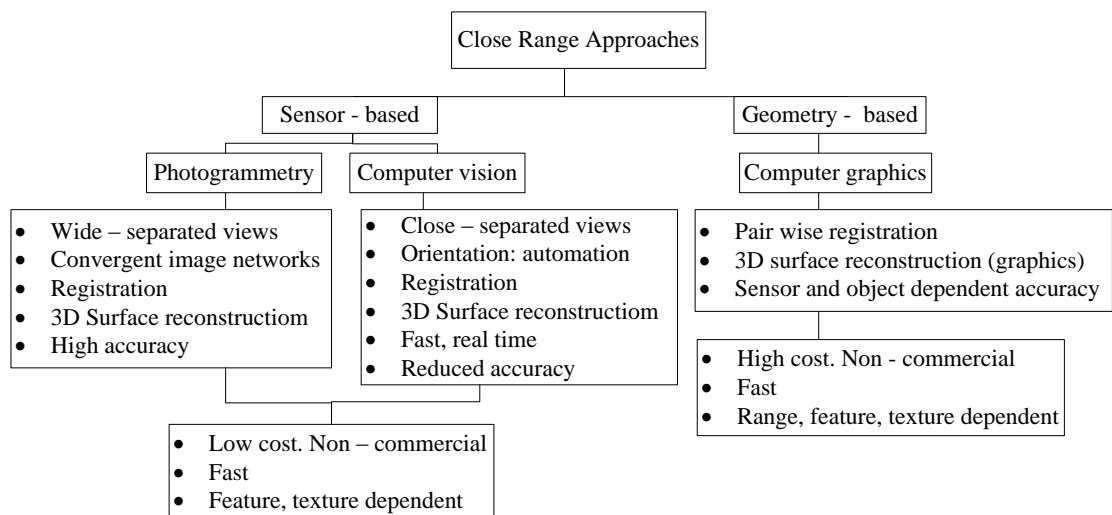
Figure 2.1: Overview of approaches applied in close range object measurement.

Considering the image formation process, the image on a camera sensor is intrinsically 2D and its pattern represents the geometric correspondence between 2D image and 3D object spaces. In other words the two dimensional image is the projection pattern of a perspective-based camera and this forms the basis of the established processing approaches within photogrammetry. Computer vision aims at robust solutions (with regards to the presence of outliers) and it is open in more general cameras and uncalibrated cases following, for example, algebraic approaches. A unified goal of both photogrammetry and computer vision today is to automate solutions in the context of the specific application's requirements. Within this research the main objective is to investigate the affine sensor model originating in principle from the computer vision areas but to develop and implement a multi-view processing framework from the photogrammetric standpoint. In this regard, there exists the necessity to firstly introduce these areas of interest, in particular to revise key aspects linked with some applications and approaches. A good overview of the core connections, as well as differences, of the subject areas of photogrammetry and computer vision can be found in the literature (Foerstner, 2002; Hartley & Mundy, 1993).

2.1.1. Photogrammetry

Introducing photogrammetry the following definition is chosen:

Photogrammetry and Remote Sensing is the art, science, and technology of obtaining reliable information from non-contact imaging and other sensor systems about the Earth and its environment, and other physical objects and processes through recording, measuring, analyzing and representation.

ISPRS Statutes and Bylaws (ST&BL) - Definitions (ISPRS, 2004)

Photogrammetry aims to derive accurate, precise and reliable measurements of the world with prior focus on camera calibration, stereo measurement, 3D object modelling and navigation. It typically focuses on high accuracy levels with classical processing having its roots on geometric approaches. In close range applications there is a particular focus on the creation of efficient measurement systems that are able to

deliver precise measurements. In fact the increased use of Charge-Coupled Device (CCD) and video based cameras have resulted in a broad use of uncalibrated cameras which opened the area to new applications. Main implementation strategies are based upon the fundamental basis of perspective sensor modelling (see section 4.3.2.) and processing involves robust self-calibrated bundle adjustment systems (which is the central topic of Chapter 4). Standard photogrammetric treatment utilizes direct minimal or iterative least squares estimation solutions (see section 4.2.). The essential problem of starting value evaluation (see section 4.3.5.) is recovered on the assumption that the correspondence problem is solved, however automation in this area is limited. Whilst, in many applications, the direct linear transform (DLT) (see section 4.3.3.), the essential matrix (originating from the principles of stereo geometry) and the spatial similarity transform can provide sufficient solutions, bundle adjustments offer highly robust solutions with full statistical analysis which is critical to photogrammetric processing and assessment.

2.1.2. Computer vision

Introducing computer vision the following statement is chosen:

Computer vision has at least two aspects. It is an engineering discipline aiming at working solutions and it is a natural science discipline aiming at understanding the human visual system.

Computer Vision and Remote Sensing - Lessons Learned (Foerstner, 2009)

Computer vision is a field strongly connected with areas of mathematics and computer science and loosely connected with physics. It focuses by concept on the re-invention of silicon-based vision to imitate or even replace biological vision. Geometric computer vision refers to the description of the way the appearance of objects changes when viewed from different viewpoints as a function of the object's shape and the camera's orientations (Hartley & Zisserman, 2004a). Computer vision presents a close relation to the fields of image processing, pattern recognition and scene analysis (Trucco & Verri, 1998). In Gruen (1996) it is stated that computer vision addresses the theory and fundamental algorithms of image and scene analysis

whereas machine vision is linked with the sensor models and the associated systems including hardware issues. Horn (1986) specifies that the central issue of machine vision is to generate a symbolic description (output) from one or more images (input) as illustrated in Figure 2.2.
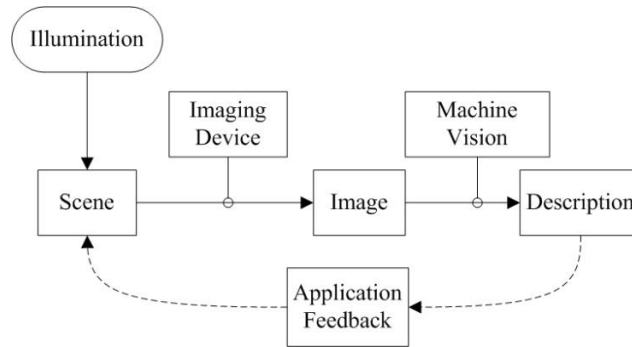
Figure 2.2: A machine vision system's task (source: Horn, 1986).

In practice, 3D computer vision shares one common task with photogrammetry and this is geometry. Its main goal is to compute 3D properties of the world from image sequences, namely recovering the cameras pose and 3D structure of the scene (structure from motion problem). The studied scenes can be geometric in shape and position (static) or include moving parts (dynamic). The typical norm is to utilize uncalibrated cameras particularly where there is no a priori knowledge about the camera (unknown internal camera geometry or interior orientation) or the cameras are equipped with zoom optical systems. For example, in robotics applications, a robot may be moving while zooming and unzooming in the absence of any internal or external camera parameters. The key advantage of employing uncalibrated cameras is that they allow exploitation of projective geometry in full. In computer vision the 3D to 2D mapping is expressed as a linear mapping of homogeneous coordinates[3]. Solutions based on linear systems are employed with the DLT method being the most straightforward solution (see section 4.3.3.). Regarding minimal solutions more robust techniques are usually employed. In the context of computer vision-based strategies, self-calibration starts from uncalibrated cameras aiming on a projective or better an upgraded subsequent Euclidean reconstruction. Algorithmic performance is usually assessed with error analysis.

2.1.3. Comparison of photogrammetry and computer vision

Photogrammetry and computer vision present a common interest in geometry and statistics. Close range photogrammetry usually involves highly controlled imaging situations with known camera calibration (although this is not the case in intelligence applications) and concentrates on a consistent global geometric description of objects. Computer vision on the other hand is more flexible regarding imaging geometries and implementation of uncalibrated cameras but usually achieves lower accuracy levels. However, in consideration of the average projection error this is small compared to the projective error from an incorrect calibration model, for example. Moreover, computer vision is view-centred based on algebraic solutions utilizing homogeneous coordinates. It is fast (real time applications) utilizing linear-based algorithms at the expense of precision. Common application paradigms are robot navigation as well as control of autonomous vehicles. On the contrary photogrammetry is world-centred, usually based on robust, error model propagated solutions aiming at high precision object measurements. Object measurement and performance evaluation of photogrammetric approaches usually require assessment over high order precision reference measurements (with the establishment of benchmarks as an example).

2.2. Applications examples

It is reiterated here (see section 1.2.) that as far as affine multi-view modelling is concerned the developed approach can be extended towards its application in any field that focuses on fine object detail measurement such as industrial metrology, cultural heritage and medicine, typical application areas of close range photogrammetry. Within this scope, such application examples are subsequently given. Particularly these are supported in the context of providing some cases that could potentially offer their measurement data acquired for application of the method. The measurement data reported here have been acquired with passive (e.g. cameras) or active (e.g. laser scanners) systems, as follows.

---

[3] Homogeneous coordinates represent a point in 3D as a four-vector $X=(X,Y,Z,T)^T$ which represents the point $(X/T, Y/T, Z/T)^T$ in non-homogeneous terms and in image space a three-vector $x=(x,y,t)^T$ representing the point $(x/t, y/t)^T$ in the associated non-homogeneous representation.

2.2.1. Spatial measurements with passive systems

In close-range engineering applications it is often the case to select the tools (hardware and software) and engineer the object of interest to make it compatible with the available measuring procedure. Figure 2.3 illustrates such a measurement example, where sparse image data are required for the purpose of digital image recording. In particular, the image on the left illustrates a geometric object with distinctive digitized natural details, whereas the image on the right shows a 3D calibration structure with coded and retro-reflective targets that cover an equivalent measurement volume in 3D.
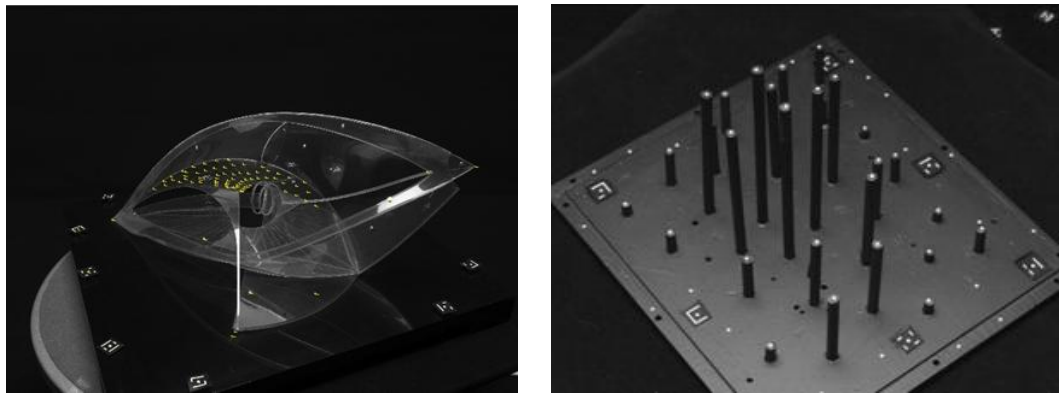
Figure 2.3: 3D object measurements. Measurement object of constructivist sculptor Naum Gabo (left) and calibration object of UCL laboratory (right) (images supplied by Tate Britain, October 2007).

In this example (see Robson et al., 2008) the object of interest is a geometric 3D structure with key characteristics being the object's transparent and plastic material which primarily mean that no natural features (textured areas) can be utilized for point identification and measurement. Yet, the object's physical geometry; that is geometric sections and edges (linear discontinuities in intensities values) could make it ideal for methods that for example utilize linear-based processing (Heuvel, 2003). For derivation of a sparse point-cloud with photogrammetric processing, the Hasselblad H2D and H3D were employed. These high resolution systems offer an analysis of ~39 M pixels (pixel size: 6.8μm and format: 7,216 x 5,412 pixels). Both objects of interest and calibration volume were located on a turntable, illuminated with controlled lights and marked with purpose-built artificial white markers, retro-reflective and coded targets (see section 3.3.1.) in order to achieve high contrast measurement features in image space. As a result the objects were imaged from a systematic range of

viewpoints including subsets of rolled images about their optical axis (imaging range: ~0.5cm). Data processing has been performed with a set of 3D technologies (cameras and laser scanners) and available software tools (VMS 8.0 and Geomagic Studio 7.0) that have been applied for the purpose of 3D measurement. Consequently, the applied multi-image bundle adjustment resulted in sparse point-cloud estimation, providing calibration, orientation and measurement parameter estimation as well as associated precisions (3D target precision= ~15.8μm; relative precision for the image network= ~1:40,000). In this particular selected example the object's lack of natural textures as well as the combination of data acquisition with the data processing chain (based on commercial systems) precluded the followed methodology from delivering fully textured 3D models. The main limitation of this approach was the conjunction of the noisy laser point data (given the object's transparent surface) together with the registration mismatches (due to software inability to handle 3D registration between different coordinate systems).

2.2.2. Spatial measurements with active systems

Quality assessment of small or large industrial products, monitoring and recording of historical monuments or organic objects in medicine and multimedia or realistic 3D models for virtual modelling in archaeology and heritage are some of the relevant applications where active laser sensors are employed. Such spatial measurements focus on the modelling of clouds of points in 3D delivering geometry or intensity values. The modelling procedure is bound to a set of processing steps (registration, modelling, texturing, visualization) that at each stage can present key deficiencies. A problem of prior significance is the 3D to 3D registration mismatches due to limitations in standard or commercial approaches usually based on a 3D similarity transform. In most cases such a transform is implemented with strategies based on the iterative closest point (ICP) algorithm and modifications (Besl & McKay, 1992). Systems utilized for spatial data processing are generally classified as triangulation or time of flight. Triangulation-based systems are applicable in ranges varying between 0.1-1m and their operational principle relies on the projection of a light spot or profile onto an object's surface which is subsequently recorded by one or two CCD cameras. Time of flight systems are applicable in ranges between a few centimetres to several kilometres and they record the range to the object by the estimation of the time that

light needs to travel from the sensor to the object and return (roundtrip). A relatively recent overview of typical active systems can be found in the literature (Blais, 2004).

The case example given here demonstrates a set of data acquired with a Metris handheld CMM[4] laser scanner. Figure 2.4 illustrates an experimental alabaster sculpture scanned with the Metris laser scanner and its derived raw point cloud which has been visualized in Raindrop Geomagic software (Geomagic Studio 7.0, 2006). Another form of data has been acquired and presented in the same figure where a wooden object has been scanned with the same system and subsequently triangulated providing a crude visualization of a local detail of the object (visualized in the Pointstream 3D Image Suite software).

Figure 2.4: Laser scan data. Point cloud of alabaster sculpture (top) and triangulated irregular network of wooden object (bottom) (data acquired in May 2006).

The Metris tool is a triangulation system which hosts a red laser light (wavelength 670nm). It projects a profile line (2D) by means of a CCD camera which forms an angle of 30° to the laser plane. The returned reflection is a function of the object's surface as well as the intensity of the laser line. It is particularly important to note that there exist particular measures that aim on system acceptance and reverification of

---

[4] Metris handheld CMM hosts the LC50 laser scanner. The scanner has a FOV 50mm, a data acquisition range of 100 mm and collects 19,200 points per second.

such optical 3D measuring systems relying on measurements of calibrated artefacts. As an example the three dimensional length measurement error which is evaluated from the difference between measured and calibrated distances between two distinctive points is reported. Specifically, the definition of the quality parameter length in such measuring systems is similar to that in ISO 10 360-2[5], which is now monitored in the guidelines VDI/VDE 2634[6] (see section 4.7.). Problems encountered with such systems are related to the sampling (see section 2.2.4.3.) and resolution of the data points as well as data voids due to surface occlusion problems. In order to generate clean, complete and registered data from scanning systems significant post-processing is required.

2.2.3. Related measurement examples

In the context of covering some case studies related to close range object measurement and applications this section reviews key paradigms that can be found in the literature.

Gruen et al. (2003) and Gruen et al. (2001) follow a data processing strategy (including phototriangulation, image matching for surface model generation, point cloud editing and view-dependent texture matching) applied into particularly difficult situations regarding surface structure and complexity with the ultimate goal being the production of textured 3D models. Whilst the authors show that their method is robust for datasets sensitive to blunders, the employed image matching method presents failure cases in imagery with large scale differences. Thus the authors applied a weighted averaging scheme to reduce the effects of radiometric differences in adjacent images. Pollefeys et al. (2003) and Pollefeys et al. (2004) proposed a very similar 3D recording approach employing an uncalibrated approach (a change in the focal length and remaining interior parameters in the video sequences is allowed) based on pixel matching, bundle adjustments and 3D model texturing. Results prove

---

[5] Established in 1994 the international standard ISO 10 360 'Acceptance and reverification tests for coordinate measuring machines (CMM)' describes test procedures for CMM applications including length measurement, form inspection, use of rotary table etc (ISO10360, 1994).

[6] The VDI / VDE 2634 guideline was drafted by the technical committee "Optical 3D measurement of the Society for measurement and automatic control (GMA) and by the working group 'Close Range Photogrammetry' of the German Association for Photogrammetry and Remote Sensing (DGPF).

that incorrect camera motion and calibration computations may enforce a global bias on the model reconstruction. Yet the method can deal with object complexity which is key, particularly where a high degree of realism is demanded, and can therefore be used as a scale model for generation of reference measurements (derivation of absolute localization and scale). Moreover, Gruen & Acka (2005) treat the registration problem as a surface matching task which in essence is based on a generalization of the least squares matching process allowing for the analysis of the final quality through statistical tools. A recent overview of the registration methods highlighting some key problems regarding mainly the 2D to 3D and 3D to 3D based approaches for object measurement in the close range can be found in (Remondino et al., 2009). Finally, Betham et al. (2009) present a 3D free-form surface measurement system built upon a strategy employing stereo-image matching with focus on the measurement of dynamic surface deformations. The implemented strategy is flexible in that it handles the problem of visibility and occlusions on the knowledge of object shape and position in 3D. Common factors to the success of the employed approach are related to the application's requirements (specified point density and accuracy) and limitations (object size, surface condition and surface characteristics like shadows and occlusions).

2.2.4. Typical problems in close range object measurement

Considering the referenced applications examples (see section 2.2.) it follows that the problems dominating close range object measurement have their source in the physical formation process. The image formation process starts from a light source which emits light energy falling onto an object's surface (irradiance given in Wm^{-2}) and is back reflected to act as an incoming ray (radiance given in $Wm^{-2}sr^{-1}$) through the angular aperture of the optical system hitting the image plane where the camera's photosensitive device is located. Figure 2.5 illustrates the fundamental imaging process drawing a vector of incident light (I), the surface normal vector (N) as well as the vector of which forms the direction of the scene irradiance (R) at an object's surface point (P). For example the radiance of opaque objects (e.g. mirror and carbon black) that do not emit their own energy depends on the strength, position, orientation, type (point or diffuse) of the light sources, and ability of the object surface to reflect energy as well as the local orientation of the surface (with relation to the surface

normal). Radiometry is fundamental to the imaging procedure, however it is regarded as complex and numerically instable. The reader is pointed to further literature for detailed coverage of the critical concepts of radiometry and associated methods (see Sonka et al., 1999c as an example) as here the main approach and hence research focus is driven by geometry.
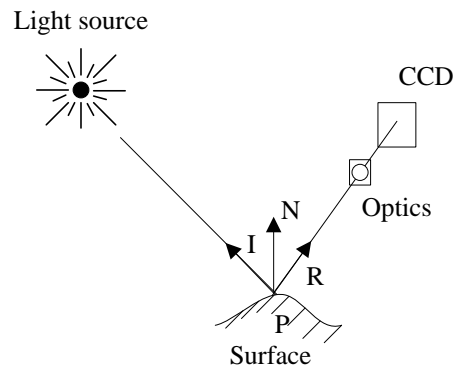
Figure 2.5: Illustration of the fundamental imaging process.

The main practical problems in the imaging process can arise from the system's optical parameters (e.g. lens, focal length, FOV and angular properties in the case of a photogrammetric camera), photometric properties (e.g. illumination, reflectance and physical effects of the sensor) and geometric parameters (e.g. camera projections, 3D pose and geometric internal distortions). Additionally to these, equally fundamental factors include the discrete nature and quantization of the intensity scale (see section 3.1.1. for imaging sensors). Viewing projections, scale and sampling are three concepts that are fundamentally linked with the geometric transformation between 2D and 3D space (see sections 2.2.4.1., 2.2.4.2. and 2.2.4.3.). A key factor is however how these properties interact with real objects that can often present high geometric complexity, occlusions, shadows or surface reflections.

2.2.4.1. Viewing projections

A digital image is a discrete 2D array of numbers (light intensities or distances). When considering the imaging sensor, besides its geometric and radiometric characteristics, a prior factor that identifies data processing and method implementation is projection. In geometric terms projection is the result of the image acquisition process. Following the classification of cameras after Mugnier et al.

(2004), and assuming that the image to object space correspondence is realized through a definite projection center, the following definitions are given (see Figure 2.6). An Euclidean camera, the typical camera met in photogrammetry, is generally identified with a principal point and it can be reduced to a normalized camera which is given with a principal distance c=1 and the rotation matrix being the identity matrix (R=I). An ideal camera is subsequently a camera ascribed with its camera constant considering that the image coordinate system coincides with the principal point. According to the same author, the basic property of an affine or projective coordinate system camera is the invariance of straight lines including optionally a principal point offset, a shear and a non-isotropic differential scaling. Finally the concept of the general camera is introduced as a camera that does not preserve any straight lines and additional parameters may be incorporated for modelling additional sensor distortions.
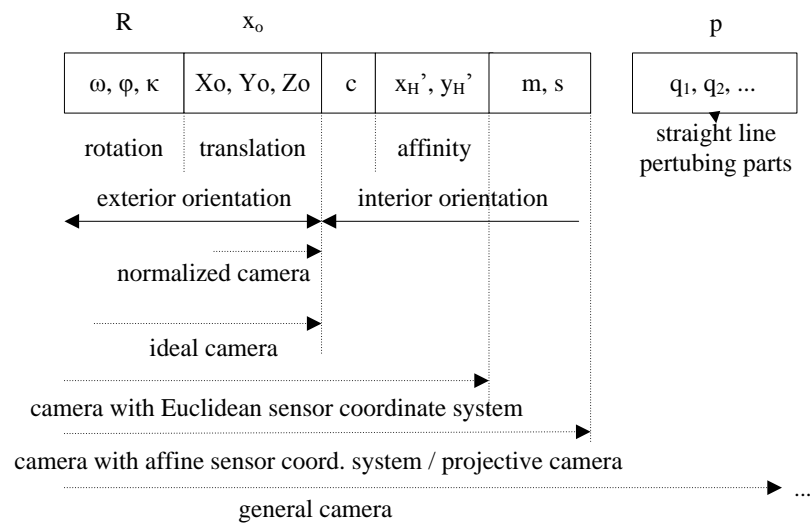
Figure 2.6: Camera models (classification follows Mugnier et al., 2004).

Imaging an object with a perspective based sensor will in general introduce geometric distortions (due to convergence imaging geometries, large object surface slopes and differences in depth) (see section 3.2.2.) which can be significantly large in the close range and will necessitate appropriate sensor modelling. The alternative affine sensor has a projection centre at infinity. It is the generalization of orthographic, scaled orthographic or parallel projection cameras and it realizes the image to object space correspondence minimizing perspective distortions. However, this projection does not connect image and object spaces in a one-to-one relation; the projection is realized through a constant scale factor. In fact a real affine camera can include identical

physical parameters when a natural camera is considered (Hartley & Zisserman, 2004b) (see section 3.4. for a detail reference of camera models).

Figure 2.7 provides an overview image of an alabaster object together with two close-ups of a selected object detail and their associated projections when imaging a 3D volume (drawn as a cube). The test alabaster object has been imaged with one affine and one perspective sensor from an identical range (~175mm). The employed affine system in this case is a Sony DFW-SX900 (pixel size: 4.78μm, format: 1,024x768 pixels) fitted with an optical telecentric system (MVO® TMLTM/0.16x lens) whereas the perspective system is comprised of the same Sony sensor fitted with a Fujinon, f:16mm lens (see sections 3.1.2. and 3.1.3. for technical systems characteristics).
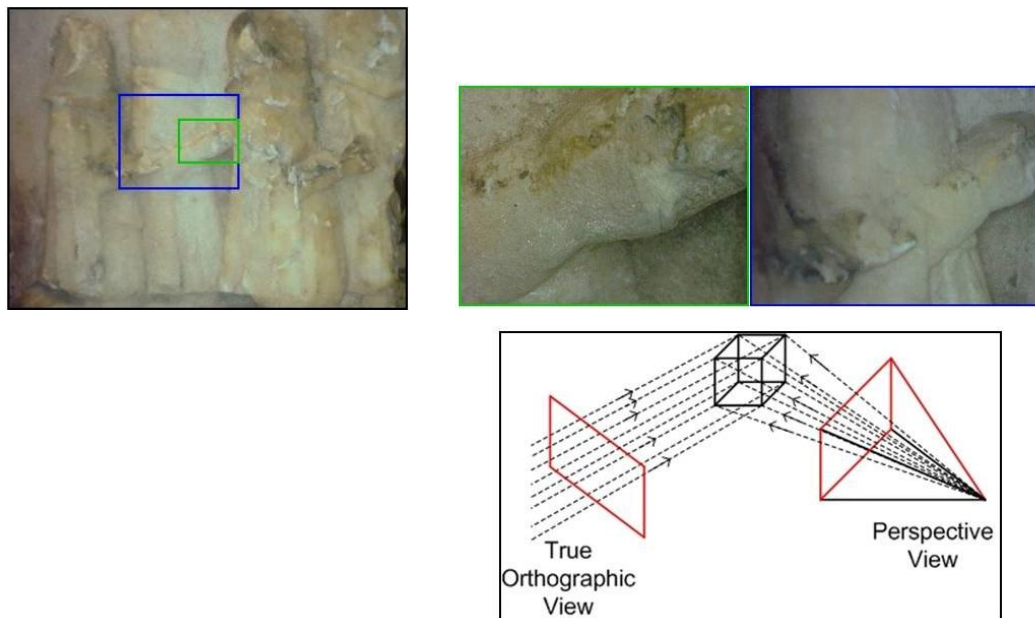
Figure 2.7: Image data of an alabaster sculpture. Affine and perspective image sensors (top left and top right respectively) and corresponding projections (bottom).

Considering the Sony sensor characteristics as well as the need to establish an identical imaging viewpoint, these views illustrate the differences in object coverage and perspective distortion between the two lens types. It is evident that the affine view has a significantly constrained footprint as well as depth of field which would pose it impractical in instances of measuring objects of large sizes exceeding for example the viewing capabilities of the utilized sensor. Particularly the selected detail covers an area of 41.67mm x 31.28mm (blue framed) when imaged with the perspective sensor and it is reduced to 23.81mm x 17.87mm (green framed) when this view is acquired

with the affine sensor (complete object covers a 3D volume of 246mm x 223mm x 40mm). As a trade-off the parallel projection image can be apparently ideal for local object modelling due to its uniform magnification factor. As can be seen, the perspective sensor (besides its geometric distortion particularly at the edges of the image) can not enhance very fine object details (small in size) that could be otherwise modelled with an affine camera. An example of objects presenting very fine details are given in (Remondino & Zhang, 2006), where it has been shown that stereo or multi-image least squares matching approaches can be applied but again these are based on the established perspective sensor geometry.

2.2.4.2. Scale

Scale is key factor in every projection problem particularly recovering the 2D to 3D image space to object space relation. In close range object measurement it is usually the case to present the inability to recover 3D scaled models or link 3D model space with a 3D object space coordinate system. In pure geometric terms scale is in essence the product of the transform acting between different processing spaces. Scale factor calculation for perspective sensors is given by the ratio $1 / k = c / h$ (where: $k=$ scale factor, $c=$ camera constant and $h=$ range measured from the mean object's depth).

To illustrate image scale in the 2D image formation process the typical 2D system is illustrated as follows (see Figure 2.8). Origin of this system is the upper left corner whereas the restitution of the relation between the image space coordinate systems is given in Appendix A.
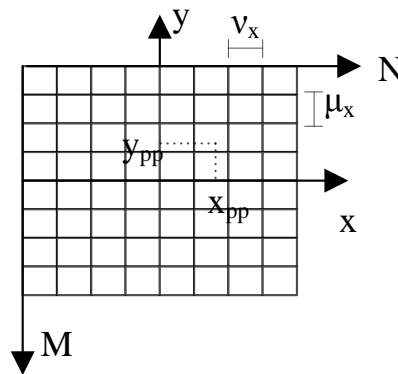
Figure 2.8: Digital image coordinate system.

CCD based cameras (see sections 3.1. and 3.2.) generate a NxM numerical array of pixels (where: N= rows and M= columns). The location of a point in image plane can differ when this is measured with CCD elements (image pixels). More precisely it is noted that n/N and m/M (where: n= horizontal light sensitive photosensors and m= vertical light sensitive photosensors, in a nxm CCD rectangular grid) are not the only parameters that are held responsible for a different scaling of the image with respect to the CCD array in both horizontal and vertical directions. The same effect is realized due to the ratio of the horizontal to vertical sizes of the CCD array (Trucco & Verri, 1998). This parameter is the aspect ratio or affinity which is usually modelled as the in plane distortions (see section 3.2.2.3.).

Within a system, scale recovery relates to the design of the employed algorithms to recover scale. For example 2D approaches are vision-based and 3D approaches are graphics-based (see section 2.1.). In image analysis scale is highly important especially in methods based on edge detection (see for example the typical Canny edge detector) or multiple scale description (Sonka et al., 1999a) specifically in strategies that make use of the reduced resolution datasets (image pyramids). Operations on the image scale space (Koenderink, 1984) apply Gaussian filters with varying standard deviations which can be very useful to extract features by isolating them for example at lower resolutions (processing at individual description levels) and subsequently locate then at higher layers (Agouris et al., 2004). Implementing image pyramids in image matching has already been discussed in Baltsavias (1991). To place scale in context particularly implementation on a fine to coarse analysis basis, it is pointed that in the direction of decreasing value of standard deviation large scale events are localized (Sonka et al., 1999a). Shape invariants (e.g. cross ratio) and invariant descriptors are beyond the scope of this text. The reader can find more information on these topics in the computer vision literature (e.g. see Sonka et al., 1999b for an overview).

2.2.4.3. Sampling

In the spatial quantization process of image formation (see section 3.2.1.1.) whilst uniform aspect ratio is usually assumed the resultant pixels will in general be rectangular. The sampling theorem specifies the highest spatial frequencies v_c

($v_c = 1/2d$, where: d= distance between adjacent CCD elements) whereas the diffraction theory states that spatial frequencies v_c' exceeding $v_c' = \alpha/\lambda f$ (where: α= linear size of the angular aperture, λ= wavelength of light, f= focal length) are filtered out not contributing to the spatial spectrum of the image. With spatial frequencies v_c nearly an order of magnitude less than v_c', aliasing effects can be present in the case where the imaged pattern contains spatial frequencies exceeding v_c (Trucco & Verri, 1998). In detail the concepts of image formation and sensor characteristics will be given in the subsequent chapter.

2.3. Summary

In summary this chapter introduces the reader to non-contact object measurement in the close range. First the areas of interest, in particular photogrammetry and computer vision, with some key principles, methods and solutions have been introduced. Subsequently, some application examples are reported, firstly linked with case studies from the current literature, as well as some ad-hoc to this work are reported at a following stage. These pose the context of the developed method in relation to object measurement applications. In addition typical problems that occur in relevant situations (characterized by the viewing projections, scale and sampling) are illustrated.

3. Digital close range image formation

Implementing image measurements on digital images requires that the background to the imaging process is firstly introduced. This chapter provides a review of the main concepts, principles and methods that were utilized as a basis of this work, reviewing or pointing where necessary to the related literature.

Although there exist different sensors modalities (photosensitive materials with different spectral sensitivities) an image in the context of this text is generated by conventional optical means sensed in the visible spectrum (λ= 400-700 nm). The camera systems (see section 3.1.) are decomposed with the description of CCD sensor technology, close range cameras and machine vision optics. The output of the image formation process, the digital image (see section 3.2.) is inherently linked with the employed sensor. Its geometry can be reduced as a result of the internal geometry of the optical system in combination with physical instabilities within the camera body. The resultant digital image is then the fundamental source where measurement methods (see section 3.3.) are utilized for data generation and initialization applied to artificially signalized point-based data structures. The chapter closes with two image projection models (see section 3.4.): the projective and the affine cameras; the investigation of which introduces the central topic of this work.

3.1. Digital camera systems

Developments in sensor technology have resulted in electronic imaging systems that utilize CCD and more recently Complementary Metal Oxide Semiconductor (CMOS) -based sensors in combination with powerful local Central Processing Units (CPUs) to increase performance, particularly flexibility at a lower cost. As a result, digital camera systems open new applications in the domain of close range imaging. Figure 3.1 illustrates the basic functionality of an imaging system outlining its main units. The generated analogue image collected at the sensor is converted into digital form by an analogue to digital (A/D) conversion taking place within the camera in which case a digital interface is utilized (RS422, camera link, USB, firewire protocol).
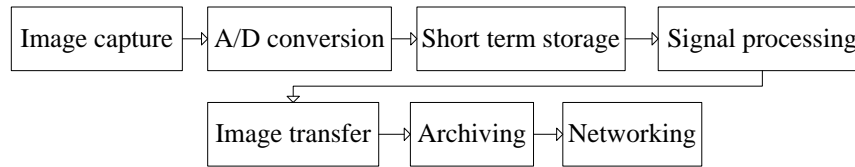
Figure 3.1: Functionality of an electronic imaging system (source of schematic structure: Schenk, 1999).

3.1.1. Imaging sensors

CCD technology was originally developed in 1970 as a memory device by Boyle with digital imaging systems based on CCD sensors effectively utilized in the eighties when they replaced vidicon tube cameras (Smith, 2009). Since then, CCD electronic cameras have been routinely utilized in camcorders, electronic still cameras to the more recent machine vision and scientific specific systems.

3.1.1.1. CCD Principle of operation

Following the comprehensive studies of Lenz (1989), Luhmann et. al. (2006), Robson & Kyle (2004), Schenk (1999) and Shortis & Beyer (1996) the operational principle of CCDs is described here. Solid state cameras utilize a sensor composed of photodiodes (with a positive region beneath the surface layer) or Metal Oxide Semiconductor (MOS) capacitors (with a metal or polysilicon electrode layer). The building block of the sensor is the semiconductor substrate which is silicon including a silicon dioxide insulator layer at its top surface (see Figure 3.2). Light photons with greater energy than the band gap energy of the semiconductor can be absorbed below the sensor's surface (depletion region) generating an electron-hole pair at each sensor element (photosensitive detector). The electrons are attracted by the positive charge and accumulated in the depletion region while the mobile holes move towards the electrical ground. The charge accumulates at opposite sides of the insulator and the actual charge is proportional to the number of absorbed photons under the electrode. In the case of lower energy photons (that exceed the band gap with $\lambda = 1.1\mu m$ wavelength) these may penetrate the depletion region and absorbed outside resulting in a potential that the electron-hole pair may recombine before reaches the depletion layer noting that not every photon generates an electron that is accumulated at the capacitor side. Hence, a CCD array requires an increased number of capacitors.
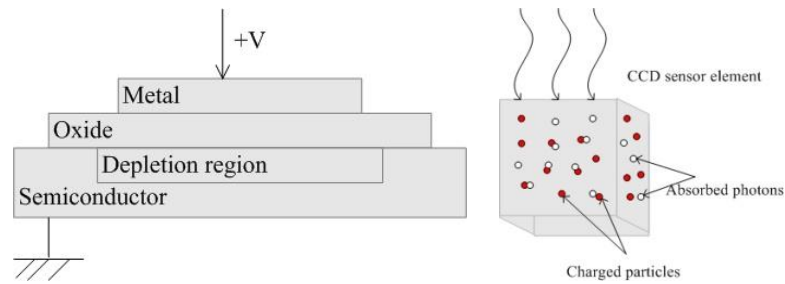
Figure 3.2: The fundamental MOS structure (left) and a CCD sensor element (right).

Outside the spectral range of λ= 400-1,100 nm silicon presents an opaque behaviour in ultraviolet light and is transparent to infrared. The intrinsic absorption on the material is limited with optimal energy to liberate electrons and therefore detect radiation in the visible and near-infrared. Dark-current is the result in background noise due to thermal effects. Longer wavelength radiation penetrates at deeper levels allowing impurities to be introduced within the sensor. The generated number of photons is linearly related to the number of electron-hole pairs and hence to the charge level. Each sensor type has a finite potential well capacity with the result that charge can overflow into the neighbour sensor elements causing blooming (see section 3.1.1.2.). This is stopped by isolation of the sensor rows by electrodes, oxide steps or channel stops. Sensor elements are typically arranged as one dimensional (1D) or 2D arrays. Line sensors connect the active sensor element to a serial read-out register to output the generated charge. In contrast, bilinear arranged CCD lines can be coupled with two read out registers. The most common matrix-based sensors transfer principles are the frame, full frame, interline and time delay integration methods (see Figure 3.3).
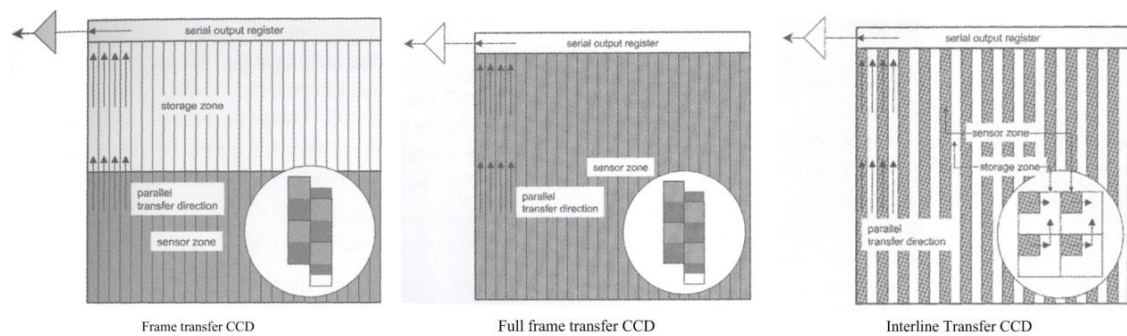
Figure 3.3: CCD sensors architecture (source: Luhmann et al., 2006).

- Frame transfer sensors (FT): Are composed of an active array that accumulates charges, a storage zone and a horizontal read out register. Charges are moved vertically from the sensor to the storage area which is read out serially per line to the register. This architecture often utilizes a mechanical shutter that covers the sensor during readout to reduce smear (continuous charge integration during read out).

- Full frame transfer sensors (FFT): Are a variation of the FT sensors where the sensor consists of an active array (imaging area) and a serial read out register. FFT sensors may present significant vertical smearing due to the need for long transfer lines.

- Interline transfer sensors (IT): Comprise one column of active detectors with a column of transfer register. The accumulated charges in the sensor columns are shifted into the transfer register columns and then read out serially (horizontal register). The structure of IT and the discrete nature of sensor elements prevent interlacing. Additionally aliasing (high frequency patterns imaged at lower frequencies) is reduced by increasing the fill factor of each pixel (for example by utilization of microlenses).

- Time delay and integration (TDI): This sensor follows the forward motion compensation logic where the sensor allows electronic linear motion compensation. The charge is transferred during an integration interval where charge is accumulated continuously during the next interval. On completion sequential image read out is performed.

The read out method of a frame array based method can be interlaced as traditionally used in television systems where the frame consists of odd and even fields where each of these correspond to odd and even lines. Contrary to the interlaced method progressive scan sensors record the whole frame at one instant. Progressive scan sensors present higher vertical resolution in the absence of interlace artefacts.

3.1.1.2. CCD main characteristics

The basic attributes that characterize image sensor performance are related to image quality (see 3.2.1.2.), and its reduction mainly due to spurious signals with various effects. For example the more dominant effects can be the dark current (thermal

generation of electrons generated on the CCD both during illumination and read out), blooming (intense light falling causes the generating photons to spill over neighbouring capacitors), smear (given by the ratio of the change in brightness above or below a bright area which covers the 10% of the sensor extent in the column direction) and so forth. The most significant parameters that characterize a sensor are summarized as follows (Edmund, 2006; Shortis & Beyer, 1996):

- Quantum efficiency: Is the ratio of the electron flux over the incident photon flux and it is characterized by the spectral sensitivity. It depends on the energy of the incident photon (dominated by the wavelength λ), the material (usually silicon) and method utilized to collect the freed electrons.

- Resolution: Expressed by the Modulation Transfer Function (MTF) (given by the contrast transfer against the resolving power, expressing the degree of contrast degradation with spatial frequency) or the Point Spread Function (PSF) (showing the dispersion of an imaged point of light through an imaging lens). The limiting resolution of a system can be identified by imaging test targets of varying or continuous series of frequencies.

- Fill factor: Is the ratio of the obtainable to the theoretical power or the ratio of light sensitive area to the total pixel size.

- Spectral response: Directed by the quantum efficiency of silicon (semiconductor material) and represented as a step function in the ideal case. The spectral sensitivity of a CCD sensor can be extended by back side thinning and illumination (astronomic applications).

- Linearity: Expresses the ratio of the maximum departure from linearity over the full range of signal level to the maximum signal level.

- Signal to noise ratio (SNR): Stated as the ratio of the signal and its noise. It is expressed in decibels as $SNR_{dB} = 20 \log_{10} (s/\sigma_s)$ (where s is the signal amplitude and σ_s is the standard deviation of the signal expressing the noise caused by photon shot noise, dark current and circuit noise). High SNR is an indicative measure of 'good' image quality.

- Dynamic range: Defined as the ratio between the peak signal level and the system noise level. Given their large sensor element capacity, FT CCDs with

large sensor elements will have greater radiometric sensitivity (increased SNR and dynamic range).

3.1.1.3. CMOS sensor overview

An alternative sensor technology is the CMOS sensor. Contrary to the CCD sensor, where each pixel's charge packet is transferred sequentially to an output structure for charge to voltage conversion, buffering and readout, in a CMOS imager the charge to voltage conversion is implemented within each pixel. This is the key difference which differentiates both technologies with relation to the sensor architecture, its advantages and shortcomings. Both CCD and CMOS sensors are equally reliable in consumer and industrial applications. The general functionality of CMOS image sensors relies on all circuit functions being positioned on a single integrated circuit chip. Timing, signal processing A/D conversion interface are placed on the imager chip. As a consequence, CMOS-based sensor systems have a reduced size. Main features constitute their inherent anti-blooming ability, potential to readout Regions Of Interest (ROI) (windowing), increased speed, operation with a single bias voltage and clock level and less power consumption (Blanc, 2001; Butler, 2003; Litwiller, 2001; Litwiller 2005; Seitz et al., 1995). In contrast, CCD technology is characterized by high quantum efficiency, low dark current, reduced pixel size, reduced operating voltages (power dissipation) and improved signal handling with significant improvements regarding performance, power consumption and sensor sizes.

The cost of CMOS imagers (silicon wafer fabrication material) can be considered with relation to integration, adaptability and flexibility. The general acceptance that CMOS imagers perform better is not always the case at high speeds and cost needs always to refer to the application's purpose. With CCD imagers dominating in general purpose applications as well as in high performance applications (scientific, industrial, medical, security and aerospace), CMOS are regarded as consumer specific devices. CCDs can also be adjusted with relation to their functionalities (readout, speed, dynamic range, digitizing depth and so forth) to fit the application's requirements. CMOS-based sensors can be considered as less expensive than CCDs when judged as systems regarding circuit functions (timing, biasing, analog signal processing,

interface and feedback circuitry) but not at a component level for the image sensor function itself.

3.1.1.4. Colour methods

Acquisition methods for colour images are generally classified as spatial multiplex, parallel acquisition, temporal multiplex as well as methods where the sensor is of true colour type. The colour cameras utilized within this work are of the first listed type; hence greater emphasis is given (see section 3.1.2.1. for camera's systems characteristics).

- Spatial multiplex system is a single exposure system. It utilizes strip or mosaic colour filter arrays (CFA) with most common the Bayer mosaic mask, typically arranged as GRGB or RGGB. Area sensors are based on pixel interpolation where the digital value of the colour band sensed by the photosite is assigned directly from the received signal, the other two colours required to form the red green blue (RGB) images are derived from the surrounding pixels. However, subsequent demosaicing and resampling of pixel intensities can reduce image quality.

- Parallel acquisition system is a single exposure system. It is based on a colour filtered prism arrangement or beam splitter that simultaneously projects incident light onto three sensors with each sensor registering intensity of one colour channel. The generated analogue signals are digitized in parallel. Based on the beam splitting principle, these systems are freed from the pixel interpolation method at the cost of increased complexity and physical size.

- Temporal multiplex system is a three exposure system. Colour is recorded employing a single sensor introducing a red, green or blue filter into the optical system. Temporal sampling of the signal generates the digital RGB equivalent of the three colour bands.

- True colour sensor is a single exposure, single chip system. Foveon X3 sensor is a CMOS high resolution colour sensor (Foveon, 2009). It consists of three stacked layers each of which has a different spectral sensitivity curve noting that different wavelengths of light penetrate silicon at different levels. The processed signals are registered to generate the RGB colour.

3.1.2. Close range cameras

The selected camera system is strongly related to the application's requirements. The Kodak DCS (Digital Camera System) series can be considered as the first high image quality single lens reflex (SLR) type cameras employing an area array CCD sensor since the early nineties (Graham, 1998). Following the technological developments, small format digital single lens reflex cameras (DSLR) with matrix sensors and an integrated to the camera's body storage device are commonly deployed. Close range systems can be video (resolution: 780 x 580 - 1,900 x 1,100 pixels), high resolution (resolution: 1,000 x 1,000 - 4,000 x 4,000 pixels) or scanning (resolution: 3,000 x 3,000 pixels - 20,000 x 20,000 pixels) cameras (Luhmann et al., 2006).

3.1.2.1. Systems characteristics

This section reviews the systems characteristics of three CCD-based digital camera systems that were utilized within this research work as illustrated in the following Table 3.1. In particular the listed camera systems were employed for the generation of datasets utilized for initialization and measurement (Nikon D100 and Kodak Megaplus ES1.0 camera systems) as well as for method testing and analysis (Kodak Megaplus ES1.0 and Sony DFW-SX900 camera systems).

	Camera system		
	Nikon D100	Kodak Megaplus ES1.0	Sony DFW-SX900
Sensor	Nikon DX CCD	KAI-1010M CCD	½ CCD
Transfer, readout	IT	IT, progressive scan	IT, progressive scan
Colour filters	Primary GRGB [Bayer mosaic]	Monochrome	Colour [Bayer mosaic]
Effective pixels	3,008 x 2,000	1,008 x 1,018	1,280 x 960
Unit cell size (μm)	7.8 x 7.8	9.0 x 9.0	4.65 x 4.65
Bit depth	12	8 / 10	24
Frame rate (fps) (single channel / dual channel)	3.06 / 5.09	15 / 30	3.75 / 7.5
Interface	USB 1.1	RS-422	IEEE 1394 - 1995

Table 3.1: Synoptic specifications for CCD-based close range camera systems.

The Nikon D100 DSLR camera (introduced in 2002) is appropriate for Nikon F mount lenses. It is equipped with a 23.7mm x 15.6mm 12-bit RGB CCD sensor rendering 6.1 million effective pixels. Data transfer is based on the interline method and the GRGB Bayer mosaic filter (see 3.1.1.4.) is utilized for generation of colour imagery (Nikon 2002; Nikon, 2009). As an example of the rapid development in DSLR cameras technology, the Nikon D3X 24.5 megapixel camera distributed by Nikon in late 2008 is given (see Figure 3.4 left). This camera system is the successor of the D3 and D700. Based on a CMOS architecture (sensor size: 35.9 x 24.0 mm, FX format, continuous shooting: 5.5 frames per second, lens type: Nikon F bayonet mount) its high image quality and good dynamic range render sufficient colour accuracy. The second listed camera is the Kodak Megaplus ES1.0 (predecessor of Redlake imaging) which is a C mount monochrome video camera. The solid state CCD sensor is an interline progressive scan sensor (see 3.1.1.2.) (analysis: 1 M pixel, active image area: 9.1 x 9.2 mm) (Kodak, 1996). The camera is utilizing an RS-422 (Recommended Standard-422) twisted pair bus interface standard for data transfer to the host computer. This camera system belongs to the class of high resolution Megaplus cameras distributed by Kodak in the late nineties. An example is the Kodak Megaplus 1.6i (sensor: Kodak KAF-1600, solid state FFT CCD) (see Figure 3.4 right) which is constructed in a rugged, compact design (Robson & Kyle, 2004). Its 1,024 gray levels (readout method: progressive scan, analysis: 1.6 M pixels, continuous shooting: 5.5 fps) and minimal dark current ensure such dynamic range and sensitivity that in combination with its square pixels cover the demanding requirements in industrial and machine vision applications, particularly when considering the time they were manufactured.

Figure 3.4: Close range cameras. Nikon D3X (left) and Kodak Megaplus 1.6i (right).

The third camera system is a Sony DFW-SX900 which is a C mount digital video colour camera utilizing a ½ type interline progressive scan CCD. Nominal operational values for this system are given (pixel size: 4.65µm at full resolution: 1,280 x 960

pixels); yet in practice data were captured at a resolution of 1,024 x 768 pixels with this sensor (see section 6.2.2.2.). Colour imagery is generated based on the Bayer mosaic pattern and there is the ability to adjust the gain of the video signal amplifier. Sony's primary colour filter CCD for colour reproduction and its square pixel CCD eliminates the need for aspect ratio conversion in the image sensor (Sony, 2001). Figure 3.5 illustrates the typical spectral characteristics curves of the Sony's CCD image sensor in the visible spectrum. Although high speed data transfer rate can be realized (IEEE 1394 serial bus interface standard); for the purpose of data generation single frame images were obtained. In all three utilized camera systems the raw acquired image data were subsequently saved in the camera's file format as defined within the accompanied software.
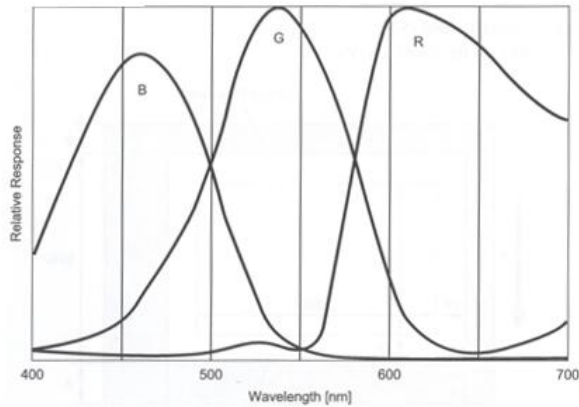
Figure 3.5: Spectral characteristics curves - Sony DFW-SX900 CCD (source: Sony, 2001).

3.1.3. Digital cameras optics

Optical lenses are a key unit of a camera system. The standard approach is to utilize off-the-shelf optics which in most machine vision applications can suffice over their customized optics counterparts. The main considerations regarding the selection of an optical system are a function of the levels of accuracy and reliability to be directed by the application's purpose. The key factors are related to the FOV (given as a range for zoom lenses, angular magnification for lenses working over a range of distances or as a fixed value for fixed focal length lenses), the primary magnification (identified as the ratio of sensor's size over the lens's FOV) and the sensor's format (specified as a maximum format or diagonal that can be covered by the selected lens) which is typically identified for the standard 4/3 aspect ratio as 1/4'', 1/3'', 1/2'', 2/3'' and 1''

(where '': denotes an inch). Additional parameters are the imaging range (measured from the front lens's diameter), resolution and reproduction of contrast levels and depth of field (DOF) (specified with a single value from the diffraction limit).

3.1.3.1. Perspective projection lenses

Most close range optical systems employed utilize standard central perspective projection lenses (see Figure 3.6). Their type can vary according to the application's purpose and selected equivalent camera (see 3.1.2.1.). Besides the imaging geometry the light source can insert additional geometric distortions (see section 3.2.2.) to the image formation process. In this work external electronic ring flash lights, light emitting diode (LED) rings and fluorescent high frequency ring illumination (green and white) sources were selected to illuminate the scene to be measured. Conventional optical systems were selected according to their nominal properties to initialize image networks and generate reference measurement data. Table 3.2 summarizes the nominal specifications of the utilized camera systems.
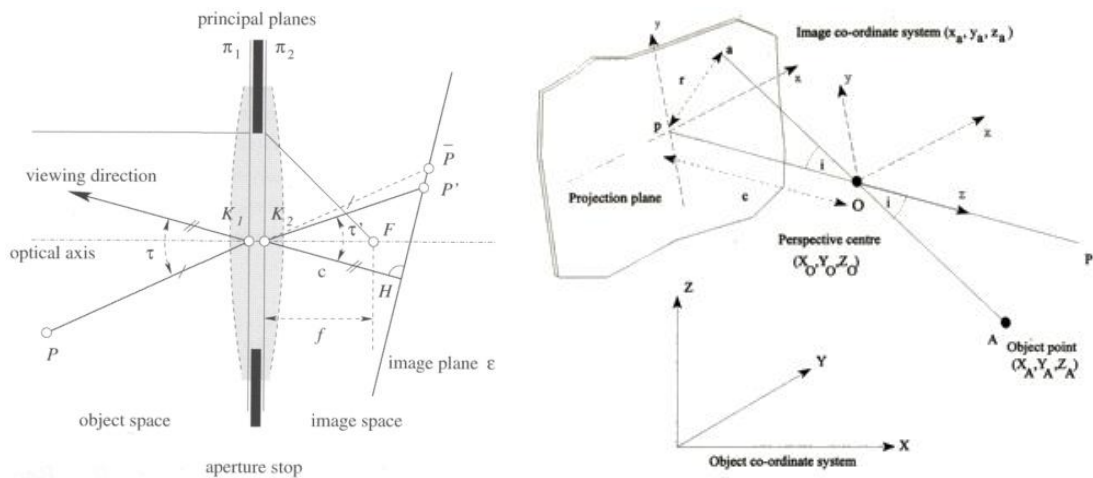
Figure 3.6: Central projection model. Optical imaging geometry (source: Mugnier et al., 2004) (left) and principle of collinearity (source: Cooper & Robson, 1996) (right).

Model	Focal	F#	AOV	Focus	Mount
Nikkor AF	28mm	2.8-22	N/A	∞-0.25m	F
Fujinon TV CF12.5HA-l	12.5mm	1.4-22	1 inch 54°13'x42^{o}01' 2/3 inches 38°47'x29^{o}35' 1/2 inches 28°43'x21^{o}44'	∞-0.1m [A] 16.07mm [B] -101mm [C]	C
Kern Switar H16RX No. 1066951	10mm	1.6-22	N/A	∞-8'' [A]	C

Table 3.2: Nominal specifications of employed optical systems. Table notation: focal = focal length, F#= aperture range, AOV= angle of view, Focus [A]= front lens diameter, [B]= back focal distance in air and [C]= exit pupil position.

The NikkorAF is a 28mm (fixed focal length) wide angle lens. This early lens model is a typical F mount lens featuring a bayonet type suited for Nikon's 35mm SLR cameras. The Fujinon TV lens is a 12.5mm lens and the Kern Switar is a 10mm similar C mount, fixed focal length lens (no zoom or autofocus) fitted to 1 inch sensors. The lens models utilized here are designed following the retrofocus[7] principle, which according to Ray (1988) can result in significant geometric distortions when compared to the more symmetric optically short focus lenses.

3.1.3.2. Affine projection lenses

The non-conventional optical configuration which preserves magnification within the DOF is referred to as telecentric and it is usually characterized by its telecentricity[8]. In the literature three different types appear; namely image-based, object-based and double-sided (bi-telecentric) optics (Lenhardt & Kreuznach, 2006). Whilst most commercial lenses are object-based, conventional lenses can be converted to telecentric by the insertion of an additional aperture (Watanabe & Nayar, 1997). Single-sided telecentric lenses maintain their properties according to their fabrication. As an example, in the object-sided case the entrance pupil is located at infinity; hence the principal rays enter into the lens in parallel to the optical axis (Konrath & Schroder, 2002). The image formation is realized under parallel projection. On the

---

[7] Retrofocus lenses: Resolve the short focus limitation characterized by the small separation between the vertex of the rear element of the lens and the focal plane. They can present problems; especially in the case of 35mm SLR type bayonet mounts cameras (Ray, 1988).

[8] Telecentricity: Determines the amount of the magnification variation within the lens's DOF at the specified imaging range (Melles Griot, 2006).

contrary, in the image-sided case the exit pupil is placed at infinity, therefore the perspective model is maintained whilst the magnification remains constant with relation to the placement of the image detector behind the lens. Double-sided telecentric lenses can be thought of as a combination of two single-sided lenses. Figure 3.7 illustrates an example of two commercial telecentric optical configurations.
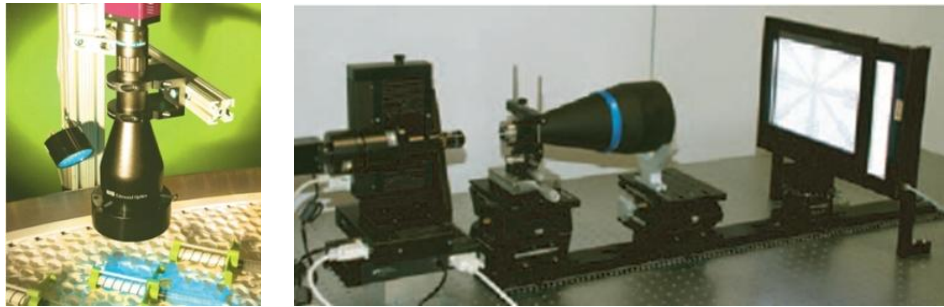
Figure 3.7: Telecentric optical systems. Techspec silver series (source: Edmund, 2009) (left) and telecetric system for contrast transfer function (CTF) measurements (source: Optoengineering, 2009) (right).

Table 3.3 lists the main differences between conventional, perspective and non-conventional, telecentric imaging configurations.

	Optical system	
	Perspective	Telecentric
Projection centre	Finite	At infinity
Scale	Variable	Constant
Distortion	Geometric (extended model)	Insignificant (radial)
FOV	Wide (~f)	Narrow
Processing	Bundle adjustment (PG)	Geometric SVD / affine (CV)
Applications	Measurement, registration, texturing, etc.	Metrology, inspection

Table 3.3: Comparison of perspective and telecentric optical systems. Table notation: PG= photogrammetry, SVD= singular value decomposition and CV= computer vision.

Telecentric optical systems are advantageous in that they offer minimal perspective distortions, constant magnification within the image format, image quality enhancement and even illumination. However they do not increase DOF and accuracy or correct illumination problems which are inherent to the imaging process. In practice, object sided telecentric lenses can be utilized in measurements where the image space to object space are correlated up to a scale factor. Specifically, 2D checking, image quality improvement are a few of the applied on-line or off-line metrological examples. Additionally, telecentric lenses can enhance the performance

of centroiding algorithms for target identification and measurement. Image sided telecentric optics can be effective in the location of three image planes in colour sensors as they result in a uniform image space illumination. The reader is directed in the literature for further analysis related to some examples of utilization of telecentric optical systems (Ahn et al., 1997; Fournel et al., 2003; Watanabe & Nayar, 1997). Within this work, a telecentric optical system was selected to approximate the affine sensor model, the investigation and analysis of which is discussed throughout this thesis particularly for method testing. The employed system was an MVO® TMLTM/0.16x (supplied by Edmund Optics in March 2006) and can be now found as TECHSPEC® SILVER series telecentric lenses in Edmund (2009). The specifications (see Table 3.4) ensure that the telecentric lens is fitted for a maximum CCD sensor format of ½ inch (6.4 x 4.8mm) and that according to the nominal values it realizes parallel projection imagery under a constant magnification of 0.16x at an imaging range of 175mm, allowing variations within a volume of DOF: ±19.7mm. It presents a radial pincushion distortion at the order of 0.3% with a 40mm FOV which is limited by the 65mm front lens diameter (for example a field of 1000 pixels will image a point 3 pixels far from the optical axis). Figure 3.8 illustrates the system which is comprised of an MVO® TMLTM/0.16x telecentric lens mounted on a progressive scan monochrome Kodak Megaplus ES1.0 camera. The illustrated telecentric lens is characterized by the drawn dimensions (where A: maximum outer diameter= 65mm, B: mounting diameter= 30mm, C: length= 191mm, D: mounting length= 50mm, E: mounting offset= 43mm and F: filter size= M62x 0.75mm).

Magn.	Rg.	Res. (image@ F10)	Telec.	Dist. (Max.)	DOF (10% @20lp/mm)	Apert. (f/#)
0.16x	175mm	>40% @40 lp/mm	<0.1°	<0.3%	±19.7mm @ F10	F6-closed

Table 3.4: Specifications for TECHSPEC® SILVER telecentric lens series. Table notation: Magn.= magnification factor, Rg.= imaging range, Res.= resolution, Telec.= telecentricity, Dist.= distortion, DOF= depth of field, Apert.= aperture (source: Edmund, 2009).
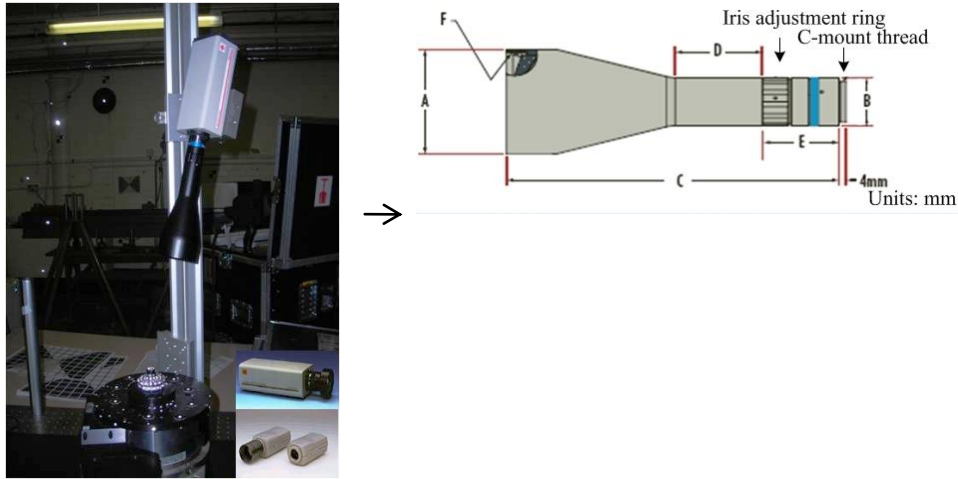
Figure 3.8: Affine machine vision system. Kodak Megaplus ES1.0 camera with MVO®
TMLTM/0.16x telecentric (left) and schematic view of telecentric lens (right).

3.2. Digital image characteristics

The imaging process can degrade the image quality of the generated imagery in both
radiometric and geometric terms. The digital image characteristics are therefore
critical to the measurement process and in this context these will be outlined with
relation to (a) the digital image properties, (b) internal geometric distortions and (c)
geometric camera stability.

3.2.1. Digital image properties

This section is concerned with the digital image and its properties. These are
discussed in the context of two considerations: (a) the digital image formation process
and (b) the quality of digital images, as follows.

3.2.1.1. Digital image formation

According to section 3.1.1.1. a sensor forms an image by the collected electrons when
photons hit a photo-sensitive material. The developed analogue in the sensor is
subsequently quantized (A/D conversion) for digital reading and processing. The
continuous image function $g(x,y)$ in a 2D plane, where x, y are its spatial variables
and the function amplitude is the density, is generated by sampling (each continuous
sample is assigned an integer value) the spatial variables and quantizing (dividing the
initial continuous range into k intervals resulting in $k=2^b$ brightness levels for b bits
per gray level) the gray levels (amplitude). Improving sampling and quantization

levels achieves a closest approximation of the continuous image function. The data structure representing a digital image is a matrix and the corresponding sampling point is a pixel usually represented as either square or rectangular cells (in which case resulting in different horizontal and vertical resolutions). A pixel collects information about the brightness at a specific location in the image. The highest spatial frequency that can be preserved without loss of information is defined by Shannon's sampling theorem. According to this, the Nyquist frequency ($f_N = 1/2\Delta x_S$) states that the smallest pixel size should be less than half of the continuous function's highest frequency. Additionally, a practical limitation is that data volume and processing times dictate suitable sampling rates to achieve the desirable resolution and accuracy (Schenk, 1999). Measurement and processing of digital images are commonly implemented with operations that take into account basic measures such as brightness, contrast and histogram.

3.2.1.2. Digital image quality

Digital image quality is a measure of the degradation which can happen during the image formation stages of capture, transmission, processing or representation. The degree of degradation is assessed by measures that compare a given image against a reference image based on mean or absolute differences or correlation methods for example (Sonka et al., 1999d). Image quality is affected by factors such as sensor dynamic range, contrast, sharpness, geometric aberrations and equally significantly photographic effects like vignetting[9] and exposure settings. Image quality is directly related to the utilized sensor, hence its properties (quantum efficiency, resolution, SNR) are the main sources affecting the quality of the measured image (see section 3.1.1.2.).

3.2.2. Internal geometric distortions

In the geometric context, image formation is the process where the bundle of rays travel through an optical lens to reach the image plane. This physical reality is modelled with what it is termed in photogrammetry interior and exterior orientations

---

[9] Vignetting: Is the effect where pixels closer to the image frame borders appear darker due to the property of optical rays with large span - off angle from the optical axis to present increased attenuation (Sonka et al., 1999a).

(IO and EO) of the camera-lens system. Yet, in the real imaging case the formed image can be 'reduced' from its theoretically exact model due to aberrations that may degrade the image quality or geometric aberrations that can alter the position of the image. These perturbations are a factor of the nature of the camera system and it is generally accepted that in CCD-based systems symmetric radial distortion, decentering distortion, focal plane unflatness (chip bowing or crinkling) and in plane distortions (electronic effects like line jitter) are common (Fraser, 2001).

3.2.2.1. Radial distortion

Radial lens distortion is the result of the Seidel aberrations and it is given as an odd-powered polynomial (see equation (3.1)). Whilst in most instances the third order term will suffice, in the case of demanding accuracies or wide angled lenses higher order terms are needed to model in full a lens's potential distortion. Judgement of the inclusion of these parameters into the calibration model is a function of their statistical significance and performance of correlation checks on the implemented parameters (see section 4.3.4.).

$$dR = k_3 r^3 + k_5 r^5 + k_7 r^7$$

$$dx_r = x_c (k_2 r^2 + k_4 r^4 + k_6 r^6)$$

$$dy_r = y_c (k_2 r^2 + k_4 r^4 + k_6 r^6)$$

$$(r = \sqrt{(x - x_o)^2 + (y - y_o)^2})$$

(3.1)

Where:

dR= radial lens distortion (mm, expressed in μm)

r= radial distance (mm)

x, y= image coordinates (mm)

x_o, y_o= principal point components (mm)

k_i= i^{th} power terms of radial lens distortion polynomial (unitless)

The radial distortion profile is formed from the set of the distortion values (μm) over a range of radial distances (mm) which correspond to the targets coverage within the

frame and it can be of Gaussian (when it is referred to the nominal camera constant or a specific camera constant for zoom lenses) or balanced form (when obtaining the mathematical equivalent of the curve by shifting the camera constant by δc) (Fryer, 1996; Fraser, 2001). In the balanced radial lens distortion, the influence of the linear term $k_o r$ refers to a uniform change in image scale equivalently to a $k_o c$ change into the camera's constant. Every point is shifted symmetrically from the principal point (dr) hence the points that lie on the same circle have undergone the same radial distortion (see Figure 3.9). Calculation of radial distortion at two focus settings (close up and infinity) allows the determination of its coefficients at any other focus setting (Magill, 1955).
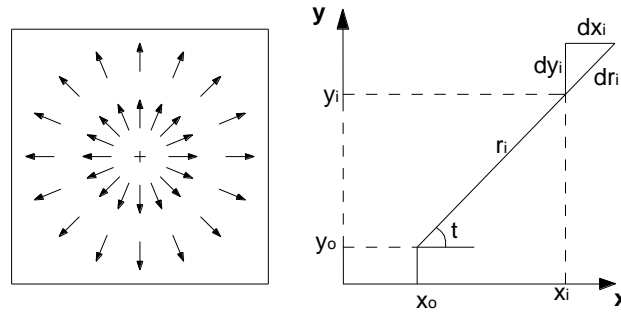
Figure 3.9: Radial lens distortion and its effect on an image frame.

3.2.2.2. Decentering distortion

The misalignment of the lens elements with relation to the optical axis results in a geometric effect known as decentering lens distortion having both radial asymmetric and tangential components (Mugnier et al., 2004). Decentering distortion is given in two components and it is represented by its profile function (see equations (3.2) and (3.4)) (Fryer, 1996; Fraser, 2001). Its parameters are highly correlated with the principal point components, noting that decentering is an order of magnitude less than radial lens distortion (order of a few tens of µm). To resolve this projective coupling 3D calibration arrays, strong intersection angles and full format coverage are normally employed.

$$dD = r^2 \sqrt{p_1^2 + p_2^2} \qquad (3.2)$$

$$dDx = p_1(r^2 + 2x) + 2(p_2 xy)$$

$$dDy = 2p_1 xy + p_2(r^2 + 2y)$$

Where:

dD= decentering lens distortion (mm, expressed in μm)

dDx= decentering lens distortion x component (mm, expressed in μm)

dDy= decentering lens distortion y component (mm, expressed in μm)

p_1, p_2= decentering distortion coefficients (unitless)

r= radial distance calculated from the principal point (mm)

x,y= image meaurements (mm)

3.2.2.3. In plane distortions

In plane distortions are usually manifested in differential scaling between the horizontal and vertical pixel spacing and introduce a non-orthogonality between image axes. In other words the affinity (scale factor between x axis and y axis) and orthogonality (deviation from 90° between x axis and y axis) terms are usually denoted as affine deformations, they are mathematically inserted into the internal calibration model and are treated with robust bundle adjustment approaches. It is important to note that when employing such extended models, singularities due to system overparameterization may occur. To avoid the system's overparameterization the parameters are implemented within the system upon the examination of their associated precisions and correlation coefficients. Specifically, insignificant parameters are generally suppressed whereas as acceptable correlations are considered those that are equal or less than 0.5 within image networks (VMS, 2009) (see section 4.3.4.).

3.2.3. Geometric stability

The geometric stability of digital cameras is held responsible for the suitability (photogrammetry and vision applications) and the accuracy that can be achieved. As a result, understanding the causes that impinge on the geometric stability in combination with the models and the measures utilized to accommodate for potential geometric instabilities is of high significance.

Digital cameras present instabilities related to the fixed CCD array with relation to the optical lens. Camera parameters are not considered to be stable due to a number of physical causes. Geometric camera stability needs to consider the camera's features in combination with the inserted mechanical influences while operating the camera. The main features that can influence the geometric stability of a camera are the resolution, zoom, focus and aperture settings. Mechanical effects include gravity (the torque of the lens mount is the product of the lens' weight multiplied by the distance of the lens' centre of gravity from the lens mount) with different viewing directions, camera heating due to long image acquisition periods and potential physical strains while operating the camera (Hastedt et al., 2002).

3.2.3.1. Stability modelling

Geometric camera stability needs therefore to consider the degree of influence of the above factors to the variations of the calibration parameters over time. The influences of geometric instabilities are accommodated with parameterization (image variant interior orientation) or mechanical stabilisation (sensor placement within the camera) of parameters (Zapp et al., 2009). Parameterization models are based on analytical correction methods with most appropriate the self-calibrating bundle adjustment. Fundamental elements concern the intrinsic elements camera constant, principal point with the geometric distortions measures (see section 3.2.2.). The image variant interior orientation parameters are modelled within a bundle adjustment procedure. One implementation is to introduce the camera constant and the principal point variations as observed unknowns within the adjustment weighted to appropriate values. This is advantageous over the model's over-parameterization and minimal correlation effects, especially with relation to the perspective centre. Calibrating an image variant interior orientation based on a common parameterization for distortion, affinity and shear is one approach. Alternative methods account for a balanced form of parameters describing radial symmetric distortion where the remaining image errors (e.g. unflatness) in sensor space are modelled using a finite elements correction grid[10] (Tecklenburg et al., 2001).

---

[10] The finite elements correction method is based on a raster type corrections grid, where each of the grid points is associated with correction values that are subsequently computed by point interpolation.

3.2.3.2. Stability evaluation

Geometric camera stability is covered in the literature with papers focusing on comparative testing of different camera systems, methods and performance measures. Camera systems fitted with zoom lenses at variant resolution or compression rate settings have been investigated with temporal variations of calibration parameters (camera constant, principal point, lens distortions, resolution and influences of zoom, focus and aperture settings) extracting estimates that are normalized to the image's width (Labe & Foerstner, 2004). Other comparative camera systems (CCD or CMOS) fitted with zoom and fixed lenses have been used to test stability with block invariant or photo invariant methods based on internal and external statistical measures (Shortis et al., 2006). Line-based calibration procedures for stability evaluation that address the degree of similarity between reconstructed bundles using different interior orientation parameters over time have also been discussed (Habib & Morgan, 2005). Calibration tests to evaluate the object space accuracy and its potential with image variant parameters and mechanical stabilization are investigated using a measuring testfield designed in compliance to the guidelines for the acceptance and reverification of optical 3D measuring systems (Zapp et al., 2009; Zapp et al., 2008). In these studies it has been proved that fixation of the focusing tube as well as preventing gravitational loads on the lens or the mount yields accuracies that are optimal for high precision surveys. The effect of lens movement due to gravity and unstable fixings has been studied and mathematically compensated (Haig et al., 2006). It is reiterated here that the criteria utilized to assess object space accuracy are based on the quality parameter length measurement error (difference between measurement and calibrated distances) as monitored in VDI/VDE guidelines that have been referenced in section 2.2.2. and will be additionally pointed subsequently in section 4.7.

3.3. Image measurement method

Digital image measurements are generated as the product of the imaging process involving the characteristics of the deployed imaging system, the properties of the measured features and the reliability of the measurement method. This section deals with the aspects of point-based features and applied measurement method utilized for sub-pixel target location.

3.3.1. Point-based features

A number of point based features can be utilized for 3D object measurement and these must be well identified. These point types can be man-made targets or natural features of high contrast and distinctiveness from the background. The quality of digital features is linked to the geometric and radiometric characteristics of the sensor and the characteristics of the feature (size, shape and texture). Good features are considered to be those that have high spatial frequency and are distinct from the background, geometrically and radiometrically invariant, interpretable, stable to noise and unique (Foerstner & Gulch, 1987). Feature quality is an important subject; dissimilarity measures (RMS residuals of mismatches) that enable, for example, judgement of feature matching between different frames based on affine motion models have been utilized (Shi & Tomasi, 1994).

Besides image-based measurements that utilize natural features (texture content and geometry), there are instances where artificial features need to be used. For example 2D or 3D artificial features can be defined in instances where there are insufficient natural locators (points, edges, regions) or where there is a need for unique point utilization for the establishment of reference measurements (rotation invariant spheres for scanning systems), engineering control or benchmarking, automation systems in metrology or accuracy enhancement. Such features can be manual (retro-reflective, coded, colour, white diffuse spheres, black on white naturally reflecting targets, eccentric, LEDs) or projected light (laser or other type of light projectors) with relation to their form and passive or active with relation to their illumination (Clarke, 1994; Luhmann et al., 2006). Figure 3.10 illustrates a sample of targets utilized for close range measurement.
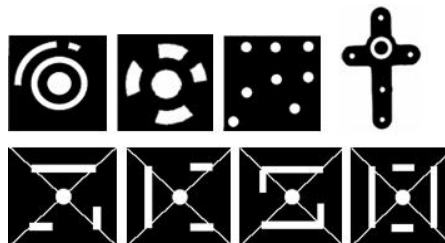
Figure 3.10: Artificial coded targets. Coded targets and exterior orientation devices (top) (source: Fraser, 1997) and coded targets example (source: VMS 8.0) (bottom).

Manmade point-based features for sub-pixel image measurement would require a minimum size of 5 pixels in the image space. Yet circular point features can be reduced due to eccentricity effects. Eccentricity is a deficiency in the image space where an ideal circle is projected to an ellipse and it can be significant with increased image scale and viewing angle. Whilst eccentricity effects can reduce image measurement quality, it has been shown that in multi-view processing frameworks they are considered to be absorbed within the process (Otepka, 2004).

3.3.1.1. Retro-reflective point features

A common type of point locators is the retro-reflective. Retro-reflective point features are adhesive targets made of retro-reflective material or an array of microprisms. The utilized material is called Schotchlite and it is constructed by 3M (Scotchlite, 2009). The building block of retro-reflective targets are 50 μm diameter spheres located on a layer and they can act as a cat's eye or as a retro-reflective prism provided they are illuminated from the camera's viewing direction in the ideal case. However, in real imaging situations the returned light will not be strictly parallel to the incident light due to a number of different factors (geometry of spheres, viewing direction, illumination and mechanical stress). Additionally, increasing the viewing angle from the normal can occlude the returned illumination by the adjacent spheres. Clarke (1994) has reported that for sub-pixel point location the targets will have to be located within a range of ±50 degrees on the object of interest and that for highest light return the light source can deviate within a cone of 0.5 degrees. Figure 3.11 illustrates a sample of different retro-reflective targets of varying diameter in the object space.

Figure 3.11: Retro-reflective point targets. Single retro-reflective (left) and single masked retro-reflective (right) (source: Geodetic, 2009).

Retro-reflective targets return high SNR and given optimal imaging (range, viewing direction, illumination) and geometric conditions (perspective image distortion) can result in 'good' measurements in the image space. Yet, the high cost of these targets

can become significant, especially in the case of coded targets or extensive signalization to delineate an object's surface.

Within this work 2mm object space diameter circular retro-reflective targets were produced from 7610 type Scotchlite™[11] high gain reflective sheeting generated with a circular hole-puncher. These targets present significant overexposure and saturation with changes in aperture settings. Moreover, their limitations regarding physical object space deformations and their large image footprints at very close imaging ranges under both perspective and affine projections, disabled them from targeting the designed structures and hence reduced their usage to sparse target objects mainly for reference purposes from sparse data (see section 6.2.2.1).

3.3.1.2. Non retro-reflective point features

Non retro-reflective point features are considered passive features that do not present retro-reflective properties. These can vary from artificial locators, including for example circular white markers on a dark background, encoded targets or natural features like distinctive edge intersections on high intensity imagery (see Figure 3.12).
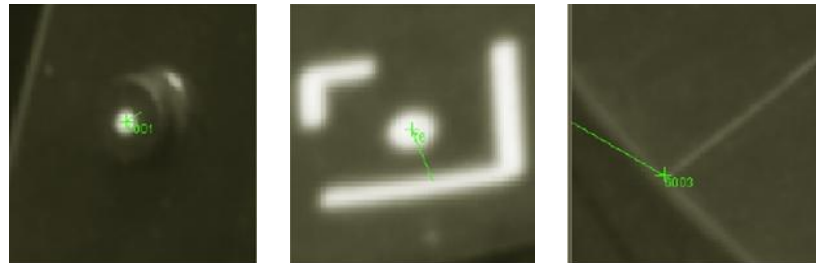
Figure 3.12: Point feature measurements. White marker on a dark background (left), coded target (middle) and natural feature on intersected edges (right). Images acquired with the Hasselblad H2D and H3D (f= 50mm, 7,216 x 5,412 pixels, pixel size 6.8μm) camera systems.

Circular point features are usually employed in instances where features of other types are impractical due to limitations regarding geometry, radiometry and texture. The centre of the target is considered as rotation invariant. However in real imagery influences of sensor, model projection and imaging conditions (viewpoint variation and imaging range) reduce an ideal circular point location to the determination of the

---

[11] Scotchlite™ High Gain Reflective Sheeting 7610: This exposed reflective lens with adhesive and liner material returns the highest brightness under wide imaging angles.

centre of an ellipse and in parallel identify the SNR of the identified locator. Point location and quality of location are factors of the employed feature detection method (see section 3.3.2.). Within this work the deployed structures are metal or plastic rigid geometrical objects that were signalized to achieve distinctive point source measurement features. The purpose-built self-adhesive targets constructed in the Rhinokeros 4.0 CAD software tool as circular white points (diameter: 0.5mm and 1.0mm) on a black background which were subsequently printed (onto adhesive laser, inkjet paper) and attached on the objects of interest. Figure 3.13 visualizes such a target imaged under affine projection geometry within a 40x40 magnification window.
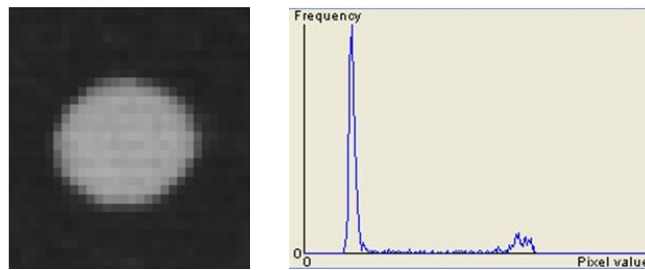
Figure 3.13: Self-adhesive white marker. Marker on a black background (left) and brightness histogram (with two local maxima) (right).

Encoded with a unique point identification number, coded targets are formed by patterns (lines, regions) surrounding the central feature point. These patterns are specific to the utilized measurement method embedded in the software (VMS, 2009). Key characteristics of these locators are their scale invariance, robust detection over rotation and model projection for recognition and measurement enabling image analysis processing methods (Shortis et al., 2003). They are usually applied in automatic orientation procedures, establishment of control and object space scale. Natural features rely on local image content and usually occur in applications ranging in scale from aerial to close range situations with most common points, lines and regions. These features are identified with algorithms based on detectors and descriptors; a review of which can be found in Remondino (2006). Here object space target occupancy, projection scale, close-up imaging ranges in combination with the scope of method testing (algorithm, model and geometry behaviour) excluded coded or natural features as a selection for data acquisition and testing.

3.3.2. Point-based measurement method

Feature measurement methods aim at the recognition and location of features in digital images based on automatic or interactive approaches for sub-pixel point location dependent on the associated application (for example measurement, registration, surface generation and texturing). A significant number of interest point methods exist in the literature based on detectors that analyze the image's signal or fit the image signal on a template. Here, point-based images were generated as input data to the deployed process. It is therefore clear that the employed point-based measurement method will be reported in the context of its application on digital images for subsequent processing and testing.

3.3.2.1. Centroid location method

Point-based structures were generated in a photogrammetric measurement tool that deploys an embedded centroid location method (VMS, 2009). The tool computes the 2D centre of an image centroid within a ROI (4-64 pixels) and it leads to a successful answer where high contrast target images occur allowing a manual point location in an alternative case. Considering that a window of size n x m pixels is placed around the target to be located, the centroid of the target is given as follows:

$$\begin{bmatrix} x_n \\ y_m \end{bmatrix} = \frac{\sum\limits_{i=1}^{n} \sum\limits_{j=1}^{m} g_{ij} \begin{pmatrix} x_i \\ y_j \end{pmatrix}}{\sum\limits_{i=1}^{n} \sum\limits_{j=1}^{m} g_{ij}} \tag{3.3}$$

Where:

x_n, y_m = centroid of the target coordinates (mm)

x_i, y_j = pixel coordinates

$g_{ij} \begin{pmatrix} x_i \\ y_j \end{pmatrix} = $ digital value of light intensity at location (x_i, y_j)

n x m = window dimensions

The method follows established approaches that have been similarly reported in Fraser (1997) and Shortis et al. (1994). Prior to the actual centroid computation the

applied thresholding method classifies pixels as target or background within the ROI. The method requires identification of the location method, threshold type and passes of relevant geometric tests. In addition a set of empirical factors (sigma value for random threshold and ellipse fit, minimum gray level range within the window and minimum span of the target image) are identified to apply each of these tests within the method. A centroid is identified as binary, weighted, square weighted or ellipse fit. The locally identified threshold within the ROI is set as the mean of the intensity values between the two peaks (high and low) on the intensity histogram. Besides standard thresholding, the tool can utilize an additive (setting a robust threshold in the presence of significant background intensities) or a random (on the assumption of background image noise) method. The target image is located with a series of geometric tests (ratio test for circular targets and target region ratio test) based on the knowledge of the location and extent of the target image within the ROI together with expected size and shape.

3.3.2.2. Epipolar geometry and back-projection

Point based measurement generation as input to a multi-view algorithm demands establishment of point correspondences. Given the knowledge of a point on an image its homologue can be identified along the epipolar line (see Figure 3.14) which is usually curved on the presence of large perspective distortions. Correspondence solutions utilize search areas close to the points of intersection of epipolar lines and can assist the point location method described above (see 3.3.2.1.).
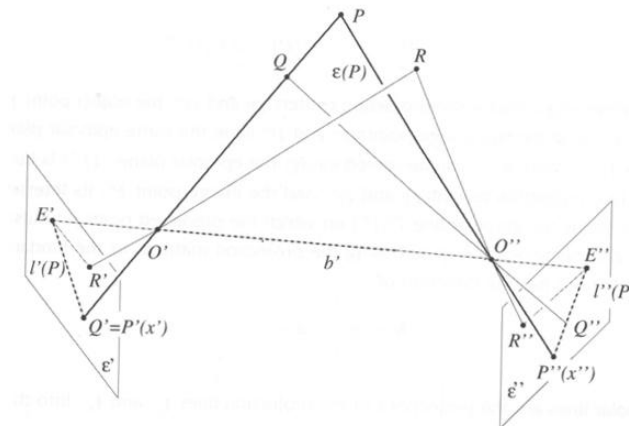
Figure 3.14: Epipolar geometry. Notation after (Mugnier et al., 2004): $\varepsilon(P)=$ epipolar plane, E', E''= epipoles, l'(P), l''(P)= epipolar lines, b'= baseline, O', O''= projection centres and P, R= objects points, Q' R', Q'' R''= imaged points.

The measurement method applied here can be listed in the following sequential steps:

1. Establish point correspondences on image data based on centroid or manual location assisted by epipolar geometry.
2. Estimate the camera's exterior parameters on an initial subset of four control point (CP) data (initialize exterior orientation).
3. Back-project on a set of known points (predetermined CP or pre-triangulated tie point (TP) data) to retrieve remaining image measurement locations and update camera's exterior parameters (backproject on resection).
4. Refine image measurements if necessary to enter the subsequent computation.

This process has been applied throughout this thesis in order to generate the implemented datasets, resulting in image networks that handle 17 (minimum) to 85 (maximum) images where 20 (minimum) to 178 (maximum) target points have been measured (CPs and TPs) (see section 6.2.2.). It is recalled here that data generation was implemented within the tool VMS 8.0. All data forms were subsequently read as ASCII files within the developed method.

3.4. Camera models

A camera is the medium which performs the mapping that projects 3D spatial information onto a 2D plane. The geometric underlying principle which establishes the 3D to 2D correspondence is the pinhole model which forms the basis of the established central perspective projection. Camera modelling is studied in analytical terms utilizing the fundamental elements of Euclidean and projective geometries[12]. Algebraic approaches are ideal for automation of image analysis and problems that implement direct solutions, computing for example the (SVD) of matrices utilized mainly from the computer vision community (Foerstner & Wrobel, 2004). Yet this work is looking at the geometry governing the camera in non-homogeneous vector terms. The reason for this approach is related to the focus on design and implementation of a multi-view problem for the affine projection case which includes the ability to apply statistical error propagation. It is noted though that where

---

[12] Projective geometry overcomes the limitations of Euclidean geometry by placing points, lines and planes at infinity as natural entities, unifying transformations such as similarity, affinity under the class of projective transformation (Foerstner & Wrobel, 2004).

necessary computer vision literature will be reviewed as part of the problem and its implementation studied here.

3.4.1. Projective camera

The projective camera is the most general camera model and it can be considered as a generalization of the perspective camera. Figure 3.15 illustrates the recovery of the perspective projection process linking 3D object with a 2D image spaces through the projection centre O.
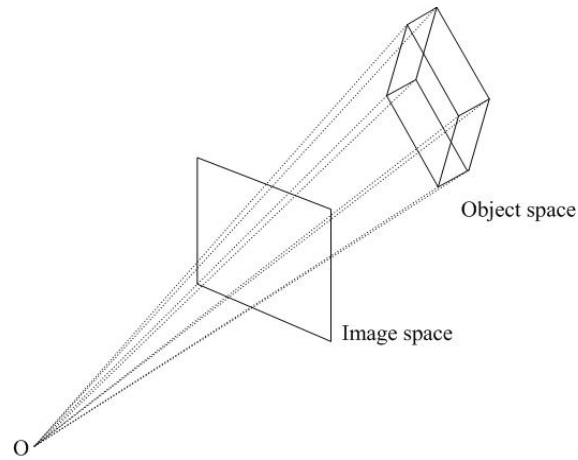
Figure 3.15: Perspective projection model.

In Euclidean geometry a 3D point $X = (X, Y, Z)^T$ is projected to a 2D image point where the line starting from the projection point to the 3D point intersects the image plane. If the relationship between 3D space and 2D space is expressed in homogenous terms then the projective camera can be written as it is given in equation (3.4).

$$\underset{(3x1)}{x} = \underset{(3x4)}{P} \underset{(4x1)}{X} \qquad (3.4)$$

$$\begin{bmatrix} fX \\ fY \\ Z \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Where:

x (3x1)= 2D image coordinates vector (mm)

P (3x4) = camera projection matrix

3. Digital close range image formation

X (4x1)= 3D object coordinates vector (mm)

f= focal length (mm)

p_{11}-...-p_{34}= projection matrix elements

The relationship between homogeneous and non-homogeneous coordinates is given as $(x, y)^T = (fX/Z, fY/Z)^T$ expressing the 2D image space vector and $(X, Y, Z)^T = (X/1, Y/1, Z/1)^T$ stating the 3D object space vector. The projection matrix P can be decomposed into a calibration matrix which includes the intrinsic camera parameters together with a rotation and a remaining matrix that are utilized to encode the extrinsic orientation parameters. These mathematics are equivalent to the established collinearity condition (see section 4.3.2.) used as the fundamental basis of the image-to-object space correspondence in photogrammetry. Weinhaus & Devich (1999) point that the main difference between the projective and the collinearity approach is the treatment of the focal length and the camera's projection center. In the collinearity condition the focal length is used as a physical parameter and no special projection matrix is required. Additionally, the projection center is maintained as a separate vector, allowing the solution of its coefficients, whereas in the homogeneous coordinate approach it is folded in the four-dimensional matrix. It is noted that intrinsic orientation refers to the elements of interior orientation (principal point and camera constant) extended by additional internal geometric terms (see section 3.2.2.) the parameters of which can vary for different camera models. The extrinsic orientation represents the elements of the exterior orientation of cameras which include the camera's attitude and position in the 3D object space coordinate system. According to Hartley & Zisserman (2004) a general projective camera is represented by a homogeneous (3x4) matrix of rank 3 with 11 degrees of freedom as scale is arbitrary. The rank requisite is essential to define the matrix mapping as an image which is a 2D plane and not a point or a line. A projective camera can be reduced to the well known perspective projection when the image and 3D object space coordinate systems are linked with a rigid transformation (Shapiro, 1995). The 2D to 3D mapping is a non-linear problem to solve and given the input data handling can become computationally expensive. When the perspective projection camera model

- 76 -

does not map the true geometry of a camera the problem can be ill-conditioned[13], hence an appropriate camera or sensor model needs to be employed.

3.4.1.1. Comparison of projective and perspective camera

Considering the main attributes of projective and perspective camera models a comparative table is given here (see Wrobel, 2001) representing the underlying geometry behind the basic orientation tasks to show in parallel the model suitability.

	Orientation model	
	Projective (linear model)	Perspective (non-linear model)
Model	$x' = P \quad X$ $(3x1) \quad (3x4)(4x1)$	$x' = \lambda \; R \quad X + T$ $(3x1) \quad (3x3)(3x1) \; (3x1)$
Parameters	11 (projection matrix P)	6 EO per image; 3-5++ IO per camera
Correspondences	6 per image	3 per image
Behaviour	Low stability, critical configurations for object space planar configurations, focus & zoom optics accomodated	High stability, object space stable configurations, stable camera & optics

Table 3.5: Projective and perspective orientation models. Notation: x'= image space coordinates vector, X= object space coordinates vector, P= projection matrix, λ= scale factor of vector x', T= object space coordinates of perspective center.

The projective camera model is highly advantageous due to its linearity. Yet its instability, increased model parameters (and correspondences) as well as inability for lens distortions accommodations can introduce significant limitations in comparison to the standard perspective camera model. These are the main reasons for the broad utilization of standard perspective models in photogrammetry. In this text, analytical modelling of perspective cameras considering only points is discussed further in section 4.3. It is however noted that the projective camera model is highly adapted to diverse situations and transformations to perspective, posing this as an ideal model in different photogrammetric tasks.

---

[13] Ill-conditioning can be a situation where small data variations can cause significant result variations.

3.4.2. Affine camera

An affine camera can be defined as a camera with a projection centre at infinity that generalizes parallel projection. The geometric interpretation of the affine image can be given considering for example the 3D to 2D image formation of a 3D object point onto a 2D image plane. Figure 3.16 provides a 1D illustration of this relationship for the perspective, parallel and orthographic projection cases.
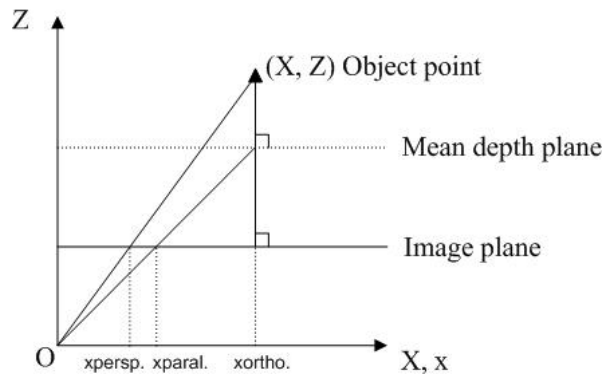
Figure 3.16: 3D-2D relation. Perspective, parallel and orthographic projections.

The 3D object space is projected onto the image plane following a line passing through the projection centre forming its perspective image (denoted as xpersp.). Departing from this, the line of sight that hits orthogonally the mean depth object plane and then perspectively projected onto the image plane forms a parallel image (location xparal.). Different terms for a parallel projection image are the scaled orthographic and weak perspective projections. It is evident that in the special case where the line of sight enters the image plane perpendicularly, this results in the strict orthographic projection with a unit scale. Shapiro (1995) illustrates the imaging case where the line of sight enters at an angle θ at the mean depth plane to subsequently form an image perspectively on the image plane and names this as a paraperspective projection. Figure 3.17 illustrates two close-up views of the parallel and orthographic projection models respectively.
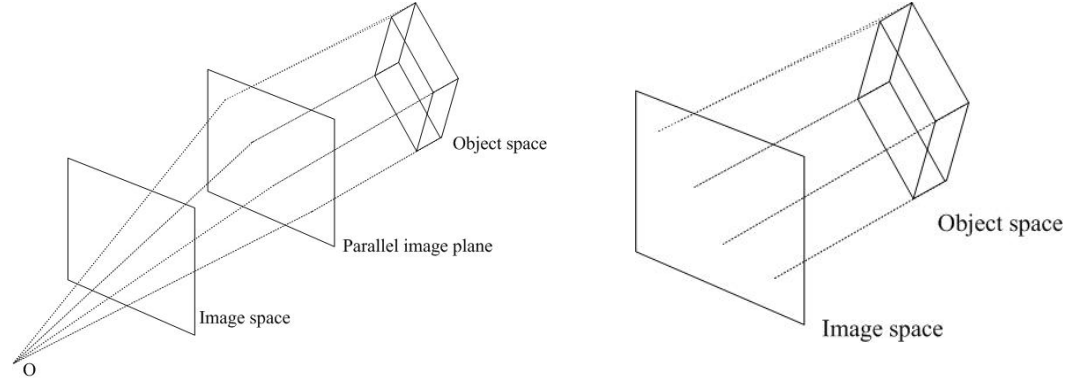
Figure 3.17: Weak perspective projection (left) and orthographic projection (right).

This work is concerned the affine camera modelling in a multi-view framework with statistical error propagation. The model will be treated in the non-homogeneous coordinate system case and in this context the model can be derived from perspective when adding two assumptions. These form the case where the object of interest presents a small FOV and a small depth variation. The affine image can then be written as it is given in equation (3.5).

$$\underset{(2\times1)}{x} = s\underset{(2\times3)}{R}\underset{(3\times1)}{X} + \underset{(2\times1)}{t} \qquad (3.5)$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = s\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} tx \\ ty \end{bmatrix}$$

Where:

x= 2D image coordinates vector (mm)

s= image scale factor (unitless)

r_{11}-r_{23}= elements of the (3x3) rotation matrix R

X= 3D object coordinates vector (mm)

t= 2D projective translations vector (mm)

An affine camera can be considered as an uncalibrated scaled orthographic camera requiring no calibration of internal camera parameters such as camera constant and principal point. This property in combination with its ability to preserve parallelism of lines enables the utilization of affine epipolar geometry in multi-view location problems. An analytical description of the affine camera, its interpretation and

algorithmic implementation can be found in the literature (Faugeras & Luong, 2001; Hartley & Zisserman, 2004a; Shapiro, 1995; Xu & Zhang, 1996). Model treatment and implementation will be the central topics of the subsequent chapters 4 and 5.

3.4.2.1. Magnification

A particular property of affine projection is its invariant scale factor. This is of particular interest as the scale factor plays the role of camera constant in the real perspective camera case. Model development, implementation and testing have been based on the ability to generate true parallel projection imagery with the deployed systems described in section 3.1.2. Following the affine projection formation reported above, a plane located parallel to the image at a range $z = z_o$ defines the lateral magnification between the distance measured in the image $(dx_I, dy_I, 0)^T$ over the corresponding distance on the image plane $(dx_P, dy_P, 0)^T$. The magnification factor will be uniform for all points that lie on the same plane and this will be estimated as $m < 1$ for the general affine case or $m = 1$ in the strict orthographic projection case. The magnification is constant when the depth range of the imaged object is relatively small in comparison to the range to the camera. The projection equations are thus simplified as given in equations (3.6).

$$x' = -mx; \quad y' = -my \tag{3.6}$$

$$m = \frac{\sqrt{(dx_I)^2 + (dy_I)^2}}{\sqrt{(dx_P)^2 + (dy_P)^2}} = \frac{f'}{-z_0}$$

Where:

x', y' = 2D image coordinates (mm)

x, y = transformed coordinates (mm)

m = magnification factor (unitless)

To prove that the test camera systems generate geometrically true affine images, the following example calculates the associated magnification factor. The deployed camera system was composed of the available Kodak Megaplus ES1.0 monochrome camera (see section 3.1.2.1.) attached with a perspective (Fujinon TV; f= 12.5mm) and a telecentric (MVO® TML; m= 0.16x) lens interchangeably (see section 3.1.3.). Subsequently, a typical calibration arrangement was designed to enable imaging of a

square grid jointly with perspective and parallel projection camera systems. The purpose built grid comprises a planar pattern (168 x 168 mm) of 22 x 22 white circular markers on a black background. It is noted that the planar pattern (constructed in the Rhinokeros 4.0 CAD tool) was printed on a laser printer and attached on a planar metal board. To generate measurable image point features in the parallel projection imagery (nominal magnification factor= 0.16x), the targets were designed with a finite size of 0.5mm diameter in the object space, resulting in 9 pixel blobs in image space. Given that the location where the telecentric lens realizes sharp parallel projection imagery is a range of 175mm from the front lens diameter, this was set as the mean imaging range. The experiment was implemented by shifting the calibration grid at regular intervals (1mm separation) within ±30mm from the mean position (DOF= ±19.7mm). To illuminate the scene two LED green ring flashlights were utilized. These were positioned at near 45^0 angles from the normal direction for balanced directional illumination (see Figure 3.18).
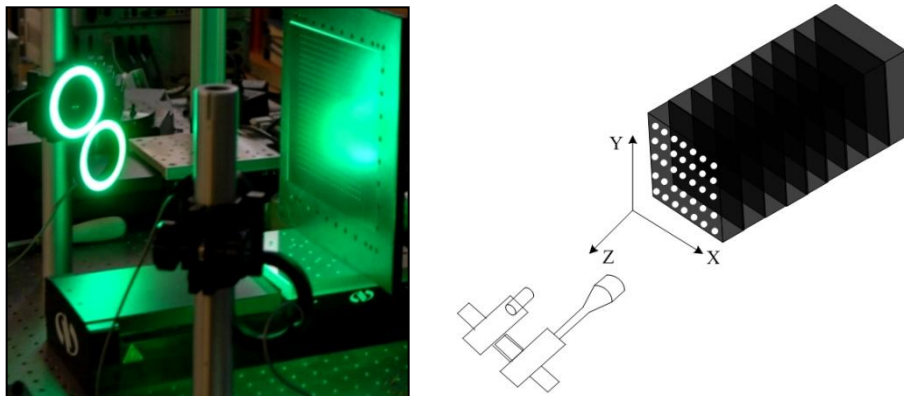
Figure 3.18: Single view calibration. Calibration arrangement in a laboratory environment (left) and experimental data capture design (right).

To generate a set of 2D image measurements, the acquired data were inserted in VMS 8.0 where a sequence processing project was initialized. The image measurement method involved application of a centroid estimation on a set of four sparse corner grid points initializing the photo orientation parameters. Next, a 'resection on backdriving' backprojects the remaining CPs on the pre-estimated exterior orientation and updates the photo locations and rotations on the total measured CPs (see section 3.3.2.2.). The resection closed with an RMS image residual of 2.78μm.

Figure 3.19 illustrates an example of the grid image supported by the selected parallel projection image (with its image residuals) to show the result of the sequence measurement process.
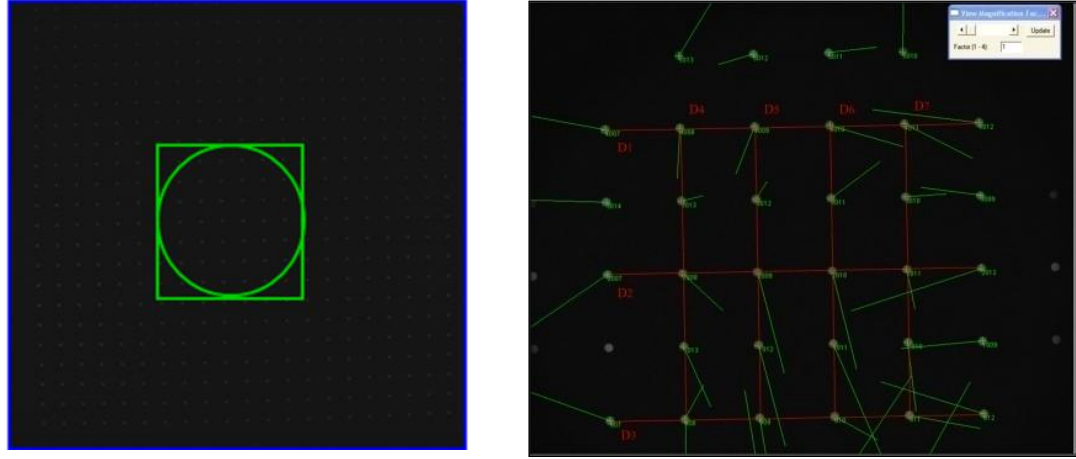
Figure 3.19: Calibration grid pattern. Pattern imaged with the Kodak Megaplus ES1.0 camera with Fujinon TV lens / f= 12.5mm (left) and Kodak Megaplus ES1.0 camera with MVO® TML 0.16x (right).

The experiment was executed with a set of seven distances (three horizontal and four vertical) selected to estimate the magnification factor within the image. According to the calculation of this scalar (as defined in equation (3.6)), m equals to 0.15x (nominal m= 0.16x) with a mean discrepancy from the nominal equal to 5.74μm (0.6 pixel) in the vertical direction and 6.30μm (0.7 pixel) in the horizontal direction.
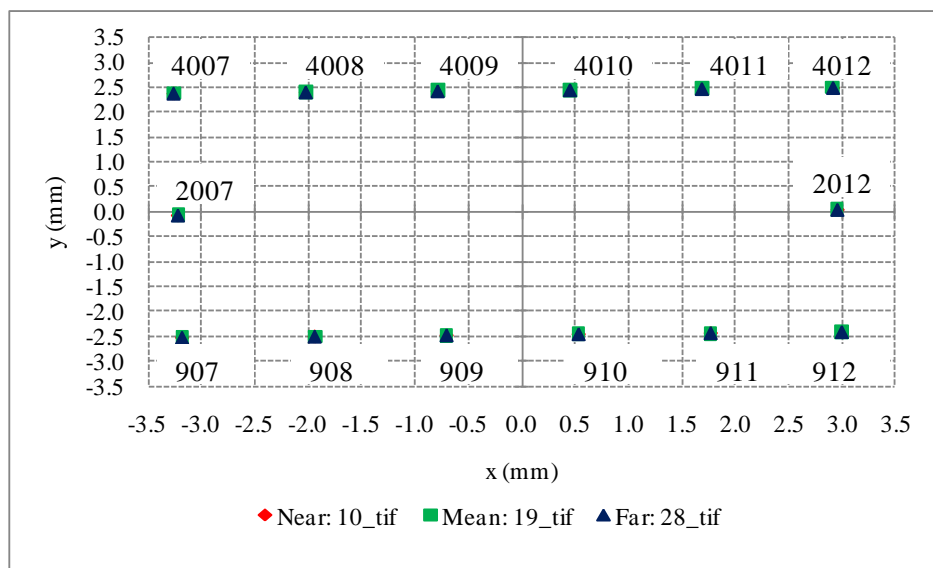
Figure 3.20: Point measurements at near, mean and far ranges.

Range	Min.(μm)	Max.(μm)	Mean.(μm)
Mean:19_tif - Near:10_tif	0.70	4.19	2.47
Mean: 19_tif - Far:28_tif	0.57	4.10	2.37

Table 3.6: Absolute differences on selected measured point locations.

Figure 3.20 illustrates the 2D image point locations on the estimation image denoted as 'Mean: 19_tif' together with the seven length measurements. The additional images 'Near: 10_tif' and 'Far: 28_tif' are located at corresponding ranges of -/+ 9mm from the mean location given the system's DOF. To estimate the 2D image discrepancies of the mean image position over the near and far range images, the corresponding absolute differences were calculated (see Table 3.6). It is evident that the deviations from the 'ideal' locations range between 0.70-4.19μm in the near range and 0.57-4.10μm in the far range with a mean discrepancy of 2.40μm. The order of these differences can be attributed to the fact that the imaging system deviates from its 'optimal' geometry and radiometry that in effect reduce the image measurement locations that would in nominal terms lead to a perfect parallel projection.

3.5. Summary

This chapter has reviewed the main concepts and principles that form the fundamental elements of digital close range image formation. In particular the building block of data acquisition and processing has been outlined with relation to the employed sensor elements and systems supported by digital image characteristics, their basic geometric elements and applied measurement method for data generation, initialization and testing. This chapter closed with the two fundamental camera models that establish the 3D to 2D correspondence which introduces the subsequent chapters that describe the developed methodology and testing for the purposes of this research work.

4. Modelling from multiple views

The exploitation of the affine sensor from the photogrammetric standpoint demands that the fundamental background of multi-view modelling is introduced prior to the analysis of the affine algorithm. In measurement sciences like photogrammetry, industrial metrology, geodesy and computer vision it is established to base computations on least squares estimation (LSE) theory as a result of its advantage to provide statistical analysis for quality assessment. Modelling from multiple views can therefore be considered in the contexts of (a) sensor modelling from multiple viewpoints, (b) aspects of geometry as well as (c) statistical quality.

This chapter reviews the basic principles behind established multi-view modelling approaches. These concepts are outlined in order to provide a review of the well established methods applied in frame cameras. In particular the self-calibrating bundle adjustment is outlined with consideration of its current state of the art (see section 4.1.). Subsequently fundamental elements of LSE theory are given to provide the basis for problem analysis as well as development (see section 4.2.). The chapter provides the background in modelling perspective cameras from the photogrammetric standpoint, covering the aspects of mathematical model formulation, self-calibration and initialization (see section 4.3.). Following the well established background of perspective approaches, the affine sensor is presented in the contexts of existing methods originating from the computer vision literature (see section 4.4.). The chapter concludes with a review of the datum problem (see section 4.5.), image network geometry (see section 4.6.) as well as quality control (see section 4.7.); core concepts in photogrammetric analysis.

4.1. Bundle adjustment method

The bundle adjustment is a very old method well known in photogrammetry whereas gradually adopted within the vision community. Triggs et al. (2000) define: *bundle adjustment is the problem that refines a visual reconstruction in order to optimize jointly 3D structure and viewing parameter estimates*. It is a geometric statistical estimation problem of simultaneous intersections of image rays linking 2D to 3D spaces through cameras poses (positions and orientations) integrating 3D object space positions (3D coordinates) with the potential to allow camera calibration (recovery of

interior orientation elements). It is conventionally formulated as a non-linear LSE problem that minimizes a quadratic form cost function of the feature reprojection errors between observed and computed image observations with the potential to enable robust outlier detection and elimination within the method.

4.1.1. Background

The bundle method has its roots in LSE theory which in combination with the evolution of computers in the fifties and continuous technological developments is today utilized as a robust tool in photogrammetry and vision applications (see section 4.1.3.) Figure 4.1 illustrates a diagrammatic form of the historical development behind the method of bundle adjustment. The fundamental principle of the bundle method developed by Schmid's single photo least squares resection based on the collinearity condition (see section 4.3.2.) whilst the complete theory was set by D. C. Brown (Brown, 1974). It has its roots at large scale aerial-triangulation problems, its basic measurement unit is the bundle of image rays and in the close range it is termed as network adjustment or phototriangulation. The bundle method follows the developments in image processing tasks which are key to image least squares matching (LSM) and robust statistics methods allowing its ability to add-in self-diagnosis in order to reach its main focus today; and this is large volume data processing as well as system automation.

Development of a bundle method entails a set of issues: the minimization of the cost function (see section 4.2.1.); starting value estimation (see section 4.3.5.); solving large normal equation systems considering matrix structure and sparsity; datum (or gauge) definition (see section 4.5.); quality control (see section 4.7.). These critical issues can be designed at the stage of method implementation whilst satisfying problem requirements and purpose of application.
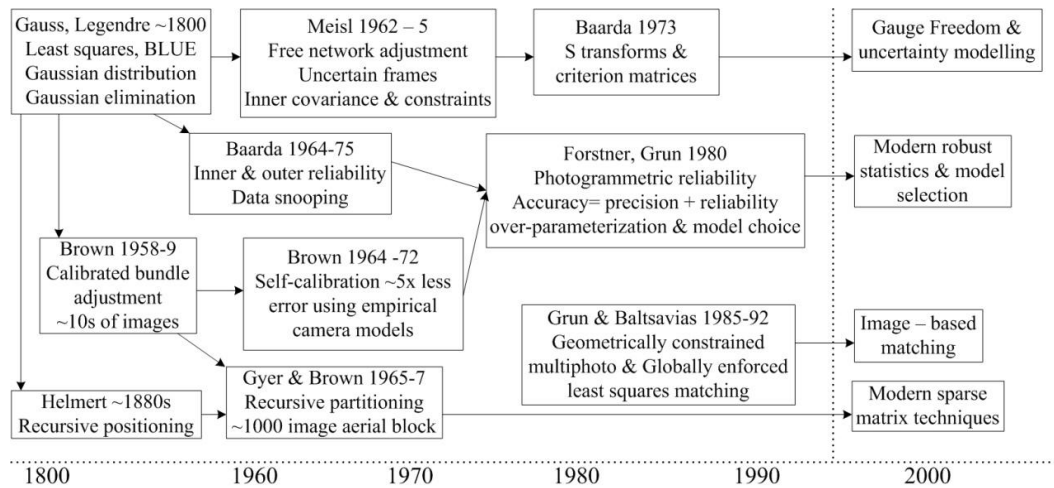
Figure 4.1: Development of the bundle method (source: Triggs et al., 2000).

In photogrammetry a routine application of bundle adjustment systems comprises the 3D measurement of engineering structures. These systems have their roots on established robust strategies already reported in Brown (1971), Clarke & Fryer (1998) and Granshaw (1980) whilst employing internal calibration models with extended parameter sets (Fraser, 2001) (see section 4.3.4.). Regarding model formulation and structure (see section 4.3.2.), these can be modified in the context of implementation of linear features as measurement entities (Hrabacek & Van den Heuvel, 2000), employment of different sensor models such as panoramic cameras (Parian, 2007) or enforcement of camera constraints as in cases of combined stereo-imaging geometries (King, 1995). Although modifications regarding method implementation and treatment exist, it is generally accepted that bundle approaches are optimal in that they offer robust solutions for system treatment.

In vision, Tsai (1987) proposes a two stage calibration technique for 3D vision metrology that recovers the camera's exterior orientation as well as camera constant, radial lens distortion and image scanning parameters (shear and aspect ratio) by applying a set of four geometric constraints obtained from the implemented model parameters. A flexible camera calibration technique for planar patterns based on an initial closed-form solution and a subsequent non-linear maximum-likelihood estimation refinement modelling focal length, radial lens distortion and aspect ratio has been suggested by Zhang (1999). In fact this tool can be found online. Additionally, Triggs (1998) developed a self-calibration approach of a moving projective camera from five views of a planar pattern treating recovery of parameters

up to a scale factor. Another example is proposed by Mayer (2005) who employs a strategy based on point extraction and LSM for precise point estimation combined with a projective bundle adjustment (Pollefeys et al., 2004) and applied to wide-based image sequences.

4.1.2. Main attributes

The main reasons for selecting the method of bundle adjustment as an appropriate approach for multi-view modelling are its flexibility, efficiency and quality control (Cooper & Robson, 1996; Luhmann et al., 2006; Triggs et al., 2000).

1. Flexibility: Refers to the ability to implement different information elements regarding camera models, 3D features (points, lines and surfaces), geometric constraints, sources (2D, 3D features and intensities) and error models.
2. Efficiency: Indicates the method's capacity to utilize economical and convergent numerical methods that take advantage of the problem's sparseness.
3. Quality control: Indicates the evaluation of accuracy, precision and reliability measures (see section 4.7.). Statistical error modelling is critical to the analysis of the estimated parameters.

Close range bundle adjustment, as opposed to the 'classical' aerial triangulation, softwares are advantageous in that they can, for example, deal with often difficult image configurations, arbitrary coordinate systems and a variety of camera systems. In addition they can handle complex, structured, large systems of normals equations. In this sense the bundle method can be considered a highly universal solution. Thus it provides a simultaneous and very effective solution to sensor modelling as well as 3D object measurement.

4.1.3. State of the Art

Further to its wide application in standard photogrammetric case studies, the method of bundle adjustment has now reached the stage of automation; therefore it comprises a very active topic of interest within the vision community. To make this point clear some examples are reported. Pollefeys et al. (2004) employed uncalibrated image

sequences acquired with a hand-held camera and computed 3D structure based on a multi-view stereo matching approach. Strecha et al. (2008) exploited image based modelling with a focus on benchmarking[14] which in their case was performed in relation to ground truth reference scan data. Some examples of evaluation datasets for both camera calibration, stereo and multi-view stereo can now be found on-line (ISPRS, 2009). Prior considerations of bundle adjustment implementation are related to parameterization, error models, linearization, optimization (utilizing the sparse structure of the normals equations coefficient matrix) and robustification techniques. PC-based commercial software packages and tools are available and widely utilized for photogrammetric applications tasks (Kruck's software BINGO; Geodetic, 2009; iwitness, 2009; Photomodeler, 2009; VMS, 2009). An open camera calibration toolbox for Matlab has been released from Bouget (2009) and has been included in the open source computer vision library distributed by Intel and can be found online (Intel 2001; Intel, 2009). Another tool is Zhang's algorithm applied in planar patterns which can be found online at Microsoft's webpage (Zhang, 2009). A recently developed open source generic sparse bundle adjustment has also been distributed, an updated version of which can be found in Lourakis & Argyros (2009). Further, Lourakis & Argyros (2005) run through benchmark tests to address performance with relation to speed and reprojection errors. In Dickscheid et al. (2008) a complete benchmarking scheme for assessment of automatic bundle adjustment results is proposed to assess orientation frame parameters in a statistical manner based on a well defined coordinate system. A good critical review on bundle adjustment methods with references in photogrammetry and vision can be found in Triggs et al. (2000).

4.2. Least squares estimation

The basic error model utilized in a bundle adjustment scheme is LSE. The method seeks in principle to derive a unique set of estimates of variables of certain properties minimizing the cost function of the weighted sum of squared residuals. The background conceptual scheme starts from a mathematical model[15] that approximates

---

[14] Benchmarking schemes require the definition of a reference dataset with superior precision and a set of statistical measures extracted from covariance analysis. A report on performance evaluation and benchmarking can be found in Foerstner (2005).

[15] The mathematical model is composed of the functional (describing the deterministic properties of the physical situation) and stochastic (describing the nondeterministic properties of the variables) (Mikhail, 1976).

the physical problem to be solved. This model then enters the LSE technique where on completion statistical testing is applied to the derived results to judge whether or not the initial model approximates the reality based on certain assumptions and rules or remodelling is needed. For a comprehensive coverage of the LSE theory the reader is guided in standard textbooks and reports (Cooper, 1987; Dermanis, 1990; Mikhail, 1976; NPL, 2001). Following the notation described in Appendix A, the least squares estimation formulation is given below.

4.2.1. Least squares mathematics

The functional model that connects the observations l (n x 1) with the unknown parameters to be estimated x (m x 1) is considered to be given as $f(x, l) = 0$ (4.1) (where f: denotes the total functions) with a stochastic model given as C_l. Now expressing the previous relationship for the associated true values this becomes:

$$f(\bar{x}, \bar{l}) = 0 \qquad (4.2)$$

The functional model is non-linear. These equations must be linearized, they need to be replaced by their approximations which are derived from the Taylor series expansion (linearization scheme). With regards to the measurements l and parameters to be estimated x, equation (4.3) is derived to first order accuracy as follows:

$$f(\bar{x}, \bar{l}) = f(x_o, l_o) + A(\bar{x} - x_o) + B(\bar{l} - l_o) = 0 \qquad (4.3)$$

Where:

$f(x_o, l_o)$ = functional vector of the first order approximations

$\underset{(cxu)}{A} = \left(\frac{\partial f}{\partial x}\right)_o$ = design or Jacobian matrix of the unknowns to be estimated

$\underset{(cxm)}{B} = \left(\frac{\partial f}{\partial l}\right)_o$ = design or Jacobian matrix of the observations

$\bar{x}, \bar{l}$ = true values that fit the functional model exactly

x_o, l_o = first order approximations to x, l

$\overline{x} - x_o$ = vector of true minus approximate values for the unknown parameters

$\overline{1} - 1_o = v$ = vector of residuals, corrections to measurements necessary to satisfy the functional model exactly.

By substitution of:

$v = \overline{1} - 1_o$, $x = \overline{x} - x_o$ and $b = -f(x_o, 1_o)$

The linearised form is obtained to be equivalent to: Ax + Bv= b (4.4). It is noted that linearization formulation is treated for the model's observations, parameters and constants.

The unique least squares estimates of $\overline{x}, \overline{1}$, denoted as $\hat{x}, \hat{1}$, are those that satisfy the least squares criterion which in the general case is written as follows:

$$\varphi(v) = v^T W v \rightarrow min \Rightarrow \varphi(v) = v^T W v + 2k^T (Ax + Bv - b) \rightarrow min \qquad (4.5)$$

Where:

W= weight matrix

k= vector of Lagrange multipliers[16] (introduced in order the estimates of $\hat{x}$ and $\hat{v}$ can be found).

Subsequently the estimates $\hat{x}, \hat{v}$ and $\hat{1}$ are given as follows:

$$\hat{x} = [A^T (BW^{-1}B^T)^{-1}A]^{-1}A^T (BW^{-1}B^T)^{-1}b \qquad (4.6)$$

---

[16] Lagrange multipliers are utilized to find the extrema (maximum or mimimum) of a multivariate function f ($x_1, x_2, \ldots, x_n$) subject to the constraint g($x_1, x_2, \ldots, x_n$)=0, where f and g are functions with continuous first order derivatives on the open set containing the curve g($x_1, x_2, \ldots, x_n$)=0, and $\nabla g \neq 0$ at any point in the curve (where ∇ is the gradient) (Wolfram, 2009a).

$$\hat{v} = W^{-1}B^T(BW^{-1}B^T)^{-1}\{I - A[A^T(BW^{-1}B^T)^{-1}A]^{-1}A^T(BW^{-1}B^T)^{-1}\}b \tag{4.7}$$

$$\hat{l} = l + \hat{v} \tag{4.8}$$

Expressing the functional relationships between measured and unknown elements to be explicit in the measured elements poses least squares more flexible. Hence, equations (4.1) and (4.4) become respectively:

$$f(x,l) \equiv f(x) - l = 0 \tag{4.9}$$

$$Ax = l_o - f(x_o) + v \tag{4.10}$$

Equation (4.10) together with W (representing the stochastic model of the observations) comprises a special case known as the linearized observation equations case. For uncorrelated image observations the weight matrix W_l is the inverse of the covariance matrix C^{-1} and it is of diagonal form with weights w_i calculated as:

$$w_i = \frac{\sigma_o^2}{\sigma_i^2} \tag{4.11}$$

Where σ_o^2, σ_i^2 are the variance of unit weight and a priori variance of observations respectively. Hence, the linearized observation equations follow the Gauss Markov theorem which states that for linear models least squares result in the Best (of minimum variance) Linear Unbiased Estimator (BLUE). It follows that minimization of the cost function $\varphi(v) = v^T W v \rightarrow$ min results in the calculation of the normal equations $N = A^T WA$ (4.12). Assuming that the normals equations matrix is of of full rank[17], that is N is non-singular, the following equations are obtained:

$$A^T WA \hat{x} = A^T Wl \tag{4.12}$$

$$\hat{x} = (A^T WA)^{-1} A^T Wl \tag{4.13}$$

---

[17] Rank is the order of linearly independent rows or columns of a matrix. The statement that matrix N is of full rank refers to the condition where the number of parameters to be estimated is equal to the parametric order of the system and that these parameters contain the necessary system information. Matrix N is invertible when W is a positive definite matrix.

$$C_{\hat{x}} = \hat{\sigma}_o^2 (A^T W A)^{-1} \qquad (4.14)$$

$$\hat{\sigma}_o^2 = \frac{\hat{v}^T W \hat{v}}{r} \qquad (4.15)$$

Where:

$\hat{x}$ = a posteriori parameters vector

$C_{\hat{x}}$ = a posteriori covariance matrix of parameters to be estimated

$\hat{\sigma}_o^2$ = a posteriori variance factor given a priori variance factor σ_o^2

r= degrees of freedom

Applying the law of propagation of covariances in the estimation procedure the covariance matrix of the estimated observations $C_{\hat{l}}$ and residuals $C_{\hat{v}}$ can be obtained.

It has been shown that the least squares principle is associated with a stochastic model. However it does not require any a priori knowledge of the residuals' distribution. In the special case where observations are normally (Gaussian) distributed the least squares will present similar properties to the maximum likelihood[18] method. Approaches based on LSE will require being capable of handling of starting values, convergence criteria as well as large volumes of data.

4.3. Perspective camera sensor

The geometric sensor model of digital camera systems is derived from the central perspective projection and it is fundamentally formulated based upon the principle of collinearity. Based on this, the functional model of the bundle method is demonstrated (see section 4.3.2.). An alternative mathematical formulation in the form of direct linear transform (DLT) is given (see section 4.3.3.). In addition, the issues of self-calibration (see section 4.3.4.) as well as starting value estimation (see section 4.3.5.) are addressed.

---

[18] The maximum likelihood method calculates the value of a set of parameters for a given statistic that results in a maximum likelihood distribution.

4.3.1. Basic definitions

From the geometric viewpoint, the basic tasks behind camera modelling can be considered as transformations between image and object spaces. Particularly these relate to 2D-2D, 2D-3D or 3D-3D transforms. This text restricts the discussion to the concepts of orientation, calibration, self-calibration and 3D similarity transform. As a result the following definitions are given:

1. Orientation: Refers to the recovery of the elements of the interior and exterior orientations.

    1.1. Interior orientation (IO): Comprises the interior (inner) camera geometry; that is principal point and camera constant (x_o, y_o, c). Extended model parameters may include radial lens distortion parameters (dR) and additional decentering (dD), together with in-plane affinity and orthogonality terms (a_1, a_2) (see sections 3.2.2. and 4.3.4.).

    1.2. Exterior orientation (EO): Determines the object space coordinates of the perspective centre and the 3D orientation angles. Position (X_o, Y_o, Z_o) and orientation $R(\omega, \varphi, \kappa)$ or quaternions (a, b, c, d) which are recovered with resection procedures (see section 4.3.5.).

2. Calibration: Determines the IO parameters. It models systematic errors (defined as physical deviations from the mathematical model) of all cameras included within a calibration network.

3. Self-calibration: Is an additional parameter estimation procedure. It accounts for the model's systematic errors (including the IO parameters) simultaneously with the system's EO and 3D point locations parameters (usually treated in a bundle estimation approach) (see section 4.3.4.).

4. 3D similarity transform: Is a seven parameter transform between two coordinate systems that spatially registers two 3D point sets. In the general case the independent transformation parameters are given through a rotation matrix $R(\omega, \varphi, \kappa)$, a translation vector $(T_X, T_Y, T_Z)^T$ and an isotropic uniform scale factor (λ). Estimation with three full (X, Y, Z) known reference CPs results in a least squares solution with a redundancy of two degrees of freedom whereas in the case of one height and two full CPs a closed form solution can be obtained.

4.3.2. Mathematical model

The bundle method is formulated as a system of equations. It is derived from perspective collineation connecting 2D image and 3D object spaces through the parameters to be estimated and it is expressed as a component pair of equations (4.16). In geometric terms the collinearity condition can be described as the parametric representation of a line in 3D constrained by the orthogonal distance of the image plane from the camera's image center, that is the focal length (Weinhaus & Devich, 1999). The scale specifies the distance from the projection center along the imaging ray through the point on the image that hits the 3D object point with the rotation matrix describing the direction of this line. This mapping models the elements of the IO and EO parameters, or otherwise the camera's position and orientation (pose), in a 3D coordinate system (datum). Analytically the collinearity condition is given as the x, y pair of equations:

$$x = x_o - c\frac{u}{w} = x_o - c\left[\frac{r_{11}(X-X_o)+r_{12}(Y-Y_o)+r_{13}(Z-Z_o)}{r_{31}(X-X_o)+r_{32}(Y-Y_o)+r_{33}(Z-Z_o)}\right] \quad (4.16)$$

$$y = y_o - c\frac{v}{w} = y_o - c\left[\frac{r_{21}(X-X_o)+r_{22}(Y-Y_o)+r_{23}(Z-Z_o)}{r_{31}(X-X_o)+r_{32}(Y-Y_o)+r_{33}(Z-Z_o)}\right]$$

Where:

x, y = image measurements (mm)

x_o, y_o = principal point locations (mm)

c: camera constant of CCD frame (mm)

u, v, w = numerator and denominator components in the collinearity condition

X, Y, Z = object point coordinates (mm)

X_o, Y_o, Z_o = projection centre OS coordinates (mm)

r_{ij} = (i, j = 1-3) elements of 3D rotation matrix R

Here, in these equations the elements r_{ij} (where i, j = 1-3) express the relative orientation between the image space and the object space coordinate systems. These are the elements of the 3D orthogonal rotation matrix R representing the applied

rotations per view (or frame) related to the 3D object space co-ordinate system. The rotation matrix expressed with regards to its trigonometric functions is given in Appendix A. The functional system of equations can now be formed considering a set of j views (frames or photos), i object points and k cameras within the image network. The bundle method is structured after the notation given in Brown (1974), Granshaw (1980) and Dermanis (1991). Equation (4.17) provides a vector-matrix representation considering one view (six EO parameters), one control point (three spatial coordinates) and one camera (five or more IO parameters) within the functional model.

$$
\begin{bmatrix} x' - x^o \\ y' - y^o \end{bmatrix}_{jik} = \begin{bmatrix} \dfrac{\partial x}{\partial X_o} & \dfrac{\partial x}{\partial Y_o} & \dfrac{\partial x}{\partial Z_o} & \dfrac{\partial x}{\partial \omega} & \dfrac{\partial x}{\partial \varphi} & \dfrac{\partial x}{\partial \kappa} \\ \dfrac{\partial y}{\partial X_o} & \dfrac{\partial y}{\partial Y_o} & \dfrac{\partial y}{\partial Z_o} & \dfrac{\partial y}{\partial \omega} & \dfrac{\partial y}{\partial \varphi} & \dfrac{\partial y}{\partial \kappa} \end{bmatrix}_{jik} \begin{bmatrix} \delta X_o \\ \delta Y_o \\ \delta Z_o \\ \delta \omega \\ \delta \phi \\ \delta \kappa \end{bmatrix}_{j} +
$$

$$
+ \begin{bmatrix} \dfrac{\partial x}{\partial X} & \dfrac{\partial x}{\partial Y} & \dfrac{\partial x}{\partial Z} \\ \dfrac{\partial y}{\partial X} & \dfrac{\partial y}{\partial Y} & \dfrac{\partial y}{\partial Z} \end{bmatrix}_{jik} \begin{bmatrix} \delta X \\ \delta Y \\ \delta Z \end{bmatrix}_{i} + \begin{bmatrix} \dfrac{\partial x}{\partial x_o} & \dfrac{\partial x}{\partial y_o} & \dfrac{\partial x}{\partial c} & \dfrac{\partial x}{\partial k_1} & \dfrac{\partial x}{\partial k_3} \\ \dfrac{\partial y}{\partial x_o} & \dfrac{\partial y}{\partial y_o} & \dfrac{\partial y}{\partial c} & \dfrac{\partial y}{\partial k_1} & \dfrac{\partial y}{\partial k_3} \end{bmatrix}_{jik} \begin{bmatrix} \delta x_o \\ \delta y_o \\ \delta c \\ \delta k_1 \\ \delta k_3 \end{bmatrix}_{k} + \begin{bmatrix} v_x \\ v_y \end{bmatrix}_{jik}
$$

(4.17)

Where:

x', y': image measurements (mm)

x^o, y^o: approximate values to the image measurements (mm)

x_o, y_o: principal point image locations (mm)

c: camera constant of CCD frame (mm)

k_1, k_3: 1st and 3rd order radial lens distortion's coefficients (unitless)

X_o, Y_o, Z_o: exterior orientation elements (mm)

X, Y, Z: 3D point co-ordinates (mm)

ω, φ, κ: 3D rotations (degrees)

The structure of the basic arrays (vectors and matrices) is considered as follows:

b (n x 1): Is the vector of the reduced (observed minus computed) image observations.
A (n x (6j+3i+(5++)k)): Is the desing or Jacobian matrix.

x ((6j+3i+(5++)k) x 1): Is the vector of unkown parameters.

It is noted that the fundamental model arrays are implemented according to structure modelling (parameterization and sparsity of arrays) as well as geometric factors (object space constraints and datum deficiencies).

In symbolic vector form equation (4.17) becomes:

$$b_{ji} = \dot{A}_{ji}\, \dot{x}_j + \ddot{A}_{ji}\, \ddot{x}_i + \dddot{A}\, \dddot{x} + v_{ji} \quad \Rightarrow \quad b = Ax + v \tag{4.18}$$

Where:

$b_{ji}=$ image measurements vector

$\dot{A}_{ji}, \ddot{A}_{ji}, \dddot{A} =$ design matrix of EO, XYZ and IO parameters

$\dot{x}_j, \ddot{x}_i, \dddot{x} =$ vector of exterior, 3D positions and interior parameters

$b=$ image measurements vector

$A=$ design matrix

$x=$ vector of parameters to be estimated

$v=$ residuals vector

4.3.3. Direct linear transform

The collinearity condition is the most flexible analytical functional model. Yet, an alternative is offered by the DLT method (Abdel – Aziz & Karara, 1971). Main advantages comprise the method's ability to handle uncalibrated cameras as well as its independency related to the recovery of starting values. As a result, the projective DLT model has gained ground in general purpose or consumer market CCD video cameras fitted with variable focus and zoom optics, as broadly utilized in computer vision applications. The projective equations of the DLT method are given as follows:

$$x + \Delta x = \frac{L_1 X + L_2 Y + L_3 Z + L_4}{L_9 X + L_{10} Y + L_{11} Z + 1} \tag{4.19}$$

$$y + \Delta y = \frac{L_5 X + L_6 Y + L_7 Z + L_8}{L_9 X + L_{10} Y + L_{11} Z + 1}$$

Where:

x, y= 2D image (mm) or pixel coordinates

L_1-…-L_{11}= algebraic parameters (unitless)

In this equation parameters L_i (where i= 1-11) express the algebraic equivalents of the EO (6 parameters) and IO parameters (3 calibration and 2 affine parameters) and can be calculated by decomposition of the eleven parameter transformation matrix. These, according to Wrobel (2001), can result in high numerical stability when the calculation is based on the knowledge of principal point coordinates and focal length. The model can be treated as a direct or iterative estimation procedure (McGlone, 1989) and can additionally allow treatment of orthogality constraints within the model (Bopp & Krauss, 1978). Its direct, non-iterative implementation can offer a fast computation stage for starting value generation through space resection (see section 4.3.5.) as an example. The offset of this problem is the demand for an increased and well distributed number of CPs (6 per image) and the demand for an increased number of correspondences. In self-calibration problems the DLT is computationally expensive regarding numerical stability and convergence when compared to the collinearity model (Fraser, 2001). It is recalled here that a comparison of the perspective and projective models has been given in section 3.4.1.1. It is generally accepted that the potential for robust estimation, good behaviour in the presence of noise and high precision levels within the bundle method are held responsible for its preference over direct estimation procedures.

4.3.4. Self-calibration

Self-calibration simultaneously estimates IO, EO and 3D object space point coordinates by relatively orienting all bundles of rays without the requirement of any a priori 3D object space knowledge (point coordinates or scale) (Gruen & Beyer, 2001). The model behind self-calibration is an extended parameter bundle estimation method.

A simplified internal projection model is considered to include the 2D image coordinates of the principal point, the camera constant as well as the radial lens distortion terms. These parameters form the IO parameters of a CCD frame as defined

above (see section 4.3.1.). Any departures from 'ideal' projection may result in a systematic error budget which needs to be compensated within the system. It has already been reported (see section 3.2.3.) that camera stability is critical in camera calibration, noting in particular that in CCD systems temporal variations can emerge in principal point location (due to warm up effects), camera constant as well as lens distortion with a change in focus settings; factors that reduce the camera's stability as a result. It follows that selection of the additional parameter model is key as this is the 'natural' model representation. However, it can be sensitive to over-parameterization which is held responsible for ill-conditioning or singularities in normal equation systems. Computation of self-calibration parameters within a bundle method is rigorous and flexible but factors of network geometry and scale variation play a significant role.

4.3.4.1. Additional parameter model

Self-calibration demands the definition of a 'physical model' that describes the internal camera geometry error sources. This is achieved by augmentation of equations (4.16) by a pair of departure functions Δx and Δy which are critical to the self-calibration success and need to be determinable in a given network configuration. A general description of these is given according to the following equations:

$$x_{ijk} - x_{ojk} = -c_{jk} f_{ijk}^{x} + \Delta x_{ijk} \tag{4.20}$$

$$y_{ijk} - y_{ojk} = -c_{jk} f_{ijk}^{y} + \Delta y_{ijk}$$

Where:

ijk= number of points, photos (views, frames), cameras

x_{ijk}, y_{ijk}= image measurements (mm)

x_{ojk}, y_{ojk}= principal point locations (mm)

c_{jk}: camera constant of CCD frame (mm)

f_{ijk}^{x}, f_{ijk}^{y}= numerator and denominator components of collinearity condition

Δx_{ijk}, Δy_{ijk}= departure functions (mm)

Gruen & Beyer (2001) provide a self-calibration model for close range camera calibration. Here, the adopted self-calibration model (utilized for the purpose initialization, pre-measurement and generation of reference image networks) is the one deployed within the in-house photogrammetric bundle adjustment tool (VMS 8.0). This model is based on a ten additional parameter set which can be optionally extended to handle up to twenty internal parameters and it is given as follows:

$$x = x + (\frac{x}{r})dR + dDx + a_1y + a_2x \tag{4.21}$$

$$y = y + (\frac{y}{r})dR + dDy$$

Where:

x, y= 2D image measurements with regards to the principal point location (mm)

$r = \sqrt{x^2 + y^2}$ = radial distance with regards to the principal point location (mm)

dR= radial lens distortion computation model. Polynomial accommodates the 3rd, 5th and 7th power order terms (mm expressed in μm)

dDx and dDy= x and y components of decentering lens distortion (mm)

a_1, a_2= orthogonality and affinity terms of the image correction systems (unitless)

In most CCD camera systems where point location is undefined, the affinity (scale) x factor (attributed to imprecise sensor element spacing) and orthogonality (shear) may be present; hence these equations are expected to be effective.

4.3.4.2. Implementation of additional parameter model

Treatment of self-calibration is based on model purpose (e.g. 3D object location, position and orientation of a moving camera as well as systematic error analysis). It is noted that the system applied here allows inclusion or exclusion of specific terms within the calibration file by adjustment of a parameter's standard deviation (one standard deviation, normal distribution 68%) to a binary value of 1 for parameter estimation or 0 to fix a parameter within the system. In situations where insignificant parameters are observed, these need to be removed from the system and to subsequently check whether their removal reduces the overall network precision in a

progressive approach. To avoid a quasi-column rank deficiency of the design matrix (noting that the identified weights decide the degree of the constraints) Gruen & Beyer (2001) treat additional terms as observed variables. It is common that the additional terms are tested through the analysis of the a posteriori matrix of correlations (see section 4.7.1.). High correlation coefficients (for example those that present correlations $\rho>0.9$) can be damaging if they occur between additional parameters and object space coordinates; such an indication points that these parameters need to be removed from the system. In addition, it is useful to examine the trace of the 3D locations covariance matrix. It is noted that in the case where suppression of a parameter results in an overall RMS increase, this parameter will need to be re-instated.

4.3.5. Starting values

Least squares approaches require knowledge of starting values of the unknown parameters to be estimated. Starting value estimation is critical to the success of the implemented algorithm; however there is not an absolutely correct answer. The nature of starting value estimation is problem dependent. For example starting values can be derived to match a desired answer but fail severely under difficult conditions (e.g. geometric situations of collinearity and coplanarity). It is generally considered that in bundle solutions weak starting values can be absorbed by subsequent estimations. However, it is advisable to avoid such assumptions as similar computations can potentially lead to a slow convergence or extreme solutions. Starting values can be derived based upon initial orientation devices as well as assumptions initiated from geometric considerations; however it is common to utilize sub-estimations (e.g. space resection and forward intersection methods) that base their calculations on a minimum subset or search through the observations. An ideal goal of starting value estimation is system automation. However this is not achieved in generic terms, hence it comprises a very attractive research topic.

4.3.5.1. Space resection and forward intersection

Common initialization procedures refer to the initialization of cameras pose as well as estimation of 3D point coordinates. For perspective sensors these are rendered through classical space resection and forward intersection approaches. The process of

space resection is a 2D to 3D orientation computation of the defined EO parameters (see section 4.3.1.). One example is to perform a closed form estimation from a minimum number of 3 non-collinear CPs. This generally results in four real out of eight discrete solutions (positive and negative) located symmetrically on either side of a plane passes through the given CPs (Wrobel, 2001). To resolve the ambiguity and obtain a unique, correct solution a fourth point is introduced in the computation. Regarding forward intersection, this calculates the 3D object point coordinates given the knowledge of the orientation parameters (IO and EO). The minimum requirement is a pair of two views resulting in a redundancy of one and it is commonly treated as an iterative estimation procedure. Initialization is an old geometric problem; however it has gained attention in the computer vision literature. In fact Dicksheid et al. (2008) utilize a RANSAC[19] sub-procedure to eliminate wrong correspondences and to generate orientation estimates which form the basis of benchmark tests with real bundle adjustment data.

4.3.5.2. Estimation of exterior orientation and 3D point coordinates

To initialize the bundle adjustment an approach based on a three-stage procedure, utilizing an initial EO, a Zeng-Wang (ZW) resection and a forward intersection, has been followed here (see Figure 4.2).
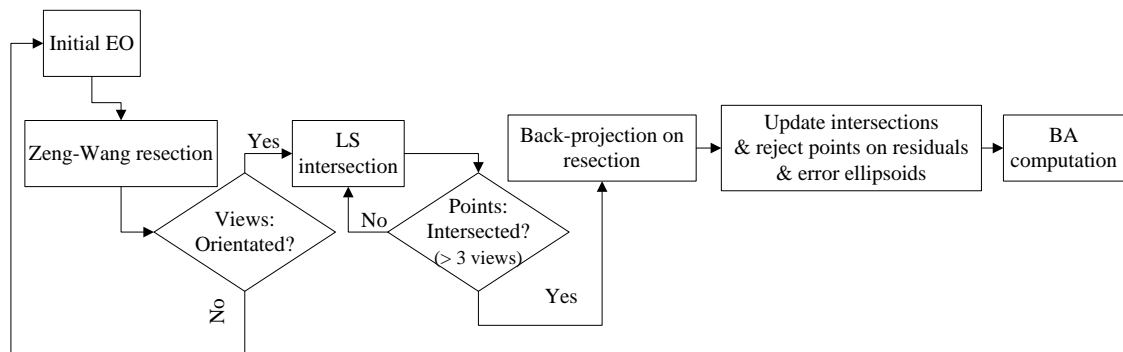
Figure 4.2: Summary of initialization strategy within the VMS 8.0 tool.

The EO parameters are estimated in a local coordinate system which is defined from the measured CPs applying an initial EO estimation which is subsequently updated by

---

[19] RANSAC (RANdom Sample Consensus) algorithm: Originally developed to interpret or smooth data contaminated by gross errors and utilized to solve the Location Determination Problem (LDP) (Fischler & Bolles, 1981).

a resection algorithm (Zeng & Wang, 1992). The stage of forward intersection is initialized as a geometric solution which is updated by a least squares estimation procedure. It is noted that resection and intersection computations use an L1[20] norm robust estimation to reliably remove outliers in the target image measurements. A description of the main stages of the initial exterior orientation (stage 1), Zeng-Wang resection (stage 2) and forward intersection (stage 3) are described as follows:

- Stage 1 - initial exterior orientation: The process is a modified closed form solution that demands a minimum number of three valid CPs in the data. It outputs the initialized EO parameters (Xo, Yo, Zo, ω, φ, κ) and a successful solution allows transfer to the subsequent stage.

- Stage 2 - Zeng-Wang (ZW) resection: It starts from a subset of four CPs locations. The solution is based on three points, utilizing the fourth to resolve the best root ambiguity (storing all answers on a stack). The LSE resection is an L2 norm procedure. As a rejection criterion for outlier rejection, it utilizes either a constant factor (which is set by the user) or a factor scaled to the RMS image residual. A successful solution exports a .log file which updates the EO parameters similarly with the 'stage 1' described above, the average image measurement residual RMS and a summary of the computations (successful, failure, insufficient, non-convergent).

- Stage 3 - forward intersection: A forward intersection procedure is performed estimating the 3D target locations as a combination of a geometric solution which is based on two rays intersection and a subsequent L2 norm least squares multi-view algorithm. The least squares method utilizes an outlier rejection factor based on the image residuals (by default set to 5.0 within the software tool). Implementation of the solution returns a convergent, non-convergent or an erroneously RMS evaluated solution. Forward intersection outputs a .log file which includes the 3D point coordinates with their associated precisions (X, Y, Z, σX, σY, σZ) as well as a descriptive summary of the target counts per view and average RMS residuals per view (frame or photo).

---

[20] L1 norm solution utilizes the magnitude (absolute values) of the residuals, whereas the L2 norm solution minimizes the sum of squares of the residuals.

4.4. Affine camera sensor

A geometric description of the affine camera has been given in section 3.4.2. It is a camera initially introduced by the computer vision community and it establishes the image to object space correspondence linearly. Common treatment has been based upon the homogeneous notation of the involved matrices. Whilst linear algorithms impose assumptions into the projective model, reducing for example solution stability, affine systems are closer to the Euclidean reconstruction. As a result, it can be considered that their linearity allows for 'simpler' system implementation. In addition, affine-based processing has attracted interest in image analysis with focus on automation. Further, in problems that approximate a surface by a plane and in cases where perspective effects are typically small on a local scale, the affine model can provide a good modelling solution.

4.4.1. Background methods

To report methods that implement the affine sensor the following examples are given. Koenderink & Doorn (1991) utilize a geometric approach based on local coordinate frame estimations from a set of known points whereas Tomasi & Kanade (1992) propose a non-local coordinate frame method that utilizes the total number of points within the scene but their solution needs to calculate cases related to rank requirements of matrices. Coordinate datums are defined as the geometric centroid of the cluster of points. Shapiro (1995) extends the previous approaches to a multiple-view solution that it is based on SVD of the matrices involved within the system. From the photogrammetric standpoint the affine sensor has been utilized to initialize orientation (resection) procedures based on the properties of the rotation matrix (Kyle, 2004) or performs long range measurements (Ono et al., 2004) as examples. Weinhaus & Devich (1999) more interestingly utilize a hybrid joint perspective and parallel projection mathematical model, of homogeneous and non-homogeneous equations, that can be optionally adapted (between perspective and affine) for the absolute purpose of mapping textures onto planar polygons. Finally in image analysis Mikolajczyk et al. (2005) give an overview of the methods that perform affine covariant region[21] detection. It has been shown that whilst the affine regions are

---

[21] Region: Is a set of pixels on any image subset. According to Mikolajczyk et al. (2005) region boundaries need not to correspond to changes in colour or texture.

variant, the normalized image pattern they cover together with the derived feature descriptors are typically invariant. In particular, the authors studied a number of six methods for detecting such regions on images and they subsequently assessed their performance related to texture, different transforms, variation in viewpoint as well as illumination.

4.4.2. Mathematical model

In non-homogeneous vector notation the affine sensor model recovers the image to object space correspondence through a mathematical relation (see section 3.4.2. - equation (3.5)) that can be thought of as a simplified collinearity condition (see equation (4.16)). In these equations no principal point exists; the projection centre is located at infinity. The mathematical model involves the parameters: 2D projective translations (tx, ty), 3D photo rotations (ω, φ, κ), and scale factor (s) regarding orientation and 3D point coordinates (X, Y, Z) regarding object space recovery. It is linear in relation to the unknown parameters and it can be derived from the standard collinearity condition by substitution of the variant image scale with an invariant scale factor which is the approximation of the system's nominal magnification factor. The initial mathematical formulation for the affine sensor can be augmented to accommodate departures from the ideal projection. To initially build and test the model, a simplified third-power radial lens distortion polynomial term (see section 3.2.2.1.) is inserted for the purpose of camera calibration. The model is implemented as a system of multiple views that are arranged in a network configuration in order to process the bundle algorithm.

The mathematical system is formed considering j views (frames or photos), i points and k cameras. Let b (n x 1) be the vector of reduced (observed minus computed) image observations, A (n x 5j+3i+2k) be the design or Jacobian matrix and x (5j+3i+2k x 1) be the vector of unknown parameters. In vector-matrix representation the affine mathematical model is given as follows:

$$\begin{bmatrix} x - x^o \\ y - y^o \end{bmatrix}_{jik} = \begin{bmatrix} \dfrac{\partial x}{\partial s} & \dfrac{\partial x}{\partial \omega} & \dfrac{\partial x}{\partial \varphi} & \dfrac{\partial x}{\partial \kappa} \\ \dfrac{\partial y}{\partial s} & \dfrac{\partial y}{\partial \omega} & \dfrac{\partial y}{\partial \varphi} & \dfrac{\partial y}{\partial \kappa} \end{bmatrix}_{jik} \begin{bmatrix} \delta s \\ \delta \omega \\ \delta \varphi \\ \delta \kappa \end{bmatrix}_{j} + \begin{bmatrix} \dfrac{\partial x}{\partial tx} \\ \dfrac{\partial y}{\partial ty} \end{bmatrix}_{jik} \begin{bmatrix} \delta tx \\ \delta ty \end{bmatrix}_{j} + \tag{4.22}$$

$$+ \begin{bmatrix} \dfrac{\partial x}{\partial X} & \dfrac{\partial x}{\partial Y} & \dfrac{\partial x}{\partial Z} \\ \dfrac{\partial y}{\partial X} & \dfrac{\partial y}{\partial Y} & \dfrac{\partial y}{\partial Z} \end{bmatrix}_{jik} \begin{bmatrix} \delta X \\ \delta Y \\ \delta Z \end{bmatrix}_{i} + \begin{bmatrix} \dfrac{\partial x}{\partial k_3} \\ \dfrac{\partial y}{\partial k_3} \end{bmatrix}_{jik} [\delta k_3]_k + \begin{bmatrix} v_x \\ v_y \end{bmatrix}_{jik}$$

Where:

x, y= image measurements (mm)

x^o, y^o= approximate values of the image measurements (mm)

s= scale factor (unitless)

ω, φ, κ= 3D rotations (degrees)

tx, ty= 2D projective translations (mm)

X, Y, Z= 3D point co-ordinates (mm)

k_3= 3rd power term of radial lens distortion polynomial (unitless)

v_x, v_y= x, y components of vectors of residuals (mm)

This is the basic formulation behind the affine sensor model. It is noted that to implement the bundle algorithm in an iterative LSE approach (see section 4.2.) the design matrix is formed from the partial derivatives of the parameters to be estimated. The calculation of these is given in Appendix A. Regarding model structure as well as method development and implementation these comprise the core theme of chapter 5 where they will be studied extensively.

4.5. Datum constraints

From the photogrammetric standpoint, image networks perform on pure image measurements. As a result, it is generally required that a datum is determined from seven coordinate system parameters (see Figure 4.3 for an example of a datum definition for a small cluster of points).
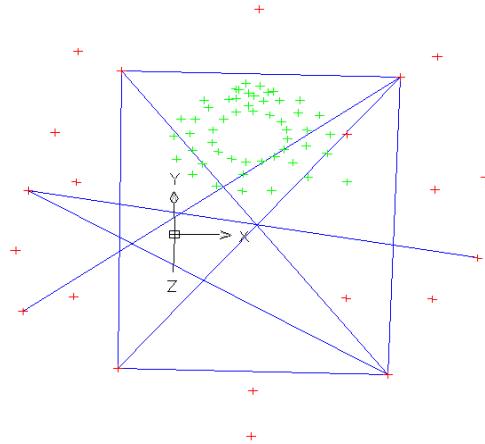
Figure 4.3: Local datum of a geometric centroid (red: CPs, green: TPs, blue: scales).

Particularly in the case where no configuration defects are present within a photogrammetric system, datum defects can be resolved by computation of the following coordinate system parameters:

λ= isotropic scale (unitless)

X_o, Y_o, Z_o: 3D translations of the origin (mm)

ω, φ, κ: 3D orientation angles (degrees)

Datum identification is a problem of a 3D similarity transform (see section 4.3.1.). In particular it results in a linear dependence of the columns of the design matrix A and thus in a linear dependence of the columns with the rows in the normal equations matrix. To solve the datum problem minimum, inner or stochastic (external) constraints are introduced in the functional model. For example minimum constraints generally require fixation of two control points (X, Y, Z) with an additional point known in the depth direction (Z) that form what is termed a zero variance reference base (Cooper & Robson, 1996). Inner constraints, which are considered as a special case of minimum constraints, define the seven datum elements with a geometric centroid (X_G, Y_G, Z_G) (see section 4.5.1.) whereas stochastic or external constraints augment the observation equations with a set of pseudo-observations which are formed from the identified control, the quality of which is given by their stochastic model (see section 4.5.2.).

4.5.1. Inner constraints

The inner constraints method (Granshaw, 1980; Cooper & Cross, 1988; Cooper & Cross, 1991) identifies a geometric centroid to be the center of mass of the defined control point data. The constraint equations augment the functional model by seven additional equations. Particularly the 3D centroid is identified as follows:

$$X_G = \Sigma \frac{X_i}{n}, \; Y_G = \Sigma \frac{Y_i}{n}, \; Z_G = \Sigma \frac{Z_i}{n} \tag{4.23}$$

Where:

X_G, Y_G, Z_G = 3D coordinates of the datum centroid (mm)

X_i, Y_i, Z_i= 3D coordinates of point i (mm)

n= number of CPs

The constraint equations state:

1. Constraint-position: The position of the centroid remains constant.
2. Constraint-rotation: The average direction of all points from the centroid remains constant.
3. Constraint-scale: The average distance of all points from the centroid remains constant.

In mathematical form:

$$\Sigma\delta X_i = 0, \; \Sigma\delta Y_i = 0, \; \Sigma\delta Z_i = 0 \tag{4.24}$$

$$\Sigma(Z_i\delta Y_i - Y_i\delta Z_i) = \Sigma(X_i\delta Z_i - Z_i\delta X_i) = \Sigma(Y_i\delta X_i - X_i\delta Y_i) = 0 \tag{4.25}$$

$$\Sigma(X_i\delta X_i + Y_i\delta Y_i + Z_i\delta Z_i) = 0 \tag{4.26}$$

Subsequently, the constraint matrix G is constructed by linearization of the seven datum (3 translations, 3 rotations, 1 scale) equations as follows:

$$G_P^T = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & . & . & . & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & . & . & . & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & . & . & . & 0 & 0 & 1 \\ 0 & -Z_1 & Y_1 & 0 & -Z_2 & Y_2 & . & . & . & 0 & -Z_P & Y_P \\ Z_1 & 0 & -X_1 & Z_2 & 0 & -X_2 & . & . & . & Z_P & 0 & -X_P \\ -Y_1 & X_1 & 0 & -Y_2 & X_2 & 0 & . & . & . & -Y_P & X_P & 0 \\ X_1 & Y_1 & Z_1 & X_2 & Y_2 & Z_2 & . & . & . & X_P & Y_P & Z_P \end{bmatrix}$$

(4.27)

Where:

$X_1, Y_1, Z_1, \ldots, X_P, Y_P, Z_P$ = 3D coordinates of CPs (mm)

The linearized form is now given:

$$\begin{bmatrix} A_1 & A_2 \\ G^T & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} l \\ 0 \end{bmatrix}$$

(4.28)

Where:

A_1 = design or Jacobian matrix corresponding to 3D point coefficients

A_2 = design or Jacobian matrix corresponding to exterior orientation

G^T = constraint matrix

x_1 = vector of the unknown 3D point parameters within the model

x_2 = vector of exterior orientation parameters within the model

l = observations vector

The resultant normal equations matrix N is symmetric but not positive definite. Inner constraints demand appropriate routines for inverse matrix computation (see section 5.5.2.2.). The generalised matrix inverse known as the Moore-Penrose inverse or pseudo-inverse offers the solution to this problem (Cooper & Cross, 1991). Inner constraints result in the minimization of the trace of the a posteriori covariance matrix $C_{\hat{x}}$.

4.5.2. External constraints

The method of external constraints requires that a set of control point data is known with a high degree of precision. Such reference data can be derived from a

measurement procedure that is considered to be of high order precision (such as a prior robust bundle estimation or CMM measurement).

Given the following definitions:

x_P= x prior vector of control point data.

C_{xP}= covariance matrix of control point data.

x^o_P= vector of approximate values (pseudo-observations) of control point data.

Equation (4.29) is now formed as follows:

$$\begin{bmatrix} A_n & A_P \\ 0 & I \end{bmatrix} \begin{bmatrix} x_n \\ x_P \end{bmatrix} = \begin{bmatrix} 1 \\ x^o_P \end{bmatrix} \& Q = \begin{bmatrix} Q_l & 0 \\ 0 & Q_{xP} \end{bmatrix} \qquad (4.29)$$

Where:

A_n= design matrix of new elements

A_P= design matrix of prior estimated elements

x_n= vector of new elements

x_P= vector of prior estimated elements

x_P^o= vector of approximate values of prior estimated elements

l= observed minus computed vector (reduced observations vector)

Q_{xP}= cofactor matrix associated to the covariance matrix C_{xP}

The normals equations matrix N is positive definite (see section 5.5.2.2) and in the first iteration, the right-hand vector of the pseudo-observations for control points (formed by the first order increments to the approximate values) becomes null. It is noted that identification of control is considered as highly critical in the precision aspect; precision can be degraded for poorly identified control point data.

4.6. Image network geometry

Image network geometry design is the problem of planning the camera's locations, image and object space configurations enclosing the object of interest in order to generate accurate measurements. Figure 4.4 illustrates an example of a convergent

image network geometry that is formed for six views imaging a 3D volume for the perspective camera sensor.
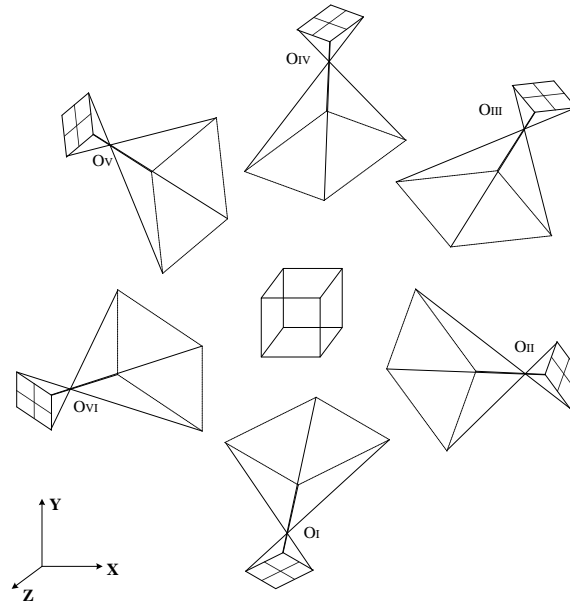
Figure 4.4: Representation of a close range image network.

4.6.1. Network geometry factors

The main factors that govern the design of an imaging network relate to: geometric strength (such as base-to-depth (B/D) ratio, angle of convergence, scale recovery), object space properties (such as spatial resolution, visibility, occlusions, incident viewing angle on features), image space properties (such as resolution) as well as number of intersected angles and redundancy factors. Design of image networks is an established problem and it is commonly treated with heuristic simulation (Fraser, 1996). To avoid singularities and instabilities several approaches exist and these may include utilization of expert systems (Mason, 1995) or establishment of optimal configurations such the discrete camera placement based on a viewing sphere for example (Sakane et al., 1987). Fraser (1996) follows the classification of the basic orders of network design as introduced by Grafarend (1974):

1. Zero order design (ZOD): Is related to the establishment of a datum given a configuration matrix (design matrix) associated with its stochastic model (covariance matrix).

2. First order design (FOD): Defines a configuration matrix (design matrix), that for a given weight matrix, yields a cofactor matrix which satisfies specified precision criteria.

3. Second order design (SOD): Is related to the optimization of the image observation precision problem.

4. Third order design (TOD): Refers to the densification problem.

Given an image measurement system and its precision, the network geometry is initialized to test the configuration and the achieved degree of precision. The network is tested against its requirements and in case of failure the covariance matrix is re-scaled or the configuration (FOD design) is modified to resolve the ambiguity, otherwise redesign is required.

4.6.2. Network geometry examples

Design 'rules' for optimal network geometry usually refer to geometric stability (see section 3.2.3.), wide baselines, wide-angled lenses and large CCDs in order to enhance object space coverage and measurement precision. Additional requirements relate to increased point redundancy and four imaging rays visibility to ensure high internal reliability over the three views case; as four rays enable detection and localization of gross image measurement errors. 3D test-fields imaged under strong convergent geometries are considered as optimal; however not absolutely established. In instances where convergent geometries do not hold, 3D arrays are essential for the recovery of the camera constant (Gruen & Beyer, 2001). In high resolution, controlled imaging engineering cases, convergent image networks within an imaging cone of 45° centred on the object and supported from 90° rotated views on their optical axis are usually built. In such cases typical accuracies at the order of 1:100,000 of the object's primary dimension are reported (Fraser et al., 1995). A range of practical network design examples can be found in El-Hakim et al. (2003).

In most instances practical limitations like physical obstructions (occlusion, visibility, direction of lines of sight with relation to features location and features distribution) and imaging constraints (resolution, DOF and FOV) can bound the network configurations to empirical establishment. For example in architectural large scale

projects (Van den Heuvel, 2003) where buildings are composed of planar primitives, corner images are acquired to tie and strengthen the sub-networks particularly when overview images are included in the computations. In sequence projects Guidi et al. (2003) utilize for example free-form objects placed on a turntable applying controlled imaging configurations. An example of such a measurement object is given in Figure 4.5 where an alabaster sculpture is illustrated (see section 2.2.2.). Image network geometry configurations that were designed for the purpose of method development and testing will be outlined throughout chapter 6 whereas some descriptive networks are given in section 6.2.2.4.

Figure 4.5: An alabaster sculpture located on a turntable.

4.7. Quality control

To assess the quality of adjustment problems a series of tests that evaluate the goodness of the solutions are usually performed. In the case of LSE cost functions, quality estimation is related to three fundamental considerations; these are defined as follows:

1. Accuracy: This term refers to the degree of closeness of an estimate to its parameter. In conventional terms, accuracy is considered as the degree of closeness to the 'true' value (Mikhail, 1976). As a result, accuracy is connected with the degree of systematic error sources, if any, in the data. Measures that address accuracy are usually expressed with regards to ground truth (e.g. high precision reference data). VDI/VDE 2634 guideline comprises such an example for accuracy evaluation of optical 3D measuring systems. It is composed of parts 1, 2 and 3 that correspond to the imaging systems with point by point probing, optical-based scanning systems and multiple view area-based scanning systems respectively (VDI/VDE, 2009). Within this text

accuracy measures are determined according to the purpose of assessment and they are given with regards to reference measurements.

2. Precision: This term refers to the degree of conformity among a set of observations of the same random variable. The spread (or dispersion) of the probability distribution is an indication of precision (Mikhail, 1976). As a result, it is linked with the presence of random errors in the measurements and it is associated with the covariance matrix analysis. The direct estimator associated with precision is the variance or standard deviation of a random variable. The a posteriori covariance matrix of the estimated parameters is a highly rich source of quality information and allows the assessment of model precision as well as detection of systematic errors within the system. Within this text the term precision (or quality) denotes the standard deviation of the associated entity.

3. Reliability: This term refers to the presence of gross errors (blunders) in the data (Cooper & Robson, 1996). Such errors are difficult to detect due to the nature of the minimization of the quadratic cost function (see section 4.2.). For example if a gross error occurs in the measurement data this will contaminate the measurements in order to minimize the target function, posing its identification difficult. Internal reliability relates to the ability of performing self-consistency checks for outlier detection. External reliability specifies the degree to which undetected outliers can affect the estimated parameters. Within this work outlier detection and elimination have not been part of the up-to date implemented algorithm (see section 5.5.).

The above three factors are critical to the evaluation and the applicability of the solution applied to a measurement problem; hence a set of sufficient measures are commonly computed and analyzed to ascertain accuracy, precision and reliability as reported in section 4.7.1. It is stated that this text follows the terminology followed in photogrammetry. For clarification purposes two additional definitions are given: 'Robustness' is referenced with regards to the presence of outliers within a system and 'consistency' characterizes an estimator when it is said to converge in probability to the parameter. However these terms have been used only in implicit terms (without an explicit utilization or derivation of associated measures) throughout this text.

4.7.1. Quality indicators

A review on the quality indicators and performance evaluation applied in bundle adjustment methods is given in Triggs et al., (2000) and Foerstner, (2005). Here the problem of quality assessment metrics is discussed outlining the typical estimators obtained from a LSE procedure.

Accuracy is addressed through measures that ensure that systematic effects are not present in the data. Examples of accuracy evaluation measures are the residual vectors (their magnitude and direction need not to present undesirable systematic patterns). A measurement is considered as 'accurate' with relation to a reference dataset, hence comparative differences need to refer to standards designed for this purpose (see for example VDI/VDE, 2009).

Precision is expressed as a quality measure associated with the covariance matrix analysis. Starting from the model formation an a priori variance factor σ_o^2 is defined to be commonly equal to unity; this implies that the contribution of the image measurements in the estimation procedure is pointed by their stochastic model σ_i^2. An example of the order (magnitude) of this value is given in the bundle software VMS 8.0 where image networks are processed with a specified image observations quality $\sigma_i^2 = 0.25\mu m$ ($\sigma_i = 0.5\mu m => \sigma_i^2 = 0.25\mu m$). The a posteriori variance factor can then be evaluated (see section 4.2.) to check if the a priori stochastic model meets the requirements. This test refers to the T statistic ($T = v^T W v$) that tests the null hypothesis where the variance factor is equal to unity against an alternative hypothesis. If the null hypothesis is rejected the precision of the measurements has not met the design, there is a gross error on at least one measurement or a systematic error occurs (Cooper & Robson, 1996). In the absence of gross and systematic errors the a posteriori covariance matrix scales the computed a posteriori covariance matrix of the estimated parameters. This is a significantly self-contained matrix in that it encloses highly rich information for quality evaluation. It describes each of the components that constitute a camera problem such as additional internal parameters, external camera elements (position and attitude) and finally 3D point coordinates. Sub-matrices can be easily derived if each of these three components need to be considered. To evaluate object

space quality the square root of the mean variance is calculated from a set of 3D points within the image network according to:

$$\bar{\sigma}_{\hat{x}i} = \sqrt{\left[\frac{\left(\text{trace } C_{\hat{x}i}\right)}{3i}\right]}$$

(4.30)

Where:

$\bar{\sigma}_{\hat{x}i}$ = mean a posteriori standard deviation of a number of i 3D CP data

$C_{\hat{x}i}$ = a posteriori covariance matrix of a number of i 3D CP data

In addition, object space can be inspected from the ratio of the largest to the smallest eigenvalue ($\lambda_{max}/\lambda_{min}$) in relation to unity. The 3D standard error ellipsoid can be calculated from the eigenvalues (magnitude) and eigenvectors (directions). In an alternative case computation of the 2D point ellipses from an algebraic calculation is performed. Regarding the internal calibration parameters, these can be checked against their significance from zero based again on the T statistic. If over-parameterization occurs then potential insignificant parameters need to be removed from the data and LSE needs to be re-evaluated. Further to these, analysis of the matrix of correlation coefficients is highly important in that it indicates the mutual variation between two random variables. Parameters that present high correlation coefficients (>0.9.) will need to be removed from the model.

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = E\{\frac{(\bar{x} - E(\bar{x}))}{\sigma_x} \frac{(\bar{y} - E(\bar{y}))}{\sigma_y}\}$$

(4.31)

Where:

ρ_{xy} = correlations coefficient of parameters x, y (unitless)

σ_{xy} = covariance between two parameters x, y (mm; converted in μm)

σ_x = standard deviation of parameter x (mm; converted in μm)

σ_y= standard deviation of parameter y (mm; converted in μm)

Reliability measures usually refer to the detection and elimination of outliers from the system. Flagging image measurements that are affected by gross errors is one possibility but alternative approaches are implemented by applying a down-weighting scheme followed by statistical tests that check the a posteriori variance factor, subsequent measurements corrections and re-estimation. This procedure is repeated until the a posteriori variance factor is unity and all contaminated measurements have been assigned with large standard deviations that in effect minimize their contribution in the LSE process. The method of outlier elimination based on residuals testing is known as data snooping (Baarda, 1968).

4.8. Summary

In summary this chapter reviews the method of multi-view modelling from the photogrammetric perspective. It starts with an overview of the bundle adjustment method supported by its main features related to the method's background, main attributes as well as state of the art. Subsequently the least squares estimation and associated mathematics are given. The perspective camera sensor is presented from the geometric viewpoint linked with the important aspects of self-calibration and starting value estimation. Subsequently the affine camera sensor which comprises the central topic of this thesis is covered with a brief description of background methods and its mathematical model formulation. The chapter additionally covers three important considerations for modelling multi-view problems; that is the aspects of datum, image network geometry and quality control.

5. Affine multi-view modelling

This chapter presents the 'affine multi-view modelling' algorithm developed within this research work. The chapter is organized into six main sections starting with a general overview as well as method description. Subsequently, the implementation tasks are listed in a methodological approach as the key stages of the algorithm are developed covering model structure, derivation of starting values and bundle adjustment framework in detail. Finally, the chapter closes with the design and initiation of the implemented method reporting a simulation example.

5.1. Method overview

The algorithm presented here is concerned with the investigation of the affine sensor in the close range. It is designed and developed for the multi-view case that solves the fundamental photogrammetric tasks; that is (a) calibrates, (b) orientates the cameras and (c) simultaneously estimates 3D point geometry. The algorithm needs to be capable to allow statistical error propagation and therefore to assess the method with the typical quality measures utilized for the purpose of performance evaluation (see section 4.7.). It is developed for close range convergent imagery arranged in a network configuration (see section 4.6.) and it is based upon the assumption that initialization is performed from sparse, artificially targeted, point-based geometric structures prior to method processing and evaluation. In the contexts of design, development and implementation of such a multi-view framework, there are some key considerations that need to be accommodated; these are formulated as follows:

1. Input data handling: Is related to the reading of the input data files with regards to affine sensor modelling as well as the design of the basic model structures.
2. Algorithm modelling: Depicts the ability of the algorithm to accommodate camera modelling factors (internal calibration and external orientation parameters) and to deal with a number of geometric factors which arise from point data treatment such as visibility handling, point coordinates contribution (CPs or TPs) in the network computations, particularly in the geometric datum

estimation (see section 4.5.). These are the primary geometric factors that are related to the 3D object positioning problem from multiple viewpoints.

3. Affine sensor modelling: Signifies the ability of the algorithm to represent, given its theoretical basis, the physical reality, particularly regarding model scale and internal camera geometry.

4. Checking and interpretation: Are issues related to the ability of the system to incorporate statistical error propagation within the method which will subsequently enable data quality assessment through the associated quality measures.

Overall decomposing this chapter, the above factors are addressed by firstly illustrating an outline of the developed algorithm in a methodological approach (see section 5.2.). Method description is followed by four fundamental sections. The structure of the model is described with the affine model formation and stochastic model initialization (see section 5.3.). Noting that the key factors presented above are valid given that the method has been initialized with 'good' starting values, another issue is the investigation of starting value derivation based on direct or indirect estimation procedures (see section 5.4.). Subsequently, the bundle adjustment [22] framework, which is the core method of this thesis, is presented with its implemented intermediate stages (see section 5.5 and subsections). The chapter closes with a description of the implemented method based upon a simulation example that was created for the purpose of initial model formation and development. This synthetic example also supports theoretical proof of the method based upon three geometric viewing cases; that is two-view, three-view and seven-view geometric arrangements that are covered in sections 5.6.2., 5.6.3. and 5.6.4.

5.2. Method description

The method of multi-view modelling was treated in the context of developing a bundle adjustment framework that seeks the answers to the primarily stated key questions (see section 1.3.). It is restated here that prior concern is (a) to derive precise 3D object measurements, (b) to check the effect of parallel projection, scale factor modelling and local coordinate frame definition within the method and (c) to

---

[22] Within this text the term bundle adjustment is denoted as BA only in tables of results.

assess method quality over precise conventional reference measurements. The bundle adjustment starts from a set of input data that are generated externally to the developed method (see section 6.4.1.). For the purpose of initialization a set of sparse 3D targeted structures can be utilized provided that they present sufficient content (intensity) as required from the employed 2D image measurement method as well as that they are highly geometric in their nature to avoid potential geometric degeneracies. In addition, the significant matters of point visibility, redundancy within network geometry are key to the problem as they affect the strength and adaptation of the algorithm developed here to the application of complex object measurement, for example.

First consideration is the initialization of the method through the derivation of suitable approximate values that are optimal in that they assist the algorithm to converge rapidly and closely to its true answer. Next, the affine bundle adjustment is run accounting for modelling issues like visibility, validity, parameter inclusion (scale and interior orientation parameters). The iterative process terminates when the appropriately set criteria have been reached. Consequently, the successful solution extracts an overall report derived from the algorithm providing the affine sensor parameters accompanied with their statistical estimators.
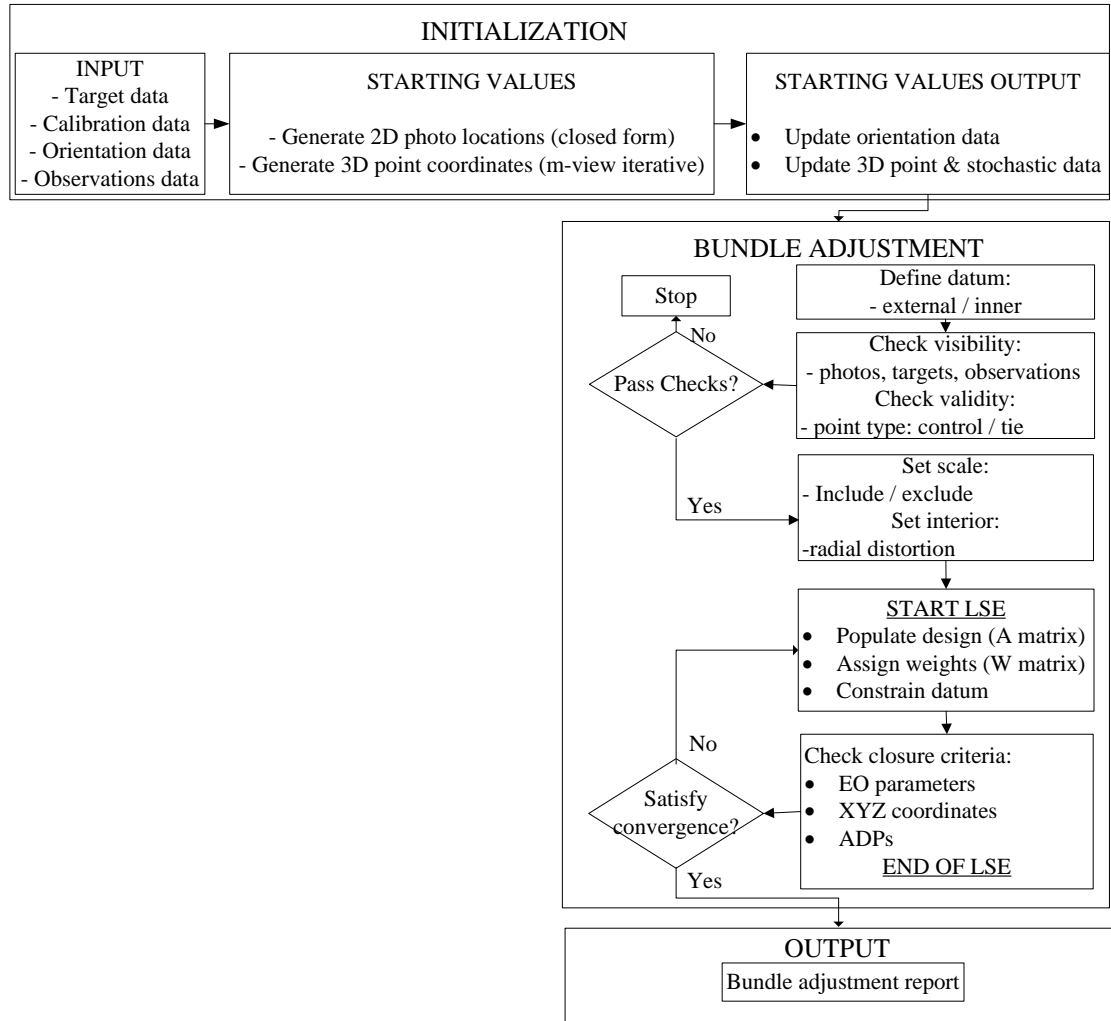
Figure 5.1: Descriptive overview of affine multi-view algorithm.

Figure 5.1 illustrates an outline of the overall method description according to its design, development and implementation applied within this research work. It is composed of three building blocks. The first, initialization of the method, stage starts from three general steps that read input data files, compute starting values and on success update the initialization data files. Subsequently, the key steps of the new affine bundle adjustment algorithm involve input data file handling, datum definition, internal checking of entities (visibility and validity), model parameter set-up, calculations and set up of sensor scale and simplified interior orientation, iterative weighted least-squares estimation procedure, convergence checking and finally output of the associated method report.

To develop and implement the affine algorithm a set of data were read as input or exported as output (.log files). These data files were of ASCII format and a sample of

these can be found in Appendix B. Supported menus that draw the processing framework have been outlined in the same Appendix B. Accordingly, the individual stages of the algorithm regarding initialization strategy as well as bundle adjustment will be addressed analytically in the following sections (see section 5.4. and section 5.5.).

The algorithm has been designed, developed, implemented and subsequently tested for the purposes of this research work. It is noted that no similar method that processes multi-view data generated from affine imagery in the close range is currently available. This point emerges the significance of the initial investigation and successive algorithm treatment presented here, for a potential future complete modelling method from affine images. It is re-iterated here that image measurement and initialization were performed in the in-house photogrammetric processing tool VMS 8.0 (see section 1.5.) externally to the developed framework. For the purpose of this research work code has been developed in C/C++ within Microsoft Visual Studio 2005 (subsequently upgraded to Visual Studio 2008).

5.3. Model structure

On the basis of implementation of a multi-view algorithm for the purposes of camera calibration and orientation and simultaneous location of sparse targeted 3D objects with statistical error propagation, algorithm structure and model geometry are of key significance. Model structure is related to the organization of the main model arrays for data population and location as well as treatment of model geometry. As a reminder, the fundamental mathematical model behind the affine sensor model has already been given in the section that renders a descriptive overview of affine camera modelling methods (see section 4.4.2. for mathematical model formation). This section provides the design of the affine model structure in relation to its fundamental arrangement.

5.3.1. Affine model structure

The prior objective of such a multiple-view task is to intersect in 3D space the lines of sight generated from imagery acquired with an affine close range camera sensor. A simplified example of such an imaging network is illustrated in Figure 5.2. This

example utilizes six views that were generated from an affine camera for the particular case that its scale factor is equal to unity. To structure the affine model in a bundle adjustment framework of prior importance is the design matrix structure. It is noted that the input data are organized as four data files (see Appendix B) that correspond to the input 3D target data, calibration data, photo orientation data and observations data initialized at a prior stage to the main developed framework (see section 5.3.3 for data structures outline).
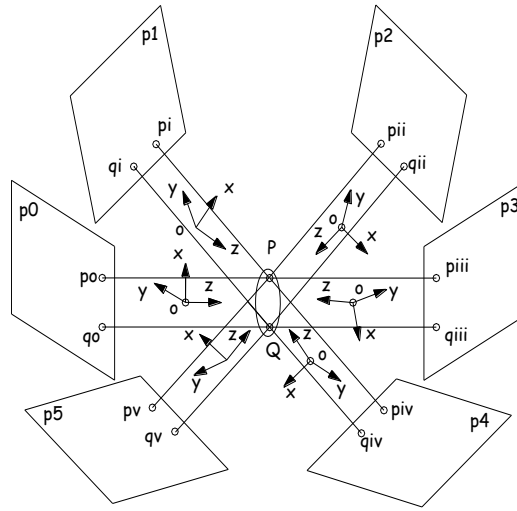
Figure 5.2: Multi-view intersection of affine images. Parallel lines of sight link 2D image and 3D object spaces. Notation: PQ= object points, p_0-p_5: photo$_0$ – photo$_5$, $p_{i-v}q_{i-v}$= image measurements, oxyz= 3D coordinate system of camera.

Following the structure of the input data, the model was structured starting from the population of the number of physical cameras k within the model followed by the number of views (images or frames) j and then by the number of targets (given their initialized 3D coordinates) i that contribute within the image network. Whilst the general problem requires that the parameters are populated and located on the mathematical basis of the affine sensor model (see section 4.4.2.), a key issue is the choice of the system parameters that correspond to the physical reality and do not for example result in an over-parameterized system. As a result, here the affine sensor is modelled accommodating a simplified interior orientation model (third power term of radial lens distortion polynomial).

The parameters to be estimated are grouped according to their type in order to assist population of the required arrays in a columnwise order. The parameters are grouped

together as pseudo-exterior orientation parameters, that is projective equivalents of the principal point components (tx_j, ty_j) populated per view (two parameters per view j) and 3D orientation angles $(\omega_j, \varphi_j, \kappa_j)$ populated again per view (three parameters per view j). The next group of parameters includes the 3D target coordinates (X_i, Y_i, Z_i) which are populated according to their visibility on each view (frame or image) (three parameters per target i). In cases where a target is not visible or its measurement is absent from an image its associated location is left void. In addition, the system is populated with one global scale factor (s_k) and the third-power term of the radial lens distortion polynomial (k_{3k}) which is modelled centrally from the computed image centre (tx, ty) per camera k. These are image invariant parameters provided that one physical camera system has been employed for data acquisition. It follows that scale factor and additional interior sensor geometry parameters are camera specific. It is re-iterated here (see section 4.4. and section 4.5.) that the model structure of the design matrix A for any set of j views (frames or photos), i points and k cameras is formulated as: A $[[(2*j*i*k) + c$ (where: c= $7_{(inner\_datum)}$ or $3*CP_{(external\_datum)})] \times [(2*j) + (3*j) + (3*i) + (1*k) + (1*k)]]$. The outline of the model parameters is drawn in Figure 5.3 whereas the structure of the design matrix (see Appendix A) with regards to parameter location for a specific example is illustrated in the subsequent Figure 5.4.

	Photo orientation parameters						3D point coordinates				Scale	ADP
$Ph_{j=1-m}$ $T_{i=1-n}$	$\dfrac{\partial fx}{\partial tx}$			$\dfrac{\partial fx}{\partial t\omega}$	$\dfrac{\partial fx}{\partial t\varphi}$	$\dfrac{\partial fx}{\partial t\kappa}$	$\dfrac{\partial fx}{\partial X}$	$\dfrac{\partial fx}{\partial Y}$	$\dfrac{\partial fx}{\partial Z}$		$\dfrac{\partial fx}{\partial s}$	$\dfrac{\partial fx}{\partial k_3}$
		$\dfrac{\partial fy}{\partial ty}$		$\dfrac{\partial fy}{\partial t\omega}$	$\dfrac{\partial fy}{\partial t\varphi}$	$\dfrac{\partial fy}{\partial t\kappa}$	$\dfrac{\partial fy}{\partial X}$	$\dfrac{\partial fy}{\partial Y}$	$\dfrac{\partial fy}{\partial Z}$		$\dfrac{\partial fy}{\partial s}$	$\dfrac{\partial fy}{\partial k_3}$

Figure 5.3: Affine sensor parameters.
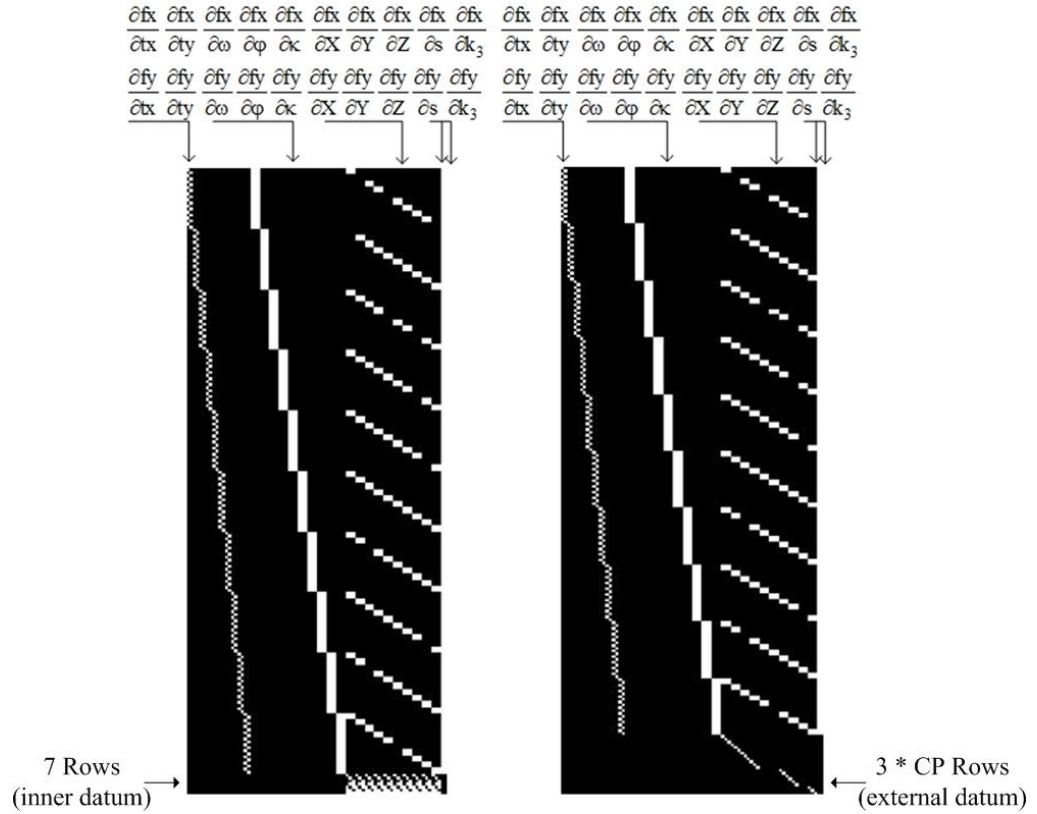
Figure 5.4: Design matrix structure - inner datum (left) and external datum (right).

Specifically Figure 5.4 illustrates the design matrix configuration for a synthetic small sample of j= 10 views, i= 10 3D targets and k= 1 cameras to illustrate model parameter location and structure. For this set the inner constraints method is composed of a data size of 207 equations x 82 unknown parameters [A (207 x 82)] whereas the external constraint datum is comprised of a data size of 221 equations x 82 unknown parameters [A (221 x 82)]. The matrix visualizes the numerical data in a binary representation. Where for example data exist the selected location is plotted in white (intensity value= 100) whilst black indicates absence of data (intensity value= 0), hence the corresponding targets are not visible on the specified image (the targets are occluded or missing). For example the presented data demonstrate minimum visibility where one image contains seven measured points (for initialization and computation of photo locations) whereas ten point data are visible or otherwise located in two views (for initialization and computation of 3D point locations). Maximum visibility is present in the case where seven images include nine measured point targets noting that nine point target data are viewed and hence located from seven views. These numbers are specific to this example. It becomes apparent that the model needs to be flexible with regards to its adaptation to include or exclude those

parameters that may have a physical effect on sensor modelling (sensor scale and internal geometric distortions) affecting recovery of 3D geometry.

The model is also capable of dealing with the important datum issue. The inner constraints method (see section 4.5.1.) augments the system's equations by seven additional rows whilst the external constraints method (see section 4.5.2.) extends the system's equations by three (where the index three corresponds to the known number of 3D coordinates for example) multiplied by the number of valid CPs pre-set (flagged) for datum definition. In the example presented here seven CPs have been assumed, resulting in twenty-one additional equations. It is further noted that population needs to treat the implemented arrays with validity checks; that is to check the number of flagged CPs and TPs that are valid according to their measurement for example in a minimum set of two views (forward intersection requirement). Flagging for parameter inclusion or exclusion within the system as well as identification of data type (CP flag = 7 or TP flag = 0) are central to the sections 5.4. and 5.5. where these will be addressed.

5.3.2. Stochastic model initialization

In the case of implementing 2D image measurements, the applied measurement method (see section 3.3.) is key to the initialized stochastic model. Although, in general there exists the ability to use natural point features within the method, testing within this work was performed on high contrast retro-reflective or passive white target features which were illuminated to enable high contrast measurement (see section 6.2.2.3. for a sample of target data quality). In similar image network computations, the input data quality is regarded to be characterized with a standard deviation of $\sigma = 0.5$ μm (which is the default precision in VMS 8.0 tool). With regards to the presence of blunders, it is noted that the method is treated with pre-processed data to ensure that the data are clean from any erroneous point correspondences as outlier detection and elimination is not part of the up-to-date solution. It follows that, a priori data quality is dictated by the empirically set image quality factor which can be optionally adjusted to the testing requirements. It is apparent, that considering magnification factor (see section 3.4.2.1.) as an inherent affine sensor property as well as projection distortions, these factors are likely to additionally affect the success of

image measurement and hence input data quality. For example it will be shown that a typical target diameter for close ranged image targets can reach the order of 36 pixels in image space under affine projection when 2 mm in diameter object space targets are imaged for measurement purposes (see section 6.2.2.3.).

The affine model implements the a priori precisions as set for the input image measurements. These form the stochastic model that weights the least squares estimation procedure (see section 4.2.). In the case of an inner constraints datum, the weight matrix is augmented based on the normalized precision factor ($\sigma_i=1.0$) whilst external constraints utilize the control point data quality to form the augmented weight matrix (σ_i pointed by 3D targets precision). The quality of the 3D targets is pointed by their pre-measurement method which in general needs to generate high order precision reference measurements. It is noted that estimation of the a posteriori σ_o gives an overall fidelity check of the initialized stochastic model. It is recalled here that Gauss-Markov based procedures are evaluated with the extracted a posteriori quality measures; here quality evaluation is extracted from the scaled to the σ_o (standard deviation) a posteriori covariance matrix. Stochastic modelling is key to quality assessment (see section 4.7.1.) particularly in order to determine method precision and reliability measures that truly reflect the quality of the computed elements and can support further analysis.

5.3.3. Data structure outline

For clarification purposes it is stated that the following set of input data are handled within the computations. Specifically 3D target coordinates (.tar data file), camera calibration information (.cal data file), exterior orientation parameters (.pho data file) and 2D image observations (.obs data file) when these are attributed to the perspective sensor model whilst ORIENT data (.pho data file) and similarly 3D target data (.tar data file) point to the affine sensor model. Here, the following notation TAR, CAL, PHOTO, OBS and ORIENT apply to the utilized data structures a description of which is given in Appendix B. In addition, two structures were designed for problem handling. The first PHOTO_ORTO structure is required for the definition of the photo pseudo-exterior orientations entities that in the case of the affine sensor enclose the positions and orientations of the photo parameters (tx, ty, s, ω, φ, κ) and the second

structure is needed for least squares problem solving (1D and 2D arrays) (see Appendix B for a description).

5.4. Derivation of starting values

A prior task to the bundle adjustment algorithm is the estimation of starting values in order to initiate the process. Initialization procedures for the established multi-view problems met in photogrammetry have been reported earlier (see section 4.3.5.). Analytical approximate value estimation can be derived from direct (without the requirement of a priori estimates) or through iterative solution. Retrieval of orientation parameters is a highly critical subject and entails significant research interest in photogrammetry. Good starting values are those that ensure high stability in full parameter space. It is essential that starting value procedures provide knowledge on the critical configurations of 3D points and that they define cases of indeterminacy, instability or multiplicity which are accommodated within the derived solutions. Here, the approximate values are retrieved from a combination of established perspective-based and derived affine based solutions that have been investigated for the purpose of this research work. Since starting values are recovered photogrammetrically and not through external methods (e.g. orientation devices or coordinate system transformations) it is important that both perspective and affine cameras cover simultaneously an identical patch of the 3D volume to be measured. Such an imaging geometry is illustrated in Figure 5.5 where a joint projection model (affine and perspective sensors) recover the image space to object space correspondence relation and locate an object volume in 3D.
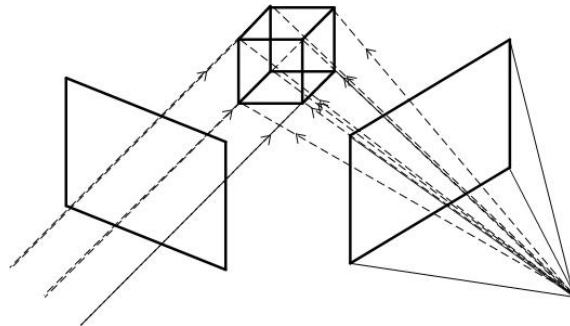
Figure 5.5: Two-view geometry of a pair of affine and perspective views.

5.4.1. Initialization structure

For the purpose of this work the problem of parameter initialization was treated as a combination of procedures through established perspective solutions regarded here as indirect procedures as well as two affine procedures one direct and one iterative written specifically for this purpose. Figure 5.6 demonstrates a general description of the initialization structure procedure.
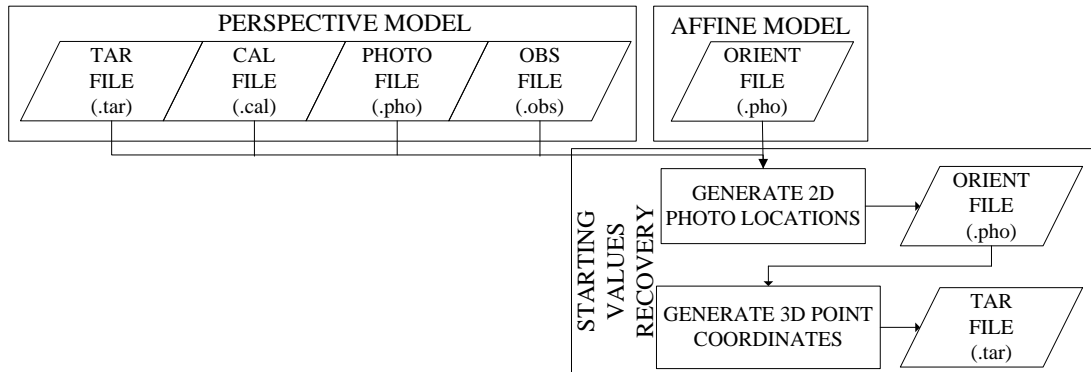
Figure 5.6: Affine model initialization.

The method starts from a set of reference control which is known in the object space. Whilst it is made clear that the subject of minimum control requirements for starting value recovery has not been investigated as part of this work, it is emphasized that this is critical in order to avoid situations of geometric degeneracies that for example can originate from indirect estimation procedures (and associated assumptions) or weak cases regarding geometry (image to object space recovery under parallel projection). However, repeated tests proved that a minimum number of nine well distributed CPs are required to result in a successful estimation of orientation angles (see section 6.4.1.2.). These data together with their associated precisions (.tar file) are inserted with the defined (.cal file) in the software tool VMS 8.0 where a set of image measurements are generated according to the method described in section 3.3. Applying established perspective procedures for initialization, requires that the physical affine sensor is approximated by an equivalent perspective sensor. This approximation is achieved based on the assumption that a perspective camera with a very long focal length is a good approximation to a parallel camera. Here it is assumed that such a guess is realized considering a nominal value for the focal length to be equal to 100,000mm which is significantly large over the measured volumes noting that the physical cameras footprint is a few tenths of mm (~40mm) in the

object space. This step is therefore completed by derivation of the 3D orientation angles (ω, φ, κ) applying the initialization methods. The output of this procedure is a set of ASCII data files target data (.tar), calibration data (.cal), photo orientations data (.pho) and image observations data (.obs) that have been updated and filled with the input required data (see Appendix B for a detailed format description).

Subsequently, these data are read in the bundle adjustment framework where two procedures take place. The first generates the 2D projective photo locations directly from the affine camera sensor through back-substitution (see section 5.4.2.2.) in the ideal case (absence of any internal distortions). On completion this procedure updates the orientation file (.pho) which now includes, and in this particular order, the 2D projective translations (tx_j, ty_j), nominal scale factor (s_i) and 3D orientation parameters (ω_j, φ_j, κ_j) for each measured valid view. These can then be inserted into the next step which computes the 3D target locations through an iterative weighted least squares procedure which is written for the affine sensor model (see section 5.4.3.). On convergence the process outputs the 3D target coordinates with their associated precisions and updates the 3D file (.tar file).

Derivation of approximate values for the affine sensor model required the design and implementation of a new structure that identifies the required members and it is for this purpose embedded within the available header file (see Appendix B for further outline). The formats of the ASCII files described and utilized here are also given (see again Appendix B). It is again stated that the problem of starting value derivation has been theoretically given in section 4.3.5. (with initial approximations for model formation are given in Appendix A).

5.4.2. Pseudo–exterior orientation parameters: Stage 1

The pseudo-exterior orientation parameters are recovered as a first stage of the overall initialization method. The method is comprised of two steps that estimate the 3D orientation angles (ω_j, φ_j, κ_j) indirectly through perspective (see section 5.4.2.1.) followed by estimation of projective translations through back-substitution (tx_j, ty_j) as a direct approach from affine projection (see section 5.4.2.2.).

5.4.2.1. Indirect 3D orientations

To initialize the orientation angles (ω, φ, κ) a modified Zeng, Wang closed form resection procedure (Zeng & Wang, 1992) is applied (see section 4.3.5.2.) on the assumption that a perspective camera with a very long focal length (e.g. assumed nominal focal length f=100,000 mm) is a good approximation to the affine camera. Prior to the established resection an initial exterior orientation procedure, again utilizing the same assumed focal length, is computed to enable initial estimation of the exterior orientation parameters as input to the resection procedure. It is therefore understood that resection acts as a refinement of the initialized estimation. The procedure demands an initial set of known CPs in the object space that are well distributed in 3D to avoid for example geometric degeneracies that can be the result of employment for example of coplanar or collinear configurations. A description of this procedure is given here.

Step 1 - Generate 3D orientations:
-   Start with a set of sparse reference target point (CPs). These provide reference control and simultaneously establish the datum for subsequent computations.
-   Estimate 3D rotations from an initial exterior orientation procedure updated by a modified closed form resection assuming a physical perspective camera with a very long focal length (e.g. here: f~100,000mm).

On success, the utilized resection procedure outputs the 3D orientation angles (ω, φ, κ) together with an overall mean image measurement residual RMS (in μm) and mean valid target image observations (see Appendix B for output files). The exterior orientation parameters regarding the 3D photo positions will be subsequently (see section 5.4.2.2.) updated by the computed affine model parameters and hence their initial determinations from this step will be ignored.

5.4.2.2. Projective translations through back-substitution

The 2D projective translations (equivalents of the well known principal point located on perspective sensors) are estimated on the knowledge of the partial exterior orientation given above (step 1). In particular, this back-substitution procedure starts from the input observations list (OBS data) where for each valid photo (search on PHOTO data identifiers (IDs)) the target points (TARGET data) that sit on each photo

are found. Provided that a point acts as a CP and that it is measured on at least two views (four rays are generally preferred), tx and ty are calculated as the accumulated sum of the reduced observations vector which is subsequently divided by the number of valid rays. This is performed per view, noting that the reduced observation vector is calculated as the observed minus computed image measurements. Computed observations are regarded those that are estimated from the affine mathematical model initialized from the starting values of the parameters and in the absence of any internal camera distortions (ideal geometric case). It is noted that this procedure is executed for each valid photo; that is when at least three rays have fallen on it (formed by valid CPs). The computed parameters update the orientations data file.

Step 2 - Mathematical formulation:

$$t_x = \frac{\sum x_{OBS} - x_{COMP}}{n_{RAYS}} \qquad (5.1)$$

$$t_y = \frac{\sum y_{OBS} - y_{COMP}}{n_{RAYS}}$$

Where:

t_x, t_y = 2D projective translations (mm)

x_{OBS}, y_{OBS} = 2D observed image coordinates (mm)

x_{COMP}, y_{COMP} = 2D computed image coordinates (from the affine sensor model) (mm)

n_{RAYS} = number of rays occurring per photo (counted from the valid CPs)

Step 2 - Projective translations through back-substitution:

- Back-substitute the 2D projective translation parameters utilizing the affine camera model and update the orientations data file.

5.4.3. Object space 3D coordinates: Stage 2

To derive 3D target coordinates, a forward intersection procedure for the affine sensor model was written. This was implemented on the knowledge of the recovered pseudo-exterior orientation parameters (see section 5.4.2.) and in the absence of any internal geometric camera distortions. The procedure starts with two basic checks to obtain the number of rays that correspond to each target (checking on TARs against OBS) and

views (checking on PHOTOS against OBS). Next the 3D target centroid is estimated from the total number of control point data similarly to the calculation of the 3D centroid that has been described in the inner constraints method (see section 4.5.1.). Subsequently, the starting values for all 3D control and tie point data are set to be equal to the (X_c, Y_c, Z_c) coordinates of the 3D centroid. A weighted LSE (forward intersection) procedure for the affine sensor model recovers the 3D point coordinates. This procedure minimizes the cost function in the object space and it results in the computation of the 3D control and tie point coordinates together with their 3D precisions. The tolerance of the intersection procedure is pre-set at 10μm in object space whereas minimum target visibility for the intersection is two rays by default unless otherwise modified. On convergence the original 3D target file is updated by the intersected 3D coordinates and their associated precisions (see Appendix B for a sample data file).

Stage 2 - Object space 3D coordinates:

- Perform a multi-view forward intersection (weighted LSE) procedure to coordinate new 3D points (CPs and TPs). The intersection closure tolerance is pre-set at 10μm in the object space for the 3D target coordinates corrections.

5.5. Bundle adjustment framework

It has already been stated (see section 4.1.) that a bundle adjustment framework provides an optimal method for processing multi-view problems to simultaneously estimate calibration, orientation parameters and 3D object geometry. To answer the question of selecting a least squares bundle approach to process universally such a linear model such as that of the affine sensor, it is recalled here that the choice of a least-squares solution (see section 4.2.) is attributed to its ability to produce a unique, unbiased and objective solution resulting in the minimum variance of estimated parameters. Moreover it delivers a quantifiable quality assessment and embeds full covariance matrix analysis within the system. To make this statement relevant to the multi-view problem examined here it is highlighted that, for instance, to combat noise derived from corresponding point measurement even the case of a direct approach should produce a least squares solution from redundant information.

To design such an iterative adjustment system, there are some important practical considerations that need to be considered which are outlined here giving an overview of the algorithm at the outset. Subsequently analysis of the intermediate steps according to design and implementation are reported.

5.5.1. Framework structure overview

The structure of the affine bundle adjustment framework is outlined in Figure 5.7. The method will be described from the stage where the input data (specifically target, calibration, photo and observations data; see section 5.3.3.) initialized from perspective-based procedures have been read within the process. Additionally, it is assumed that the explicitly generated data based on the affine-sensor model (specifically orientations and updated 3D target data) have been already initialized as previously described. For initialization of data structures see section 5.4.
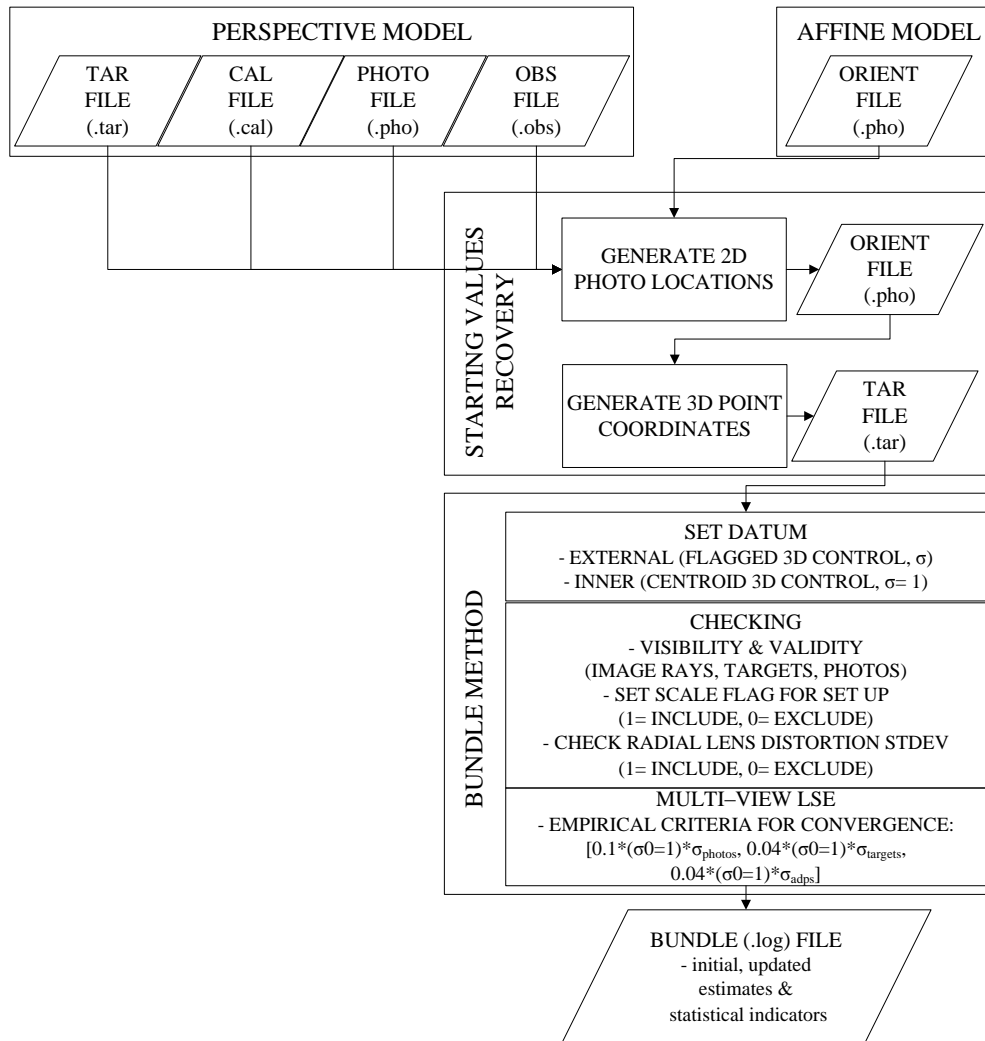
Figure 5.7: Affine multi-view model processing.

The implemented bundle method is blocked into three main parts, each one representing an implementation stage. The first stage requires that the datum is defined in order to enable data processing on either an external or inner datum definition. Next the measurement data are inserted within a list of checks related to issues like data visibility and validity, parameter set-up for inclusion or exclusion from the system (scale factor and radial lens distortion coefficient). The final stage involves the implementation of the iterative multi-view least squares estimation process. Here the issues of array handling for inversion as well as convergence criteria are reported. On convergence the algorithm outputs a descriptive report listing, on a statistical basis, the initial as well as estimated parameters together with their associated quality factors.

5.5.2. Algorithm implementation

Implementation of the algorithm is detailed according to the three individual stages given above. The affine bundle adjustment is described as a series of algorithmic steps: (a) prior to the iterative process (see section 5.5.2.1.), (b) least squares estimation process (see section 5.5.2.2.), (c) convergence (see section 5.5.2.3). Sections 5.5.3 and 5.5.4 outline the closure criteria and quality estimates.

5.5.2.1. Prior to the iterative process

Step 1 - Datum definition:
- Set up datum for coordinate system definition. Two options are provided. External constraints establish the 3D datum based on the identified CPs (datum_flag= 'e') where CPs are considered those targets that are indexed with a flag= 7. Alternatively, inner constraints (datum_flag= 'i') establish the 3D datum based on a 3D centroid identified again from the flagged CP data.

Step 2 - Visibility:
- The total number of target points is checked against their occurrence onto the valid number of images within the image network to retrieve the total number of rays corresponding to each target point. This derives the point visibility from photos.

- The total number of photos is checked against the valid image observations to retrieve the total number of rays corresponding to each photo. This derives the number of target points measured on each photo.
- The total number of photos is checked against the number of physical cameras. It is noted that each physical camera specifies a camera calibration set of parameters.

Step 3 - Validity:

- The number of image observations that are valid within the model are obtained and stored in a counter (n_obs).
- The number of target points (CPs and TPs) that are valid within the model are obtained and stored in a counter (n_targ).

Step 4 - Starting values:

- The system is initialized with the pre-determined starting values corresponding to photo orientations (PHOTOS), 3D target coordinates (TARs) and additional parameters (ADPs).

Step 5 - Set up scale factor:

- Scale factor is specified from the nominal scale factor associated with the camera sensor (image invariant factor). Scale estimation is embedded within the system according to an identified flag. Particularly, scale_flag= 1 includes the scale factor within the system computations and scale_flag= 0 excludes the scale factor from the system estimations.

Step 6 - Parameter locations:

- The individual parameter locations are set for the 2D projective translations, 3D orientation angles and 3D target coordinates with an implemented check.

Step 7 - Parameter estimations:

- Calculation of the estimated parameters is based upon the valid system parameters.

Step 8 - Constraints:

- This step performs a calculation of the number of constraints for both datum methods. The external constraints will augment the system by a number of three (3D coordinates) times the number of valid CPs, whereas the inner constraints will augment the system by seven additional equations.

Step 9 - Radial lens distortion:

- The active image frame (maximum extent) is calculated from the image observations within the system. Accordingly, the frame radius is estimated from the equations given below (see equations (5.2) and (5.3)). Subsequently the additional parameter set (here the third power term of the radial lens distortion polynomial) is calculated from a function initialized for the radial lens distortion part only (without the inclusion of the extended parameters model that compensates for example affinity and orthogonality terms). It is evident that the simplified radial lens distortion is calculated from the valid image observations that clearly identify the active image frame extents as opposed to utilization of the total frame (given in pixels within the .cal file).

$$\text{frame} = \sqrt{x_r^2 + y_r^2} \tag{5.2}$$

$$r = \frac{\text{frame}}{\sqrt{2}} \tag{5.3}$$

Where:

frame= active image frame calculated from the image observations x_r, y_r (mm)

r= image radius for radial lens distortion estimation (mm)

5.5.2.2. Least squares estimation process

Step 10 - Iterative process:

- The standard LSE arrays are populated within the main LSE loop. Within this implementation, the algorithm sets up the ADPs, computes the reduced (observed minus computed) vector of observations and it calculates and locates the partial differentials that correspond to the model parameters within the design matrix (see Appendix A).

Step 11 - Derivation of system normals:

- This step first populates the weight vector given the input image observation precisions. It then populates the augmented arrays for the external or inner constraints datum cases (augmented design matrix and observations vector).

- The normal equations matrix N= A^TWA and normal equations vector A^TWb are successively obtained with the inversion problem [23] treated on the vectorized normal equations matrix as required from the available inversion routines. It is noted here that the inner constraints inversion method has the property that the resultant covariance matrix has a minimum trace and hence the standard deviations of the object points are estimated with minimum quantities. The 3D centroid becomes the origin of the datum which is a fixed point with a standard deviation equal to unity. The inverted array is back-stored as a matrix for subsequent extraction of the quality measures.

Step 12 - Estimation of parameters:

- Subsequent population of the corrections to the parameters vector allows assignment of the final parameters (2D projective translations, rotations, targets, scale and k_3).

- The parameter data are reset according to the starting values derived from the estimated parameters (again given here as 2D projective translations, rotations, targets, scale and k_3).

5.5.2.3. Convergence

Step 13 - Convergence:

- The iterative procedure terminates when the convergence criteria have been reached (see section 5.5.3.). This is performed by a partial check of the individual absolute correction values parameter set (2D projective translations, rotations, targets, scale and k_3) over the defined criteria scaled to

---

[23] The inversion method applied in the external constraints case utilizes a Cholesky decomposition sub-routine which is suitable for positive symmetric matrix inversion whereas the inner constraints method utilizes a general matrix Moore-Penrose subroutine based on the Gauss Jordan elimination.

a pre-set unit weight of unity for the convergence check ($\sigma 0 = 1$) times the corresponding parameter precisions.

Step 14 - Termination:

- If the convergence criteria are met then the process stops and the solution is guided to the output stage. The LSE loop has reached to its end; the quality estimates output is given in section 5.5.4.

5.5.3. Closure criteria

As in every iterative adjustment process, the developed bundle method terminates when appropriate criteria have been reached. In theory, such an iterative process may converge to the desired solution, diverge, oscillate or repeat in certain cycles. Unless the system is degenerate, or for example erroneously established, convergence can be quickly achieved provided numerical instabilities are not encountered. In measurement applications it is important that the best precision is achieved. Best precision is that which derives an answer as close to the theoretical solution up to potential insignificant discrepancies (due to modelling as well as machine precision factors). At the same time the obtained solution needs to ensure an economical answer. For example an iterative procedure that would perform numerous iterations for a relatively small gain in accuracy is considered as impractical and therefore tends to be avoided.

Given the above considerations the selected criteria need to be representative of indicators of convergence and evaluate the precision of the system. Therefore, the strategy applied here groups different types of parameters according to the expected precision levels in terms of convergence. Specifically, the convergence limits for photo orientations are set to $0.1 \times \sigma 0 \times \sigma_i$ whereas the 3D target coordinates and additional parameters are ascribed a tolerance of $0.04 \times \sigma 0 \times \sigma_i$ (where σ_i is the associated parameter precision). These limits are given as follows:

- Tolerance for orientation parameters: crit_photo (= 0.1) * $\sigma 0$ (= 1) * σ_{PHOTO}
- Tolerance for 3D point parameters: crit_tar (= 0.04) * $\sigma 0$ (= 1) * σ_{TAR}
- Tolerance for additional parameters: crit_adps (= 0.04) * $\sigma 0$ (= 1) * σ_{ADPs}

Checking based on these tolerances is performed on the corrections to the estimated parameters. The reason for generally scaling the photo orientations at a larger limit is that photo orientations are generally weakly determined and they can in some instances lead to an oscillating solution presenting convergence difficulty. These empirical factors are equivalent to those used within a conventional robust bundle adjustment that has been primarily tested with simulated and real data (within the VMS 8.0 tool).

5.5.4. Quality estimates output

On a successfully convergent solution the algorithm exports a report that includes an analytical description of the estimated parameters together with a set of statistical quality measures. In summary, the report outputs the initiated parameters grouped as initial photo parameters (2D positions in mm and 3D orientations in degrees) and initial target locations (in mm) with associated precisions (in μm). Next a set of synoptic descriptors are given to describe numerically the estimation procedure. These include the number of equations, number of unknowns, iteration of convergence, redundancy and a posteriori standard deviation (σ_o). The computed image residuals (x, y and mean in μm) are given as well as the radial lens distortion profile for camera calibration (that corresponds to the simplified calibration model). Subsequently, the updated orientations parameters (in mm and degrees) are provided with their estimated precisions (in μm) as well as the updated target locations (in mm) again with their corresponding precisions (in μm). Calculated ray visibility and point flags for data identification (e.g. indexed CP or TP data) are additionally supported. An update of additional parameters with associated precisions is given, including an indication of their significance (where significance= correction / precision). Successively, a summary of the 3D target precisions (in μm) as obtained from the a posteriori covariance matrix is exported. Finally, the report is completed with the extracted full correlation coefficients matrix as this is considered as highly useful for further analysis and evaluation of the calibration and measurement outcome. A sample of the generated report (.log file) for both the inner and external datum cases can be found in Appendix B.

5.6. Method description by simulation

Evaluation of a novel bundle adjustment algorithm for close range photogrammetric measurement, requires that design by simulation is performed at a first development stage. The problem of network design which is well established has been reported earlier (see section 4.6.). Here, the theoretical model description and therefore method initialization is assessed by design of a simulation project to enable analytical method development. In parallel, the goal of this section is to provide an evaluation of the developed bundle adjustment method.

The simulation framework is built upon the assumption that a synthetic volumetric array comprises the test object in the object space (see Appendix C) and that it is measured from incremented viewpoints and wide intersection angles chosen for this purpose. Particularly, a cube occupying a volume of 10 x 10 x 10 mm in object space (with a specified 3D precision of 25μm) observed from two-viewpoints, three-viewpoints and subsequently seven-viewpoints was designed (see sections 5.6.2., 5.6.3. and 5.6.4. accordingly). A descriptive diagrammatic form of the simulation process is provided in Figure 5.8.
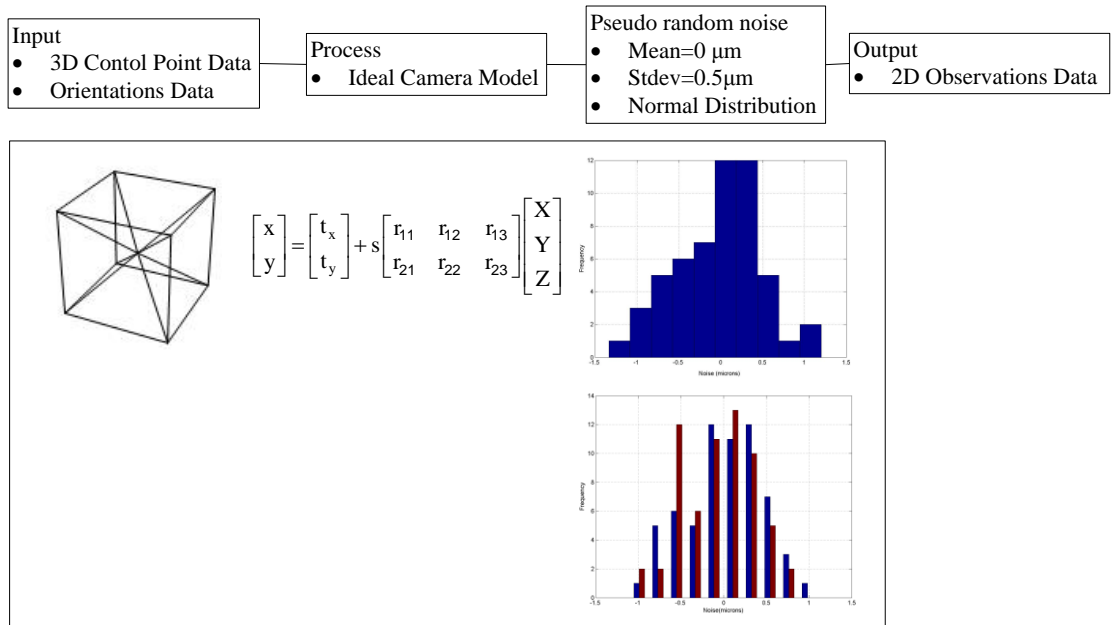
Figure 5.8: Descriptive simulation diagram applied to a synthetic cube dataset.

The simulation process starts from known CP geometry in the object space as defined by the eight cube vertices and its centroid (X_i, Y_i, Z_i, where i=1-9). Subsequently the

3D data points are measured from each viewpoint as specified per imaging case to generate the photo orientation locations (tx_j, ty_j, s_k, ω_j, φ_j, κ_j, where j= 2, 3 and 7 as associated with the three equivalent projects and k (= 1) the number of physical cameras within the system). Next 3D target and orientation data are inserted into the ideal affine camera model in the absence of internal camera distortions to generate a set of 2D image observations. A next step increments the extracted observation data by a randomly initiated set of corresponding standard deviations that follow the normal distribution (where: Mean= 0.0μm and Stdev=0.5μm). Successively, the bundle adjustment is processed with the set of input CP (TAR), calibration (CAL), orientation (ORIENT) and observations (OBS) data. It is noted that the system is based upon a local datum which is defined from the designed CPs (external datum method). A detailed representation of the simulation process is given below (see Figure 5.9).
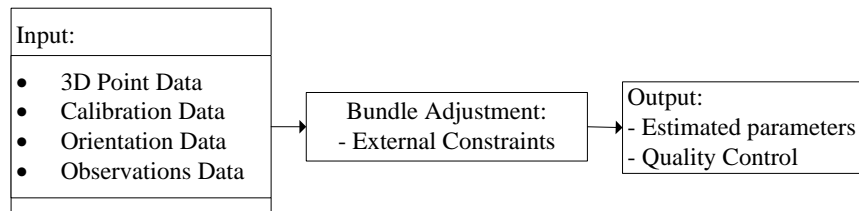
Figure 5.9: Bundle adjustment framework for simulation data.

5.6.1. Input model parameters

The input model parameters were identified to be equivalent with those that correspond to the real sensor parameters of the 1M pixel Kodak Megaplus ES1.0 camera with MVO® TMLTM/0.16x lens (see sections 3.1.2.1. and 3.1.3.2.). This system has been used for the purposes of system testing and application (see Chapter 6). It follows that the defined virtual camera is characterized by a scale factor of s= 0.16 (which identifies the image magnification), a pixel size of 9.0μm and a format size of 1,008 x 1,018 pixels. Considering the precision levels encountered in close range image networks, the stochastic model was initialized with an image observation quality of σ= 0.5μm and a 3D point coordinates precision of σ= 25μm in all simulation tests. It is re-iterated here (see section 5.5.3.) that the convergence criteria were set to $0.1 \times \sigma 0 \times \sigma_i$ for the photo orientations whereas for the 3D point coordinates and the additional parameter a tolerance of $0.04 \times \sigma 0 \times \sigma_i$ was ascribed (noting that σ0 is by default set to 1 and σ_i is the a posteriori parameter precision). The input model

parameters that were utilized to set up the simulation projects are given in the following Table 5.1.

Image scale	Pixel size (μm)	Image size (pixels)	RMS$_{xy}$ (μm)	σ$_{XYZ}$ (μm)
0.1600	9.0	1,008 x 1,018	0.5	25

Table 5.1: Input model parameters in the simulation processing tests.

5.6.2. Case A: Two-view geometry

The two-view imaging geometry is covered with two cases by altering the direction of the lines of sight intersections with the X axis direction (as defined by the local coordinate system). In particular this section examines two instances that involve two virtual viewpoints, the first one having a basis nearly parallel to the X axis and the second one forming a basis nearly vertically to the direction of the X axis respectively. The scale factor is equal to 0.16, as described above, and the image observations are generated with identical noise (see Figure 5.8 for associated histogram). The image network geometries are drawn in Figure 5.10 and the associated parameters for both datasets are given in Table 5.2.
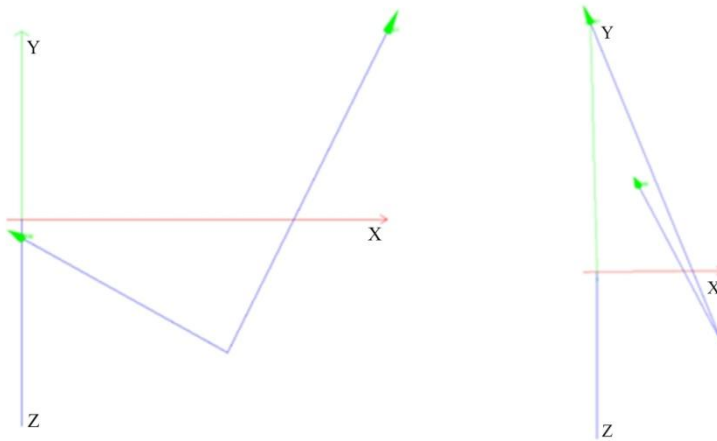
Figure 5.10: Imaging configuration of two-view geometry case. Geometry with near parallel basis to X datum axis configuration (left) and near vertical basis to the X datum axis (right). XYZ system is visualized in red, green and blue respectively.

Photo	tx (mm)	ty (mm)	s	ω (degrees)	φ (degrees)	κ (degrees)	Rays
1000_h	0.1423	-4.6332	0.16	-74.7258	-32.3675	-64.8622	9
1001_h	1.6584	1.1229	0.16	-116.3754	24.5694	133.2145	9
1000_v	-2.0177	-3.8078	0.16	-71.7631	-10.9171	-33.7956	9
1001_v	3.3984	-2.5645	0.16	-105.0204	-17.8026	-137.9776	9

Table 5.2: Orientation parameters for the two-view geometry case.

The bundle adjustments were processed based upon an external constraints datum as defined from the determined CPs (see Table 5.3 for a summary of results).

Dataset	Iter.	R.	Scale (σ_{scale})	k_3 ($\sigma_{\kappa3}$)	σ_o	RMS_{xy} (μm)	σ_{XYZ} (μm)
s_0.16h	2	24	0.1600 (0.0887)	-1.5721×10^{-5} (1.3655×10^{-5})	0.50	0.14	6.83
s_0.16h	2	24	0.1600 (0.0636)	5.4044×10^{-5} (2.2152×10^{-5})	0.31	0.09	4.42

Table 5.3: BA statistics for the two-view geometry case. Table notation: Iter.= number of iterations, R.= redundancy, Scale= scale factor, σ_o= unit weight, RMS_{xy}= image misclosure, σ_{XYZ}= 3D object space precisions.

In the case of the first dataset (near parallel basis to the X axis; s_0.16h) the solution results in an a posteriori σ_o of 0.50, whereas in the second dataset case (near vertical basis to the X axis; s_0.16v) the associated σ_o is 0.31. It is evident that these figures are not significantly different. Again, in the first case (s_0.16h) the RMS back-projection error is larger in a direction orientated nearly parallel to the basis resulting in a value of 0.25μm (1/30th of a pixel) as opposed to an RMS of 0.04μm (1/225th of a pixel) which is the equivalent value in the y direction. The situation in the second case (s_0.16v) results in balanced RMS image residuals of 0.09μm (1/100th of a pixel) in both x and y directions. The 3D point location precisions vary within 6.83μm and 4.42μm for each first and second imaging cases respectively.
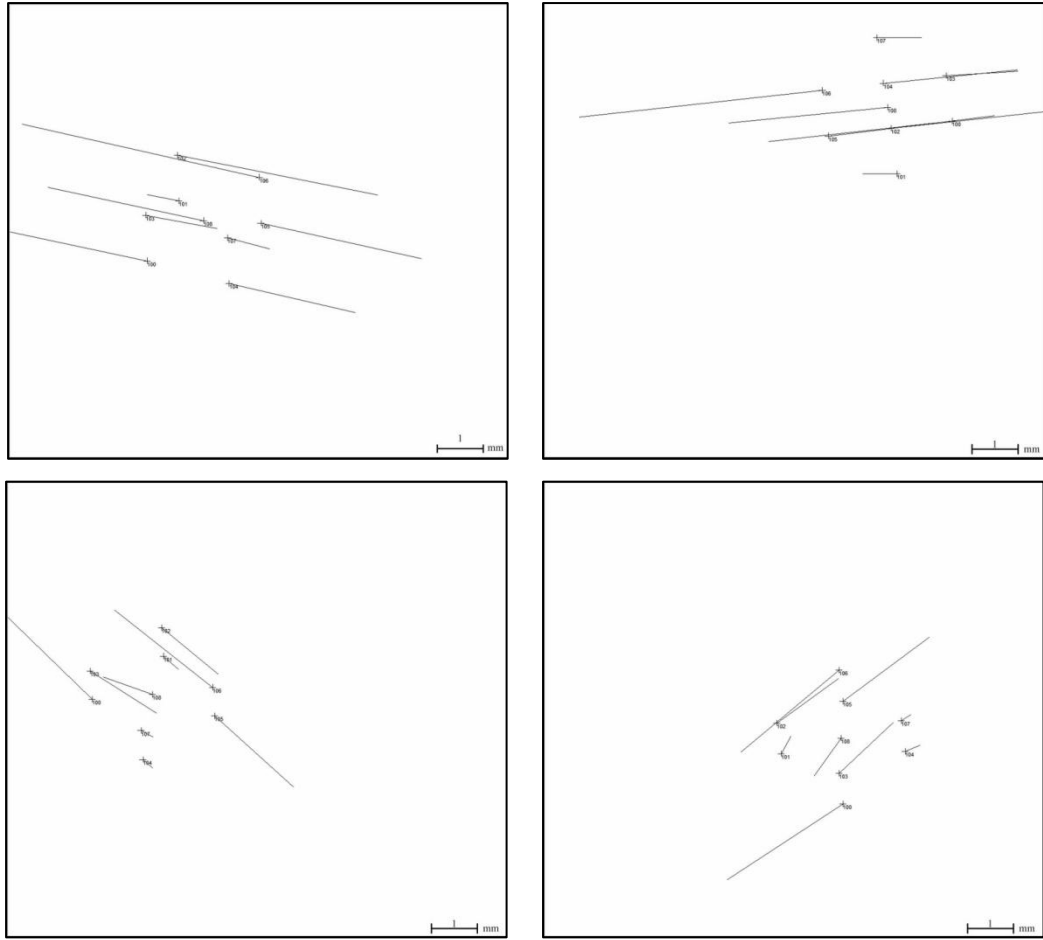
Figure 5.11: Image observation residuals for the two-view geometry case. S_0.16h (top) and s_0.16v (bottom) (magnification plot 25x).

A display of the image residuals for both imaging cases (see Figure 5.11) shows that these follow a systematic pattern in a direction parallel to the basis of the geometric configuration. This is a factor of the weak geometry recovery from two views and demonstrates that there is a geometric correlation of the two-views geometric case with the estimation of 3D point coordinates (equivalent to the perspective case). The residual plots are shown with a magnification factor of 25x for visualization purposes.

5.6.3. Case B: Three-view geometry

Analysis of the three-view geometry case is made to examine model behaviour when altering the scale factor. The initial set scale of 0.16 is changed to 0.5 and 1.0 while preserving the orientations of the photos and the 3D target positions as initially set. Figure 5.12 illustrates the geometric arrangement of the three-view geometry case and Table 5.4 indicates the orientation parameters of the designed dataset.

Photo	tx (mm)	ty (mm)	s	ω (degrees)	φ (degrees)	κ (degrees)	Rays
1000	0.1425	-4.6334	0.16; 0.5; 1.0	-74.7258	-32.3675	-64.8622	9
1001	-3.0773	0.1897	0.16; 0.5; 1.0	-82.7854	7.4404	30.1569	9
1002	1.6582	1.1229	0.16; 0.5; 1.0	-116.3754	24.5694	133.2145	9

Table 5.4: Orientation parameters for the three-view geometry case.
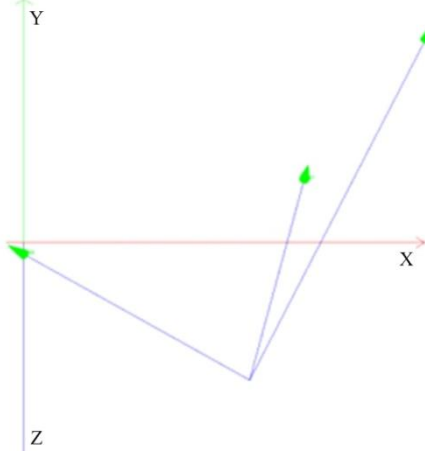
Figure 5.12: Three-view geometry configuration.

All three tests were processed with an external constraint datum based on the pre-defined CP data. Processing the image observations data makes it apparent that tuning the scale factor to the pre-set values results in a change in the generated 2D image observations by the associate magnification factor. The bundle adjustment converges with an aposteriori σ_o of 0.8. This indicates that the increased number of viewpoints strengthens the ray intersections as well as the system redundancy when compared to the two-view imaging case reported above (see section 5.6.2.). The RMS image misclosure is 0.26μm (~1 /30th of a pixel) and the mean precision of 3D points is 10μm in the object space. These figures are given in Table 5.5 whilst Figure 5.13 illustrates that the random residual patterns indicate the absence of systematic errors from the system.

Dataset	Iter.	R.	Scale (σ_{scale})	k_3 ($\sigma_{\kappa 3}$)	σ_o	RMS$_{xy}$ (μm)	σ_{XYZ} (μm)
s_0.16	2	37	0.1600 (0.1368)	-1.7485×10^{-5} (1.8235×10^{-5})	0.79	0.26	10.26
s_0.50	2	37	0.5000 (0.4120)	-1.6905×10^{-7} (5.8004×10^{-7})	0.79	0.26	10.08
s_1.00	2	37	1.0000 (0.8137)	-5.1918×10^{-9} (5.0392×10^{-8})	0.79	0.26	10.05

Table 5.5: BA statistics for the three-view geometry case.
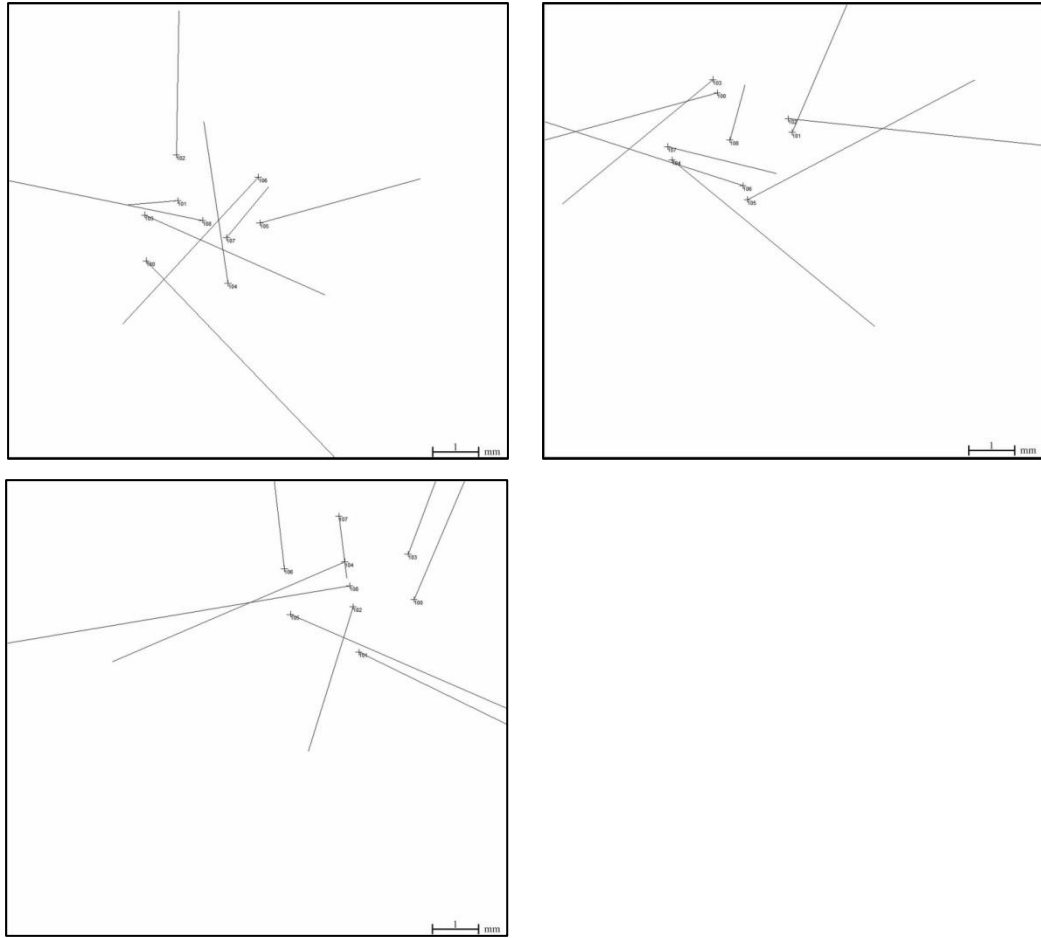
Figure 5.13: Image observation residuals for the three-view geometry case. S_0.16
(magnification plot 25x).

It is noted that when the 'complete' bundle adjustment method is run, here the term
'complete' refers to the inclusion of the radial lens distortion coefficient within the
computations; it is likely that the estimated radial lens distortion profiles are
erroneously determined. This is reasonable considering that the simulated project
comprises an approximation of the physical reality and hence presents a weakness in
the recovery of the internal camera geometry factor.

5.6.4. Case C: Seven-view geometry

The third case involves testing with a greater number of views in relation to the
minimum intersection requirements as examined in the previous cases (see sections
5.6.2. and 5.6.3.). The seven-view geometry is designed in a wide-angled network
arrangement as shown in Figure 5.14 with corresponding orientation parameters as
indicated in Table 5.6.
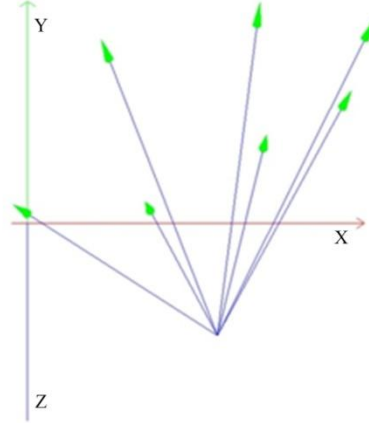
Figure 5.14: Imaging configuration of seven-view geometry case.

Photo	tx (mm)	ty (mm)	s	ω (degrees)	φ (degrees)	κ (degrees)	Rays
1000	0.1423	-4.6332	0.16	-74.7258	-32.3675	-64.8622	9
1001	-2.0177	-3.8078	0.16	-71.7631	-10.9171	-33.7956	9
1002	-3.0772	0.1895	0.16	-82.7854	7.4404	30.1569	9
1003	-0.9284	2.1266	0.16	-93.8974	20.6170	89.8087	9
1004	1.6584	1.1229	0.16	-116.3754	24.5694	133.2145	9
1005	3.3984	-2.5645	0.16	-105.0204	-17.8026	-137.9776	9
1006	2.7517	1.3587	0.16	-111.5494	6.2445	153.3023	9

Table 5.6: Orientation parameters for the seven-view geometry case.

Dataset	Iter.	R.	Scale (σ_{scale})	k_3 ($\sigma_{\kappa 3}$)	σ_o	RMS_{xy} (μm)	σ_{XYZ} (μm)
s_0.16_7	2	89	0.1601 (0.1371)	-2.1261×10^{-5} (1.4594×10^{-5})	0.81	0.27	10.37

Table 5.7: BA statistics for the seven-view geometry case.

The bundle adjustment converged after two iterations (see Table 5.7) with a σ_o of 0.81μm whereas the RMS image misclosure is 0.27μm (~ 1/ 30th of a pixel) and the mean precision of the 3D points is estimated to be equal to 10.37μm in the object space. It is evident that increasing the number of viewpoints from three to seven does not have a significant effect on the solution behaviour. Figure 5.15 supports the extracted residual vectors where again the random residuals patterns provide a further proof that the system is freed from any systematic biases within the computations.
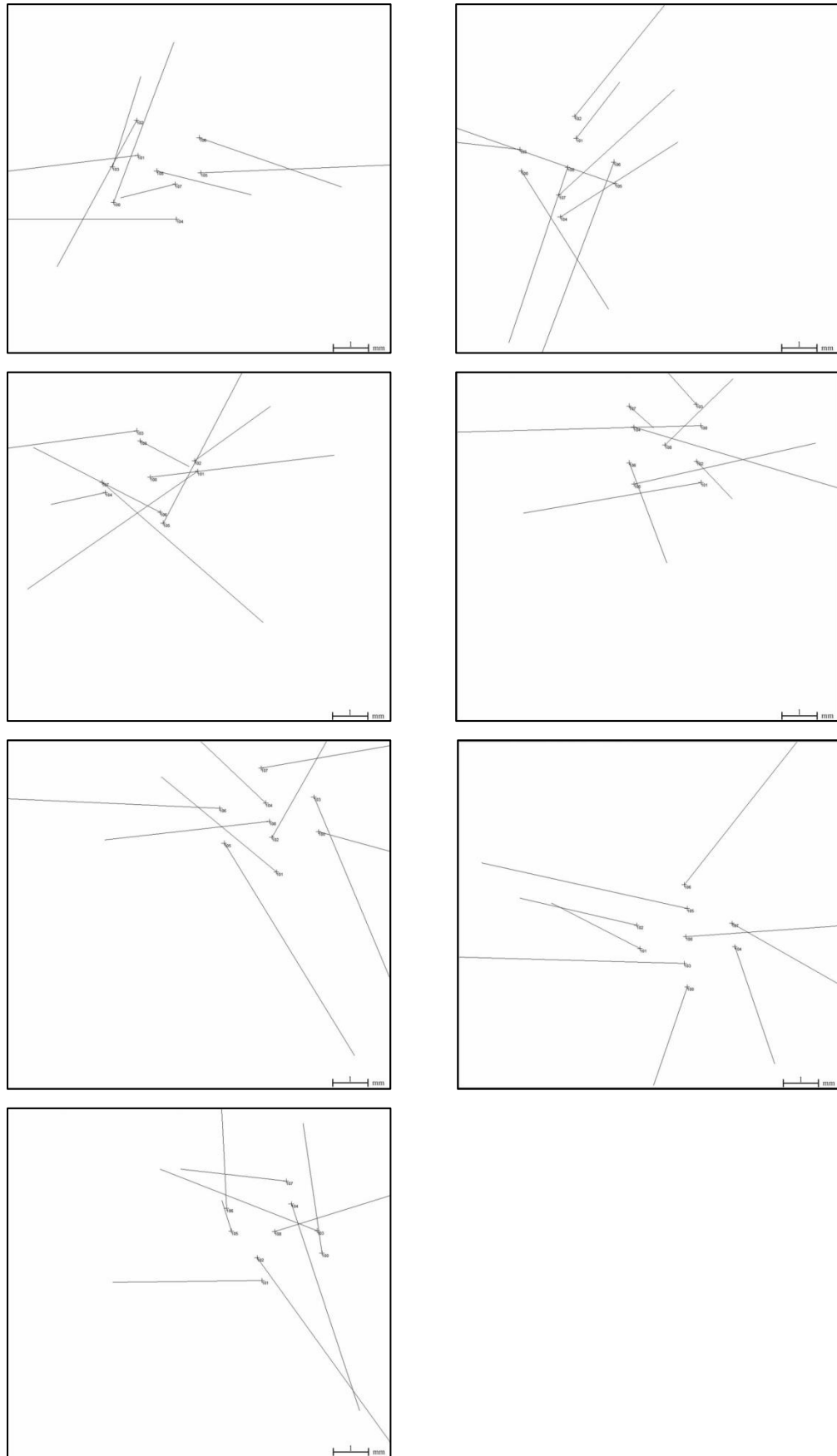
Figure 5.15: Image observation residuals for the seven-view geometry case. S_0.16_7, (magnification plot 25x).

Besides residuals display and evaluation of their magnitude, another useful measure is to check their distribution patterns. From observation of the histogram plots from the bundle adjustment image observation residuals (see Appendix C), it follows that the residuals tend to follow the normal distribution while increasing the number of geometric viewpoints ranging from two views up to seven views. It is evident that while the three-view case results in a near normal residual distribution the seven-view case provides a normal distribution with associated statistics (see Figure 5.16). This evidence proves that the implemented functional model is correct. It is now expected that incrementing the number of images for the m-view case would lead to an optimized solution whereas the designed and implemented model provides a good solution to modelling the affine multi-view sensor.
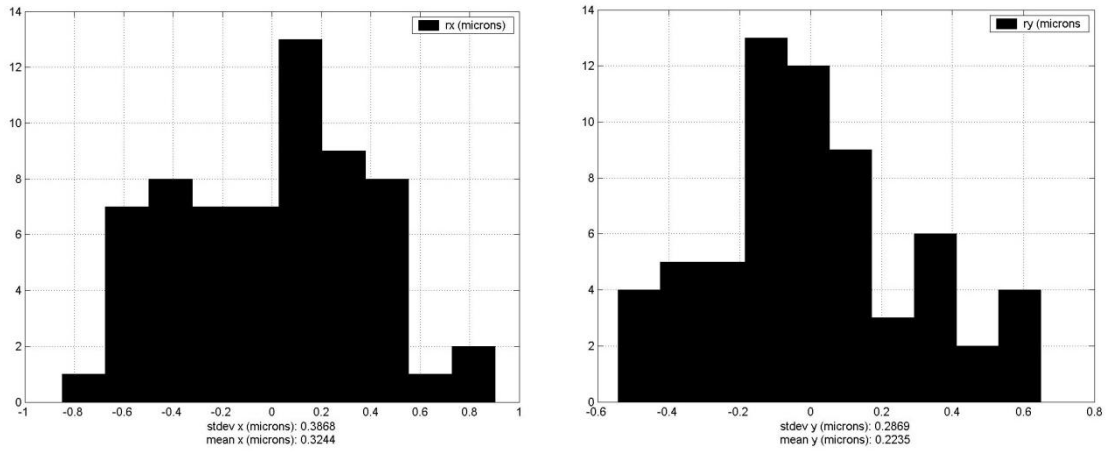
Figure 5.16: Histograms of BA residuals for the seven-view case. The extracted statistics are given per image coordinate direction x: stdev= 0.39μm, mean= 0.32μm and y: stdev= 0.29μm, mean=0.22μm.

5.7. Summary

In summary this chapter provides descriptions of the affine method design, development and implementation with simulated data. The chapter starts with an overview of the method which is subsequently linked with the fundamental model structure for algorithm treatment. Key consideration to the algorithm is the issue of derivation of approximate values based on a combination of implicit perspective and explicit affine procedures. The developed bundle adjustment framework outlines a set of factors (datum definition, visibility checks, validity checks, starting values determination, handling of scale factor, parameter locations, parameter estimation, constraint evaluations, radial lens distortions computations), least squares estimation

(iterative procedure, derivation of system normal equations, parameter estimation) and convergence criteria as well as the aspects of closure and extracted analytical indicators. Handling of these considerations within the algorithm is essential for flexibility in image-network based measurement. Subsequently, the chapter closes with analysis and design of the new affine-based multi-view framework through a simulation method that computes the implemented solution for three geometric imaging cases. The problem is decomposed with the two-view, three-view and seven-view geometry cases and the evaluation of the derived results.

The reader is now directed to Chapter 6 'Results and analysis' for method assessment with real data. Particularly the affine multi-view modelling method will be tested extensively through a series of real image-network measurements all designed, initiated, implemented and analyzed for the purpose of this research work. Method assessment is performed for each subsequent stage of the method with variant geometric data structures which are always linked with practical considerations that are expected to be encountered in real world applications.

6. Results and analysis

To assess the affine multi-view bundle adjustment algorithm a set of tests were performed. The objective was to evaluate the developed approach in the aspects of (a) method correctness and (b) effectiveness in practical situations. This goal is achieved with application of the method on a series of test datasets that were designed for testing of each of the individual algorithm stages as well as full algorithm treatment.

6.1. Main objectives

To assess the correctness of the algorithm the following considerations will be asserted:

1. Initialization of the method. Starting values were recovered from a combination of starting value generation approaches.
2. Bundle adjustment algorithm. The algorithm recovers orientation parameters, 3D point coordinates and internal calibration with the ability to define a datum (external or inner constraints), sensor scale and radial lens distortion model (k_3 term).

To evaluate practical model behaviour the following factors will be handled:

1. Assessment in object space and 3D point estimation.
2. Invariance of sensor scale within the system.
3. Independent evaluation with reference measurements. Sparse 3D reference measurements are provided by high order precision datasets. These datasets were generated from strong, convergent, redundant image networks acquired with high resolution digital SLR cameras and subsequently processed with a robust perspective bundle adjustment implemented within the software tool VMS 8.0.

The above comprise core objectives for the purpose of methodological testing. However, it is important to make it clear that the developed bundle adjustment framework was implemented based upon two assumptions:

1. Starting values were generated from implicit perspective (based on the geometric assumption that a perspective sensor with a long focal length is a good approximation to the parallel sensor) and explicit affine approaches (based on the assumption that the 3D orientation rotations have been previously recovered successfully).

2. Data processing was implemented on the hypothesis that no outliers were present in the system (based on significant pre-processing).

These two assumptions comprise the main weaknesses of the designed method. Firstly, the recovery of starting values lacks a more generalized approach where starting values would be recovered in full from pure implementation of the affine sensor mathematical model. Secondly, system implementation with significant editing ensures that the system is freed from outliers. A more realistic implementation would be to embed within the system a robust outlier detection and elimination scheme based on residuals testing adopted from conventional approaches (see section 4.7.1.).

System assessment and analysis are performed based on a set of statistical indicators that were extracted utilizing the measures reported earlier (see section 4.7.). Particularly statistical indicators will assess system accuracy and precision with most measures extracted from the a posteriori covariance analysis. Aspects of model assessment, algorithm convergence, parameter correlation, 3D object space point correlation with range, error ellipsoids, sensor scale invariance as well as accuracy checks will be part of the testing and analysis. The test datasets were generated from a series of camera systems and geometric objects; these were treated as (a) test, (b) high-order precision reference, or (c) independent test datasets. The design, set-up and data acquirement of close-up image datasets were performed at the UCL's calibration laboratory. Some published results from method development and testing can be found in Rova et al. (2008a), Rova et al. (2008b) and Rova et al. (2009).

6.2. Test datasets

The test datasets are decomposed in three sub-sections according to their (a) design, (b) components and (c) datasets as follows (see sections 6.2.1., 6.2.2. and 6.2.3.).

6.2.1. Design

To design a set of test datasets for method assessment and analysis the following aspects were considered:

- Sensor scale: To address algorithm behaviour in the aspect of scale recovery within the system, a set of different camera systems were utilized.

- 2D image quality: To generate high contrast image data, artificial signalized markers, specifically white markers on a black background and retro-reflective targets were utilized. It is recalled that (see section 3.3.) the employed image measurement method affects the input quality of the image targets. In addition, the invariant affine sensor scale factor resulted in large blobs that significantly reduced image quality (see section 6.2.2.3.). As an example 2mm diameter targets in object space produce 36 pixel diameter blobs under affine projection.

- Image network geometry: Network configurations were designed to ensure geometric strength, convergence and redundancy (see section 4.6.). Particularly wide-angle viewpoints, strong intersection angles and wide separated views were designed. Redundant datasets were generated with sets composed of 17 to 85 views and a number of targets ranging between 20 and 178 targets (see section 6.2.3.).

- 3D point geometry: Point targets were arranged in a sparse 3D distribution following the 3D object geometry, reduce occlusion problems and ensure high redundancy when imaged from variant geometric viewpoints.

- Object space recovery: To recover 3D geometry and scale, a number of different geometric structures were utilized. Specifically these will enable method assessment in 3D space and evaluation of geometry and scale recovery factors.

6.2.2. Components

The basic components that characterize the test datasets are defined as follows:

- The test objects are of geometric 3D type and they are defined with their size, shape and measurement markers (see section 6.2.2.1.).
- The camera systems are characterized by their sensor elements and the attached optical lens that realizes the image to object space projection[24] (see section 6.2.2.2.).
- The image measurement quality is affected by the imaging systems, controlled illumination as well as measured white markers and retro-reflective targets (see section 6.2.2.3.).
- The image networks are acquired to be as highly convergent, dense and geometrically strong as possible (see section 6.2.2.4.).

6.2.2.1. Test objects

Figure 6.1 and Table 6.1 provide a descriptive outline of the test objects that were utilized to generate the test datasets. These are ordered as A, B, C, D and E in a chronological order starting from the most recently acquired datasets. Their design demonstrates the representation of the different geometric test cases. In fact these are utilized to assess the method and extract useful analysis of results.

---

[24] Within this text image space and object space are denoted as IS and OS only within tables of results.
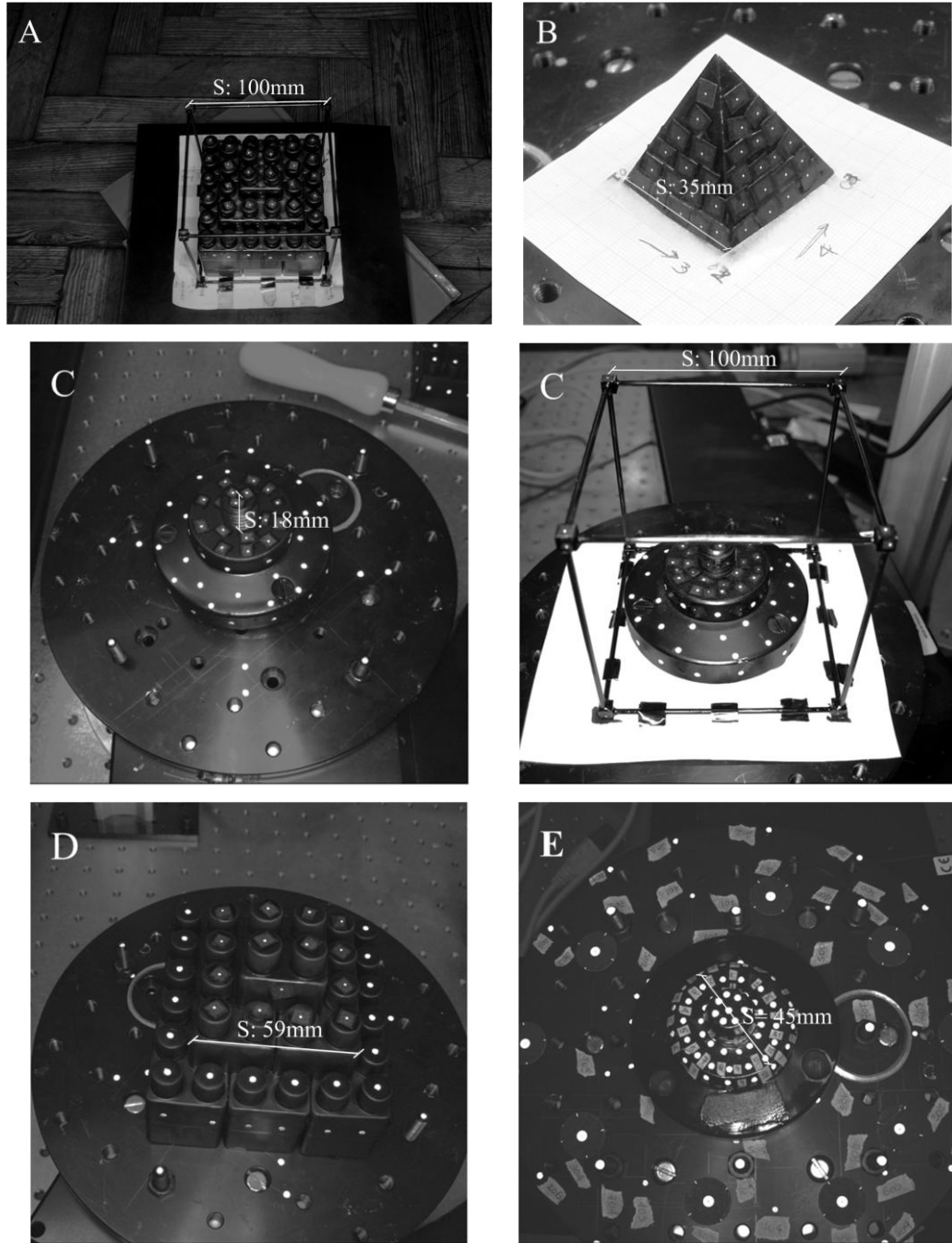
Figure 6.1: Measurement structures per dataset.

The geometric structures 'lego', 'pyramid', 'centroid', 'lego', and 'centroid' correspond to the datasets A, B, C, D and E respectively. The objects are targeted with sparse signalized retro-reflective or white markers of varying diameter that range between 0.5 and 2.0mm in the object space. The general design requires that the targets cover the 3D geometry in full for object measurement (tie points) with CPs located at the objects' edges.

Object ['y]	Geometric structure				Datasets
	Size (mm)	$\emptyset_{WM}$ (mm)	$\emptyset_{RT}$ (mm)	R (mm)	
Lego ['09]	89x104x89	0.5, 1.0	-	175	A
Pyramid ['08]	35x40x35	0.5	-	175	B
Centroid ['08]	45x18x45	0.5, 1.0	2	175	C
Cube ['08,'09]	100x100x100	0.5, 1.0	-	350	C, A
Lego ['07]	59x20x59	1.0	-	175	D
Centroid ['06]	45x18x45	-	2	175	E

Table 6.1: Measurement structures per dataset. Table notation: ['y]= year, $\emptyset_{WM}$= white marker diameter, $\emptyset_{RT}$= retro-target diameter and R= imaging range.

Following the listed objects a detail description is given here:

- Datasets A and D: The test object is of 'lego' construction structured as a stepped 3D volume that covers up to double the dimensions of a single frame and is sparsely targeted to delineate its planar sub-surfaces.

- Dataset B: The test object is a square based pyramid which covers a geometric volume of 35x40x35mm and similarly sparse targeted.

- Dataset C: The test object is a cube, designed as a calibration array to enable pre-calibration and measurement. It is a 3D wireframe (volume: 100x100x100mm) composed of six square faces, eight vertices and twelve edges. Each vertex is represented as a sub-cube (6x6x6mm) that hosts three 0.5mm diameter targets at the three front square faces leaving the remaining three to support the linking edges.

- Datasets E and C: The test object comprising datasets E and C is of centroid type. Following the definition after Wolfram (2009a) a geometric centroid is defined as the centre of mass of a 2D planar closed form surface with a given mass and density. The point data are distributed in order to cover the footprint of the test systems (field of view: ~40mm) and they are arranged in 18mm in depth (depth of field: ~20mm) according to specifications.

6.2.2.2. Camera systems

Data acquisition and testing were implemented with five different camera systems (see section 3.1.2.1. and section 3.1.3.2.). Convenient codes for the systems are given in the following table.

Camera-lens system	Sensor	Code
Nikon D100 [Nikkor f=28mm]	DSLR, Nikon DX format 23.7x15.6mm, RGB CCD, 7.8μm, 3,008 x 2,000 pixels, USB interface	CAM_P1
Kodak Megaplus ES1.0 [Fujinon TV / f=12.5mm]	Progressive scan, monochrome, 9.0μm, 1,008 x 1,018 pixels, RS-422 interface	CAM_P2
Kodak Megaplus ES1.0 [MVO® TMLTM/0.16x]	Progressive scan, monochrome, 9.0μm, 1,008 x 1,018 pixels, 8 bits / pixel 20MHZ, RS-422 interface	CAM_A3
Sony DFW-SX900 [Switar/ f=10mm]	Progressive scan, colour, 4.78μm, 1,024 x 768 pixels, 24 bits / pixel, firewire interface	CAM_P4
Sony DFW-SX900 [MVO® TMLTM/0.16x]	Progressive scan, colour, 4.78μm, 1,024 x 768 pixels, 24 bits / pixel, firewire interface	CAM_A5

Table 6.2: Technical characteristics of camera systems.

The coded systems CAM_P1 (Nikon D100 fitted with a 28mm Nikkor lens), CAM_P2 (Kodak Megaplus ES1.0 fitted with a 12.5mm Fujinon lens) and CAM_P4 (Sony DFW-SX900 fitted with a 10mm Switar lens) are perspective-based. They are utilized to generate high order precision photogrammetric measurements for the purposes of pre-calibration, pre-measurement, data initialization, generation of reference measurements and independent testing. It is recalled here that the term 'photogrammetric measurements' denotes sparse data generation through robust bundle adjustment computations within the in-house photogrammetric software tool VMS 8.0.

Equivalently, the coded systems CAM_A3 (Kodak Megaplus ES1.0 fitted with an MVO® TMLTM/0.16x lens) and CAM_A5 (Sony DFW-SX900 fitted with a MVO® TMLTM/0.16x lens) were utilized to generate the affine image datasets required for method testing. It is stated here that the deployed telecentric optical system is the physical approximation of the affine camera sensor in the close range, resulting in a 40mm footprint in the object space at a 0.16x nominal image scale.

6.2.2.3. Image quality

Image quality factors as well as target measurement method are critical to image acquisition. Image capture, transmission and processing can affect the critical image factors (such as noise, dynamic range, sharpness, contrast, distortion, vignetting, exposure and artefacts) (see section 3.2.1.2.). Here, to achieve uniform illumination external high frequency fluorescent ring lights were utilized with simultaneous control

on the imaging settings (aperture, focus, exposure, gain, brightness, gamma and stability) on each camera system and accompanied software. In such a measurement environment, the imaged point features are affected by illumination at acquisition time in case these are white markers whereas retro-reflective targets present high contrast images when illuminated from the camera's viewing direction. Image quality was controlled externally to the developed method; as a result the data were generated without applying any image pre-processing method. However, data generation resulted in a variation of the point data per dataset. According to Figure 6.2, the targets are displayed under perspective and affine image projections where it is shown that differences in resolution, imaging conditions, object space target diameters and projection scale affect 2D image measurement quality. To make this point clear, Table 6.3 gives a description of some extracted image characteristics that correspond to these target data.

PERSPECTIVE

a:2mm Ø RT/KD d_{RGB}: 1mm Ø WM /NK   e: 1mm Ø WM/KD

AFFINE

f: 2mm Ø RT/KD   i: 1mm Ø WM/KD j_{RGB}: 1mm Ø WM/SN
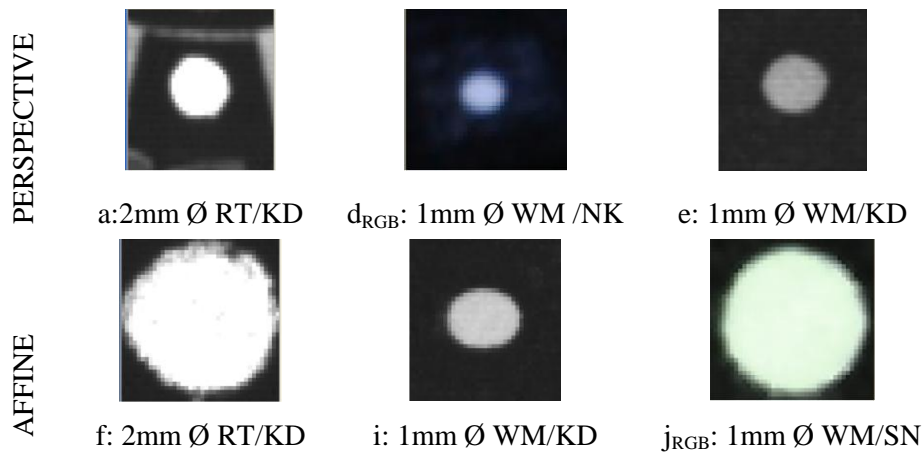
Figure 6.2: Perspective and affine image targets. Figure notation: RT= retro-reflective, WM= white marker, KD= Kodak sensor, NK= Nikon sensor, SN= Sony sensor.

Camera	PP	AP	PP	AP	PP	AP
Target	(a)	(f)	(d_{RGB})	(i)	(e)	(j_{RGB})
Ø (pixels)	16	36	10	8	8	34
FB	255	255	175, 188, 232	207	166	236, 251, 232
BB	37	37	17, 23, 35	33	37	23, 25, 22

Table 6.3: Signalized target image characteristics. Table notation: Ø= target diameter, PP= perspective projection, AP= affine projection, FB= foreground brightness, BB= background brightness, (a), (f), (d_{RGB}), (i), (e) and (j_{RGB})= coded image targets.

Overall, the employed point types are of 2mm (retro-reflective targets), 0.5mm and 1mm (white markers) diameters in object space. Individual target images range within 4-16 pixels under perspective and 17-36 pixels under affine projection respectively. It

is evident that the white markers present low contrast whilst the retro-reflective targets are saturated. For example it is evident from Table 6.3 (examples (e) and (a)) that the intensity value of the centroid equals 166 in the case of the selected white markers, whereas in the case of the shown retro-reflective targets the associated intensity value equals 255.

Besides characterizing target intensity, typical problems are sharpness and shape loss. Particularly, where a target fails to be located within the depth of field, the imaged blob appears blurred, reducing image and measurement quality. In cases, where the imaged surface is located nearly parallel to the viewing direction, circular targets appear as ellipses. These characteristics are given by an illustration of an additional set of targets with their associated histograms (see Figure 6.3). It can be particularly seen that target B is projected as an ellipse when imaged under affine projection showing the effect of the deviation of the viewing projection rays from the surface normal where its sharpness loss is attributed to its location outside the depth of field. A further description of some extracted targets that were selected to describe image quality can be found in Appendix C.
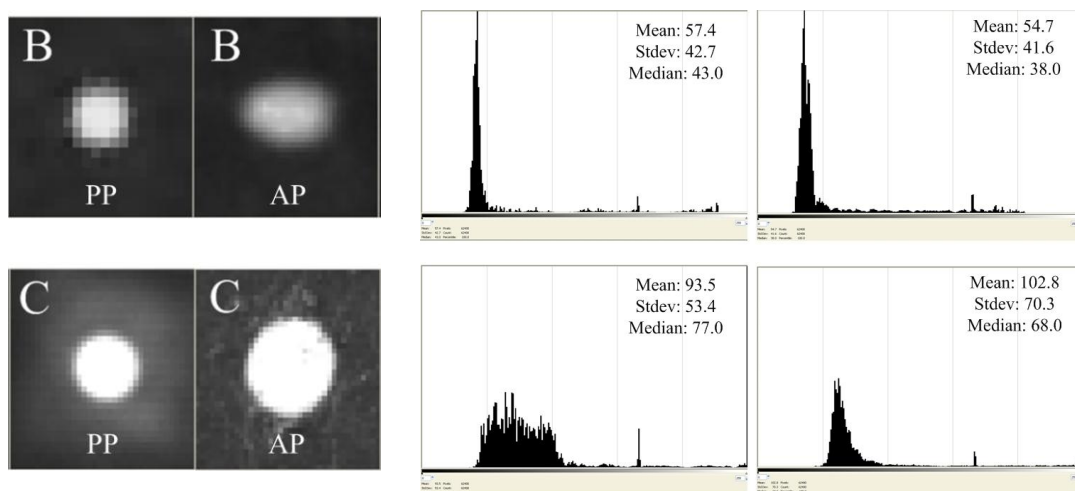
Figure 6.3: Signalized image target pairs of perspective (PP; left) and affine (AP; right) views and associated brightness histograms.

Image measurements were initialized within the software tool VMS 8.0. It has already been reported (see section 3.3.2.) that the software applies a set of point-based methods (manual, centroid, correlation, least squares matching). Particularly target measurement was performed with a combination of methods including manual point

digitization and centroid estimation with the features of epipolar drawing and backdriving turned on. Dependent on the target image quality (content), the centroid estimation method was empirically adapted to set optimal centroid measurement parameters with geometric tests that allow sub-pixel point measurement. Table 6.4 gives the empirically set parameters for the selected sample of targets (for a further description see Appendix C).

Target	Ø (pixels)	IW (pixels)	LM	TT	S	GT
B [PP]	6	20 x 20	w_centroid	histogram	circular	on
B [AP]	17	40 x 40	ellipse_fit	histogram	circular	on
C [PP]	8	36 x 36	w_centroid	histogram	circular	on
C [AP]	18	38 x 38	w_centroid	histogram	circular	on

Table 6.4: Measurement method parameters within VMS 8.0. Table notation: Ø = target diameter, IW= image window, LM= location method, TT= threshold type, S= shape, GT= geometric tests.

The software produced successful results in areas with significantly small features (diameter less than 5 pixels) and symmetrical homogeneous patterns and even in more difficult cases with dominant blurred points. However close-up network characteristics (such as imaging range, strong intersection angles, viewing direction, projection scale and depth of field) significantly increased target image diameter and degraded image quality. As a result, in cases of poor target image detection it was necessary to manually measure the point target data.

6.2.2.4. Image networks

The image networks were acquired in strong convergent configurations from multiple viewpoints in consideration of the imaging conditions, object geometry, point distribution, visibility and redundancy factors. Figure 6.4 visualizes a sample of three selected perspective and affine image networks as a 3D lattice of points and observations.
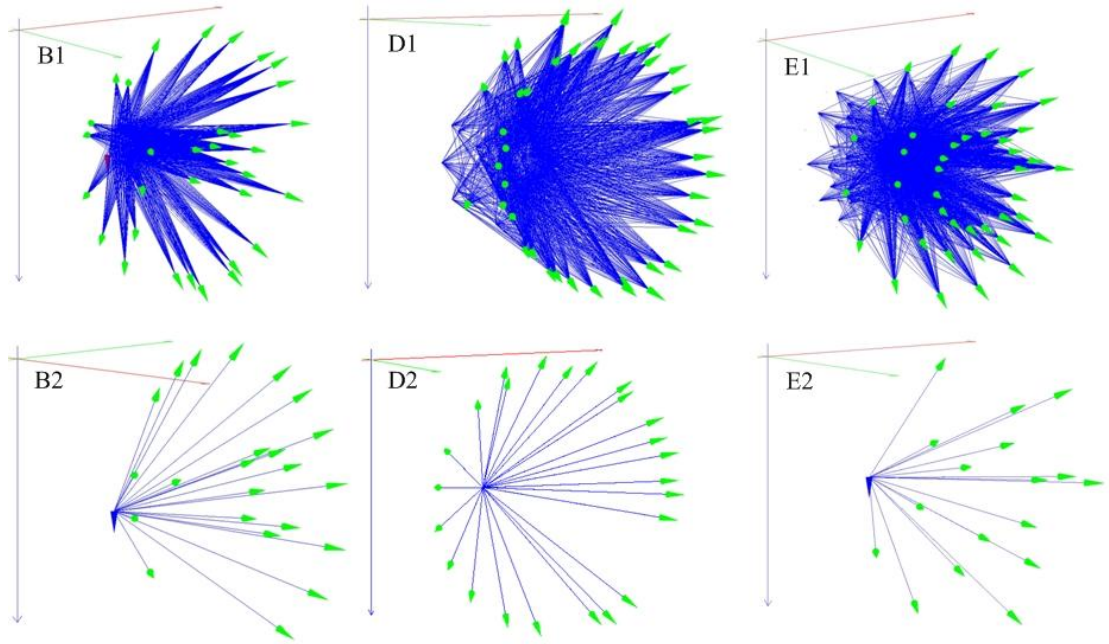
Figure 6.4: Image networks configured under perspective and affine views (green cones: photo location, blue lines: observations).

Particularly the coded datasets B1, D1 and E1 describe the geometry of the perspective image networks whereas the coded datasets B2, D2 and E2 describe the geometry of the corresponding affine image networks. It is evident that perspective and affine image networks were designed as far as possible to form consistent imaging geometry for further processing and testing.

6.2.3. Datasets

Within this work twelve different datasets were designed for processing and checking. These are grouped utilizing the selected codes A, B, C, D and E according to object type (see section 6.2.2.1.). Table 6.5 tabulates these datasets in consideration of the employed camera systems (see section 6.2.2.2.) which are now described in relation to the designed image networks (see section 6.2.2.4.) with the associated views and targets that contribute in the subsequent data measurement and testing.

Dataset	Camera system	Views	Targets
A1	CAM_P1	44	178
A2	CAM_A3	49	122
A3	CAM_A5	85	86
B1	CAM_P4	33	65
B2	CAM_A5	23	44
C1	CAM_P2	46	61
C2	CAM_P2	41	25
C3	CAM_A3	45	29
D1	CAM_P2	45	52
D2	CAM_A3	24	20
E1	CAM_P2	43	89
E2	CAM_A3	17	54

Table 6.5: Processed datasets for testing.

Specifically, dataset C3 is designed to evaluate the method's behaviour regarding full bundle adjustment treatment. Assessment is based on the extracted typical quality indicators, measures that assess the behaviour of convergence and the consistency of parameter correlations. Datasets B2, D2 and E2 are utilized to assess the method in relation to object space recovery. This test utilizes measures based on analysis of correlation behaviour with range as well as a visualization of the error ellipsoids. Image networks A2 and A3 are utilized to test the system's scale for two different camera systems again based on standard statistical indicators as well as object space accuracy checks. Finally, dataset C3 is selected to independently test the method compared to the results obtained from dataset C2 which is processed with a perspective bundle adjustment implemented within the software tool VMS 8.0. Again assessment is performed utilizing the standard statistical indicators as well as evaluation of 3D discrepancies.

6.3. Practical aspects

The application experiments focus on method assessment and analysis, particularly in evaluation of aspects of both method's benefits as well as defects. Whilst practical issues will be conditionally dependent on each sequential testing step from the data acquisition to the data processing phase, here all tests were performed for very close-up imaging situations.

Particularly, the method's working range is limited to some hundreds of millimetres (imaging range= 175mm) with a significantly small depth of field (depth of field=

±19.7mm) tolerance. Whilst these ranges set some practical limitations, network geometry and targeting design will need to allow high precision measurements. It is generally expected that small diameter targets will generate a few-tenths of pixel target images which will in effect suffice for perspective and affine imaging. However the smallest selected target size was designed as non retro-reflective 0.5mm diameter white marker due to physical construction limitations. These point targets result in 17 pixel diameter target images when for example imaged with the affine Sony camera system.

Image network geometry varies per object-case according to the requirements of object coverage, visibility and occlusion as well as strong intersection rays. In addition, projection magnification and scale invariance within the image frame in relation to the sensor's small field of view constrain object size. Specifically the test objects (see section 6.2.2.1.) had to be highly local with sufficient characteristic features and density that allowed measurement and simultaneously produced a sample that could be used to extract useful model behaviour and assessment. Natural textured objects or surfaces with point-based features were not available for testing. In fact such datasets would not introduce significant information in the scope of testing the developed approach. However datasets of different nature (e.g. dense point clouds generated from photogrammetry or laser scanning) could act as good reference datasets for independent system evaluation.

Regarding object measurement is such cases where the dimensions of an object occupy a volume larger than the camera's field of view there will be additional requirements regarding initialization as well as tie point location for object coverage and measurement. As a result, these demands increase cost related to image datasets volume, pre-processing and editing. Further issues that emerge from data quality, initialization or data processing of the test data will be addressed in the subsequent sections as these may vary per experimental case.

6.4. Model assessment

This section assesses model behaviour in the aspects of method initialization, datum constraints and calibration parameters within the system. Method assessment is

evaluated with application of the method on datasets C (see section 6.2.3.) utilizing the extracted statistical indicators, system convergence analysis and correlation measures.

6.4.1. Initialization

To initialize the method two steps were followed. First the generation of the reference image measurements is given which is subsequently followed by the stage of derivation of starting values for the affine image dataset.

6.4.1.1. Reference measurements

The first aim is to provide reference measurements. To achieve this goal the convention is to acquire perspective image data in strong, convergent, redundant image network arrangements which are subsequently processed with established robust bundle adjustments.
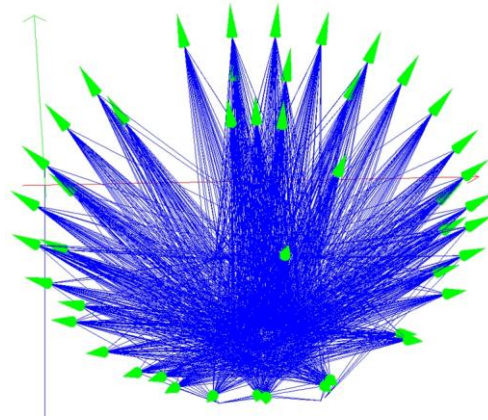
Figure 6.5: Image network geometry of reference dataset.

Figure 6.5 illustrates the designed image network geometry for dataset C1 which is composed of 46 convergent images, 61 point targets (20 CPs and 41 TPs) and 6 measurement scales that were defined for correct object scale recovery. The input data quality (for CPs and scales) was considered to be equal to 25µm after a target-to-target edge measurement of points with the available digital callipers. In object space the target data vary between 2mm diameter retro-reflective and 1mm white markers. Although control is distributed in a highly 3D configuration, image quality in conjunction with the imaging direction of the lines of sight (some targets are

occluded) result in an inability to fully locate control in 3D in some views. As a result, the control volume is reduced from 3D to 2D (control lie on a plane). This loss is tackled ensuring highly redundant measurements and introduction within the system of six object space scales (four located on the top plane and two located on the base).

The data processing framework involves a series of iterative passes. Given the knowledge of control these involve initial exterior orientations updated by resection approaches that estimate the exterior orientation parameters. Subsequent forward intersections coordinate the tie point data in 3D. When it is considered that sufficient (as complete and redundant as possible) data are measured, a bundle adjustment is processed on the final estimation step. The bundle is processed with the inner datum method and an additional parameters model which includes only the first two radial lens distortion terms (k_3 and k_5) according to their significance within the system (see Appendix C). The solution converged after 4 iterations with an RMS image residual of $1/10^{th}$ of a pixel (~0.91µm) (see Table 6.6). In object space the 3D targets are estimated with a precision of 10.75µm and a relative precision for the image network of 1:17,000. Within this thesis, relative precision is estimated as the mean estimated coordinate standard deviation divided by the maximum 3D dimension (usually bounding box diagonal) in the network. In other words, it follows that for a cube of side α= 10cm, its calculated diagonal (where diagonal is given as: $\alpha \times sqrt(3)$) equals 17.321cm, resulting in a 2D image precision of 10.20µm which confirms the above quoted result.

Dataset	Iter.	R.	σ_o	RMS$_{xy}$ (µm)	σ_{XYZ} (µm)	Relative precision	RMS scale (µm)
				IS	OS		
C1	4	2,667	1.00	0.91	10.75	1:17,000	101.32

Table 6.6: BA statistical indicators - dataset C1. Scales: 6, ADPs: x_o, y_o, c, k_3, k_5, constraints: inner. Table notation: Iter.= iteration, R.= redundancy, σ_o= unit weight, RMS$_{xy}$= image misclosure, σ_{XYZ}= 3D points standard deviation, RMS scale residual= RMS residuals in OS scale.
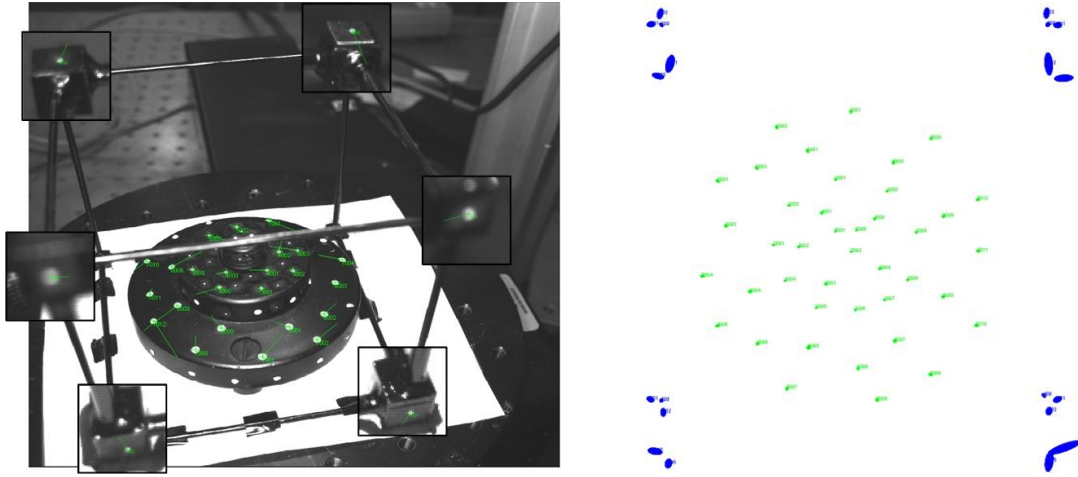
Figure 6.6: BA image residuals with highlighted CPs (left) and error ellipsoids (CPs: blue and TPs: green) (right).

An additional check is obtained from observation of the distribution patterns of the bundle adjustment residuals; their magnitude (0.91μm ~1/10th of a pixel) and random distribution assure the correctness of the conventional BA solution (see Figure 6.6 left). Moreover, significant is the fact that the inner constraints datum generally results in a uniform quality precision (uniform error ellipsoid shape) with their large semi axes pointing towards the centroid of the cloud of points and vertical axis indicating the error in the viewing direction (see Figure 6.6 right).

The introduction of the scale measurements within the image network result in a small decrease in precision for the control points which are specifically located at the base of the cube as these are coordinated from fewer viewpoints. However control is limited; this in combination with its poor quality pose scale implementation within the bundle adjustment necessary to restore scale in the object datum. To analyze the 3D target quality it is useful to check the distribution of the 3D precisions in the 'XZ' (planimetric) and 'Y' (depth) directions. Figure 6.7, Figure 6.8 and Table 6.7 provide a description of the a posteriori precisions of the data. The mean precisions are estimated to be equal to σ_{XZ}= 15.52μm and σ_Y= 29.27μm for the CP data and σ_{XZ}= 5.52μm σ_Y= 7.54μm for the TP data. It is indicated (from the highlighted point data) that the worst achieved precision in the viewing direction (Y) is 98.00μm for the CPs and 9.70μm for the TPs. In fact maximum standard deviations occur in CP104 (count 5) and CP105 (count 6) which are located on the base plane of the calibration cube and they are coordinated from 4 viewpoints, whereas minimum standard deviations

occur in CP500 (count 18) which is located on the top plane of the cube and it is coordinated from 43 viewpoints. Maximum and minimum standard deviations in the case of TP estimation were observed in TP7004 (count 23) which is located on the external ring of the centroid (2mm diameter retro-target) coordinated from 30 viewpoints and minimum standard deviations in TP5001 (count 8) which is located on the top, inner plane (1mm diameter white marker) coordinated from 39 viewpoints. These numbers indicate the measurement quality of the reference data. Particularly they play a critical role as they characterize the input data quality; that is they initialize the stochastic model in the subsequent processing stage (see section 6.4.1.2.).
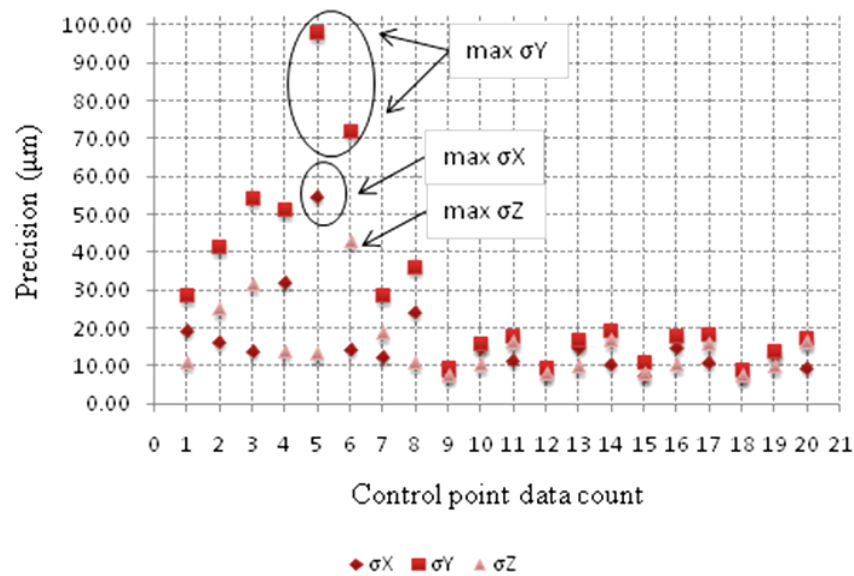
Figure 6.7: BA a posteriori 3D precisions - CPs.
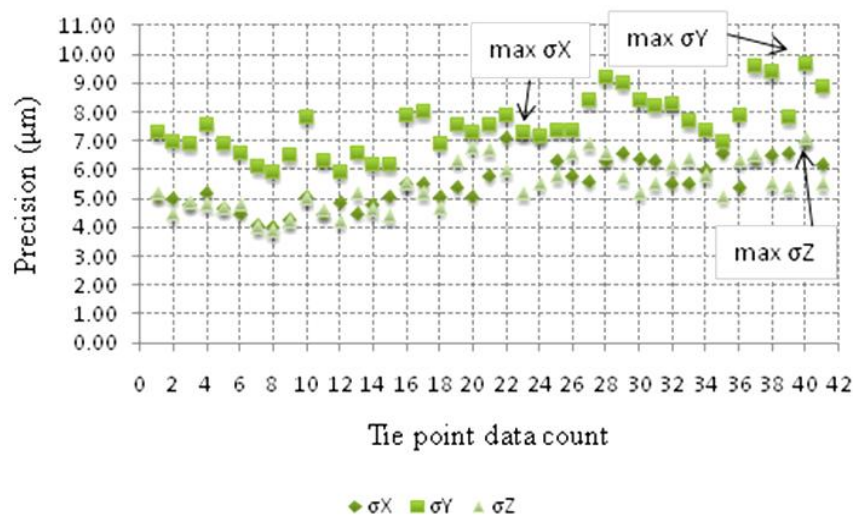
Figure 6.8: BA a posteriori 3D precisions - TPs.

	Control points			Tie points		
	σX (μm)	σY (μm)	σZ (μm)	σX(μm)	σY (μm)	σZ (μm)
Mean	15.82	29.27	15.23	5.60	7.54	5.44
Max	54.90	98.00	43.00	7.30	9.70	7.10
Min	7.30	8.80	7.50	4.00	5.90	3.90

Table 6.7: BA a posteriori 3D precisions - dataset C1.

6.4.1.2. Affine starting values estimation

The pre-measured centroid object is now utilized to generate and subsequently process the affine image dataset. Specifically 'dataset C3' is composed of 45 images (22 front views, 18 top views and 5 views oriented on their optical axis), 9 control points (6 located on the exterior basis and 3 located on the top plane of the centroid) and 20 tie points.
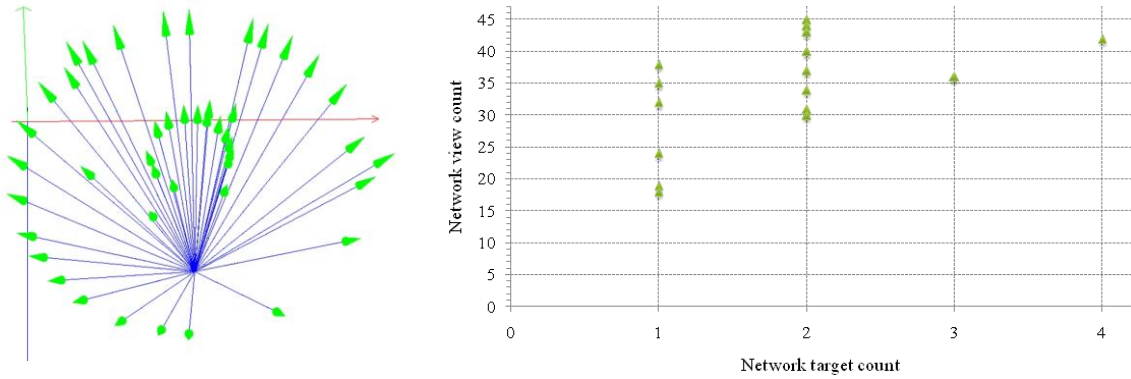
Figure 6.9: Affine image network geometry (left) and data visibility (right).

From Figure 6.9 it can be seen that the image network is geometrically strong, convergent and redundant with high target visibility. For example worst visibility case is the point TP10004 which is observed in 18 views. It is re-iterated here (see section 5.4.) that data initialization is recovered from the generated CP data (and their associated precisions) based on an initial exterior orientation and resection approaches that estimate the 3D orientation rotations. This outputs an orientation file with the 3D photo rotations and the nominal image scale (s= 0.16) for the affine sensor. Next, a direct closed form back-substitution on the affine functional model estimates the 2D projective translations and updates the orientation file. 3D target coordinates are estimated from an affine-based intersection approach. This updates the control and locates all new tie point data (TPs). The results of the intersection procedure are given

in Table 6.8 from where it is evident that the method converged after 2 iterations with a posteriori sigma of 3.18 and a mean 3D point precision of 2.31μm.

Dataset	R.	Iter.	σ_o	σ_{XYZ} (μm)
C3	2,017	2	3.18	2.31

Table 6.8: 3D intersection statistical measures - dataset C3.

Illustration of the residual vectors of both stages of resection and intersection (see Figure 6.10) makes it evident that the combination of the residuals magnitude and their random distribution indicate the correctness of the initialization approach.
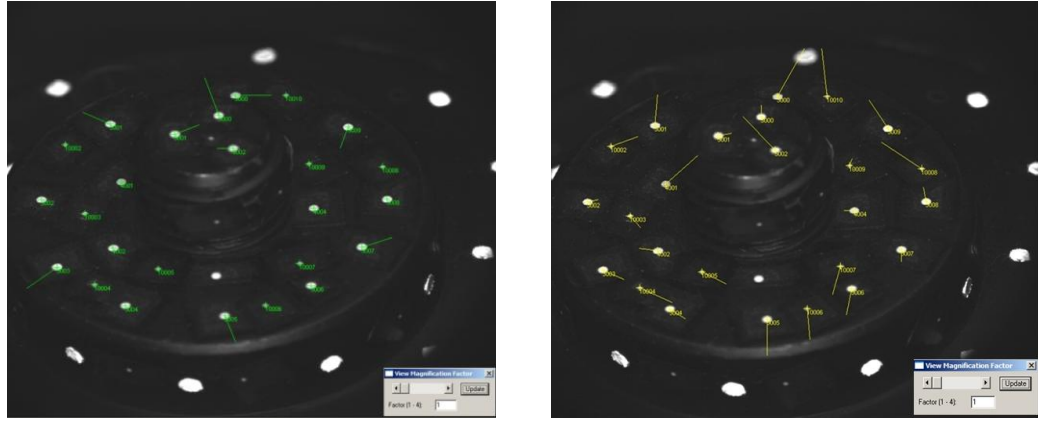
Figure 6.10: Residuals. Resection CPs (left) and intersection - CPs and TPs (right).

6.4.2. Affine bundle adjustment results

Bundle adjustment performance is assessed in the aspects of evaluation of method behaviour as well as convergence with parameter estimation. Assessment is obtained utilizing the typical statistical indicators and additional measures that are defined at each test case. It is however important to note that where precisions have been used, these are extracted from the a posteriori covariance matrix which is scaled to the a posteriori standard deviation.

6.4.2.1. Model assessment

In this experiment (dataset C3) a comparative set of bundle adjustments was run to assess both external and inner constraint datum methods. Particularly the different runs were coded after point initialization (CP or TP) and datum implementation

(external or inner) with the calibration model (inclusion or exclusion of k_3 term) also implemented within the different runs. For comparative assessment all bundle adjustments were processed with identical calibration, orientation and 3D target data. To maintain datum as defined at its initial set up, the 3D control data were set to their pre-determined coordinates together with their associated precisions as these were identified from the reference image network (mean $\sigma X= 15.82\mu m$, $\sigma Y= 29.27\mu m$, $\sigma Z= 15.23\mu m$) (see section 6.4.1.1.). It follows that the remaining tie points were updated from their intersected 3D target coordinates whereas a uniform precision of 5µm was set in the target file. Table 6.9 summarizes the statistical indicators of the bundle adjustment results. The solutions converge with an a posteriori sigma of ~2.00 which indicates that the initialization of the stochastic model is potentially over-estimated. It is additionally noted that the default input quality of the image observations is set to $\sigma_{xy}= 0.5\mu m$, whereas the 3D precisions of the control (external constraint datum) can be considered as too optimistic. The overall bundle adjustments converge rapidly after 2-3 iterations with an RMS image residual of $1/11^{th}$ pixel ($RMS_{xy}= 0.8-0.9\mu m$ for the 9.0µm pixel size Kodak sensor).

C3	Iter.	R.	Scale	$k_3 \times 10^{-4}$ ($\sigma_{\kappa 3} \times 10^{-6}$)	σ_o	RMS_{xy} (µm)	$\sigma X, \sigma Y, \sigma Z$ (µm)
CTPE	3	1,813	0.1611 (0.0000)	1.3048 (6.0521)	2.26	0.79	5.44, 7.66, 5.39
CTPE	3	1,814	0.1613 (0.0000)	-	2.53	0.88	6.10, 8.59, 6.05
CPE	3	1,203	0.1611 (0.0000)	1.1834 (7.1455)	2.24	0.77	3.94, 5.78, 3.90
CPE	3	1,204	0.1613 (0.0000)	-	2.48	0.85	4.36, 6.41, 4.33
CTPI	2	1,791	0.1611 (0.0002)	1.3012 (6.0730)	2.26	0.79	252.42, 253.24, 252.41
CTPI	2	1,792	0.1613 (0.0002)	-	2.54	0.88	283.20, 284.12, 283.20
CPI	2	1,156	0.1611 (0.0001)	1.1749 (7.2265)	2.26	0.77	125.87, 126.39, 125.89
CPI	2	1,157	0.1613 (0.0001)	-	2.50	0.85	139.61, 140.19, 139.63

Table 6.9: Affine BA statistical indicators - dataset C3. Model parameters: tx, ty, ω, φ, κ, X, Y, Z, s, k_3. Table notation: CTPE= control, tie points, external constraints, CPE= control points, external constraints, CTPI= control, tie points, inner constraints, CPI= control points, inner constraints.

The 3D targets are coordinated with an increased σ in the viewing direction (Y) between 5.78μm and 8.59μm for the solutions processed with the method of external constraints and between 126.36μm and 284.12μm for the solutions processed with the method of inner constraints. It follows that there is a significant difference in 3D precision between external and inner constraints. In fact, the inherent 3D object scale within the inner constraints method reduces precision due to a potential correlation of the uniform sensor scale with the datum scale factor which can pose inner datum scale as impractical within the system. A trial to remove object space scale from the datum equations resulted in an inversion problem of the normal equations matrix, hence this problem was not investigated further. In detail (see Table 6.9) the external method results in a mean 3D precision of 6.16μm (CTPE solution) and 4.54μm (CPE solution). Yet, inner datum resulted in a mean 3D precision of 252.69μm (CTPI solution) and 126.08μm (CPI solution). The image residuals present random patterns and a visual inspection of their histograms shows that they follow the normal distribution. Figure 6.11 illustrates an example view (CPE and CPI solutions) with their associated histograms (residuals grouped in 10 bins with associated statistics: Solution CPE: $stdev_x$= 1.00μm, $mean_x$= 2.24μm; $stdev_y$= 1.03μm, $mean_y$= 1.82μm and Solution CPI: $stdev_x$= 1.00μm, $mean_x$= 1.26μm; $stdev_y$= 1.03μm, $mean_y$= -0.42μm).
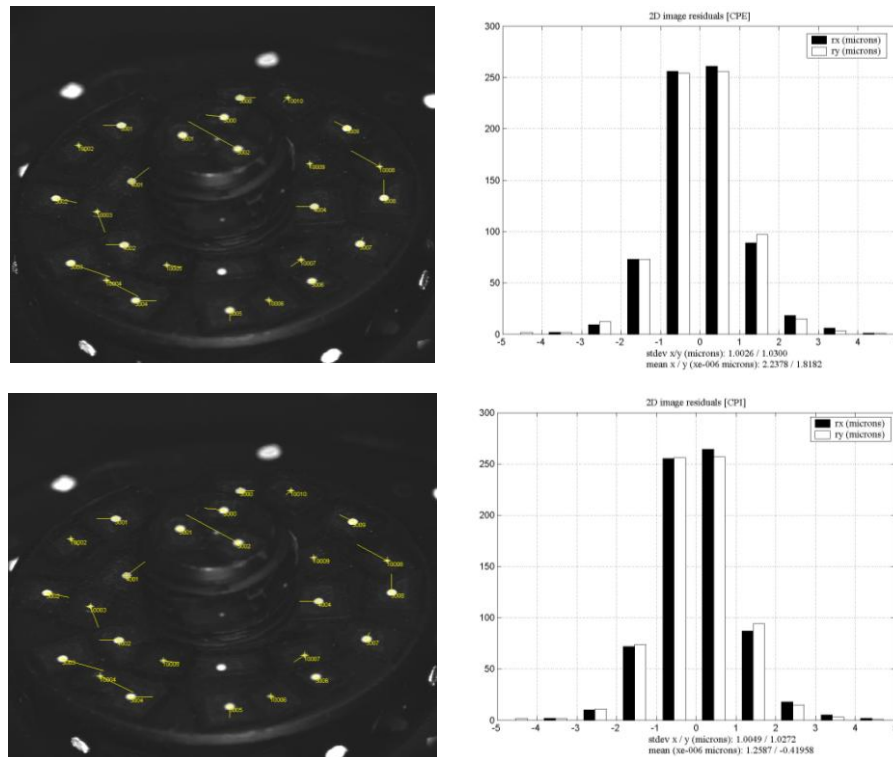
Figure 6.11: Affine BA residuals and histograms - CPE (top) and CPI (bottom).

Considering system calibration, both external and inner constraints solutions decrease 3D target precision when not-accommodating for the k_3 term (see Table 6.9 for estimated k_3 parameter and associated precision). The radial lens distortion profiles (see Figure 6.12 left) estimate a radial distortion of dr= 8.3µm (positive, pincushion distortion) at a maximum radial distance of r= 4mm of the image format which agrees with the telecentric system specifications (maximum distortion < 0.3%). In addition, the fact that these patterns are similar for both datum cases provides the confidence that the functional model is correct.
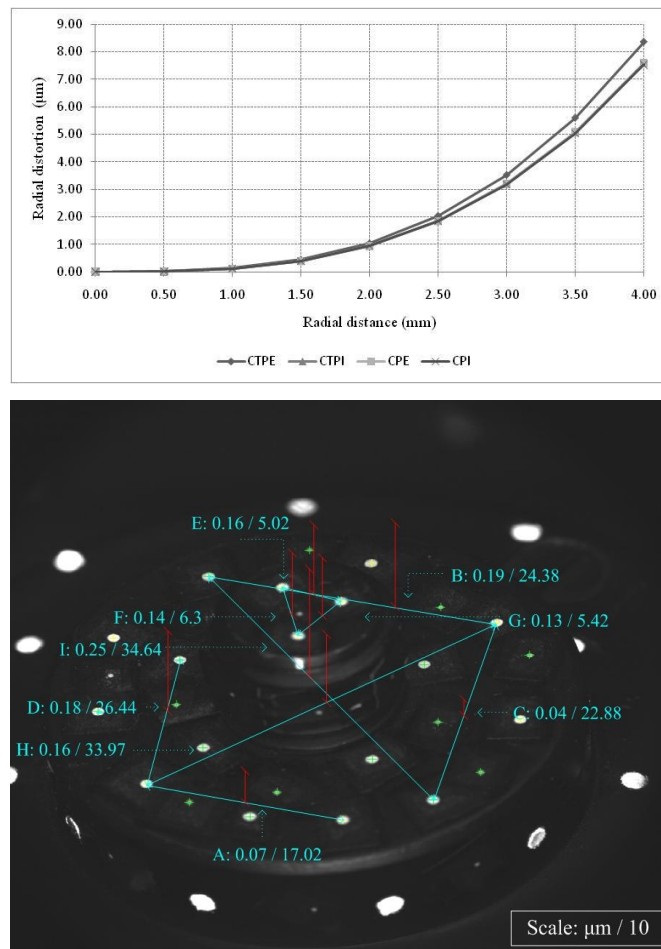
Figure 6.12: Radial lens distortion profiles (left) and 3D distances (right) - CTPE.

Finally with regards to accuracy evaluation nine checks were performed on selected 3D distances and their corresponding estimations from the bundle adjustment solutions. 'Ground truth' was generated with measurements obtained with a digital calliper considering a measurement uncertainty of ±25µm, noting though that a more realistic precision would be equal to σ= ±50µm (empirical value given the manual measurement uncertainty). The calculated 3D differences (true against evaluated from

the bundle adjustment) appear to be consistent for all processed solutions varying between 35.02µm and 250.53µm. Figure 6.12 (right) illustrates the linear 3D differences between true and evaluated dimensions as vertical error displacements (red lines) on the associated distances (cyan lines) for one selected solution (coded CTPE). For example, the 3D displacement of the distance 'H' is 0.16mm (it is scaled to 15.6 mm for the purpose of visualization) and it is displayed over a length of 33.97mm. The observed large discrepancies in the data can be attributed to the initialization of the stochastic model for the intersected 3D target data (σ_{XYZ}= 20.11µm for CPs and σ_{XYZ}= 5µm for TPs) that contribute and hence tie the measurements to the defined datum. In addition these significant differences can result from the uncertainty in measurement precision of the available callipers.

6.4.2.2. Convergence behaviour

In particular demanding situations that for example require processing of large datasets an additional aspect of bundle adjustment processing is to ensure the method's algorithmic efficiency. Here the algorithm is run as an iterative method with inversion treated with the external and inner constraints routines that have already been referenced (see section 5.5.2.2.). The test datasets were processed ensuring that no outliers were present in the measurements. The bundle adjustments converged rapidly after 2-3 iterations presenting high numerical stability.

To evaluate model behaviour with convergence, two fully controlled solutions (CPE and CPI) were utilized (see Table 6.9). The first utilized measure is the normalized span of model parameters which evaluates the change in model parameters between successive pairs of iterations (iteration n+1, iteration n) scaled to the magnitude of model parameter at iteration (n+1) (see equation (6.1)) and which is visualized over the model parameter count in the following figures.

$$NSMP = \frac{ABS(P_{n+1} - P_n)}{|P_{n+1}|}$$
(6.1)

Where:

NSMP= normalized span of model parameters

P_{n+1}= parameter value at iteration n+1

P_n= parameter value at iteration n

Figure 6.13 illustrates the derived NSMP between iteration pairs 1 and 0 for the overall parameters that correspond to the CPE solution (external constraints with full control).
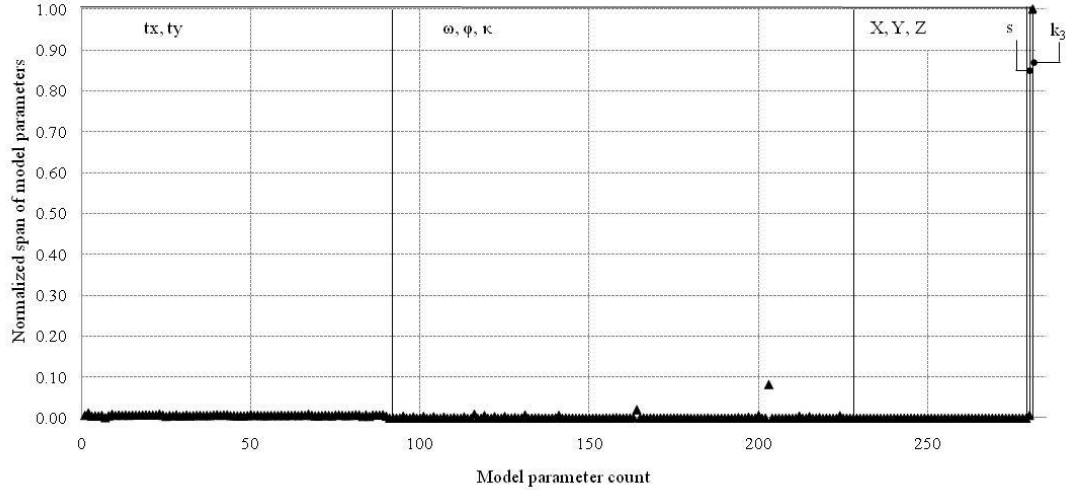
Figure 6.13: Normalized span of model parameters (full). CPE - Iter. 1-0.

To illustrate in detail model behaviour, Figure 6.14 provides the derived figures for the model parameters excluding the k_3 term from the visualization of the NSMP values as in the first pair of iterations NSMPk_3= 1.0, enhancing as a result the inter-structure of the derived pattern. It can be seen that this indicator presents a relatively stable behaviour up to 0.02. Extreme values (highlighted points in the graph) that deviate from the average pattern behaviour were observed for two rotation elements ω_{1025}= -53.2453 degrees where NSMP(ω_{1025})= 0.020 and ω_{1038}= -131.4346 degrees where NSMP(ω_{1038})= 0.082. The indicator for the scale factor s=0.1600 is given as NSMP(s)= 0.007.
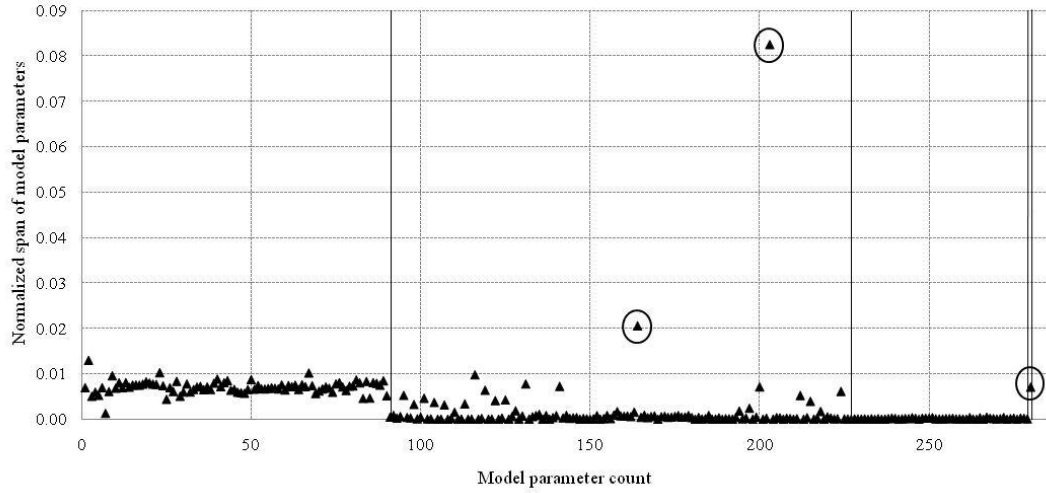
Figure 6.14: Normalized span of model parameters (-k_3). CPE - Iter. 1-0.

Accordingly Figure 6.15 and Figure 6.16 illustrate the convergence behaviour between iterations 2 and 1 as well as iterations 3 and 2. It can be seen that the model presents relatively stable behaviour in terms of data agreement. In the case of the iteration pair 2 and 1, the previous extreme cases are now given as NSMP(ω_{1025})= 3.05×10^{-4} and NSMP(ω_{1038})= 6.28×10^{-4} for the above observed rotations whereas the equivalent index for the scale factor is now NSMP(s)= 6.32×10^{-7} and radial lens distortion k_3 term is NSMP(k_3)= 5.52×10^{-4}. In the subsequent iteration pair 3 and 2 the associated values are given as NSMP(ω_{1025})= 8.71×10^{-7}, NSMP(ω_{1038})= 1.15×10^{-7}, NSMP(s)= 6.21×10^{-9} and NSMP(k_3)= 2.37×10^{-6}. It is again noted that these figures are shown in the graphs as highlighted points.
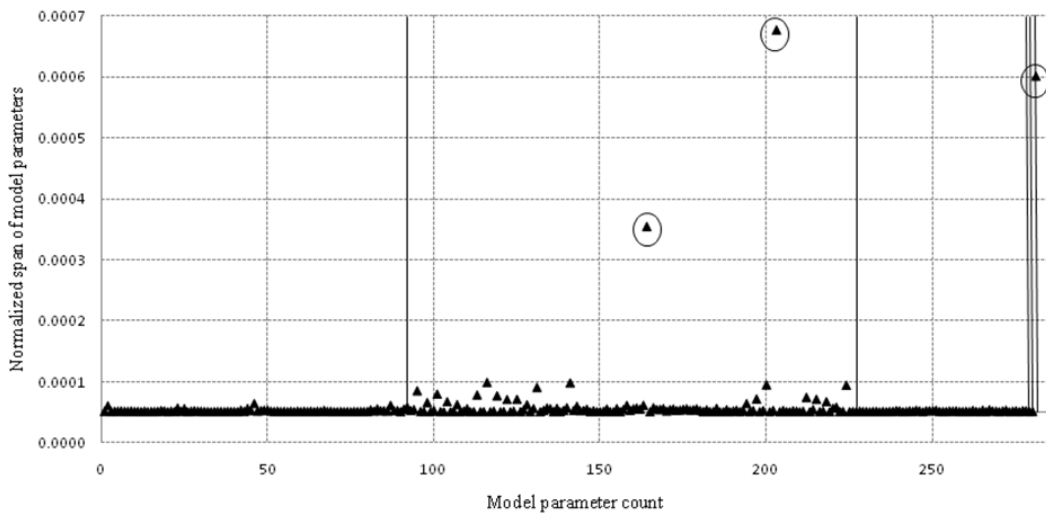
Figure 6.15: Normalized span of model parameters (full). CPE - Iter. 2 - 1.
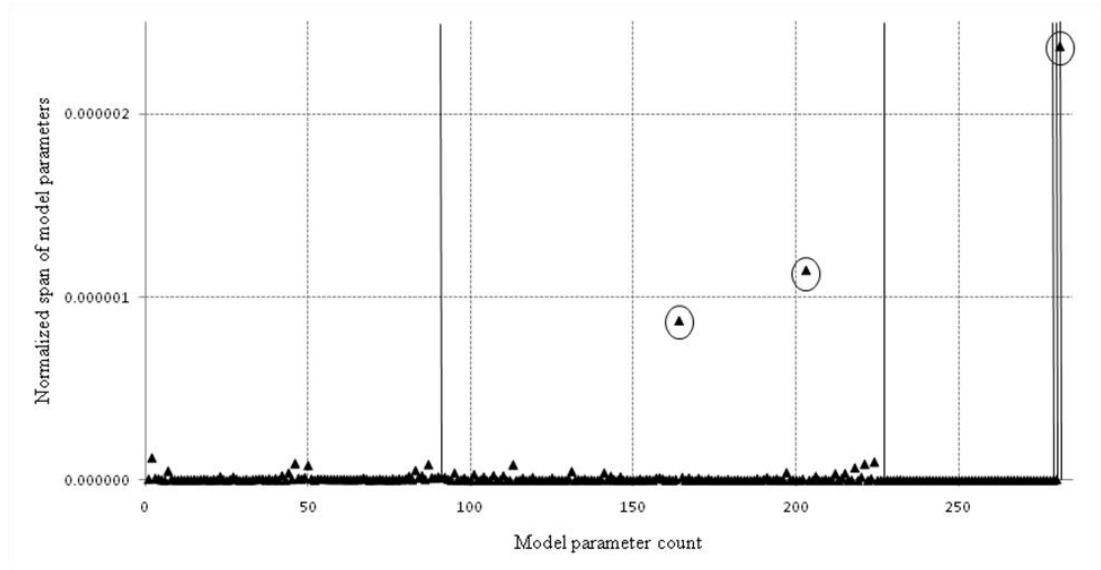
Figure 6.16: Normalized span of model parameters (full). CPE - Iter. 3 - 2.

It is evident from the above graphics that the individual pairs of iteration passes have a clear relation. Specifically the NSMP values of iteration pair 2-1 as compared to its previous iteration pair 1-0 are approximately two orders of magnitude smaller. Accordingly the iteration pair 3-2 in comparison to the iteration pair 1-0 presents a five order magnitude difference for the NSMP value. Besides convergence behaviour it is additionally important to check the precisions of the estimated parameters at the final iteration stage. Figure 6.17 illustrates the parameter precisions for the CPE solution as these are extracted at the third iteration stage for the total number of parameters.
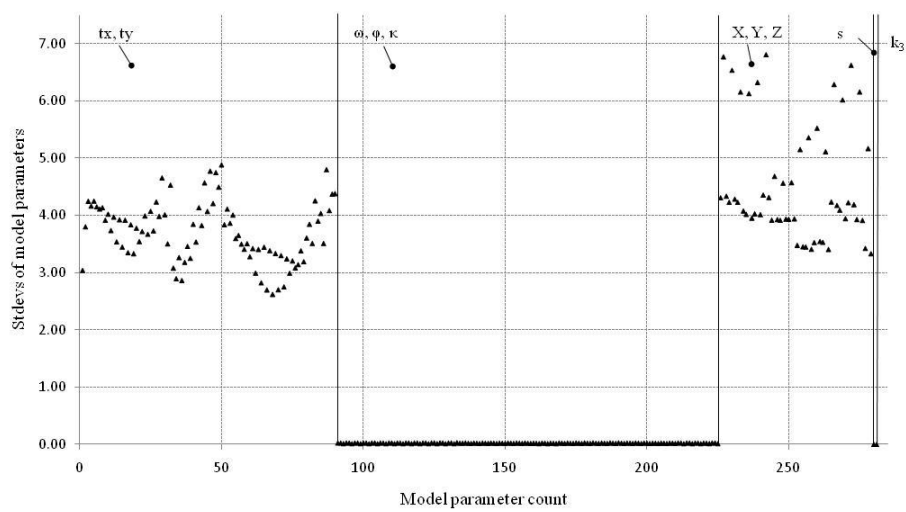
Figure 6.17: Precisions of model parameters. CPE - Iter. 3.

To better describe the above patterns the following graphics display in colour coded (red, green and blue) representation the parameter precisions. As a result these graphics isolate the patterns of the sub-groups that correspond to translations (tx, ty) (see Figure 6.18), rotations (ω,φ,κ) (see Figure 6.19), 3D target coordinates (X, Y, Z) (see Figure 6.20) as well as scale (s) and additional parameter term (k_3) jointly (see Figure 6.21). It is evident that the precisions of the 2D projective translations vary between 2.62μm and 4.89μm whereas the precisions of the 3D orientation angles range between 0.01degrees and 0.03 degrees. The precision patterns present a consistent variation in the data as these are illustrated for each estimated image (that is tx, ty and ω, φ, κ are illustrated per image).
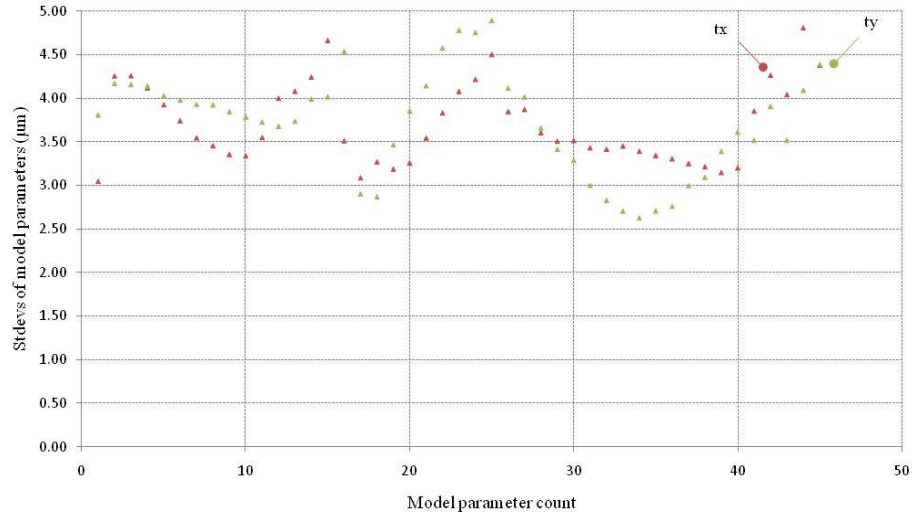
Figure 6.18: Precisions of model parameters (tx, ty). CPE - Iter.: 3.
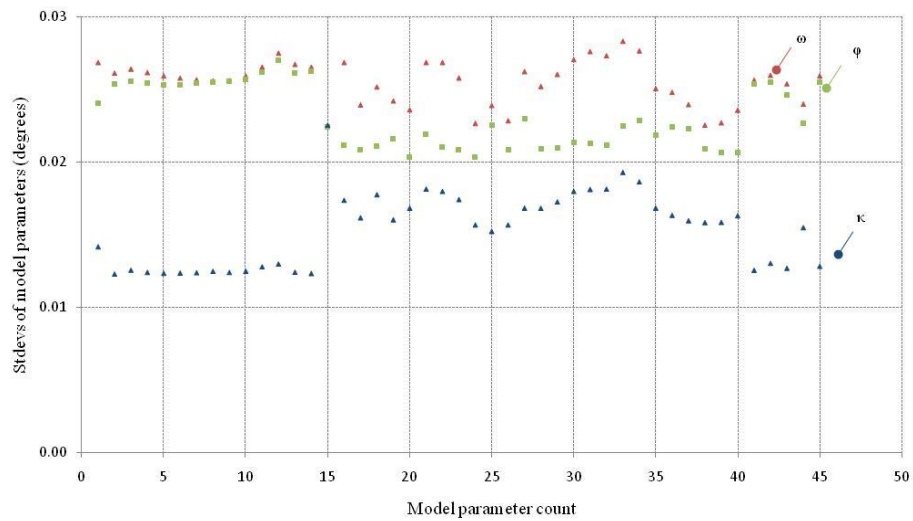
Figure 6.19: Precisions of model parameters (ω, φ, κ). CPE - Iter.: 3.

3D target coordinates are estimated with precisions that range between 3.33μm and 6.81μm noting again that these numbers correspond to the CPE solution (see Figure 6.20), whereas scale and k_3 term are estimated with associated precisions σ_s= 3.03 x 10^{-5} and σ_{k3}= 7.15x10^{-6} (see Figure 6.21).
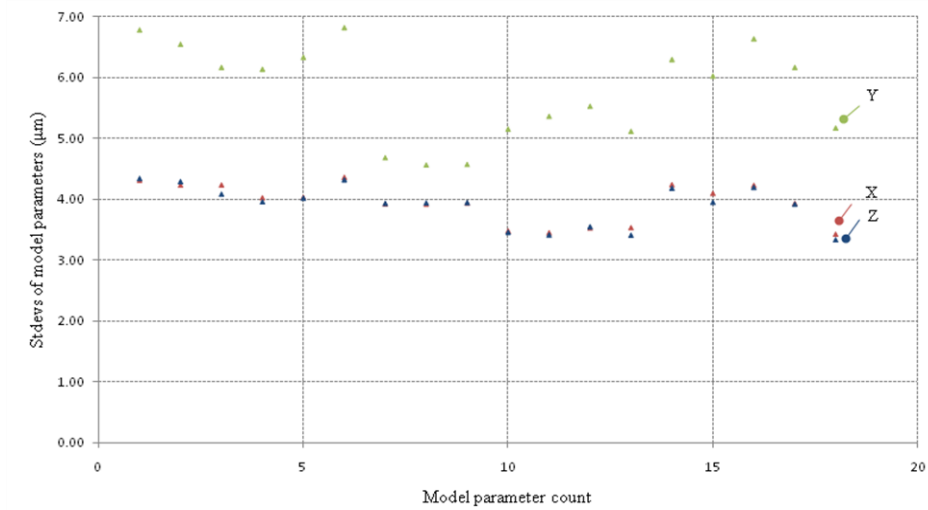
Figure 6.20: Precisions of model parameters (X,Y,Z). CPE - Iter.: 3.
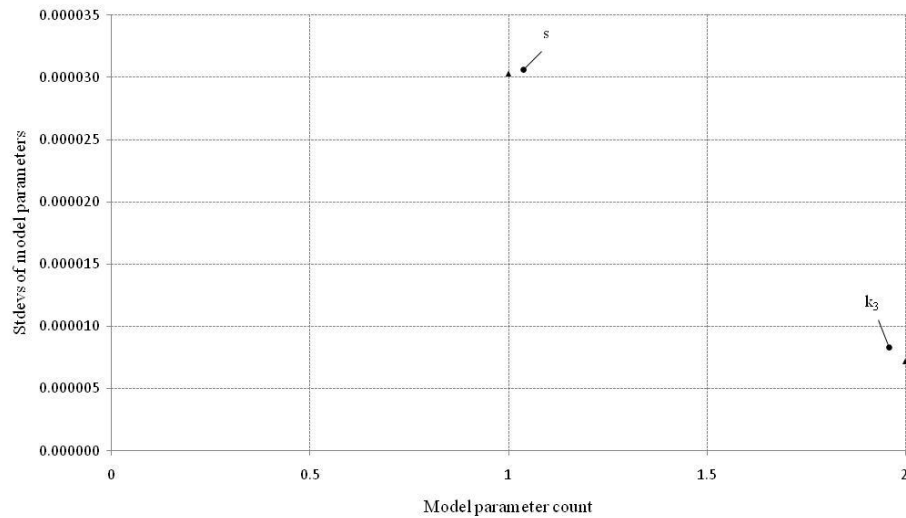
Figure 6.21: Precisions of model parameters (s, k_3). CPE - Iter.: 3.

By examining the corresponding figures for the inner constraints solution (CPI) (see Figure 6.22), it can be seen that the NSMP values that correspond to the first iteration pair 1-0 for the full estimated parameters agree in distribution and range with the previously reported CPE solution.
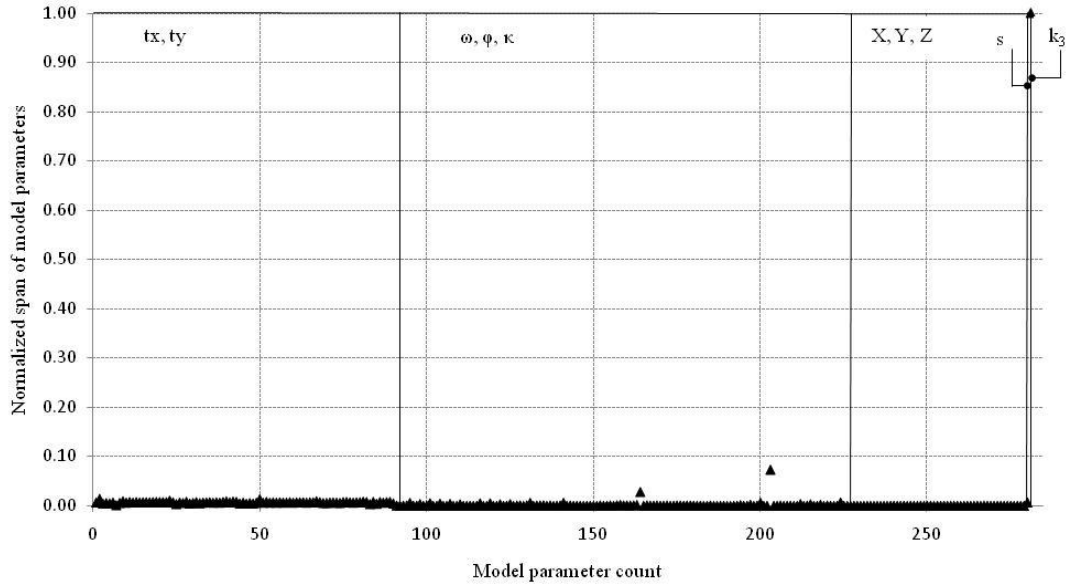
Figure 6.22: Normalized span of model parameters (full). CPI - Iter. 1-0.

To better analyze these data Figure 6.23 presents the corresponding patterns for the estimated parameters when excluding the radial lens distortion term k_3. It is evident that the NSMP values result in a similar distribution pattern in comparison to the external datum case (see Figure 6.14). Again the highlighted points NSMP(ω_{1025})=0.028, NSMP(ω_{1038})=0.073 and NSMP(s) = 0.007 present the most significant deviations.
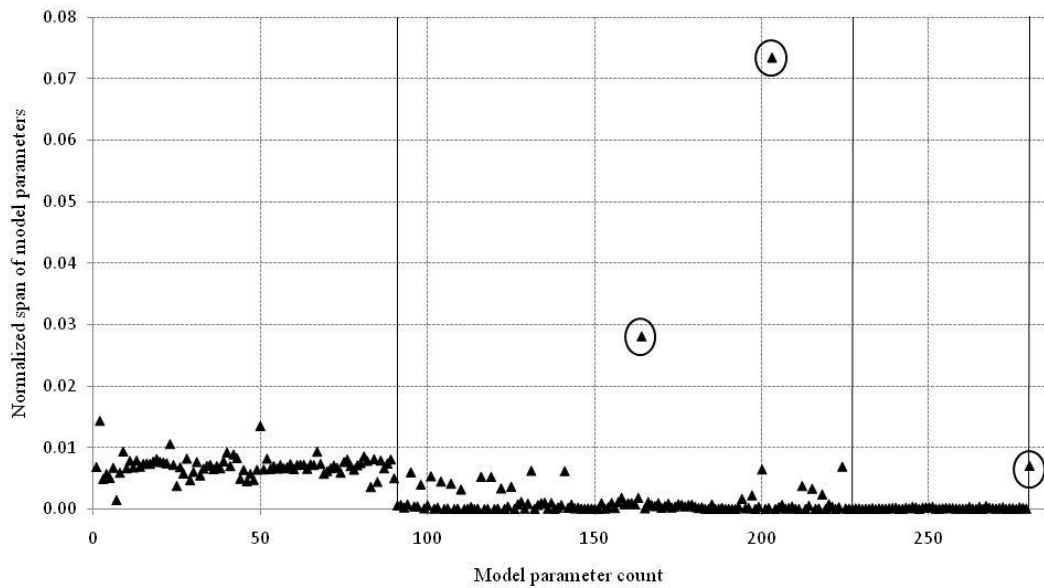
Figure 6.23: Normalized span of model parameters (-k3). CPI - Iter. 1-0.

For the iteration pair 2-1 (see Figure 6.24) these values are given as follows: $NSMP(\omega_{1025})=3.84\text{x}10^{-4}$, $NSMP(\omega_{1038})=5.40\text{x}10^{-4}$, $NSMP(s) = 2.85\text{x}10^{-7}$ and $NSMP(k_3)= 7.40\text{x}10^{-4}$.
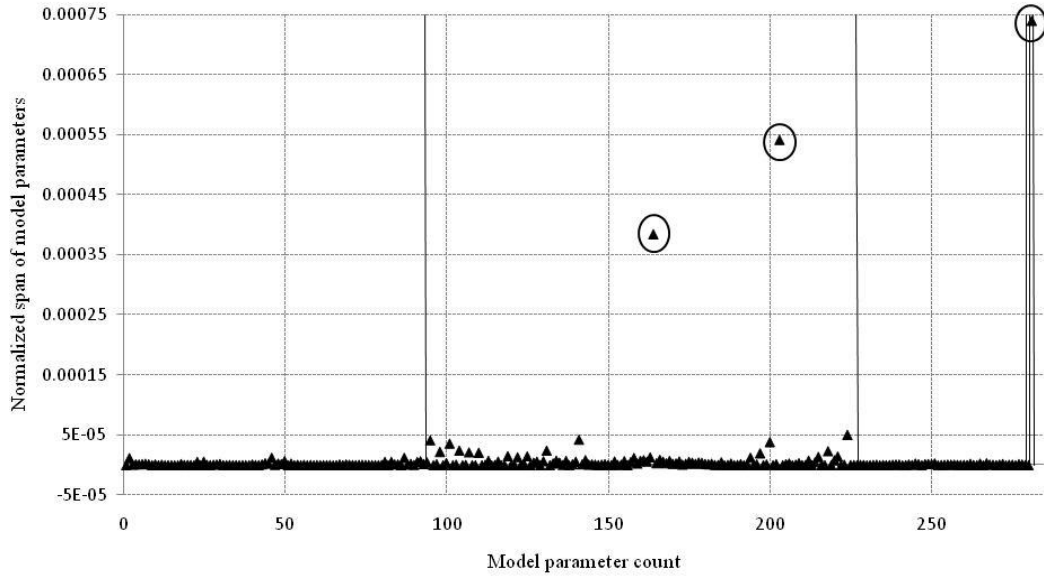
Figure 6.24: Normalized span of model parameters (full). CPI - Iter. 2-1.

Again it is noted that the two iteration pairs differ in two orders of magnitude while reaching the final convergence solution. The overall parameter precisions at the convergence stage of iteration 2 are now illustrated in Figure 6.25.
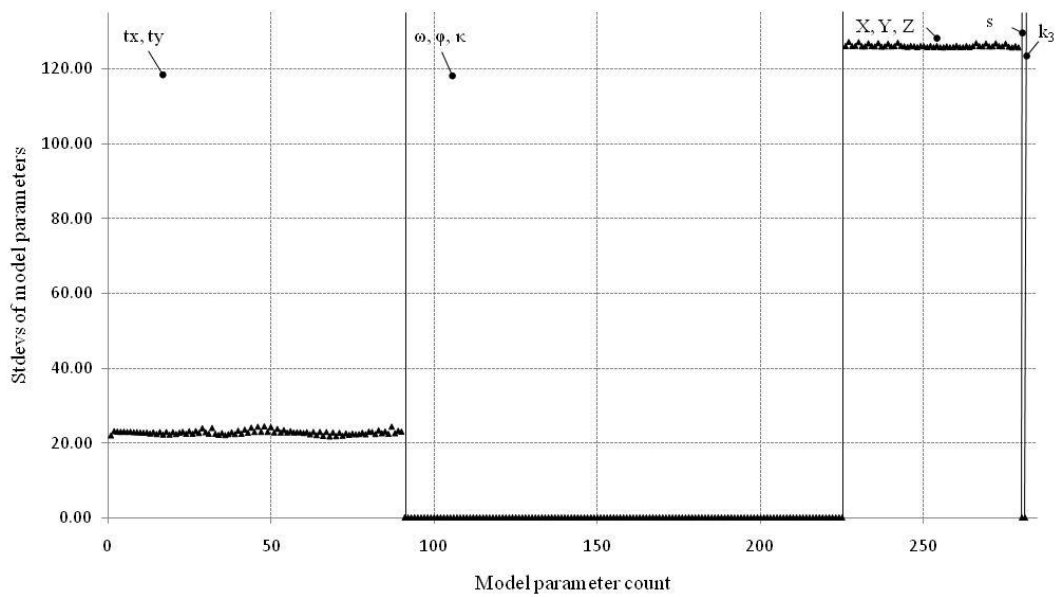
Figure 6.25: Precisions of model parameters. CPI - Iter. 2.

6. Results and analysis

To better highlight the parameter precisions four, additional sub-graphics were generated. The 2D projective translations precisions $\sigma_{tx,ty}$ range between 21.72µm and 24.36µm, 3D rotations precisions $\sigma_{\omega,\varphi,\kappa}$ range between 0.04degrees and 0.08degrees, 3D target coordinates precisions $\sigma_{X,Y,Z}$ range between 125.65µm and 127.08µm and finally σ_s for scale factor and σ_{k3} for k_3 term are given as 9.34×10^{-5} and 7.23×10^{-6} accordingly. Given the identical configuration in the data between both CPE and CPI bundle adjustment runs, it is evident that inner constraints reduce significantly the parameter precisions for the datum dependent parameters (tx, ty, ω, φ, κ, X, Y, Z, s) when these are compared with the external constraints solution. These results are illustrated in Figure 6.26, Figure 6.27, Figure 6.28 and Figure 6.29 in this specific order as follows.
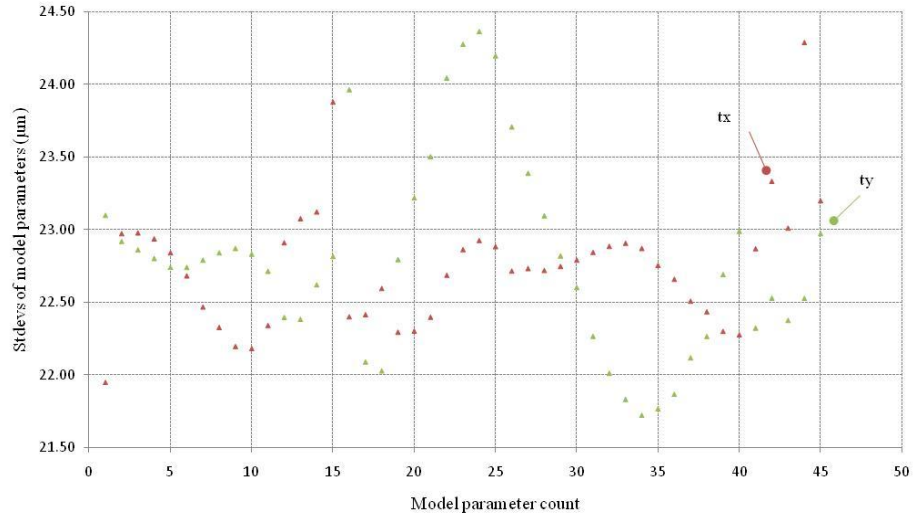
Figure 6.26: Precisions of model parameters (tx, ty). CPI - Iter. 2.
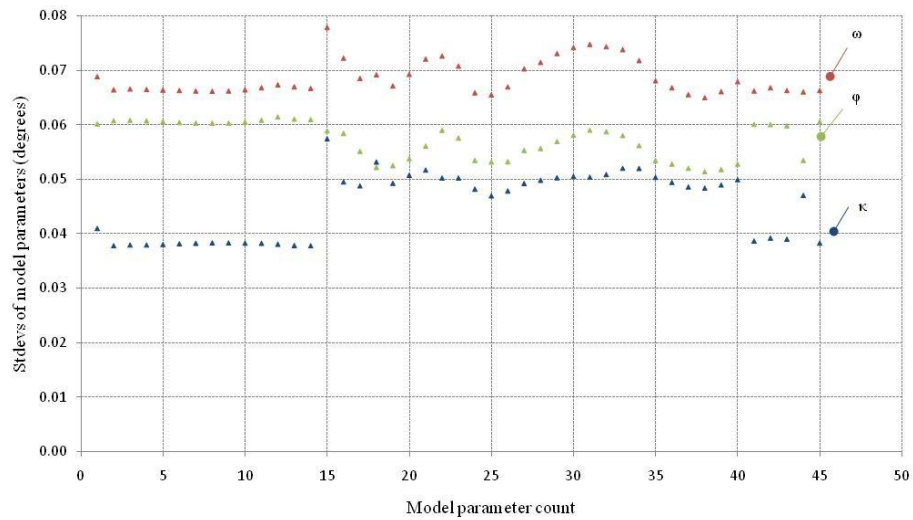
Figure 6.27: Precisions of model parameters (ω, φ, κ). CPI - Iter. 2.

- 181 -

Figure 6.28: Precisions of model parameters (X, Y, Z). CPI - Iter. 2.
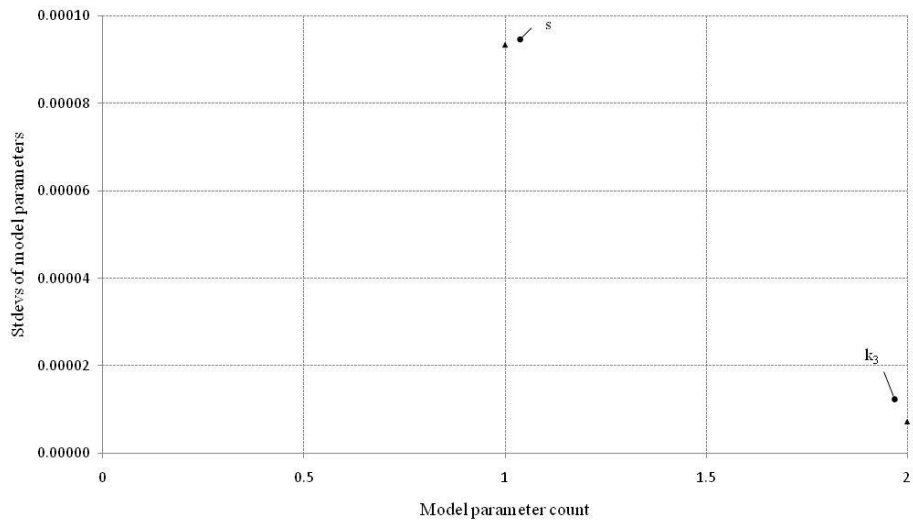
Figure 6.29: Precisions of model parameters (s, k3). CPI - Iter. 2.

It is recalled here that all precisions stated above reflect the quality of the method's behaviour scaled to the a posteriori standard deviation. It follows that the described precision patterns will be influenced by the sigma value which is at the order of 2.2 for both CPE and CPI solutions. To provide an additional description of convergence behaviour for the controlled solutions examined here, the mean of absolute differences (MAD) for both examined cases are derived. These figures describe the differences in convergence for the individual pairs of iterations as these were given above (see Table 6.10).

Solution	Iter.	MAD$_{tx,ty}$ (mm)	MAD$_{\omega,\varphi,\kappa}$ (degrees)	MAD$_{X,Y,Z}$ (mm)	MAD$_s$	MAD$_{k3}$
CPE	1	5.08×10^{-2}	2.08×10^{-2}	6.56×10^{-3}	1.14×10^{-3}	1.18×10^{-4}
CPE	2	6.69×10^{-6}	1.33×10^{-4}	4.73×10^{-5}	1.02×10^{-7}	6.53×10^{-8}
CPE	3	5.41×10^{-8}	2.24×10^{-7}	4.37×10^{-8}	1×10^{-9}	2.8×10^{-10}
CPI	1	4.98×10^{-2}	2.14×10^{-2}	6.36×10^{-3}	1.13×10^{-3}	1.17×10^{-4}
CPI	2	6.03×10^{-6}	1.46×10^{-4}	4.58×10^{-5}	4.6×10^{-8}	8.69×10^{-8}

Table 6.10: Convergence behaviour of model parameters. Table notation: MAD= mean of the absolute differences for each iteration in relation to its previous iteration.

It is evident that inner constraints degrade the quality of the model for the datum variant parameters (tx, ty, ω, φ, κ, X, Y, Z and s) but not for the datum invariant parameter (k_3) when these are compared with the corresponding precisions obtained from the external datum solution. The reduction in precision affects mostly the 3D target coordinates with a significant decrease in precision by a minimum to maximum factor of 19 and 38 whereas the reduction factor in precision for the 2D projective translations ranges between 5 and 8. These are less significant for the 3D rotations varying between 3 and 4 whereas the precision for the scale factor is reduced by approximately 3. These results are displayed in Table 6.11.

Solution	tx, ty (μm)		ω,φ,κ (degrees)		XYZ (μm)		s	k_3
	σ_{max}	σ_{min}	σ_{max}	σ_{min}	σ_{max}	σ_{min}	$\sigma \times 10^{-5}$	$\sigma \times 10^{-6}$
CPE #3	4.89	2.62	0.03	0.01	6.81	3.33	3.03	7.15
CPI #2	24.36	21.72	0.08	0.04	127.08	125.65	9.34	7.23

Table 6.11: Estimated precisions of model parameters.

6.4.2.3. Correlations consistency

Besides model assessment (see section 6.4.2.1.) and evaluation of convergence (see section 6.4.2.2.), a useful check is to inspect the correlation coefficient output comparing the different bundle adjustment runs. Calculation of the correlation coefficient has been given in equation (4.31) (see section 4.7.1.). Here the correlation coefficient measure is extracted from the scaled to the a posteriori precision covariance matrix. For interpretation purposes Figure 6.30 provides the formulation of the correlation data. This is done for a synthetic set of 2 views and 3 points within the image network where parameter correlations are blocked in red frames as it is shown below.
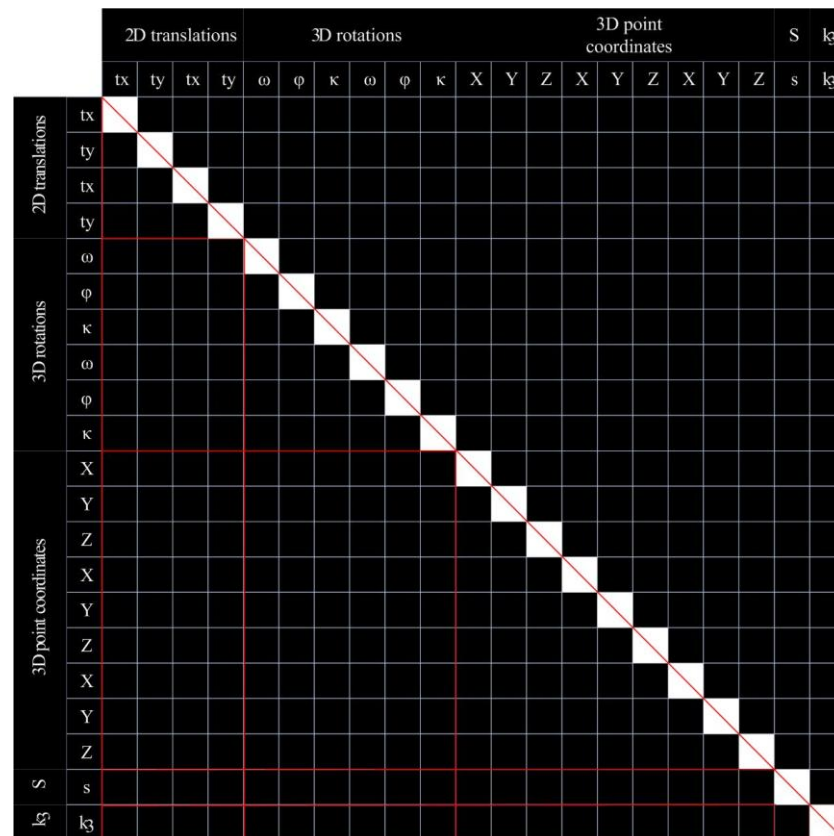
Figure 6.30: Correlation coefficient matrix structure.

For comparative evaluation the numerical data are visually represented by mapping of the scalar values within the range 0 (no correlation) to ±1 (complete correlation) to the grayscale range 0 (black) to 1 (white). Figure 6.31 visualizes this intensity ramp highlighting the lower triangular part in red framed sub-blocks for better interpretation of the patterns. The illustrated patterns present the test cases that correspond to the solutions outlined in Table 6.9 (see section 6.4.2.1).
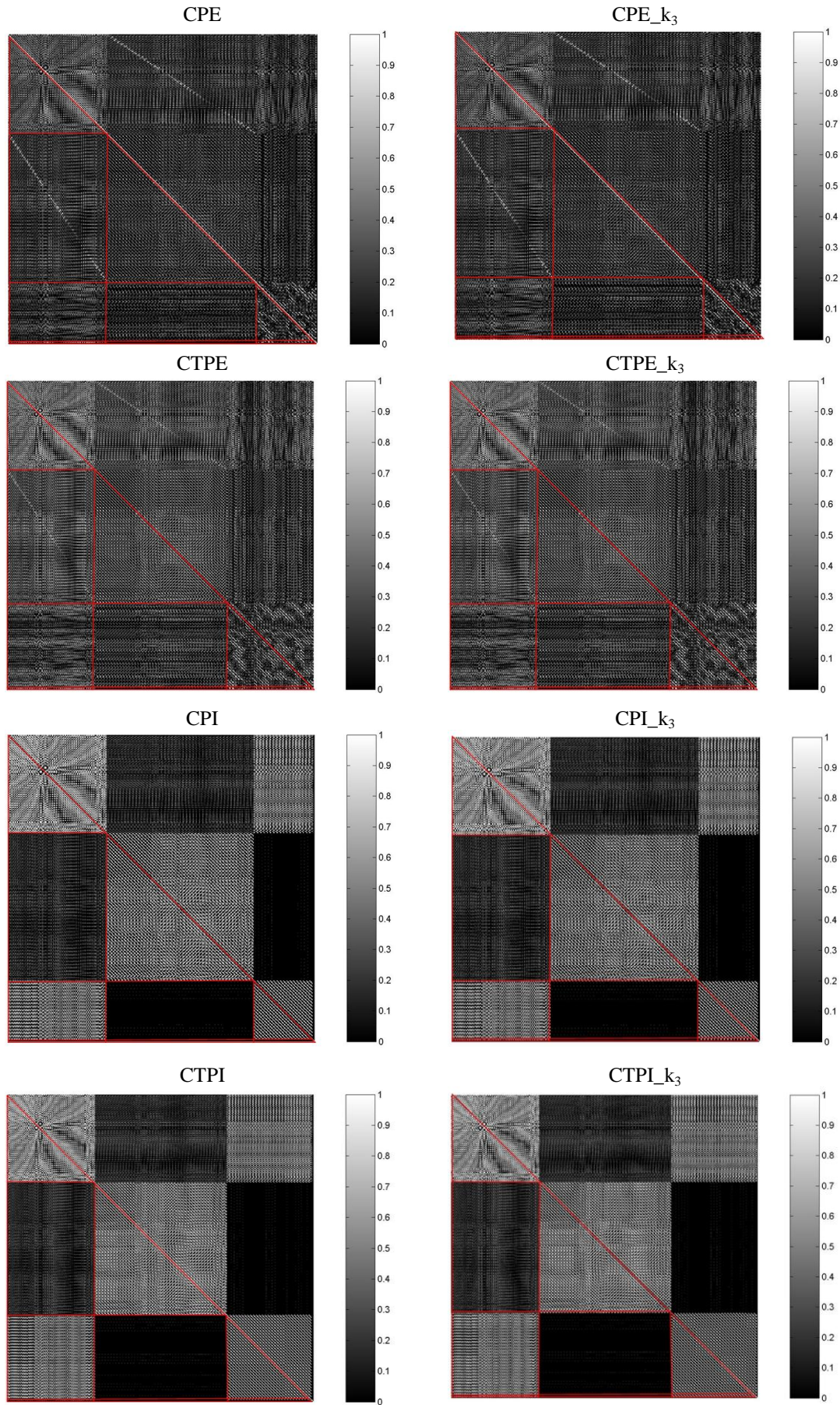
Figure 6.31: Correlation coefficient matrices for model parameters.

Both external and inner datum solutions present high correlations between parameters of the same kind; that is 2D projective translations (tx, ty), 3D rotations (ω, φ, κ), 3D target locations (X, Y, Z), scale factor (s) and radial lens distortion term (k_3). External constraints, in particular, present significant cross correlations between parameters (tx, ty) and (ω, φ, κ), (tx, ty) and (X, Y, Z) as well as (ω, φ, κ) and (X, Y, Z), whereas inner constraints present minimum correlations between (ω, φ, κ) and (tx, ty) as well as (ω, φ, κ) and (X, Y, Z).

In fact, the external constraints solutions demonstrate distinctive high correlation between (tx, ty) and (ω, φ, κ). This effect is observed along the diagonal of this sub-block. To isolate this pattern Figure 6.32 illustrates the correlation matrix for this particular block extracting the values that are considered to present high correlations; that is the locations where ρ>0.7. In addition, to interpret these patterns it is useful to visualize the image network geometry highlighting those photo locations that present high correlations (see Figure 6.32 right). Moreover, Figure 6.33 provides the spatial location of (tx, ty) with (ω, φ, κ). Specifically, the 2D projective translations are illustrated as triangles where the correlation coefficients are plotted as vertical linear displacements colour coded per rotation (red: ω, green: φ and blue: κ) and drawn in a left (tx) to right (ty) direction centred from a black index line.
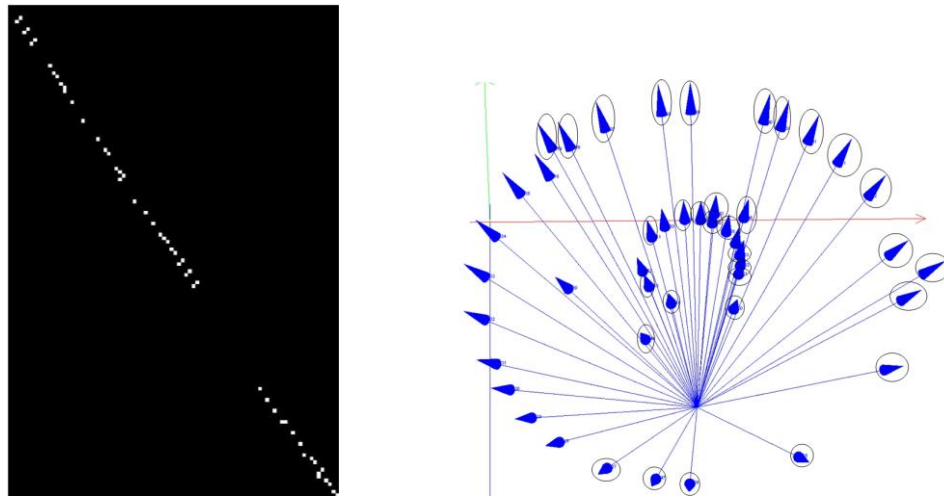
Figure 6.32: Correlation coefficient matrix between tx, ty and ω, φ, κ (left) and image network geometry (right). CPE (ρ>0.7).
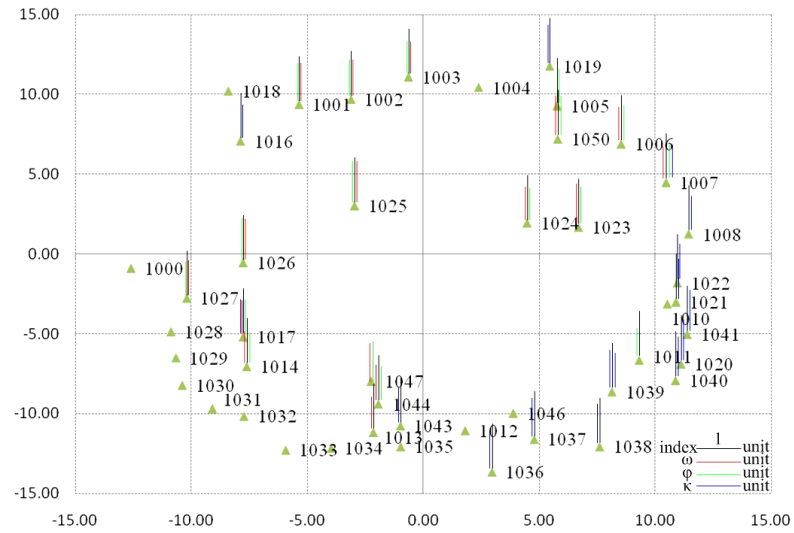
Figure 6.33: Spatial location of tx, ty with correlations between tx, ty and ω, φ, κ.

By observing these figures and provided that the photos contain sufficient CP data within the image format, the external constraints datum (defined from the CPs) results in a high correlation between the 2D projective translations and the 3D orientation angles.
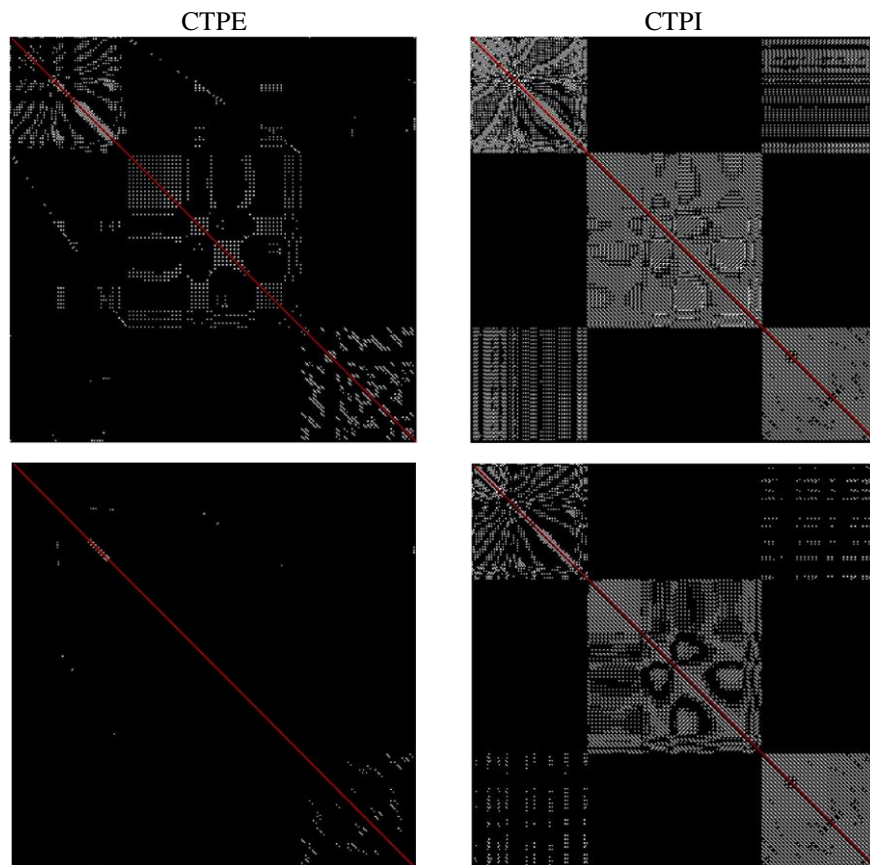
Figure 6.34: Correlation coefficient matrices for model parameters - ρ> 0.75 (top) and ρ> 0.9 (bottom).

To examine further the correlations of the model parameters, the correlation data of CTPE and CTPI solutions were filtered to derive those parameters that present a correlation coefficient of $\rho > 0.75$ and subsequently $\rho > 0.90$ (see Figure 6.34). These were visualized as 1 where a correlation exists or 0 otherwise. External constraints present high correlations for similar types of parameters with no significant cross correlations between different kinds of parameter pairs. Again some distinctive correlations are present between 2D projective translations and 3D rotations as well as some random correlations between 2D projective translations and 3D target coordinates whereas the correlation coefficients between 3D rotations and 3D target coordinates are minor. Inner constraints present high correlations between 2D projective translations and 3D target coordinates showing the clear influence of the identified datum on the quality of the 3D target data (see 6.4.2.1.). Moreover the uniform scale factor of the inner constraints method causes significant correlations in 3D space recovery, noting the high presence of correlations between 3D targets as opposed to the external datum case.

6.5. Object space assessment

To assess object space and 3D point estimation three datasets coded as B2 (pyramid), D2 (lego) and E2 (centroid) (see section 6.2.3.) were utilized. This check is evaluated with two performance measures. The first checks the object space proximities by visualization of the correlation coefficients over the 3D target distances per each X, Y and Z directions and the second plots the error ellipsoids of the estimated 3D target coordinates. The processed datasets vary in three aspects and these are target image quality, object geometry as well as distribution of control and photo orientations within the image network. These factors affect the quality evaluators utilized here. This section is covered with the reference measurements results (see section 6.5.1.), affine bundle adjustment results (see section 6.5.2.), correlations with proximities (see section 6.5.3.) and error ellipsoids (see section 6.5.4.).

6.5.1. Reference measurements

Initialization of the three test datasets was implemented based on a prior independent self-calibrating bundle adjustment (perspective, software tool VMS 8.0) similarly to

the processing method reported above (see section 6.4.1.). The overall results of the initialized bundle adjustment solutions are illustrated in Table 6.12.

Dataset	Iter.	R.	σ_o	RMS_{xy} (μm)	σ_{XYZ} (μm)	Relative precision	RMS scale (μm)
				IS		OS	
B1	5	1,899	1.00	0.26	8.95	1:6,000	125.56
D1	7	2,859	1.00	0.63	8.60	1:27,000	286.51
E1	7	4,227	1.00	0.97	6.28	1:41,000	-

Table 6.12: BA statistical indicators - datasets B1, D1 and E1.

6.5.1.1. Dataset B1 - pyramid

Dataset B1 was measured with a bundle adjustment processed with an inner constraints datum, fifth and seventh terms of the radial lens distortion polynomial regarding calibration and 12 scales (measured with a digital calliper) to provide an accurate object space scale within the network. The solution converged rapidly with a redundancy of 1,899 after five iterations with an RMS image residual of $1/10^{th}$ of a pixel for the Sony sensor (with a pixel size of 4.78μm at a resolution of 1,024x768 pixels). In object space the solution results in a relative precision for the image network of 1:6,000, 3D target coordinates precision of 8.95μm and an RMS object scale equal to 125.56μm.
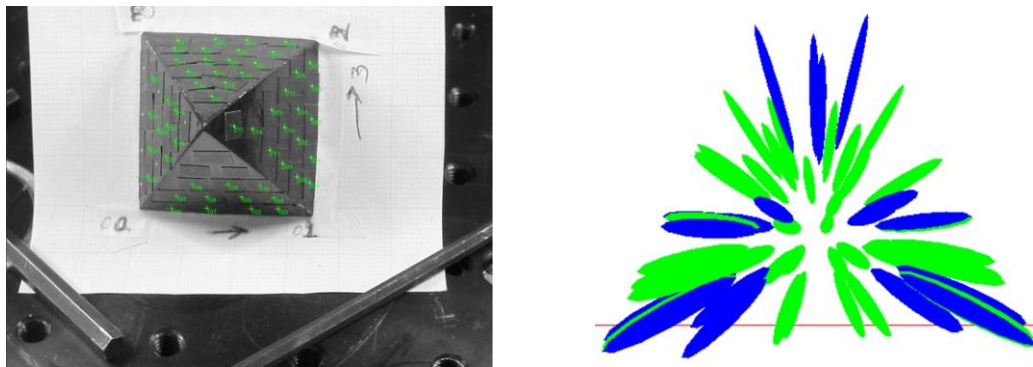
Figure 6.35: 3D target points (top view - left) and error ellipsoids (front view - right) - dataset B1.

Figure 6.35 gives an example view of the 3D target points located on the pyramid object together with their error ellipsoids (blue ellipsoids= CPs and green ellipsoids= TPs visualized in the software VMS 8.0). This illustration additionally describes the point arrangement; that is CPs are distributed at the edges of each face leaving all

other target points to be treated as TPs. The error ellipsoids for both CPs and TPs are not significantly different in magnitude which is a natural result of inner constraints. The 3D points were coordinated from an average number of 17.5 viewpoints. It is recalled that the large semi axes of the error ellipsoids point towards the centroid of the objects whereas the vertical axis indicates the error in the viewing direction. The fact that the major semi axes points towards the datum shows a weakness in scale recovery for the pyramid. The ellipsoids present a highly elongated shape noting the configuration geometry with the point location (the photos were acquired from a top angle in relation to the object's four planar facets).

6.5.1.2. Dataset D1 - lego

Dataset D1 was measured with a bundle adjustment processed with an inner constraints datum, full additional parameters within the calibration model and 7 measured scales in the object space. The bundle converged after seven iterations with an increased redundancy of 2,859 with an RMS image residual of $1/15^{th}$ of a pixel for the Kodak sensor (with a pixel size of 9.0μm at a full resolution of 1,008 x 1,018 pixels). In object space the relative precision of the image network resulted in 1:27,000, the 3D target coordinates precision was estimated to 8.60μm and the object space scale was 286.51μm.
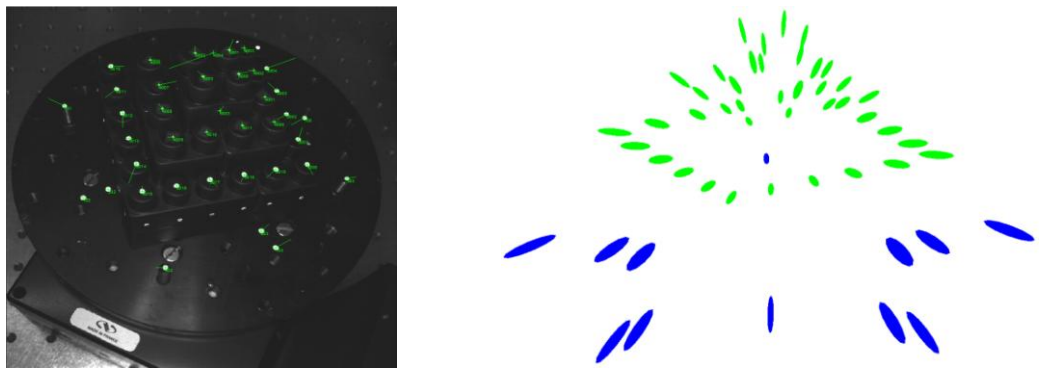
Figure 6.36: 3D target points (top view - left) and error ellipsoids (front view - right) - dataset D1.

The 3D points are estimated from a mean number of 31.5 viewpoints which is increased significantly related to the previous image network. The illustrated error ellipsoids (see Figure 6.36) show that the quality of the estimated 3D target coordinates for the TPs (green ellipsoids) are of uniform precision for each of three

planar faces. Similarly, the CPs (blue ellipsoids) that were measured as 3D targets located on a turntable present a uniform precision which is attributed to the inner datum. It is noted that points that are located further than the datum origin present reduced target quality (error ellipsoids are larger). Similarly to the pyramid object (see section 6.5.1.1.) the major axes of the ellipsoids point towards the datum. This confirms the weakness in scale recovery within the inner constraints datum.

6.5.1.3. Dataset E1 - centroid

Dataset E1 was similarly processed with a bundle adjustment run with an inner constraints datum, two radial lens distortion terms within the calibration model and without the inclusion of any object space scales. In this case the solution that was run with a redundancy of 4,227 converged successfully at the seventh iteration with an RMS image residual at the order of $1/10^{th}$ of a pixel for the monochrome Kodak sensor. Object geometry was recovered with a relative precision for the image network of 1:41,000 and a mean 3D target coordinates precision of 6.28μm.
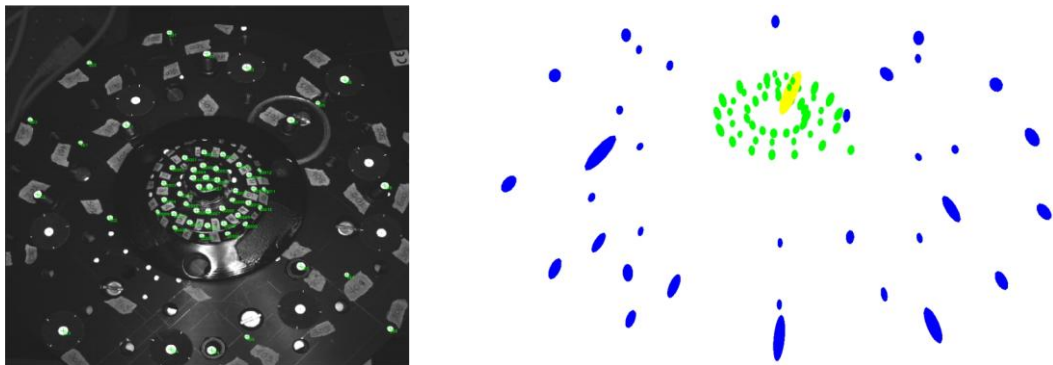
Figure 6.37: 3D target points (top view - left) and error ellipsoids (front view - right) - dataset E1.

The 3D point error ellipsoids confirm the geometric strength of the image network noting that the estimated targets are coordinated from a mean number of 26.7 image rays. Specifically (see Figure 6.37) it is evident that the error ellipsoids vary in shape for the CPs (blue ellipsoids) according to target image quality. This is observed from the difference in magnitude between the large and small diameter CPs. In the case of the estimated TPs the ellipsoids (drawn in green) have a uniform shape and are of a smaller magnitude in relation to the control (these are closer to the origin of the centroid). In fact it is evident that one target which was coordinated from four

intersected viewpoints results in a relatively weak 3D precision (σX=12.7μm, σY=22.3μm, σZ=18.1μm) and therefore it is coloured in yellow from the software.

6.5.2. Affine bundle adjustment results

Datasets B2, D2 and E2 were acquired with the affine sensor and processed with the developed affine bundle adjustment. In this processing test, the affine bundle was run with the external constraints datum and the third order radial lens distortion term (calibration model). The stochastic model was initialized with an a priori image observation quality of 0.5μm for datasets B2 and D2. Dataset E2 however presents very large target diameters in image space (target image diameters are equal to 36 pixels), therefore this dataset was initialized with an a priori image observation quality of 1.5μm. The 3D target precisions were set to 25μm for the control point data and to the associated a posteriori precisions obtained from the pre-measurements of the tie point data (see section 6.5.1.). It is noted that the signalized point data describe the objects geometry only in sparse terms.

Dataset	Iter.	R.	Scale	k_3 ($\sigma_{\kappa 3}$)	σ_o	RMS$_{xy}$ (μm)	σX, σY, σZ (μm)
B2	3	608	0.1656 (1.0632)	4.4414×10^{-4} (2.5059×10^{-5})	1.62	0.53	7.69, 9.93, 7.28
D2	2	430	0.1614 (1.2899)	4.0454×10^{-5} (8.6230×10^{-6})	2.57	0.75	27.88, 31.52, 28.03
E2	3	1,287	0.1651 (2.3232)	4.7493×10^{-4} (3.7241×10^{-5})	3.37	3.45	36.54, 43.81, 36.86

Table 6.13: Affine BA results - datasets B2, D2 and E2.

6.5.2.1. Dataset B2 - pyramid

This dataset is composed of 23 photos and 44 target points of which 20 are treated as control and 22 as tie. To analyze the bundle adjustment results, it is important to consider the quality of the data. It is noted that the target point data occupy a diameter of 17 pixels in image space (0.5mm diameter white markers in object space) (see Appendix C for a sample of target image quality). A primary requirement is that for photo orientation initialization the control had to be distributed in 3D. In addition, the minimum visibility requirement for optimal target measurement is four rays (each

target needs to be coordinated from four viewpoints). The image network geometry has already been illustrated in Figure 6.4 - network B2 (see section 6.2.2.4.). Figure 6.38 provides an example affine view from this dataset jointly with the visibility frequencies for the control and tie point data within the image network.
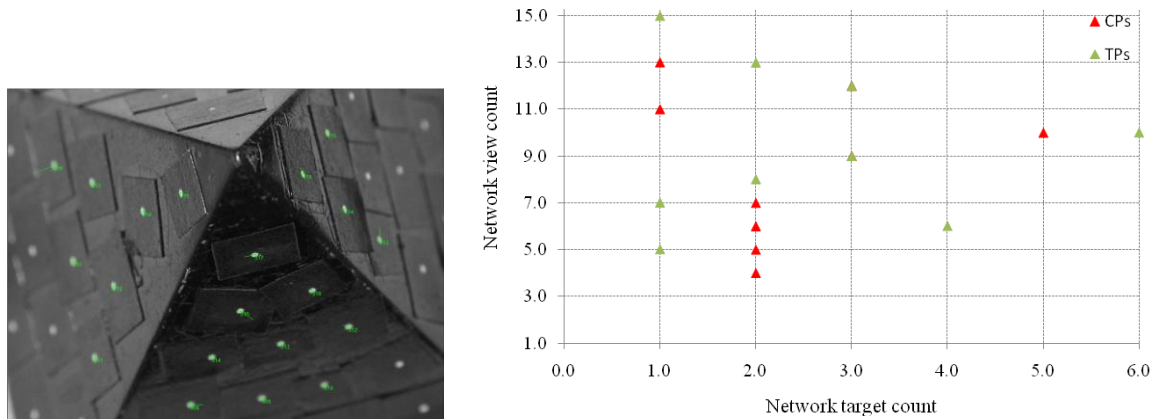
Figure 6.38: Affine (left) and CPs and TPs data visibility (right) - dataset B2.

Regarding control, two CPs were observed from four viewpoints with the best case being one target observed from 13 viewpoints. Moreover minimum visibility is reported the case where one tie point was calculated from the intersection of 5 rays whereas maximum visibility is the case where one tie point is coordinated from the intersection of 15 rays. Besides characterizing the data according to their visibility it is important to note that configuring the data for simultaneous initialization (e.g. four CPs forming a volume) and measurement (e.g. tie point overlap) was difficult. This was the case particularly considering the object's planar facets, the object's small 3D volume in combination with the physical limitations of affine imaging sensor. In fact whilst image network geometry is highly convergent, it is relatively sparse again provided that an increased number of views would be required for increased frame to frame overlap and full object coverage in the ideal case. Finally the location of the object's surface in relation to the photo locations results in an acute angle between the observation lines and each of the four planes of the pyramid. As a consequence the image measurement quality of the targets is reduced (targets imaged as elliptical blobs). This is particularly the case for those targets that do not reside within the identified (±19.7mm) field of view. The bundle adjustment was processed with a redundancy of 608 observations and converged rapidly after 3 iterations with a sigma

nought equal to 1.62 (see Table 6.13). The achieved quality in image space is equal to 0.53μm that is $1/9^{th}$ of a pixel for the Sony sensor, whereas the 3D targets are coordinated with a mean 3D precision of σ_{XYZ}=8.3μm (σX= 7.69μm, σY=9.93μm, σZ= 7.28μm). The calibrated image scale is equal to 0.1656 (with precision σ_s=0.1264). The bundle results in a radial distortion of 6.94μm at a radial distance of 2.5mm from the image centre. The calculated k_3 term is equal to 4.44×10^{-4} (with a precision of σ_{k3}= 2.5095×10^{-5} and a significance of -0.0069) (see Figure 6.39).

```
#################################################
Lens Distortion Profile for Camera Calibration
Units: microns
Radius Value
0.0000 0.0000
0.5000 0.0555
1.0000 0.4441
1.5000 1.4990
2.0000 3.5532
2.5000 6.9398
<< UPDATED ADDITIONAL PARAMETERS DATA (k3) >>
Units: mm
Value Precision Correction Significance
4.441441509e-004 2.5059e-005 -1.7265e-007 -0.0069
#################################################
```

Figure 6.39: Calibration output - dataset B2.

6.5.2.2. Dataset D2 - lego

This dataset is composed of 24 photos, 20 targets of which 10 are treated as control and 10 as tie. Figure 6.40 illustrates an example affine view together with the data visibility that characterizes this image network. The point data occupy 18 pixels in image space (1mm diameter white markers in object space). Control data present minimum visibility in the case where two targets are measured from eight viewpoints whereas regarding tie point data only one target is coordinated from 4 viewpoints. The object geometry points that all target data are sparsely arranged within each frame and they are located between two (top-bottom) square planes at regular separations.
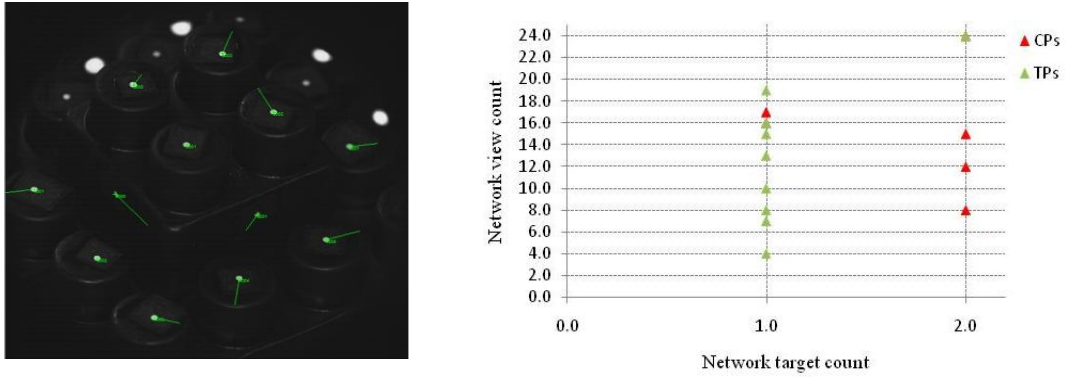
Figure 6.40: Affine view (left) and CPs and TPs data visibility (right) - dataset D2.

The bundle adjustment was processed with a redundancy of 430 observations and converged rapidly after 2 iterations with a sigma nought equal to 2.57 (see Table 6.13). In image space the triangulation misclosure is 0.75µm which is equal to 1/12th of a pixel for the Kodak sensor. In object space the 3D targets are coordinated with a mean 3D precision of σ_{XYZ}=29.14µm (σ_X=27.88µm, σ_Y=31.52µm, σ_Z=28.03µm). The calibrated image scale is equal to 0.1614 (with precision σ_s=1.2899 x10^{-4}) whereas the calculated k_3 term is equal to 4.0454x10^{-5} (with precision σ_{k3}= 8.6230x10^{-6} and a significance of 0.0066) (see Figure 6.41).

```
##############################################
Lens Distortion Profile for Camera Calibration
Units: microns
Radius Value
0.0000 0.0000
0.5000 0.0051
1.0000 0.0405
1.5000 0.1365
2.0000 0.3236
2.5000 0.6321
3.0000 1.0923
3.5000 1.7345
4.0000 2.5891
4.5000 3.6864
<< UPDATED ADDITIONAL PARAMETERS DATA (k3) >>
Units: mm
Value Precision Correction Significance
4.045420537e-005 8.6230e-006 5.6812e-008 0.0066
##############################################
```

Figure 6.41: Calibration output - dataset D2.

6.5.2.3. Dataset E2 - centroid

This dataset is composed of 17 photo and 54 point data (12 CP and 42 TP). Target image quality is pointed by the 36 pixels in diameter blobs in image space (2mm diameter retro-reflective targets in object space). To characterize this dataset it is firstly pointed out that whilst this network was originally designed in a wide separated three ring arrangement, the selected photos cover the 3D object space in a cone

arrangement and they are geometrically wide and convergent. This is a highly advantageous point of this image network particularly when this is compared to the image networks of datasets B2 and D2 (see Figure 6.4 in section 6.2.2.4.). Contrary to this, the centroid consists of concentric rings of point data (retro-reflective targets) with significantly large target diameter (2mm in object space which is equal to 36 pixels in image space) for this particular imaging range (r= 175mm) and sensor (CAM_A3). As a result, the data present a reduced image quality; therefore this image network was processed with an a priori image quality of 1.5µm.
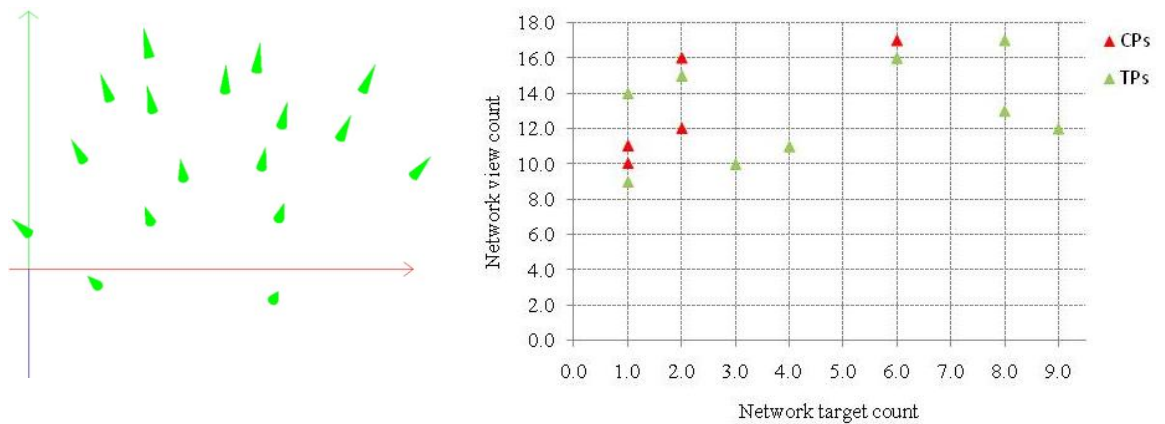
Figure 6.42: Affine image network (left) and CPs, TPs data visibility (right) - dataset E2.

Figure 6.42 provides an illustration of the image network geometry and the data visibility. Both control and tie point data are highly visible and they range between 10 and 17 views (CPs) and 9 and 17 views (TPs) accordingly. The bundle adjustment was processed with a number of 1,287 redundancies and converged rapidly after 3 iterations with an a posteriori sigma nought of 3.73 (see Table 6.13). The triangulation misclosure is 3.45µm which is approximately equal to $1/3^{rd}$ of a pixel for the Kodak sensor in image space. In object space the 3D targets are coordinated with a mean precision of σ_{XYZ}=39.07µm (σ_X= 36.54µm, σ_Y= 43.81µm, σ_Z= 36.86µm). The calibrated image scale is equal to 0.1651 (with precision σ_s=2.3232x10^{-4}) whereas the calculated k_3 term is equal to 4.7493x10^{-4} (with precision σ_{k3}=3.7241x10^{-5} and a significance of 0.0101) (see Figure 6.43). It is noted that for a complete display of the bundle adjustment results that were obtained from three datasets B2, D2 and E2 the associated radial lens distortions profiles are provided in Appendix C.

```
################################################
Lens Distortion Profile for Camera Calibration
Units: microns
Radius Value
0.0000 0.0000
0.5000 0.0594
1.0000 0.4749
1.5000 1.6029
2.0000 3.7994
2.5000 7.4208
3.0000 12.8231
3.5000 20.3626
4.0000 30.3954
4.5000 43.2779
<< UPDATED ADDITIONAL PARAMETERS DATA (k1) >>
Units: mm
Value Precision Correction Significance
4.749284375e-004 3.7241e-005 3.7683e-007 0.0101
################################################
```

Figure 6.43: Calibration output - dataset E2.

6.5.3. Correlations with proximities

For object space evaluation, the first measure that was utilized checks the relation of the 3D target correlation coefficients against object space proximity. It is noted here that the term proximity denotes the Euclidean inter-target distance for each possible target combination and which is calculated as follows.

$$D_{(n+1,n)} = \sqrt{(X_{n+1} - X_n)^2 + (Y_{n+1} - Y_n)^2 + (Z_{n+1} - Z_n)^2}$$

(6.2.)

Particularly, this is done by visualization of the target Euclidean distance (x axis) over the 3D target correlation coefficients (y axis) derived per X, Y and Z direction (RX, RY, RZ) from the a posteriori covariance matrix. It is noted that the correlation coefficient vectors RX, RY and RZ were extracted from the upper triangle of the associated correlation coefficient matrix (absolute values) which for an example for a 5x5 array is given as follows.
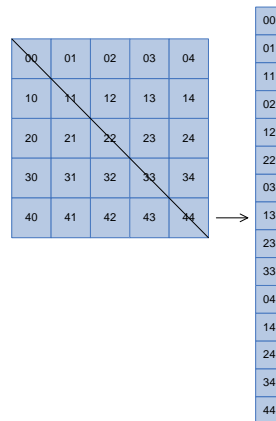
Figure 6.44: Structure of correlation coefficient array.

6.5.3.1. Dataset B2 - pyramid

Evaluation of the 3D correlations in object space for dataset B2 (pyramid object) is implemented with visualization of the absolute values of the 3D correlation coefficients in each direction RX, RY and RZ against the corresponding 3D inter-target separations (see Figure 6.45). As a result, the generated graphs show the behaviour for each inter-direction correlation separately.
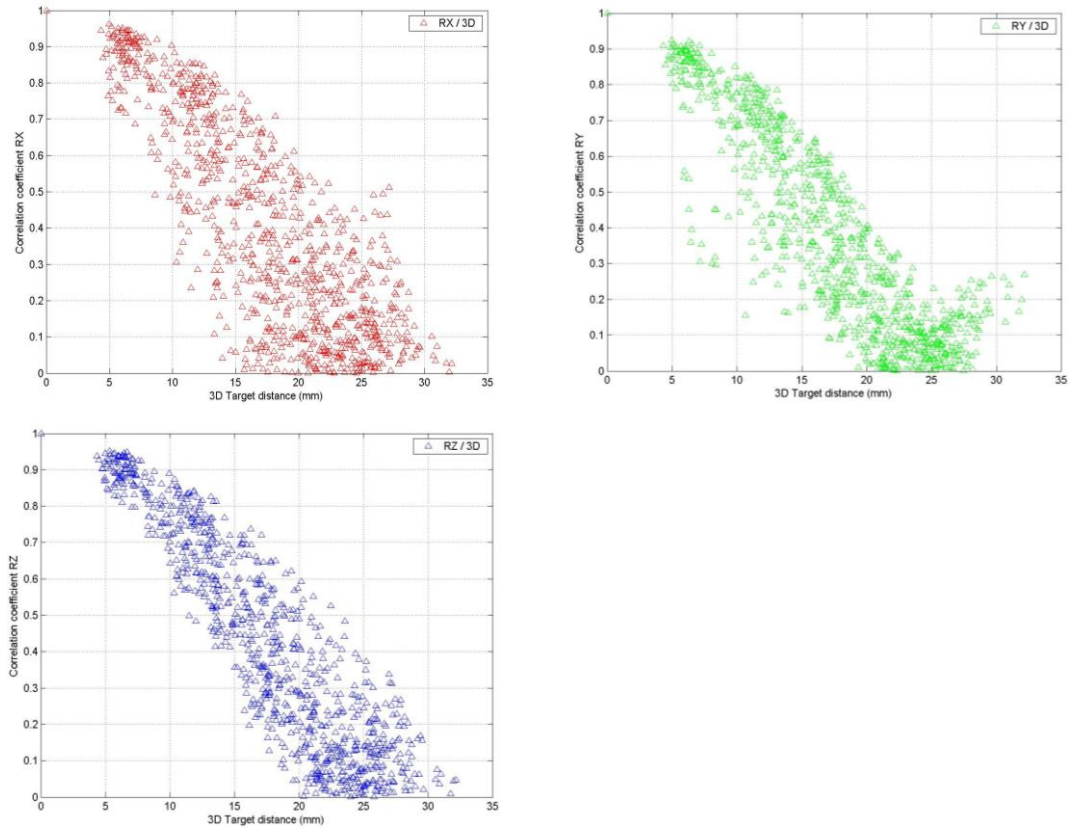
Figure 6.45: Correlation coefficients with proximities (red: RX, green: RY, blue RZ) - dataset B2.

The target-pairs differences range between 4.3mm (minimum separation) and 32.1mm (maximum separation). Correlations RX, RY and RZ increase inversely with target-distance. In fact targets that are separated with distance D= 4mm-10mm are highly correlated in all three directions X, Y, Z (ρ= ~0.7-1.0) whereas only a few targets present low correlations in Y (ρ= ~0.3-0.7) and X (ρ= ~0.5-0.7). In general Z (object depth) direction presents a smooth behaviour when compared with correlations in X and Y that present a wider spread between the ranges 10mm-32mm. An interesting point is that as opposed to the general trend (that is correlations decrease with an increase in target-separation) there exists a cluster of targets separated between D=

~27mm-32mm that tend to have increased correlations. However the magnitude in correlations is not exceeding 0.3 therefore these can be attributed to potential poor uncertainty in precision (e.g. low image quality, viewpoint intersection, low measurement redundancy) and not necessarily a problem in the data.

6.5.3.2. Dataset D2 - lego

Figure 6.46 illustrates the 3D correlation coefficients in each direction RX, RY and RZ over the corresponding 3D inter-target separations. In the case of dataset D2 the target-pairs differences range between 9.98µm (minimum separation) and 61.97µm (maximum separation). It is evident that similarly to the previous case (see section 6.5.3.1.) correlations RX, RY and RZ increase inversely with inter target-distance.
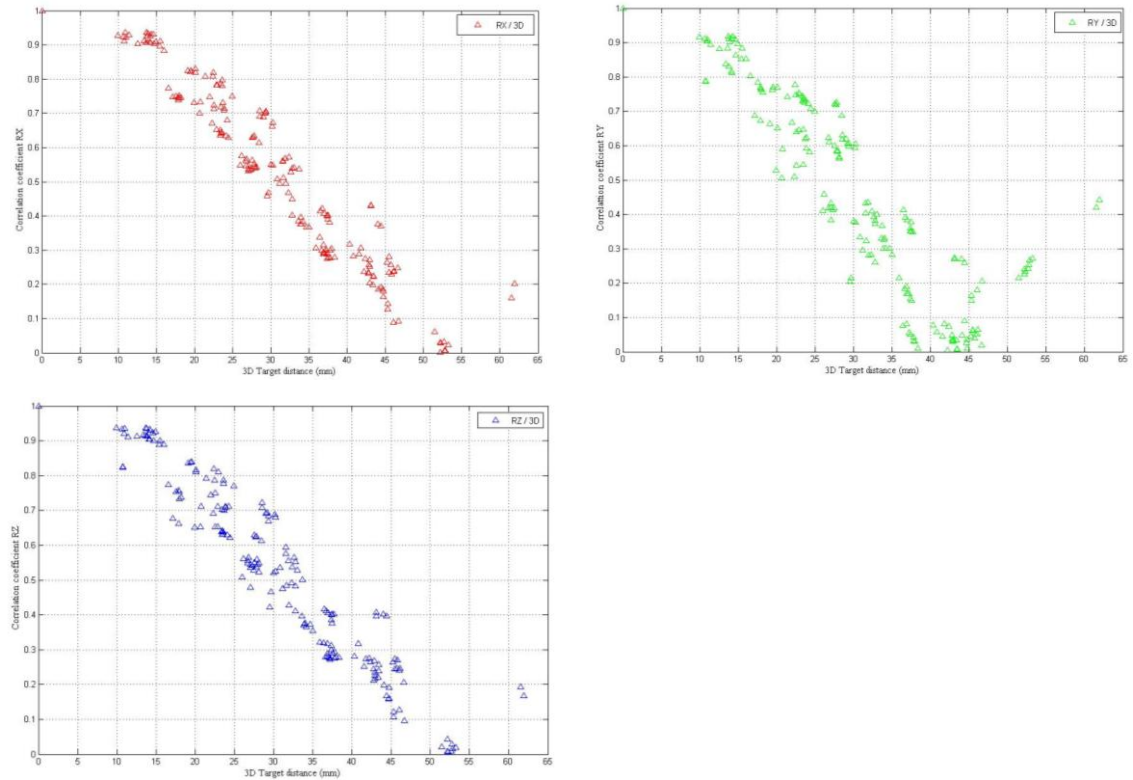
Figure 6.46: Correlation coefficients with proximities (red: RX, green: RY, blue RZ) - dataset D2.

Targets separated between D= ~10mm-20mm present high correlation (ρ= ~0.7-1.0) in X with only a few exceptions in directions Y and Z where D=~15mm-20mm with correlations being just below ρ= ~0.7. The data follow in general a smooth arrangement with only a small cluster (D= ~50mm-55mm) presenting correlations

between (ρ= ~0.2-0.3) in Y. In addition, correlations between CP4000 and CP4006 (with $D_{4000-4006}$= 61.51mm) as well as CP4000 and CP4009 (with D= 61.97mm) present an increase in correlations (with ρ_X, ρ_Z= ~0.1-0.2 and ρ_Y= ~0.4-0.5) when these should be reduced in comparison to their counterparts. However the correlation coefficients are considered as small and the relative increase in the standard deviations can only be attributed to poor target measurement quality.

6.5.3.3. Dataset E2 - centroid

Similarly to datasets B2 and D2 Figure 6.47 illustrates the absolute 3D correlation coefficients in each direction RX, RY and RZ over the corresponding 3D inter-target separations. The 3D target data range between 3.3mm (minimum separation) and 48.7mm (maximum separation).
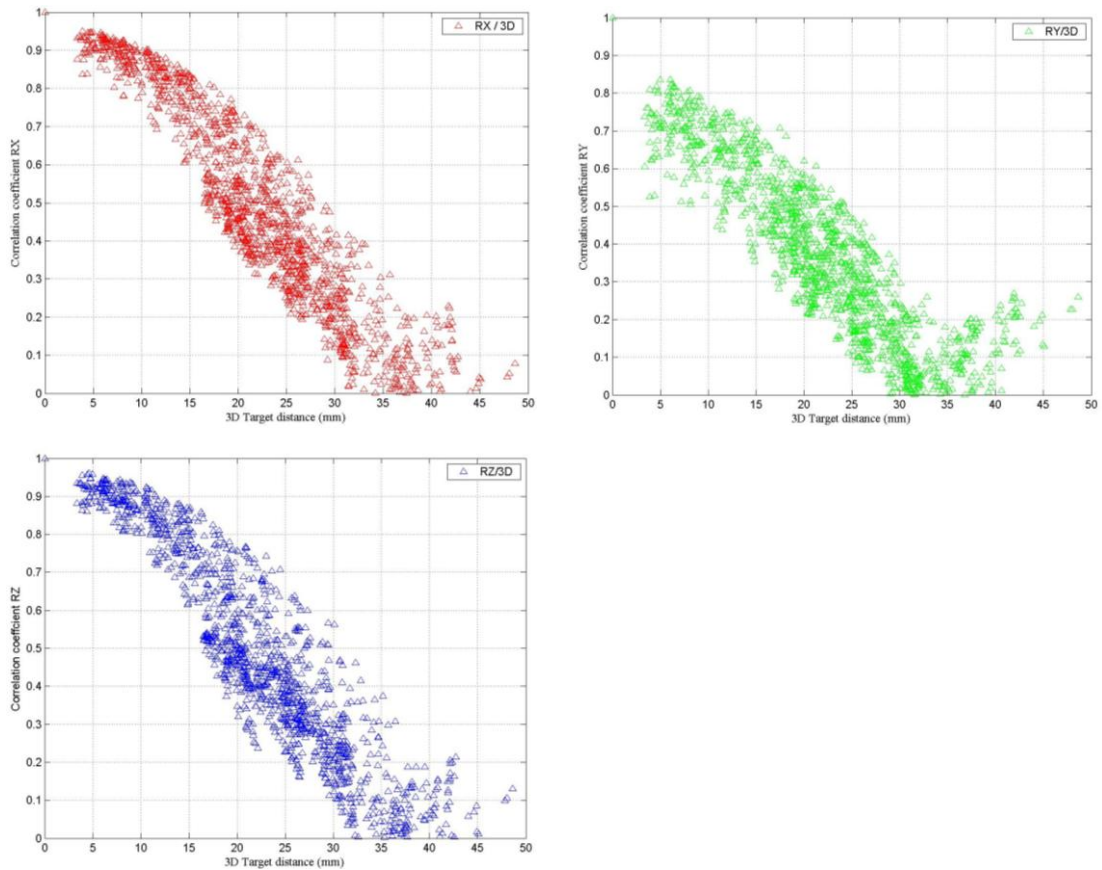
Figure 6.47: Correlation coefficients with proximities (red: RX, green: RY, blue RZ) - dataset E2.

Correlations RX, RY and RZ increase inversely with target-distance similarly to the results presented above. Significant correlations (ρ= ~0.7-1.0) occur between D=

~3.0mm - 10mm in X and Z with a relatively similar behaviour. It is interesting that correlations RY appear to be reduced in magnitude when these are compared to corresponding correlation coefficients in X and Z within the range 0-40mm (for example for distance D= 0-10mm; ρ_{XZ}> ~0.8 and ρ_Y= ~0.5-0.8, D= 10-20mm; ρ_{XZ}= ~0.4-0.9 and ρ_Y= ~0.2-0.8, D= 20-30mm; ρ_{XZ}= ~0.2-0.8 and ρ_Y= ~0.0-0.6, D=30-40mm; ρ_{XZ}= ~0.2-0.8 and ρ_Y= ~0.0-0.6). However within the range (D= 40-50mm 3D targets are correlated within ρ= ~0-0.2 in X, Y and Z). Moreover two correlation values at the far end of these figures seem not to follow the decreasing pattern of their counterparts. Specifically this occurs for the pair CP10000-TP20009 (where D=48.65mm with ρ_X= 0.08, ρ_Y= 0.26, ρ_Z= 0.13) as well as the pair CP10000-TP20008 (where D= ~48.04 with ρ_X= 0.05, ρ_Y= 0.23, ρ_Z= 0.10) but again these correlation coefficients are considered to be insignificant. It is however pointed that this repeatable pattern (see sections 6.5.3.1., 6.5.3.2. and 6.5.3.3) observed at the edge of the image format might result from a potential small uncorrected radial distortion error within the data measurement.

6.5.4. Error ellipsoids

The second measure that is utilized for object space evaluation is the visualization of the absolute 3D point error ellipsoids that are derived from the a posteriori covariance matrix (subblock for 3D targets). Error ellipsoids are considered to be highly useful for evaluation of bundle adjustment results. In particular they can characterize network orientation, scale and datum location. Here the error ellipsoids are visualized for each control and target point that contributes within the image network; these are ordered as extracted from the target data file.

6.5.4.1. Dataset B2 - pyramid

Figure 6.48 illustrates the location of the measured CP (red triangles) and TP (green triangles) and Figure 6.49 visualizes the corresponding 3D views of the error ellipsoids for dataset B2 (pyramid object) within the image network.
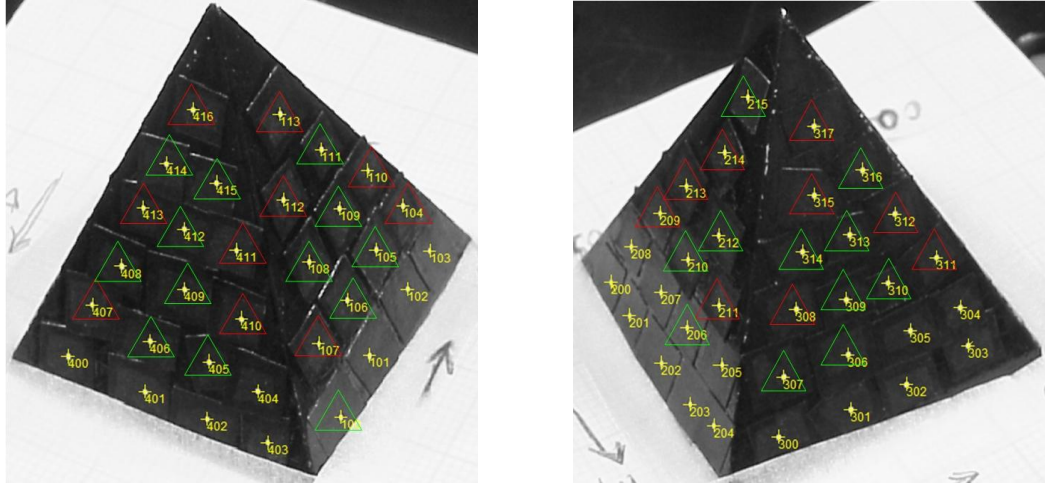
Figure 6.48: Point data distribution CP: red triangles, TP: green triangles - dataset B2.
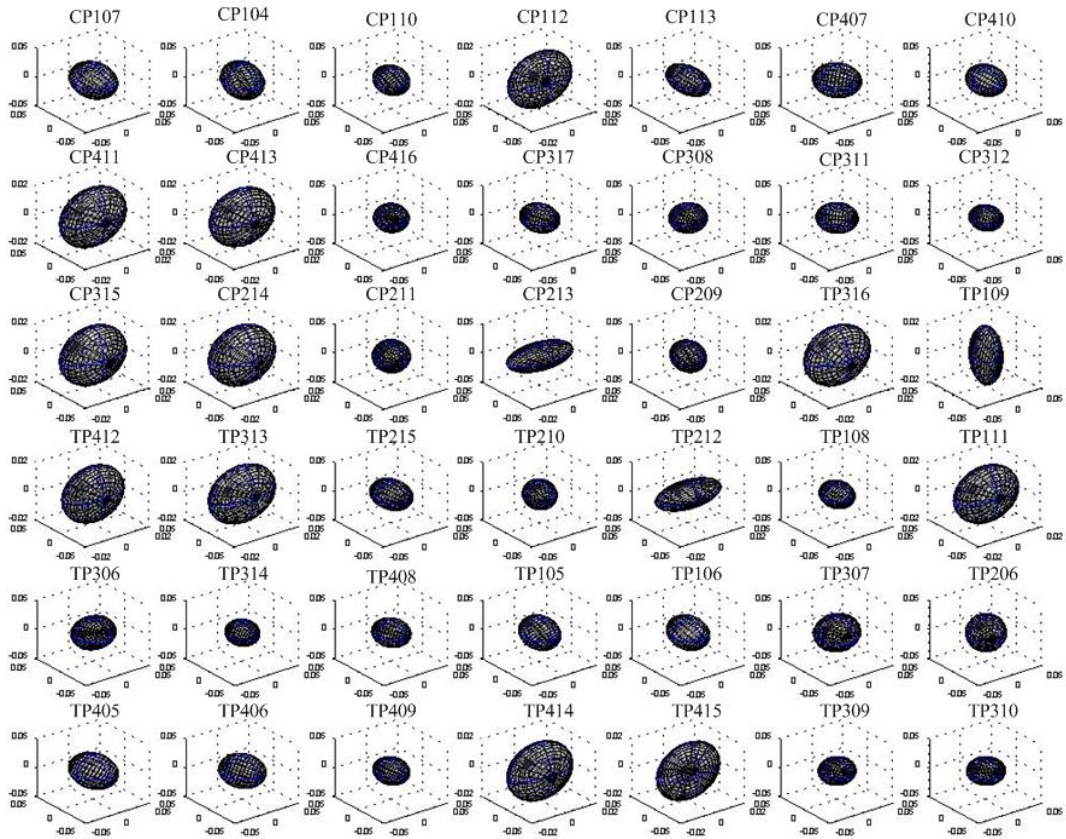
Figure 6.49: Error ellipsoids (external datum) - dataset B2.

The external constraint datum defined the control target coordinates with an overall 3D precision of $\sigma X= 7.69\mu m$, $\sigma Y=9.93\mu m$, $\sigma Z=7.28\mu m$ (see section 6.5.2.). From the ellipsoids pattern, it is evident that error ellipsoids of points that lie on a similar plane and row present similarities in shape and size. Specifically target points that are located in the fourth and fifth rows in all four facets present largest error ellipsoids (counting six rows per facet from the pyramid's basis to its peak). Such examples are

targets CP112, CP411, CP413, CP315, CP214, TP316, TP412, TP313, TP111, TP414, TP415. Opposed to these, targets that are located at the object's edges or are more spread (e.g. are coordinated from an increased number of intersection angles) present uniform and small in size error ellipsoids. Such examples are targets CP110, CP410, CP416, CP308, CP312, CP211, CP209, TP215, TP210, TP108, TP314, TP408, TP105, TP106, TP106, TP409, TP309, TP310.

6.5.4.2. Dataset D2 - lego

Figure 6.50 illustrates a pair of two views showing the arrangement of the CP (red triangles) and TP (green triangles) data within the image network and Figure 6.51 visualizes a 3D view of the corresponding error ellipsoids for dataset D2 (lego object).
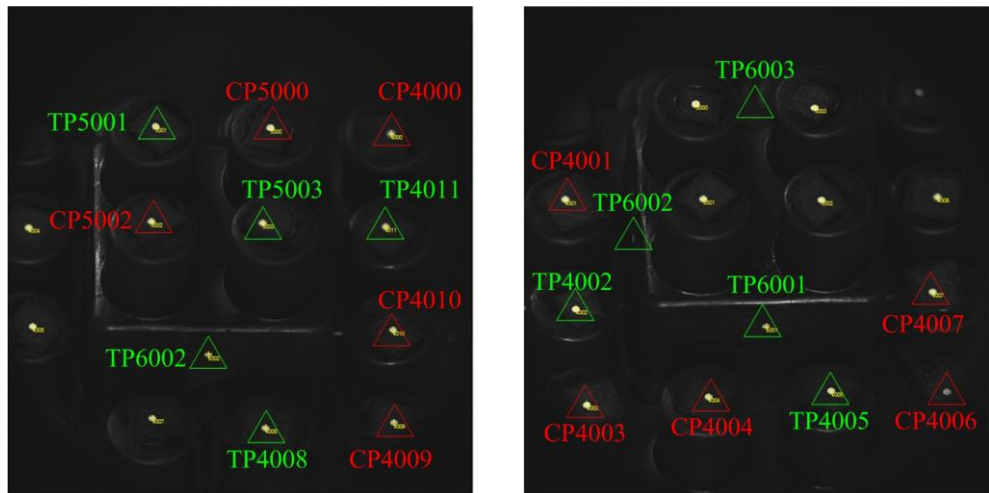
Figure 6.50: Point data distribution (CP: red triangles, TP: green triangles) - dataset D2.
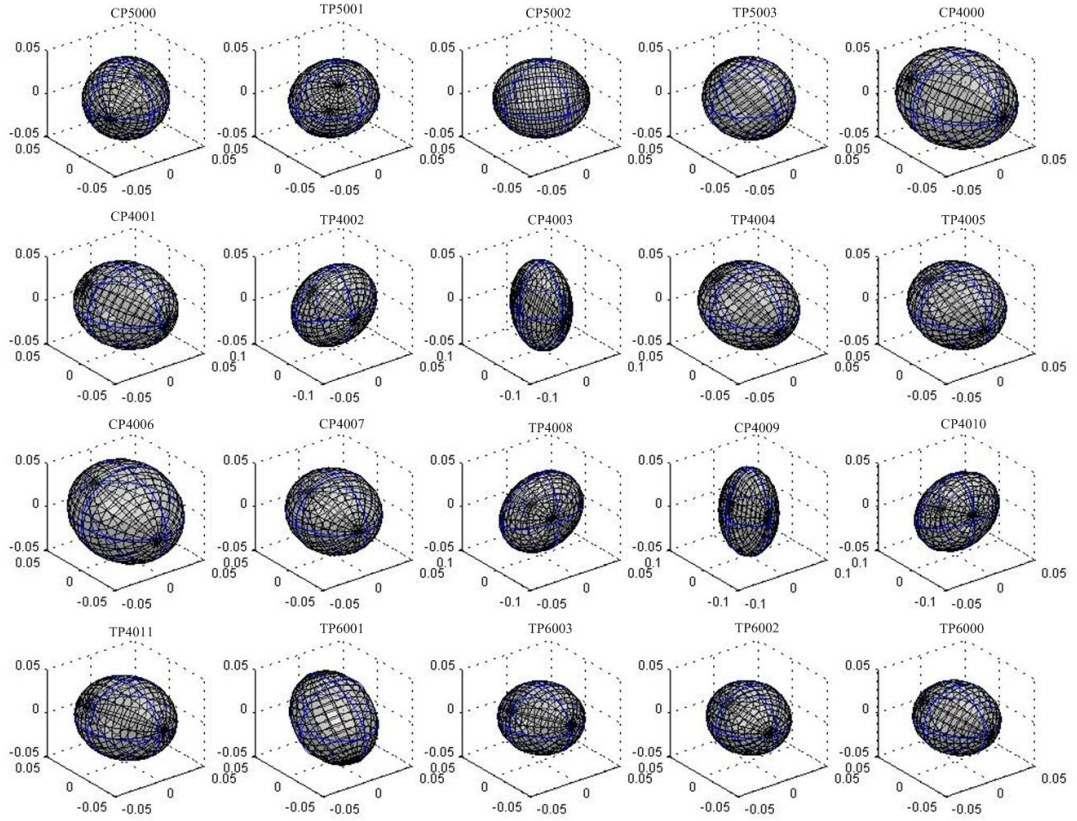
Figure 6.51: Error ellipsoids (external datum) - dataset D2.

The error ellipsoids are of similar magnitude and shape for the total number of targets which is reasonable considering the uniform image quality of the target points and simultaneously considering that the a posteriori 3D target precision was estimated as σX= 27.88μm, σY=31.52μm, σZ=28.03μm with a relatively increase in Y direction. Targets CP4000, CP4006 present the largest ellipsoid patterns. Moreover it is interesting that the error ellipsoids for targets CP4000 and CP4006 as well as targets CP4003 and CP4009 that are located on the base plane along the diagonals present high similarity in shape, orientation and magnitude. To check if this results from a weak orientation in the data (e.g. ray intersection angles) or if it characterizes the quality of the control data that were utilized to constrain the network, an obvious check was to re-run the solution with an inner constraint datum (with identical orientations, photo and target data). The inner datum solution resulted in a uniform error ellipsoid shape for all target points which proves that the network is highly homogeneous and strong (see Figure 6.52). As a result, the observed ellipsoid patterns in the external constraints case are not a function of a weakness in orientation. These can relate to poor control determination for these particular target points from data

pre-measurement (see section 6.5.1.2) or affine network computation (see section 6.5.2.2.).
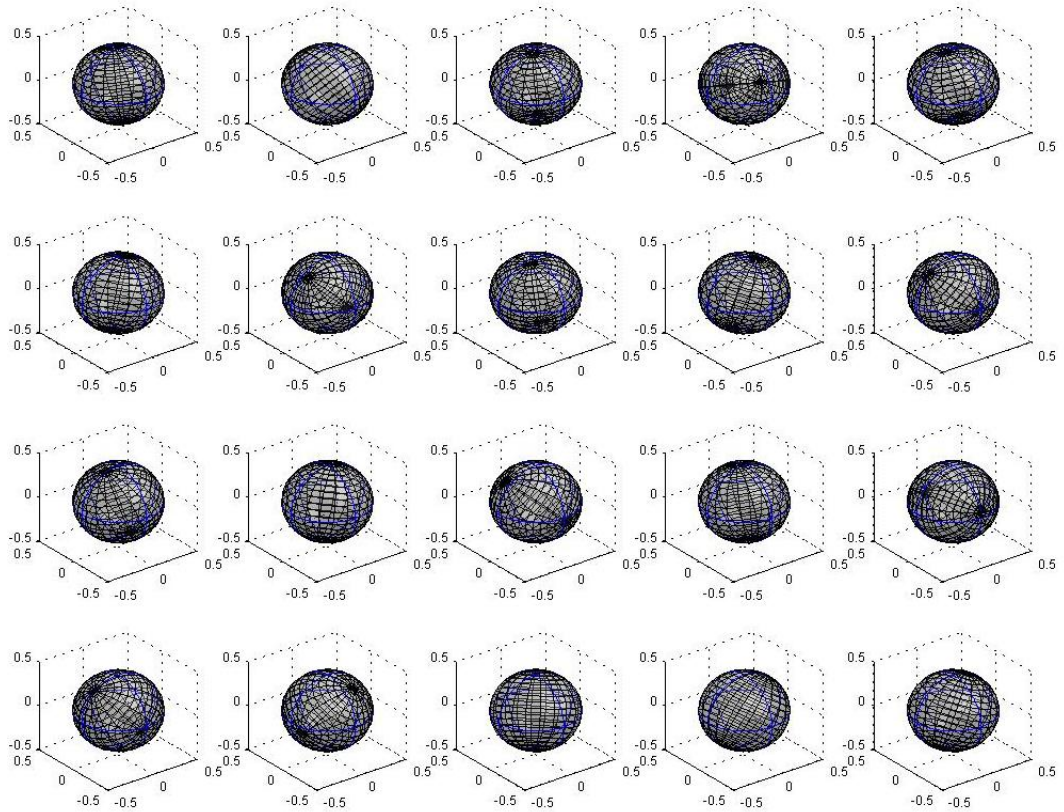
Figure 6.52: Error ellipsoids (inner datum) - dataset D2.

6.5.4.3. Dataset E2 - centroid

Figure 6.53 and Figure 6.54 illustrate the distribution of the control and tie point data within the image network as well as the 3D error ellipsoids of the estimated point data. The shape, magnitude and direction of the error ellipsoids suggest that no systematic effects were present within the affine image network. It is recalled here that the estimated 3D targets were coordinated with an a posteriori quality of $\sigma X=$ 36.54µm, $\sigma Y=$43.81µm, $\sigma Z=$36.86µm (see Table 6.13 in section 6.5.2.). To give some examples of large error ellipsoid patters; these characterize targets CP10000, TP20007, TP20008, TP20009, TP20010 that are located in the most outer ring of the structure as well as targets TP20011, TP20012, CP20013 and TP20014 that are located in the second ring (counting from the outside).
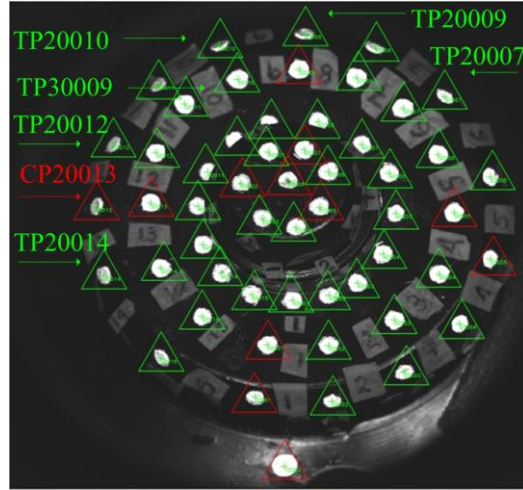
Figure 6.53: Point data distribution CP: red triangles, TP: green triangles - dataset E2.
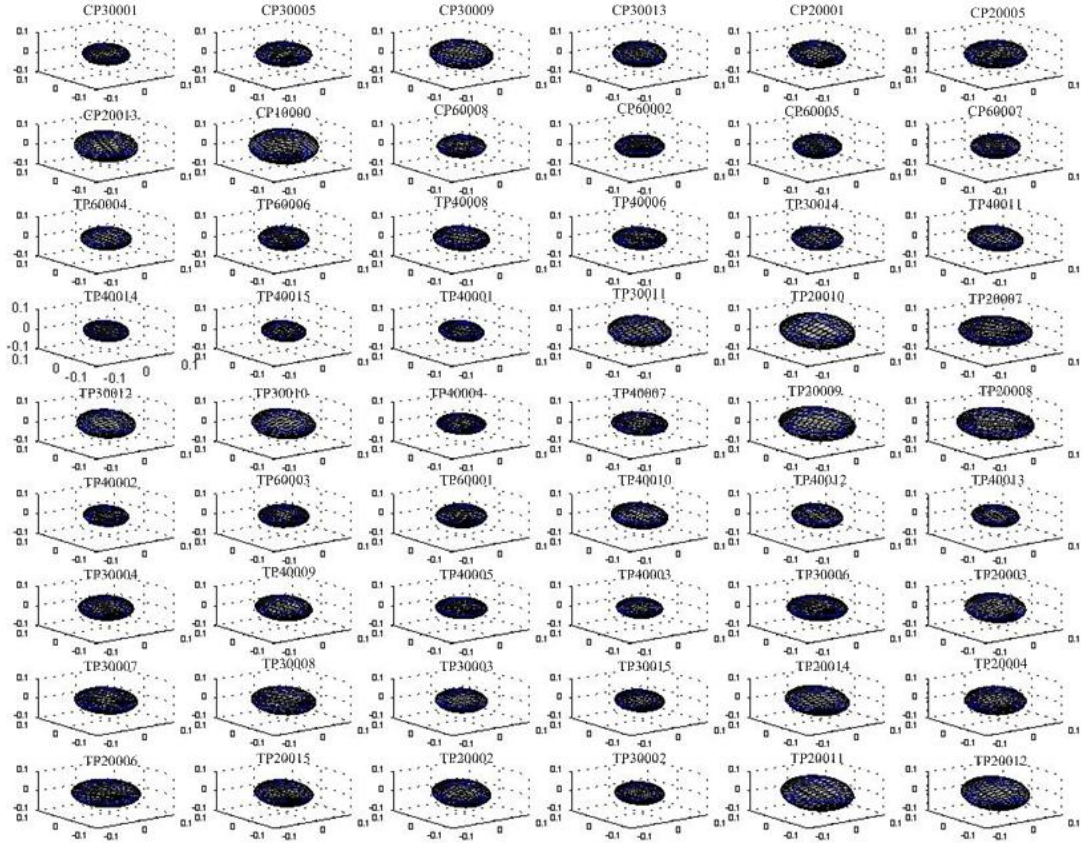
Figure 6.54: Error ellipsoids - dataset E2.

6.6. Scale invariance assessment

To assess sensor scale within the developed system, the bundle method was tested using two different datasets A2 and A3 acquired with the Kodak sensor (CAM_A3) and the Sony sensor (CAM_A5) respectively whereas initialization and target pre-measurement were performed utilizing dataset A1 captured with the available Nikon

DSLR camera (CAM_P1). Comparative results are obtained from statistical analysis of the bundle adjustment.

6.6.1. Reference measurements

Dataset A1 was established from a strong, highly convergent image network from 44 viewpoints at an imaging range of 400mm. It is noted that the rigid cube frame occupies a volume of 100 mm^3 enclosing the lego structure where 178 point targets (white markers) were observed (mean target visibility of 15.1) with an image quality pointed by a target image diameter of 4.5 - 9.0 pixels in image space. Figure 6.55 provides a description of the reference dataset. Specifically it illustrates the reference calibration object highlighting the control point data (red ellipses) as well as the 10 object space measurements (drawn in blue) that were introduced as scale measurements in the bundle adjustment (a priori precision of object scales was set to 50μm). In addition it illustrates the image network geometry and the photogrammetrically derived 3D point cloud.
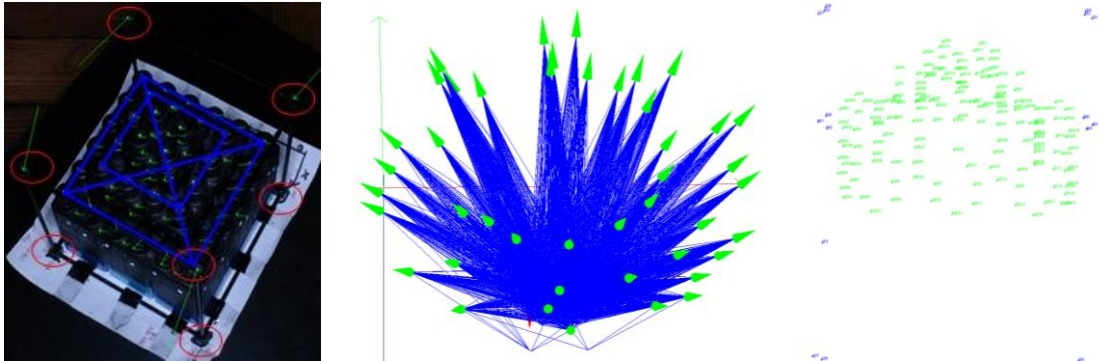
Figure 6.55: Calibration rigid structure (left), image network (middle) and photogrammetric point cloud (right).

The dataset was processed with the bundle adjustment within the VMS 8.0 tool defining an inner constraints datum and the internal calibration model (excluding insignificant decentering and affinity terms). The bundle was implemented with a redundancy of 4,564 and it converged after 10 iterations resulting in a triangulation misclosure of 0.37μm (that is 1/20th of a pixel for the Nikon sensor). Object geometry was recovered with a relative precision for the image network of 1:36,000, a mean 3D target coordinates precision of 5.66μm and an RMS object scale residual of 103.28μm (see Table 6.14).

Dataset	Iter.	R.	σ_o	RMS$_{xy}$ (µm)	σ_{XYZ} (µm)	Relative precision	RMS scale (µm)
				IS	OS		
A1	10	4,564	1.00	0.37	5.66	1:36,000	103.28

Table 6.14: BA statistical indicators - datasets B1, D1 and E1.

6.6.2. Affine bundle adjustment results

Figure 6.56 illustrates a sample perspective view with two inset views acquired from the two different camera systems with clear differences in footprint and image content. The main differences of the test datasets are related to sensor characteristics noting that the key difficulty in this particular test case was the object's dimension was larger than the footprint of the imaging system. The acquired datasets were processed with the pre-measured data and datum as defined above (see section 6.6.1.). However the object's dimensions, occlusions and image quality demanded significant pre-editing in order to remove measurements that were partially occluded in subsequent images and simultaneously did not present sufficient ray intersection (3 or 4 rays per frame) as an example.
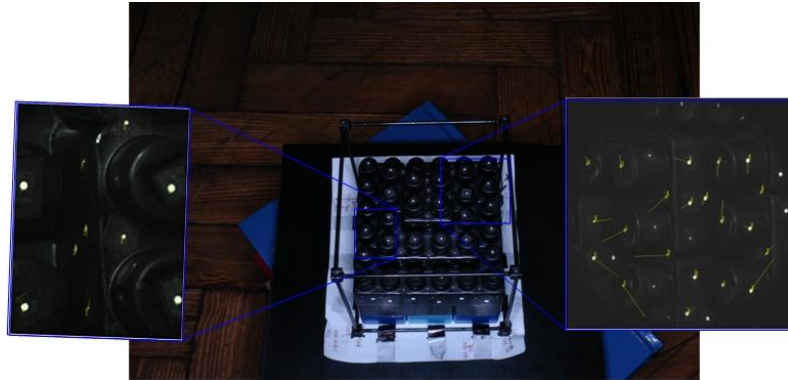
Figure 6.56: Perspective view with affine views - CAM_A5 (left) and CAM_A3 (right).

The Sony camera system (CAM_A5) is characterized by a highly narrow field of view as well as limited resolution (see section 6.2.2.2.) when this is compared to the Kodak sensor footprint (CAM_A3). As a result the number of images required to cover the object volume increase and the target dispersion within each frame become more limited. However network design and image acquisition ensured that a minimum number of control points (4-6 per view) were present for stable frame initialization and that there was sufficient multi-image coverage (through tie point measurement)

and redundancy. This consideration is critical for subsequent calibration and network stability. Table 6.15 provides the comparative bundle adjustment results for datasets A2 and A3 noting that these were processed with an external datum and the implemented calibration model.

Dataset	Iter.	R.	Scale (σ_{scale}x10^{-5})	k_3 ($\sigma_{\kappa3}$)	σ_o	RMS$_{xy}$ (μm)	σX, σY, σZ (μm)
A2	3	1,670	0.1615 (7.5282)	2.6356x10^{-5} (5.3196x10^{-6})	4.04	1.30	25.73, 30.53, 25.71
A3	3	653	0.1659 (3.4151)	2.2805x10^{-4} (2.3532x10^{-5})	2.07	0.56	17.75, 34.83, 15.94

Table 6.15: Affine BA results (external datum) - datasets A2 and A3.

6.6.2.1. Dataset A2 - Kodak sensor

In the case of dataset A2 the image network was processed with 35 CP, 87 TP and 44 photo data. The measurements present a minimum number of 4 rays (for all target points) and a maximum number of 26 and 18 visibilities (for control and tie points respectively) (see Figure 6.57). Initialization was performed with the approach described in the previous test cases (see section 6.4.1. as an example). Resection computes an RMS image measurement residual of 1.52μm (to define the 3D orientation angles). A subsequent closed form back-substitution estimates the remaining 2D projective translations. The 3D target coordinates for the control were obtained from the reference measurements (see section 6.6.1.) therefore the datum was defined from the identified 35 control point data.
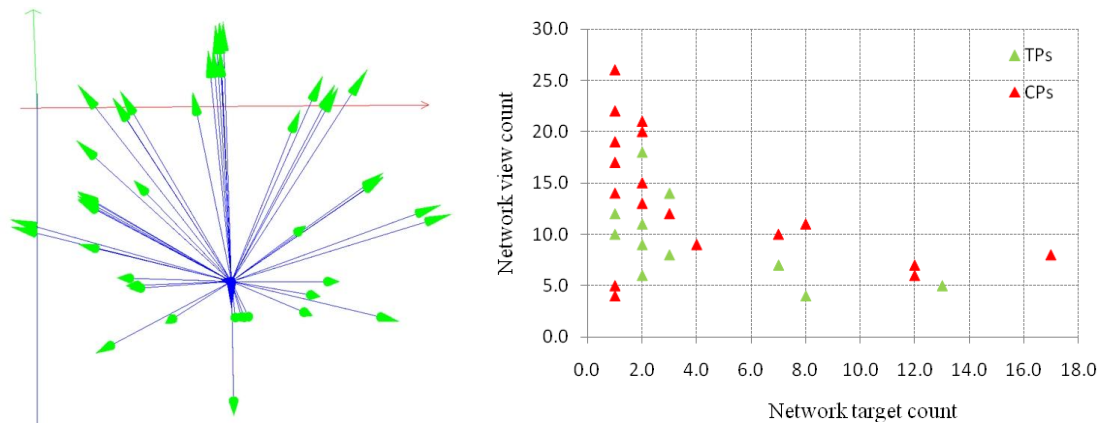
Figure 6.57: Image network geometry (left) and data visibility (right) - dataset A2.

The system was run with a redundancy of 1,670 observations and converged rapidly after 3 iterations with an a posteriori sigma nought of 4.04 and a triangulation misclosure of 1.30µm ($1/10^{th}$ of a pixel). Object points were estimated with a mean a posteriori 3D precision of 27.32µm ($\sigma X = 25.73$µm, $\sigma Y = 30.53$µm, $\sigma Z = 25.71$µm). The sensor scale is equal to 0.1615 (with a precision of 7.5282×10^{-5}) whereas the radial lens distortion was calculated as $k_3 = 2.6356 \times 10^{-5}$ (with a precision $\sigma_{k3} = 5.3196 \times 10^{-6}$ and a significance (where: significance= correction / precision) of -0.0024 which result in a correction of 1.28×10^{-8} for the radial lens distortion term).

6.6.2.2. Dataset A3 - Sony sensor

Dataset A3 is processed with 78 CP, 8 TP and 85 photo data. The data visibility ranges between 3 (for CP data) and 16 rays (for CP and TP data) with a mean number of valid target image observations of 5.72 within the image network. For complete object coverage this particular dataset presented some significant difficulties that are worthy of mention. The Sony sensor presents a very narrow field of view when used with the telecentric lens and the point data are highly sparse for these particular magnified close-ups. As a result and to stitch the images through point data measurement (given that there is significant overlap and control for frame initialization and subsequent datum definition) the image network was designed to be highly 'systematic' (with regards to the imaging range) with relatively closed separated viewpoints (narrower bundle of rays) (see Figure 6.58). The resulting 85 views within the image network ensured sufficient geometric strength and redundancy for data processing.
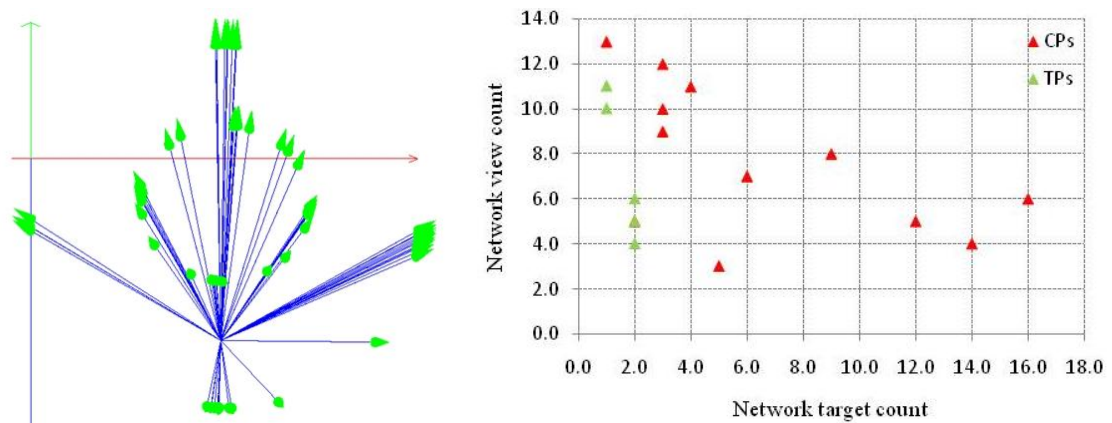
Figure 6.58: Image network geometry (left) and data visibility (right) - dataset A3.

Initialization was implemented similarly to the Kodak sensor (see section 6.6.2.1.) with a mean resection image measurement residual of 1.22μm and a closed form estimation of the 2D projective translations. The system was run with a redundancy of 653 observations and converged rapidly after 3 iterations with an a posteriori sigma nought of 2.07 and a triangulation misclosure of 0.56μm (1/9th of a pixel). 3D object coordinates were estimated with a mean a posteriori 3D precision of 22.84μm (σX=17.75μm, σY=34.83μm, σZ=15.94μm). The sensor scale was estimated to be equal to 0.1659 (with a precision of 3.4151×10^{-5}) whereas the radial lens distortion was calculated as 2.2805×10^{-4} (with a precision $\sigma_{k3}=2.3532 \times 10^{-5}$, a significance which is equal to -0.0256 and a resultant correction of -6.03×10^{-7}). Although redundancy, image quality and network geometry are reduced when compared to dataset A2, the bundle results show that this dataset provides an improvement in both estimating RMS image misclosure as well as 3D point precision. This is an obvious outcome considering the 78 control point data that were utilized to tie this image network (stable datum definition).

6.6.2.3. Object scale

To provide evidence of the system's ability for object scale recovery with both employed sensors (Kodak and Sony) two additional checks are given. The first measure calculates the absolute difference of five key distances over the estimated distances that were obtained from the bundle adjustments (see Figure 6.59 for illustration of data arrangement).
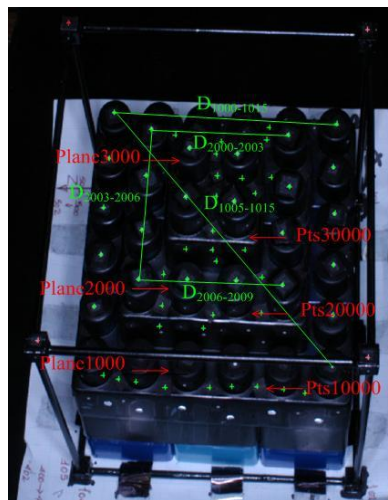
Figure 6.59: Object scales and point data arrangement.

It is evident that in comparison to dataset A3, dataset A2 is closer to truth (denoted by S which stands for scale) by 5.45μm in the case of distance $D_{2003\text{-}2006}$ (where the maximum discrepancy occurs) and 16.07μm in the case of distance $D_{2006\text{-}2009}$ (where the minimum discrepancy occurs) (see Table 6.16).

Distance	$AD_{(S\text{–}A2)}$ (μm)	$AD_{(S\text{–}A3)}$ (μm)
$D_{1000\text{-}1005}$	47.44	67.34
$D_{1005\text{-}1015}$	76.49	90.87
$D_{2000\text{-}2003}$	99.23	110.99
$D_{2003\text{-}2006}$	128.06	133.51
$D_{2006\text{-}2009}$	36.42	52.49

Table 6.16: Distance checking - datasets A2 and A3. Table notation: AD= absolute difference, S=scale, A2= dataset A2, A3= dataset A3.

The second check calculates the mean absolute discrepancy in 3D directions X, Y and Z between the reference measurements (dataset A1) that initiated the control data from premeasurement of dataset A1 and the estimated point coordinates for the individual datasets A2 and A3. In Table 6.17 these are grouped for the different object planes (coded plane1000, plane2000 and plane3000 starting from base to top plane) as well as the points located at the intermediate corresponding sides (coded pts10000, pts20000, pts30000).

Dataset	A2			A3		
	MAD_X (μm)	MAD_Y (μm)	MAD_Z (μm)	MAD_X (μm)	MAD_Y (μm)	MAD_Z (μm)
Plane1000	14.05	18.90	11.77	8.23	12.70	7.78
Plane2000	5.07	7.00	5.23	7.34	6.66	11.82
Plane3000	4.76	7.25	3.87	6.65	23.68	11.13
Pts10000	13.03	25.45	10.65	59.03	83.89	23.73
Pts20000	10.43	21.57	11.23	8.92	11.14	8.95
Pts30000	4.23	11.96	6.82	10.89	16.60	7.92

Table 6.17: Object space discrepancies - datasets A2 and A3. Table notation: $MAD_{X,Y,Z}$= mean of the absolute discrepancies between reference measurements and estimated point data.

It is evident that the most significant discrepancies in both instances occur in the Y (viewing) direction. The Kodak sensor coordinates results to a mean 3D discrepancy of MAD_{XYZ}= 10.73μm (MAD_X = 8.59μm, MAD_Y=15.35μm, MAD_Z=8.26μm) whereas the Sony sensor results in a 3D discrepancy of MAD_{XYZ}= 18.17μm (MAD_X = 16.84μm, MAD_Y=25.78μm, MAD_Z=11.89μm). The closure of the estimated discrepancies in relation to the reference data is attributed to the geometric strength,

redundancy and image quality of the Kodak sensor as well as the significant control of its competent Sony sensor. 3D points appear to be closer to the reference measurements for the data which are located in the middle and top planes of the lego structure (coded plane2000) which is a natural expectation considering geometric strength (wide ray intersection angles), increased visibility and frame coverage. Finally points that lie on the lego structure's faces (coded Pts10000 - Pts30000) present reduced accuracy due to their poor image quality (image targets vary in diameter between 9-18 pixels for the Kodak sensor and 17-33 pixels for the Sony sensor) as well as visibility (targets are occluded in relation to the network viewpoints) and therefore less redundancy in their successive measurement frames. It is noted that point targets that lie on the three planes (coded Plane1000 - Plane3000) present an improved accuracy. This is attributed to the image quality considering direction of illumination and geometric viewpoint location as well as uniform target dimensions (image target diameter of 18 pixels for Kodak sensor and 33 pixels for Sony sensor).

6.7. Independent testing

To independently test the affine bundle adjustment algorithm, the developed method is compared over a conventional bundle approach utilized within the software tool VMS 8.0. For this reason three datasets were acquired; the first provides reference measurements for point data initialization (dataset C1), the second is used for processing of the perspective bundle adjustment (dataset C2) and finally the third is generated for affine bundle adjustment processing (dataset C3) (see section 6.2.3.). It is re-iterated here that processing of dataset C1 has already been given earlier in the model assessment test case (see section 6.4.1.1.).

6.7.1. Bundle adjustment results

Both datasets were processed based on a similar processing framework. Specifically dataset C2 was processed with 47 images, 18 CP and 7 TP data and dataset C3 was processed with 45 images, 18 CP and 11 TP data. For comparative evaluation both datasets were run with identical control; that is the datum was initiated from the identified reference image network.

6.7.1.1. Image networks

Figure 6.60 provides a comparative illustration of the perspective and affine image network geometries together with two selected views per case. It can be seen that both image networks were acquired in a two-ring strong convergent configuration. In addition, to extract useful evaluation of the affine image dataset over the available bundle adjustment (within the software tool VMS 8.0) it was ensured that the point data covered identical volume in 3D object space. As a result and besides the sparse point data arrangement in the case of the perspective image dataset the targets occupy nearly 1/4 of the image frame, however these are located in the middle as far as possible for sufficient point estimation as well as calibration parameters. Regarding point visibility in the case of dataset C2 maximum visibility ranges between 40 and 32 views (two CPs are coordinated from 40 views and one TP is coordinated from 32 views) and minimum visibility ranges between 25 and 22 views (one CP is coordinated from 25 views and one TP is coordinated from 22 views). For dataset C3 the corresponding data are given: maximum visibility ranges between 45 (one CP) and 38 views (one TP) whereas minimum visibility lies between 29 (one CP) and 18 (one TP) views (see Figure 6.61).
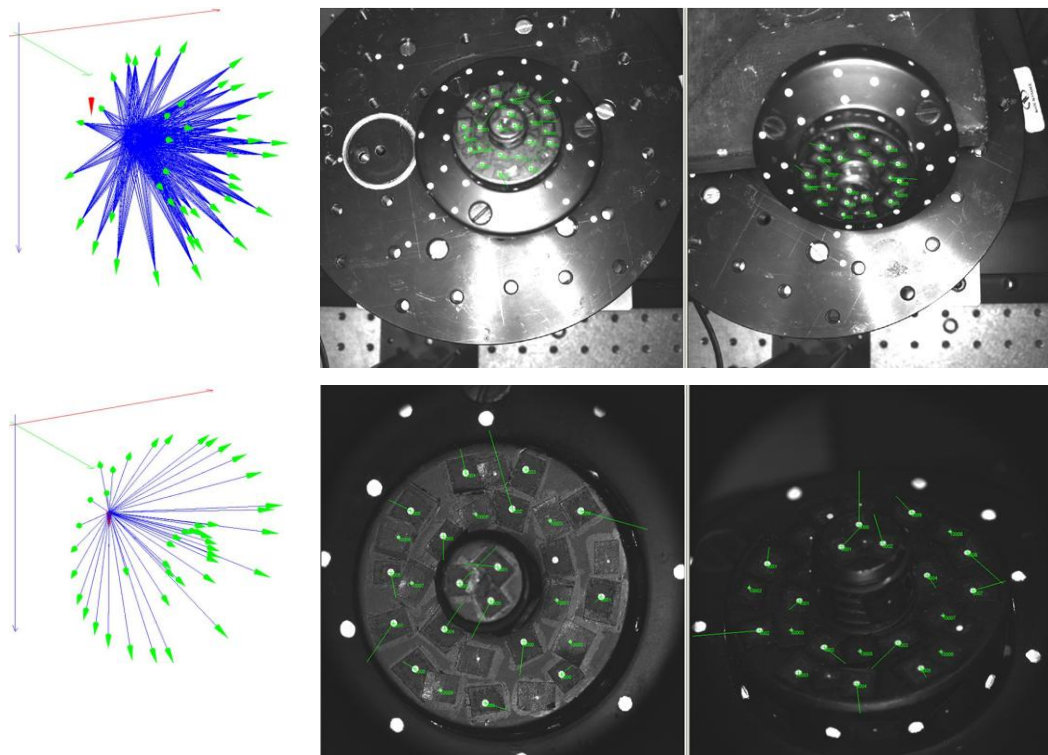
Figure 6.60: Image network geometry and views - dataset C2 (top) and dataset C3 (bottom).
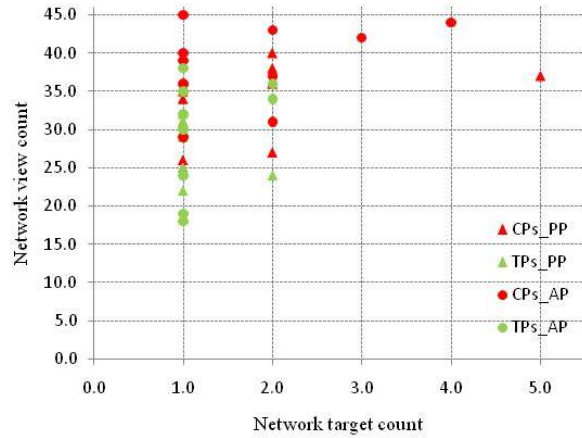
Figure 6.61: Data visibility - dataset C2 (PP: perspective projection) and dataset C3 (AP: affine projection).

6.7.1.2. Results

Dataset C2 was processed with 47 images, 18 CP and 7 TP data with an external constraints datum and full calibration parameters. Similarly dataset C3 was processed with 45 images, 18 CP and 11 TP data with an external constraints datum and inclusion of the implemented radial lens distortion term (k_3). The comparative bundle adjustment results are given in Table 6.18.

Dataset	Iter.	R.	σ_o	RMS_{xy} (µm)	σX, σY, σZ (µm)
				IS	OS
C2	4	1,343	1.00	0.78	8.92; 10.46; 8.94
C3	3	1,838	2.23	0.79	19.14; 21.93; 19.20

Table 6.18: Bundle adjustment statistical indicators - datasets C2 and C3.

In the case of dataset C2 initialization was implemented with a resection procedure (for the 47 photo data) producing an RMS image measurement residual of 0.74µm (with a mean valid target image observation number of 13.17) and a subsequent affine forward intersection procedure with σ_{XYZ}= 5.23µm. However on the final estimation stage control was updated by its reference 3D coordinates (and their associated standard deviations) to ensure that the datum remains as it was originally defined. The bundle adjustment converged after 4 iterations with a number of 1,343 redundancies resulting in a triangulation misclosure of 0.78µm (~1/12th of a pixel). In object space the bundle estimates 3D points with a mean precision of 9.44µm (σX=8.92µm, σY=10.46µm, σZ=8.94µm) whereas the relative precision for the image network is equal to 1:5000. Calibration is recovered with the 10 parameter self-calibration model

embedded within the software and the radial lens distortion at a radial distance of 3.5mm estimated to be equal to -44.65μm for the Kodak sensor.

Initialization of dataset C3 was based upon 18 control point data (reference measurements). The resection procedure (of the 45 photos within the network) resulted in an RMS image measurement residual of 1.37μm (with a mean valid target image observations of 13.45). The subsequent affine forward intersection procedure was run with a number of 2,011 redundancies resulting in a sigma nought of σ_o= 2.99 and a σ_{XYZ}= 2.18μm after 2 iterations. To keep the datum defined at its initial set up the control data were set to their initial reference measurements. The affine bundle adjustment (dataset C3) converged after 3 iterations with a number of 1,838 redundancies and an RMS image misclosure of 0.79μm (~1/12th of a pixel) presenting high similarity over dataset C2. In object space the bundle estimates 3D points with a mean precision of 20.09μm (σX=19.14μm, σY= 21.93μm, σZ= 19.20μm). Here the camera system which employs the Kodak fitted with the telecentric lens is calibrated with a scale of 0.1611 (σ_s= 1.4401x10^{-4}), radial lens distortion is 8.45μm at a radial distance of 4.00mm (k_3=1.3204x10^{-4}, σ_{k3}=6.0180x10^{-6}, significance= 0.0150).

6.7.1.3. Object space

For object space evaluation two measures are calculated. The first provides a visual display of the 3D target error ellipsoids for both datasets C2 and C3. Specifically Figure 6.62 visualizes the 3D target error ellipsoids for dataset C2 (CPs coloured in blue and TPs coloured in green) with an ellipse scale factor of 10.3 (within the software tool VMS 8.0). The error ellipsoids in the case of the affine dataset C3 are displayed with a default scale of 1.0 and they are listed according to their order (obtained from the input 3D target file and labelled in blue: CPs and green: TPs). It is noted that 3D error ellipsoids present uniform shape, magnitude and orientation with regards to object geometry – location, point type and precision. As a result the 3D points present uniform quality (as obtained from the scaled a posteriori covariance matrix) which is confirmed by the tabulated results (see Table 6.19).
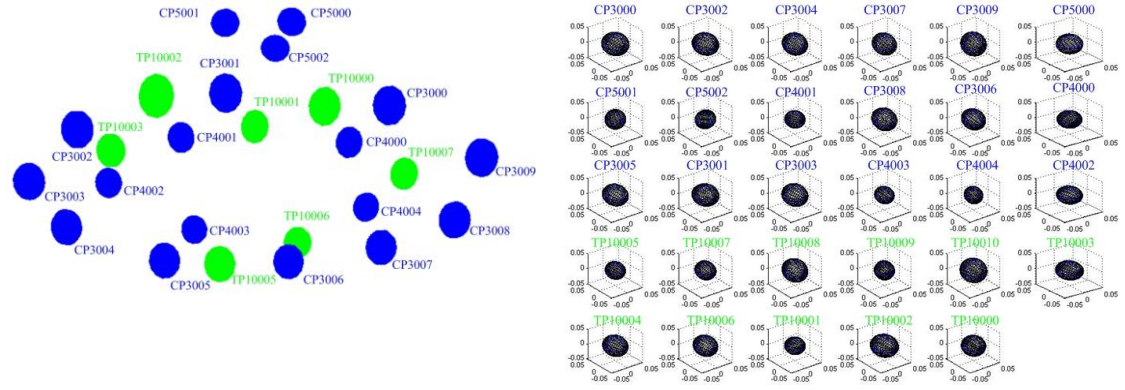
Figure 6.62: 3D target error ellipsoids - dataset C2 (left) and dataset C3 (right).

Dataset	CP			TP		
	σX (μm)	σY (μm)	σZ (μm)	σX (μm)	σY (μm)	σZ (μm)
C2	8.89	10.26	8.94	8.99	10.93	8.89
C3	19.11	21.61	19.22	19.16	22.44	19.17

Table 6.19: 3D target precisions - datasets C2 and C3.

Firstly dataset C2 estimates 3D target points with a precision which is improved by 10μm in three (X, Y, Z) directions in comparison to dataset C3 considering that the affine bundle scales the 3D point precisions to an a posteriori sigma nought of 2.23. In both cases the error ellipsoids of the control present similar shape and magnitude per concentric ring (from the outer towards the inner target rings). Particularly, CP4003 (which presents minimal σ_Z, 31 and 25 rays in 'C2' and 'C3' respectively) and CP4004 (with minimal σ_X, 31 and 26 rays in 'C2' and 'C3' respectively) as well as CP5001 (with minimal σ_Y and 40 rays in C2) and CP5002 (which presents minimal σ_Y with 44 and 40 rays in 'C2' and 'C3' respectively) are examples of small error ellipsoid shapes. TP10000 is an example of reduced quality (coordinated from 25 and 36 views and estimated with 3D precisions $\sigma_{XYZ}=$ 10.27μm and $\sigma_{XYZ}=$ 21.16μm in C2 and C3) whereas TP10001 is given as an example of a better quality (coordinated from 24 and 35 views and estimated with 3D precisions of $\sigma_{XYZ}=$ 8.83μm and $\sigma_{XYZ}=$ 17.73μm in C2 and C3 accordingly).

To evaluate accuracy Table 6.20 summarizes the mean of the absolute discrepancy values for the control point data within the image network. It can be seen that the affine image network derives a good comparative solution in comparison to the perspective dataset. In fact given its geometric strength as well as complete point coverage within the image frame, it results to an agreement of 6.27μm with the

reference data against an absolute discrepancy of 6.42μm that occurs in the case of dataset C2 (3D discrepancies of control over reference point data).

Dataset	MAD_X (μm)	MAD_Y (μm)	MAD_Z (μm)
C2	6.64	4.55	8.08
C3	6.55	6.36	5.90

Table 6.20: Control point discrepancies - datasets C2 and C3. Table notation: $MAD_{X,Y,Z}$= mean of the absolute discrepancies between reference measurements and control point data.

6.8. Summary

In summary this chapter provides an extensive assessment of the developed affine multi-view algorithm in the context of close range object measurement. The employed tests were designed in order to test and evaluate the method in relation to both considerations of correctness as well as effectiveness in practice. In particular the tests have derived method behaviour in the aspects of initialization, bundle adjustment algorithm, object space recovery, invariance of model scale as well as independent evaluation with reference data. Assessment has been performed utilizing the typical statistical indicators extracted from the bundle adjustment as well as measures that evaluate specific aspects of the method and which have been given analytically at each separate experimental case with regards to precision and accuracy aspects. The implemented bundle adjustments were run on an Intel® Core ™ Duo CPU, 2.80GHz, 1.59GHz, 1.96GB of Random Access Memory (RAM). Table 6.21 summarizes the performance characteristics of the overall datasets that were utilized to test the method. It is evident that the data sizes of the processed bundles are relatively small; these are given in the context of complete method description and not for purposes of evaluation of the method's performance.

Dataset	CPU time (hr:min:sec)	Memory usage	VM size
C3 (#CTPE)	0:00:30	23,432K	17.576K
C3 (#CTPI)	0:00:33	25,996K	21,116K
B2	0:00:06	18,304K	11,888K
D2	0:00:04	13,752K	7,320K
E2	0:00:12	19,536K	13,396K
A2	0:01:19	93,872K	87,720K
A3	0:00:52	117,056K	110,856K
C3	0:00:24	22,512K	16,132K

Table 6.21: Comparative performance characteristics.

Model assessment has proved that the affine method converges rapidly (in the absence of outliers) with a triangulation misclosure of $1/10^{th}$ - $1/11^{th}$ of a pixel for the employed camera systems (C mount progressive scan cameras fitted with an MVO® TMLTM/0.16x telecentric lens). Object space is recovered with a 3D precision which varies within a few tenths of microns for the six different affine datasets when the bundle is run with an external constraints datum and the implemented internal calibration model. Inner constraints result in a significant reduction of 3D point precision which has been concluded that this is attributed to modelling of inner constraints datum for this scale invariant sensor model. In addition it has been shown that inner constraints degrade model quality for the datum variant parameters (projective translations, 3D rotations, 3D point coordinates and scale) but not for the datum invariant parameter (third power term of the radial lens distortion polynomial). The correlation analysis check has proved that significant correlations occur between parameters of the same kind with some distinctive correlations between 2D projective translations and 3D rotations as well as 3D target coordinates. Yet inner constraints present high correlations between 2D projective translations and 3D target coordinates which result from definition of the centroid datum with a simultaneous increase in correlations of 3D target coordinates as opposed to the external datum case.

Evaluation of the 3D point coordinates correlation coefficients over the corresponding 3D target separation has shown that targets present high correlations inversely with their 3D inter-target distance. Moreover, the 3D point error ellipsoids provide a good indicator of the achieved quality in 3D space; in fact it has been shown that the estimated points are in good agreement with similar type of precisions achieved from well-known reference photogrammetrically derived measurements (through robust bundle adjustment implementation).

Testing with camera systems that employ two different sensors (Kodak Megaplus ES1.0 monochrome and Sony DFW-SX900 colour cameras fitted with the employed MVO® TMLTM/0.16x telecentric lens) has shown that the method derives accurate results at the order of 10-20µm in comparison to reference measurements. Finally an independent check has evaluated the method over a perspective-based image network that was run with the available in-house bundle adjustment tool (VMS 8.0) for

comparative assessment purposes. It has been proved that the affine bundle adjustment results in a very close quality agreement with the perspective bundle adjustment in the aspects of RMS image space (with a misclosure of $1/12^{th}$ of a pixel), 3D point estimation (with a 3D precision of ~10-20µm and a sigma nought of 1.0 and 2.23 for perspective and affine image networks accordingly) as well as accuracy (3D discrepancies over reference measurements range within 10-20µm). The central conclusions derived from development and implementation of the affine multi-view method are given in the subsequent Chapter 7.

7. Conclusions and future research

This research has investigated the problem of affine multi-view modelling for the purpose of close range object measurement. A multi-view framework has been designed, developed and tested against simulated and real datasets. This chapter first presents a summary of the approaches presented in this thesis (see section 7.1.), subsequently provides the central conclusions (see section 7.2.) and finally suggests directions for future research (see section 7.3.) and closes with a final research point (see section 7.4.).

7.1. Summary

The main focus of this research has been the investigation and development of a new method for the purpose of system calibration, orientation and 3D measurement from affine image networks. This is treated in the context of addressing the fundamental questions that have been presented in the problem statement section (see section 1.3.) and are re-iterated as follows:

- Is it realistic to generate, measure and process real affine multi-view images within a modelling framework in the context of deriving precise close range object measurements?
- In the context of such a framework, how do sensor geometry (parallel projection rays, invariant scale factor and calibration) as well as local coordinate frame (datum) influence method precision and 3D point estimation (object geometry)?
- What is the quality of affine multi-image modelling in comparison to established photogrammetric solutions?

In these aspects the thesis starts with the investigation of the research context (Chapter 1). This is done with formulation of the background in the concepts of non-contact object measurement (Chapter 2), digital close range image formation (Chapter 3) as well as multi-view modelling in the photogrammetric aspect (Chapter 4). The thesis can be considered in terms of: (a) method development and (b) application and assessment of the system. System development focuses on starting value recovery as well as design and implementation of the bundle adjustment algorithm (Chapter 5).

Application of the method considers each development stage and provides extensive assessment through testing with the designed experimental cases (Chapter 6). In the contexts of development, testing and application of the affine system, this work dealt with the aspects of starting value estimation (see section 7.1.1.), bundle adjustment algorithm (see section 7.1.2.) and application of the algorithm (see section 7.1.3.).

7.1.1. Starting value estimation

The problem of starting value generation has been recovered from a combination of implicit and explicit approaches to the affine sensor. Initialization is implemented on the assumption that a perspective sensor with a very long focal length is a close approximation to the affine sensor. Given this hypothesis the process starts from a set of well defined control points in the object space where an initial exterior orientation updated by a resection routine is applied to initialize the 3D orientation angles. Subsequently closed form back-substitution is applied on the affine model to estimate the 2D projective translations of the photo locations. In cases where there is insufficient knowledge of 3D point geometry in the form of control and inadequate tie points coordinated through pre-measurement (e.g. from prior robust perspective bundle adjustment), there is the option to run a least squares forward intersection procedure that estimates 3D points from the affine sensor model. This is performed with the initialized orientation estimates and in the absence of any internal geometric distortions (e.g. for calibrated cameras).

7.1.2. Bundle adjustment

The affine bundle adjustment approach is novel to this research work and has been designed in order to be able to handle a number of significant factors and issues. Firstly the developed framework reads a set of starting data grouped as target, calibration, photo, orientation and image observation data that are used to populate data structures needed for algorithm development. The method is implemented as a multi-view bundle adjustment with a stochastic model initialized from the input weights of image observations.

The algorithm is able to identify the datum for the system with either external or inner constraint methods. In the case of external constraints the stochastic model is initialized from the weights of control point data. Alternatively an inner constraint datum can be chosen; this calculates the centroid of data points from the identified control and ascribes a normalized precision to the point data that contribute in the image network. The system is also capable of handling variable target and photo occlusions. This is achieved by counting the number of valid image observations, control and tie points as well as photos in the network. Sensor scale factor is handled for optional inclusion or exclusion within the estimation procedure. In consideration of internal sensor distortions, a simplified radial lens distortion model constitutes the calibration model.

The system is run as an iterative least squares observation procedure; it converges when appropriate empirical criteria are satisfied (tested over the orientation, target and calibration parameters). A successful solution provides a statistical analysis of the method. Outputs include the initial and estimated parameters together with the associated quality measures as extracted from the a posteriori covariance analysis. The developed system has been evaluated at each stage of its development. As a result, both algorithm considerations (stochastic model, datum, visibility, data points, parameter, scale factor and calibration model) as well as aspects of geometry (sensor scale invariance, internal calibration and 3D object point recovery) have been addressed.

7.1.3. Application of the algorithm

Practical assessment of the method in the aspect of performance evaluation and modelling has been carried out through testing with real datasets. Specifically, twelve image network datasets have been acquired with both perspective and affine sensors. The employed perspective camera systems have been utilized for the goals of initialization, pre-measurement and generation of reference measurements; the datasets have been processed with an established robust bundle adjustment tool. Affine image networks have been acquired with an off the shelf machine vision system that utilizes a telecentric MVO® TMLTM/0.16x lens attached on two

different C mount (Kodak Megaplus ES1.0 and Sony DFW-SX900) camera bodies that have been available for testing.

The developed approach firstly tested model behaviour and evaluated initialization, overall bundle adjustment results, convergence behaviour and consistency of correlation checks with statistical quality measures (see section 6.4.). At a subsequent stage, it addressed object space behaviour based on reference image networks that provide initial 3D coordinates (see section 6.5.). Parallel projection image networks of varying geometry are processed to derive comparative bundle adjustment results which are evaluated with typical statistical indicators as well as correlation analysis in 3D object space. The problem of sensor scale evaluation and its invariance is treated employing the two different camera systems which are again initialized, tested and assessed with typical measures in image and object spaces with additional accuracy checks (see section 6.6.). Finally, the affine multi-view approach is assessed over conventional robust bundle adjustments, initialized and consequently processed independently of the method presented here (see section 6.7.). In every test case, assessment has been extracted with specified quality and accuracy metrics. The tests have proved that the method can produce sub-pixel measurement precisions which are comparable to established photogrammetric methods (see section 7.2.).

7.2. Conclusions

This section outlines the central conclusions of this research that explicitly satisfy the prior objectives of this research (see section 1.4.). Therefore this section addresses modelling (research objective '2'), object space and sensor scale (research objective '3') and independent assessment (research objective '4') considerations. The central conclusions of this research are formulated as follows:

7.2.1. Modelling analysis

The performance of the bundle adjustment algorithm has been assessed in relation to model, convergence as well as correlation behaviour. To obtain representative results tests have been run as comparative bundle adjustments with different calibration (inclusion and exclusion of radial distortion model), point data (control and tie) and datum (external and inner constraints) implementation (see section 6.4.).

1. Results prove that bundle adjustments converge rapidly (2-3 iterations) with stability offering sub-pixel image measurement precisions of the order of $1/10^{th}$ of a pixel. Visual inspection of the residuals has demonstrated random patterns; these together with their normal distribution suggest that the functional model is correct and that no systematic effects are present within the system.

2. The sensor is calibrated effectively with the inclusion of the implemented internal calibration model ($k_3 = 1.3048 \times 10^{-4}$, $\sigma_{\kappa 3} = 6.0521 \times 10^{-6}$ with an external datum solution with CP and TP data) and estimates a positive radial lens distortion (dr= 8.3μm at a maximum radial distance r=4mm for the Kodak Megaplus ES1.0 sensor). Encouragingly the radial lens distortion curves are consistent for both external and inner datum cases. Moreover, exclusion of the calibration model from the estimation reduces 3D point data precision (at the order of a few microns). This result is also valid for conventional perspective bundle methods provided that the functional model compensates systematic effects through additional parameter handling and no over-parameterization occurs; that is the additional parameters represent the physical reality.

3. The affine adjustment method successfully estimates 3D target coordinates and their quality. Results have demonstrated that the achieved 3D target precisions are of the order of ~5-7μm in the case of external constraints and ~126-283μm in the case of inner constraints. These figures are scaled to the a posteriori standard deviation (σ_o= ~2). Moreover, it has been repeatedly proven that even when employing strong convergent and redundant image networks, target coordinate precision is reduced by a magnitude of ~2 units in the viewing direction for both datum methods. Particularly reduction of 3D point precision in the case of inner constraints highlights potential 'correlation' between object scale and invariant sensor scale.

4. Accuracy assessment is naturally performed with comparisons over independent measurements generated from high order quality (resolution and precision) systems (photogrammetric cameras, laser scanners or CMMs). Accuracy checks have been shown a general agreement in geometry (point data) when these are compared with manual key measurements (acquired with a digital calliper). However, these results can not be considered as decisive

and further tests using full independent coordination of targets would be required to provide a rigid outcome.

5. The tests examined within this thesis have been acquired under controlled laboratory conditions; that is using systematic rotation and translation imaging stages to acquire regular geometric networks of images. In this regard difficult geometries or missing data have not been a part of testing. Investigation of the method's convergence has shown that successive iteration pairs result in high agreement for the estimated parameters. Some random extreme variations have been particularly observed in the case of 3D photo orientations. In fact, this check has demonstrated that the inner constraints method degrades the quality of the model for the datum variant parameters (2D projective translations, 3D photo rotations, 3D target positions and sensor scale factor) but not for the datum invariant parameter (radial lens distortion term).

6. Analysis of parameter correlations has shown that both datum methods present high order correlations between parameters of the same kind. In particular it has been highlighted (see section 6.4.2.3.) that the external datum presents distinctive correlation patterns between 2D projective translations and 3D orientation angles whereas inner constraints are dominated by minimal correlation between 3D targets and 3D photo rotations.

7.2.2. Object space analysis

To assess object space recovery from the affine bundle adjustment the method was tested with three datasets (of different object geometry) that were initialized from reference measurement data (perspective bundle adjustments). The method was run with an external constraint datum and the calibration lens model. Besides the typical statistical indicators, assessment was obtained utilizing two measures that illustrate: (a) 3D point coordinates correlation coefficients (in X, Y and Z directions) over the inter-target proximities as well as (b) 3D error ellipsoid patterns (see section 6.5.).

1. Results show that the bundle adjustments in this test converged rapidly (2-3 iterations) with an a posteriori standard daviation ranging between $\sigma_o = \sim 2.0$ and 3.0. Considering the difference in input image quality between the different datasets B2, D2 and E2 image point diameters range between 17

pixels, 18 pixels and 36 pixels for each case; the a priori observation precision was set to 0.5μm for 'B2' and 'D2' whilst 1.5μm was used for 'E2'. The RMS image misclosures were estimated at $1/9^{th}$, $1/12^{th}$ and $1/3^{rd}$ of a pixel whilst 3D points have been coordinated with precisions of 8.30μm, 29.14μm and 39.07μm again per test case. As a result, RMS image misclosure is at a sub-pixel level whereas object points are estimated with precisions in the order of a few microns.

2. Evaluation of the absolute values of the 3D correlation coefficients in each direction RX, RY and RZ against 3D inter-target separations has shown repeatability within the different datasets; that is 3D correlation coefficients increase inversely with target-distance. Particularly dataset B2 (pyramid structure) presents a smoother distribution in Z in comparison to X and Y directions. Moreover dataset E2 (centroid structure) has shown correlation coefficient values of reduced magnitude in Y as opposed to correlation coefficients in X and Z directions that present similarities in magnitude and which are separated with inter target distances within the range 0-40mm. In addition it has been observed that small clusters of targets tend to increase in correlation when their inter target distance is increased but their magnitudes are insignificant and hence have not been investigated further.

3. Visualization of the error ellipsoids (described by their magnitude, shape and orientation) show the achieved precision levels of the affine bundle adjustment. These have indicated the influence of the input target image quality, image network geometry (strength, visibility and redundancy) as well as 3D control point quality for the external constraints datum. The homogeneous nature of the observed ellipsoids builds the confidence (together with the random pattern of the residual vectors) that the functional model is correct. An additional bundle adjustment check with an inner constraints datum (dataset D2) resulting in a uniform error ellipsoids shape, has proved that the image network is highly homogeneous and strong. As a consequence, the observed ellipsoid patterns in the external datum case are not related to potential orientation weaknesses but to potentially poor control determination.

7.2.3. Sensor scale analysis

To assess the effect of scale within the system, the bundle method has been tested with two different camera systems (telecentric lens MVO® TMLTM/0.16x attached to a Kodak Megaplus ES1.0 as well as a Sony DFW-SX900). The utilized test object occupies a 3D volume larger than the sensor footprint. Initialization has been implemented from prior reference photogrammetric measurements. The test image networks were processed with pre-defined control data (to maintain the 3D coordinate datum definition) and associated 3D precisions set to 25μm. Assessment has been derived utilizing the bundle statistical indicators and recovery of object scale has been evaluated with calculations of 3D absolute discrepancies (see section 6.6.).

1. Results have shown that the bundle adjustments converge after 3 iterations with an a posteriori σ_o= ~2.0 - 4.0, a triangulation misclosure of $1/10^{th}$ - $1/9^{th}$ of a pixel and a 3D point precision up to 30μm. The calibrated image scale has been recovered to 0.1615 (with a precision of σ_S= 7.53×10^{-5}) and 0.1659 (with a precision of σ_S= 2.35×10^{-5}) for each camera system which show the difference in scale recovery when using the two Kodak and Sony sensors.

2. Object scale has been evaluated with calculation of the absolute differences (between externally measured and derived from the bundle computation distances) which reach the order of 130μm. Most significantly the overall mean 3D absolute discrepancy (between reference datum measurements and estimated 3D point coordinates) has proved to be equal to 10.7μm (Kodak camera system) and 18.2μm (Sony camera system). The small magnitude of these discrepancies can be attributed to image network geometric strength.

7.2.4. Independent evaluation

Final experimental case comprises testing of the bundle adjustment algorithm in comparison to a perspective bundle adjustment within the software tool VMS 8.0. Initialization has been recovered through pre-measurement at a prior stage of testing whereas processing has been implemented with identical control and precisions (σ_{XYZ} = 25μm). Here, assessment has been performed utilizing two measures, the first visualizes the 3D point error ellipsoids and the second the 3D absolute discrepancies for the common control data (see section 6.7.).

1. Results have demonstrated a triangulation misclosure of $1/12^{th}$ of a pixel on convergence in both network cases. In fact a posteriori standard deviation and 3D points are estimated as $\sigma_o = 1.0$ and $\sigma_{XYZ} = 9.44\mu m$ (perspective dataset) and $\sigma_o = 2.2$ and $\sigma_{XYZ} = 20.09\mu m$ (affine dataset) which provide a first evidence that the data given their input quality estimate 3D point coordinates with a good agreement.

2. The visualized error ellipsoids have shown that their patterns present similar shape and magnitude for the control located at each concentric ring whereas the error ellipsoids of tie points show the influence of the a posteriori 3D precision.

3. Regarding accuracy evaluation, the mean absolute 3D discrepancies for the control data have resulted in an agreement of ~6μm (6.27μm for the perspective dataset and 6.42μm for the affine dataset) against the initially identified reference photogrammetric measurements. Consequently the affine bundle adjustment algorithm has shown to provide results of similar quality with the well known and well understood perspective case under controlled test conditions with significant repeatability within the various tests reported within the subsequent test cases.

7.2.5. Critical assessment

Following the concluding remarks presented above, a critical evaluation of the principal defects that characterize the developed approach is now demonstrated.

- **Use of target measurements:** The current framework is initialized from discrete point features of high contrast (retro-reflective targets and white markers on a black background). In practical terms natural textured objects could be employed but such test objects would not insert any additional information for test purposes. In addition, image measurements have been initiated within the in-house bundle adjustment software externally to the developed framework. A key limitation to this process is that the affine images require significant manual digitization and processing given the sensor magnification factor in combination with the limited capabilities of object targeting. As a result, there is a significant data pre-processing cost which

makes the method inefficient. A more sophisticated, automated initialization procedure based on automatic feature extraction would be required in an ideal case.

- **Orientation value estimation:** In the aspect of modelling difficult and non-controlled image network configurations, the current initialization approach can be considered as suitable for network geometries utilized within this thesis. The system is built upon geometric assumptions that relate to the 3D orientation angles recovery through perspective as well as pre-knowledge of 3D control. A complete exploitation of the affine sensor model for direct initialization considering degenerate cases has not been considered.

- **Tie point initialization:** A more dedicated 3D intersection procedure regarding algorithm implementation would be required in an ideal case. The current method coordinates targets in 3D through forward least squares estimation directly from affine projection provided the sensor is externally oriented and internally calibrated.

- **Inner constraints datum**: Inner constraints significantly reduce system precision as a result of a unified object space scale. Inherent lack of scale variation within the affine image when compared to the perspective case has resulted in a weaker network solution. Thus, an issue of alternatively modelling scale in the case of inner constraints remains open.

- **System calibration:** The implemented calibration model has been proved to calibrate the sensor successfully (at the level of ~1/10 of a pixel); an additional issue is whether a different affine system would demand an augmented additional parameter estimation model.

- **Outlier detection:** The current least squares approach minimizes the cost function in the feature re-projection error and provides the statistical quality of results. Yet an open issue remains due to the fact that the method is prone to blunders. This inability is a key problem as in real world measurement tasks it is expected that common users do not have absolute control over the implemented data unless an expertise operator is assumed. Outlier detection and automation are essential parts of algorithm design as they enhance system reliability and practical usability; therefore an outlier detection method could be adopted within the affine method.

7.3. Future research

This research has presented an affine multi-view approach in the context of close range object measurement. This comprises a joint calibration, orientation and measurement problem and as such it can be extended to further processes that on their core implementation rely on the 2D to 3D geometric recovery. In the scope of areas of application, the presented system can be applied to any close range measurement task (such as in industrial metrology, archaeology and medicine) that aims to measure or recover fine object detail. Suggestions for future research are reported here in terms of optimization of the current algorithm that would for instance increase or improve modelling as well as extend the current framework in the context of method application.

7.3.1. Method extension

- **Initialization:** The developed method can be extended further in order to optimize starting value estimation from pure affine projection. This includes estimation of orientation angles through space resection as well as investigation of minimum conditions and critical geometric configurations for the absolute purpose of processing parallel projection images.

- **3D point estimation:** Whilst the current forward intersection method is another form of least squares multi-view solution suitable for calibrated cameras it would be of great benefit if 3D point estimation would have been treated by a refined solution. Such an approach would be for example to perform a geometric closed form solution that minimizes the ray intersections as a first stage updated subsequently by a multi-view solution that minimizes the reprojection error in image space to calculate 3D point positions. These problems are considered as solved for the fundamental perspective case and thus a more sophisticated solution can be similarly applied for the affine case.

- **Outlier detection:** An outlier detection and elimination tool that applies system self-diagnosis would be desirable. Such robust approaches are typically implemented utilizing statistical testing or a down-weighting scheme. As an example, statistical tests are applied to check the a posteriori sigma nought and

re-estimation stops when all contaminated measurements have been removed from the system and the a posteriori sigma nought becomes unity.

- **Sensor calibration:** The internal calibration model could subject to alternative system requirements (different affine sensor) demand a more sophisticated treatment. An obvious solution to this issue would be to embed an additional parameter set (radial lens, tangential lens distortions as well as affinity and orthogonality terms), statistically checked and accepted or rejected within the system.

These considerations are critical in that they would open the method in the aspects of implementation, automation regarding robustness and efficiency as well as practical usability.

7.3.2. Application extension

Following the results presented in section 7.2. sub-pixel recovery of affine projection image sequences is possible to a level appropriate for applications such as texture mapping, sensor fusion or feature automation.

- **Texture mapping:** In physical terms, the sensor by virtue of its optimal parallel projection presents minimal internal geometric distortion and a more consistent image sampling over an object surface when compared to a perspective imaging system. This combination can offer the potential to optimize the image quality of discrete signalized point targets. As a result, texture mapped models from real affine image sequences can be possible provided that correct multi-view registration and surface approximation have already been recovered. This could be of specific interest to the measurement and visualization of fine object detail where typical perspective-based modelling and texturing procedures can be ineffective. Standard algorithms can fail to capture very fine object details or can be error-prone to significant sensor distortions.

- **Sensor fusion:** The method could be extended to generate a hybrid system that will integrate perspective and affine sensors in a joint adjustment. Although such a solution has been crudely investigated within the course of this research in experimental terms, this approach has not been exploited

further in algorithmic and practical terms. An effective solution could potentially introduce geometric constraints between photo orientations provided that both sensors view the same 3D volume. As a consequence a bundle adjustment would be processed similarly to methods that process convergent stereo-pairs.

- **System automation:** A final promising application of an affine bundle adjustment would be to integrate automatic feature extraction within the system. In particular, given the inherent invariant sensor scale, it would be beneficial to further investigate the method in the aspect of automatic feature or line extraction directly from real affine imagery. This approach could possibly enhance automation given the minimal geometric distortion and thus optimize the automatic correspondence between different image frames which on success would enter the multi-view modelling algorithm presented here.

7.4. A final point

At the outset of this research it was unknown whether affine projection images would offer sufficient content and allow the establishment of a new close range approach. This research has successfully shown that the affine sensor offers a good alternative to the well-understood and well-established perspective sensor for the purpose of close range object measurement. It is possible to utilize, develop and implement more theoretical camera models that are closer to metric reconstruction and often adopted from the computer vision community. This thesis contributes in the development of a new affine bundle adjustment system applied to convergent imaging networks. Current state of the art is focused on fast algorithms as well as automation and implementation of advanced intensity and range imaging systems. Thus, such topics could be potentially embedded and applied in the case of processing affine imagery. The research community focuses on system automation; it is however expected that sensor modelling and 3D measurement will still be very active topics of research in the field of photogrammetry.

References

Abdel-Aziz, Y. I. & Karara, H. M. (1971). Direct linear transformation from comparator coordinates into object space coordinates in close range photogrammetry. ASP/UI Symposium on close-range photogrammetry, pp. 1-18.

Agouris, P., Doucette, P. & Stefanidis, A. (2004). Chapter 12: Automation and digital photogrammetric workstations. In: J. C. McGlone (Ed.), Manual of photogrammetry, V Edn., ASPRS, pp. 949-981.

Ahn, S. C., Lee, S., Meyyapan, A. & Schenker, P. (1997). Experiments on depth from magnification and blurring. Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 733-739.

Baarda, W. (1968). A testing procedure for use in geodetic networks. Netherlands geodetic commission. Publications on geodesy, 2 (5), 97 pages.

Baltsavias, E. P. (1991). Multiphoto geometrically constrained matching. PhD thesis, Institut für Geodäsie und Photogrammetrie (ETHZ), 221 pages.

Besl, P. J. & McKay, H. D. (1992). A method for registration of 3D shapes. Pattern Analysis and Machine Intelligence, IEEE Transactions, 14 (2), pp. 239-256.

Betham F., Herd H., Luhmann T. & Ohm J. (2009). Free form surface measurement with image sequences under consideration of disturbing objects. In: Gruen & Kahmen (Eds.), Optical 3-D Measurement Techniques IX, pp. 51-61.

Blais, F. (2004). Review of 20 years of range sensor development. Journal of Electronic Imaging 13 (1), pp. 231-243.

Blanc, N. (2001). CCD versus CMOS - has CCD imaging come to en end?. In: D. Fritsch & R. Spiller (Eds.), Photogrammetric Week 2001, pp. 131-137.

Bopp, H. & Krauss, H. (1978). Extension of the 11-parameter solution for on-the-job calibrations of non-metric cameras. International Archives of Photogrammetry 22 (5), pp. 7-11.

Bouguet, J.-Y. (2009). Camera calibration toolbox for matlab. http://www.vision.caltech.edu/bouguetj/calib_doc/ (accessed October 2009).

Brown, D. C. (1971). Close range camera calibration, Photogrammetric Engineering, and Remote Sensing, pp. 855-566.

Brown, D. C. (1974). Evolution, application and potential of the bundle method of photogrammetric triangulation. International symposium of Stuttgart, Commission III, International Society for Photogrammetry, 95 pages.

Butler, M. (2003). Robots and vision 2003: CMOS imaging and application in machine vision. DALSA, 30 pages.

Clarke, T. A. (1994). An analysis of the properties of targets used in digital close range photogrammetric measurement. Videometrics III. SPIE Vol. 2350, pp. 251-262.

Clarke, T. A. & Fryer, J. G. (1998). The development of camera calibration methods and models. The Photogrammetric Record, Vol. 16 Issue 91, pp. 51-66.

Cooper, M. (1987). Chapter 5: Principles of least squares estimation. In: Control surveys in civil engineering. Collins, pp. 159-193.

Cooper, M. & Cross, P. (1988). Statistical concepts and their application in photogrammetry and surveying. The Photogrammetric Record, Vol. 12. Issue 71, pp. 637-663.

Cooper, M. & Cross, P. (1991). Statistical concepts and their application in photogrammetry and surveying (continued). The Photogrammetric Record, Vol. 13. Issue 77, pp. 645-678.

Cooper, M. & Robson, S. (1996). Chapter 2: Theory of close range photogrammetry. In: Close range photogrammetry and machine vision, Whittles Publishing, pp. 9-51.

Dermanis, A. (1990). Adjustment of observations and theory of estimation. Vol. 2 Volumes, Ziti. Thessaloniki Greece, In greek, 311 pages.

Dermanis, A. (1991). Analytical photogrammetry. Ziti. Thessaloniki Greece. In greek, 429 pages.

Dickscheid, T., Labe, T. & Forstner, W. (2008). Benchmarking automatic bundle adjustment results. In: 21st Congress of the International Society for Photogrammetry and Remote Sensing (ISPRS), Beijing, China, pp. 7-12.

Edmund, O. (2006). Best of Edmund optics application notes (2000 - 2006), Technical report (downloaded March 2006), 109 pages.

Edmund (2006). Producer of optics, imaging and opto-mechanical components technology.
http://www.edmundoptics.com/onlinecatalog/displayproduct.cfm?productID=2407 (accessed July 2009).

El-Hakim, S. F., Brenner, C. & Roth, G. (1998). A multi-sensor approach to creating accurate virtual environments. ISPRS Journal of Photogrammetry and Remote Sensing 53, Vol. 53, Issue 6, pp. 379-391.

El-Hakim, S. F., Beraldin, J. A. & Blais, F. (2003). Critical factors and configurations for practical 3D image-based modelling. 6th Conference on 3D Measurement Techniques. Zurich, Switzerland. Vol. II, pp. 159-167.

El-Hakim, S. F., Li, R., Mostafa, M., tao, C. V., Toth, C., Fraser, C., Chapman, D., Patias, P., Gillen, L., Mugnier, C. J., Lucas, P., Walford, A. & Wu, S. S. C. (2004). Chapter 14: Photogrammetric applications. In: J. C. McGlone (Ed.), Manual of photogrammetry, V Edn., ASPRS, pp. 1015-1103.

Faugeras, O. & Luong, Q.-T. (2001). The geometry of multiple images. The laws that govern the formation of multiple images of a scene and some of their applications, MIT Press, Cambridge, Massachusetts, London, England, 644 pages.

Fischler M. A and Bolles R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM Vol. 24 (6), pp. 381-395. http://doi.acm.org/10.1145/358669.358692.

Foerstner, W. (2002). Computer vision and photogrammetry-mutual questions: Geometry, statistics and cognition. Swedish Society for Photogrammetry and Remote Sensing, pp. 151-164.

Foerstner, W. (2005). Performance evaluation and benchmarking of algorithms or systems for calibration, orientation and surface reconstruction, 32 pages.

Foerstner, W. (2009). Computer vision and remote sensing - lessons learned. In: Dieter Fritsch (Ed.) Photogrammetric Week '09, Wichmann, pp. 241-249.

Foerstner, W. & Gulch, E. (1987). A fast operator for detection and precise location of distinct points, corners and centres of circular features. ISPRS Intercomission Workshop, Interlaken pp. 281-305.

Foerstner, W. & Wrobel, B. (2004). Chapter 2: Mathematical concepts in photogrammetry. In: J. C. McGlone (Ed.), Manual of photogrammetry, V Edn., ASPRS, pp. 15-180.

Fournel, T., Coudert, S., Fournier, C. & Ducottet, C. (2003). Stereoscopic particle image velocimetry using telecentric lenses. Institute of physics historic archives JISC 14 (4), pp. 494-499.

Foveon (2009). Foveon Inc.: Innovator in design, development of image sensors and capture systems.
http://www.foveon.com/ (accessed August 2009).

Fraser C. S., Shortis M. R. and Ganci G. (1995). Multi-sensor system self-calibration. Videometrics IV SPIE, Vol. 2598, pp. 2-98.

Fraser, C.S. (1996). Chapter 9: Network Design. In: K. B. Atkinson (Ed.), Close Range Photogrammetry and Machine Vision. Whittles Caithness, Scotland, pp. 9-51.

Fraser, C. S. (1997). Innovations in automation for vision metrology data. Photogrammetric Record Vol. 15 (Number 90), pp. 901-911.

Fraser, C. S. (2001). Chapter 4: Photogrammetric camera component calibration: A review of analytical techniques. In: Calibration and orientation of cameras in computer vision, Springer - Verlag Berlin / Heidelberg, pp. 95-121.

Fraser, C. S., Shortis, M. R. & Ganci, G. (1995). Multi-sensor system self-calibration. Videometrics IV, SPIE Vol. 2598, pp. 2-18.

Fryer, J. G. (1996). Chapter 6: Camera calibration. In: Close range photogrammetry and machine vision, Whittles Publishing, pp. 156-179.

Geomagic Studio 7.0 (2006). Geomagic Studio 7.0 laser scanning software. http://www.geomagic.com/en/products/studio/index.shtml (accessed December 2009).

Geodetic (2009). Geodetic systems Inc. 3D industrial measurement systems. http://www.geodetic.com/ (accessed August 2009).

Grafarend , E. W., 1974. Optimization of image networks. Bolletino di Geodesia e Science Affini, 33 (4), pp. 351-406.

Graham, R. (1998). Chapter: 5 digital cameras. In: Digital imaging. Whittles Publishing, pp. 77-104.

Grammatikopoulos, L., Kalisperakis, I., Karras, G. & Petsa, E. (2005). Data fusion from multiple sources for the production of orthographic and perspective views with automatic visibility checking, in 'XVIII CIPA Int. Symposium', pp. 819-824.

Granshaw, S. I. (1980). Bundle adjustment methods in engineering photogrammetry. The Photogrammetric Record 10 (56), pp. 181-207.

Gruen, A. (1996). Chapter 4: Development of digital methodology and systems. In: Close range photogrammetry and machine vision, Whittles Publishing, pp. 78-105.

Gruen, A. & Akca, D. (2005). Least squares 3D surface and curve matching. ISPRS Journal of Photogrammetry and Remote Sensing Vol. 59, pp. 151-174.

Gruen, A. & Beyer, H. A. (2001). Chapter 7: System calibration through self-calibration. In: Calibration and orientation of cameras in computer vision, Springer - Verlag Berlin / Heidelberg, pp. 163-193.

Gruen, A., Remondino, F. & Zhang, L. (2003). Image-based automated reconstruction of the great buddha of bamiyan, Afghanistan. Proceedings of SPIE, Videometrics VII Vol. 5013, pp. 129-136.

Gruen, A., Zhang, L. & Visnovcova, J. (2001). Automatic reconstruction and visualization of a complex buddha tower of bayon, angkor, Cambodia. Proceedings 21. Wissenschaftlich-Technische Jahrestagung der DGPF, "Photogrammetrie-Fernerkundung-Geoinformation: Geodaten schaffen Verbindungen", Band 10, Konstanz, pp. 289-301.

Guidi, G., Cioci, A., Atzeni, C. & Beraldin, J.-A. (2003). Accuracy verification and enhancement in 3d modeling: Application to donatello's maddalena. International Conference on 3D Digital Imaging and Modeling, pp. 334-341.

Habib, A. & Morgan, M. (2005). Stability analysis and geometric calibration of off-the-shelf digital cameras. Photogrammetric Engineering and Remote Sensing 71 (6), pp. 733-741.

Haig, C., Heipke, C. & Wiggenhagen, M. (2006). Lens inclination due to instable fixings detected and verified with VDI/VDE 2634 Part 1, 7 pages.

Hartley, R. & Mundy, J. (1993). The relationship between photogrammetry and computer vision. In: Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision, SPIE Vol. 1944, pp. 92-105.

Hartley, R. I. & Zisserman, A. (2004a). Multiple view geometry in computer vision. Second Edn., Cambridge University Press, 646 pages.

Hartley, R. & Zisserman, A. (2004b). Part I: Camera geometry and single view geometry. Chapter 6: Camera models. In: Multiple view geometry in computer vision, 2nd Edn., Cambridge University Press.

Hastedt, H., Luhmann, T. & Tecklenburg, W. (2002). Image-variant interior orientation and sensor modelling of high-quality digital cameras. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences Vol. 34, pp. 27-32.

Heuvel, F. A. van den (2003). Automation in architectural photogrammetry (line-photogrammetry for the reconstructionfrom single and multiple images), PhD thesis, Delft, The Netherlands, 190 pages.

Horn, B. K. P. (1986). Robot vision, MIT Press, Cambridge, MA, USA, 509 pages.

Hrabacek, J. & van den Heuvel, F. A. (2000). Weighted geometric objects constraints integrated in a line - photogrammetric bundle adjustment. In: Proceedings of the XIXth Congress of International Society for Photogrammetry and Remote Sensing (ISPRS), Amsterdam, 8 pages.

Intel (2001). Open source computer vision library. Reference manual, 377 pages.

Intel (2009). Open computer vision library download.
http://sourceforge.net/projects/opencvlibrary/ (accessed October 2009).

ISO10360 (1994). Acceptance and reverication tests for coordinate measuring machines (CMMs).
http://www.leitz-metrology.com/pdf/iso_10360_for_coordinate_measuring_machines_CMMs.pdf (accessed July 2009).

ISPRS (2004). Statutes and bylaws. International Society for Photogrammetry and Remote Sensing, 19 pages.

ISPRS (2009). ISPRS working group III/1 2004-2008.
http://www.commission3.isprs.org/wg1/ (accessed September 2009).

Iwitness (2009). Iwitness and Iwitness Pro. Photogrammetry software systems.
http://www.iwitnessphoto.com/ (accessed October 2009).

King, B. (1995). Bundle adjustment of constained stereopairs - mathematical models. Geomatics Research Australasia (63), pp. 67-92.

Kodak (1996). User's manual. Kodak Megaplus camera. Motion analysis systems.

Koenderink, J. J. (1984). The structure of images. Biological Cybernetics 50 (5), pp. 363-370.
http://dx.doi.org/10.1007/BF00336961.

Koenderink, J. J. & van Doorn, A. J. (1991). Affine structure from motion. Journal of Optical Society of America Vol. 8 (2), pp. 377-385.

Konrath, R. & Schroder, W. (2002). Telecentric lenses for imaging in particle image velocimetry: A new stereoscopic approach. Experiments in fluids 33 (5), pp. 703-708.

Kyle, S. (2004). Using parallel projection mathematics to orient an object relative to a single image. The Photogrammetric Record Vol. 19 (Number 105), pp. 38-50.

Labe, T. & Foerstner, W. (2004). Geometric stability of low-cost digital consumer cameras, In: International Archives of Photogrammetry and Remote Sensing, Vol. XXXV, pp. 528-535.

Lenhardt, K. & Kreuznach, B. (2006). Optical measurement techniques with telecentric lenses.
http://www.schneiderkreuznach.com/index_e.htm (accessed July 2009).

Lenz, R. (1989). Image data acquisition with CCD cameras. Optical 3-D Measurement Techniques, pp. 22-35.

Litwiller, D. (2001). CCD vs. CMOS: Facts and fiction, Reprinted from the January 2001 Issue of PHOTONICS SPECTRA ©, Laurin Publishing Co. Inc., 4 pages.

Litwiller, D. (2005). CMOS vs. CCD: Maturing technologies, maturing markets. DALSA, Reprinted from the August 2005 Issue of PHOTONICS SPECTRA © Laurin Publishing, 5 pages.

Lourakis, M. & Argyros, A. (2005). Is levenberg-marquardt the most efficient optimization algorithm for implementing bundle adjustment?, In: 2005 IEEE International Conference on Computer Vision, (ICCV'05), Vol. 2, Beijing China, pp. 1526-1531.

Lourakis, M. & Argyros, A. (2009). SBA: A generic sparse bundle adjustment C/C++ package based on the levenberg-marquardt algorithm.
 http://www.ics.forth.gr/ lourakis/sba/ (accessed September 2009).

Luhmann, T., Robson, S. & Harley, I. (2006). Close range photogrammetry. Principles, methods and applications. Whittles Publishing, 500 pages.

Maas, H. G. (2008). Chapter 5: Close range photogrammetry sensors. In: Advances in photogrammetry, remote sensing and spatial information sciences. ISPRS congress book, Vol. 7, CRC, Taylor and Francis ISPRS book series, pp. 63-72.

Magill, A. A., (1955). Variation in Distortion with Magnification. Journal of Optical Society of America. Vol. 45, (3), pp. 148-152.

Mason, S. (1995). Expert system-based design of close-range photogrammetric networks. ISPRS Journal of Photogrammetry and Remote Sensing 50 (5), pp. 13-24.

Mathworks (2009). Object oriented programming in Matlab.
http://www.mathworks.co.uk/ (accessed October 2009).

Mayer, H. (2005). Robust least-squares adjustment based orientation and auto-calibration of wide-baseline image sequences. Proceedings of British machine vision conference, pp. 384-393.

McGlone, J. C. (1989). Chapter 4: Analytic data-reduction schemes in non-topographic photogrammetry. In: Non-topographic photogrammetry, second Edn., American Society for Photogrammetry and Remote Sensing, pp. 37-57.

MellesGriot (2006). Machine vision lens fundamentals.
http://www.mellesgriot.com/ (downloaded May 2006), 9 pages.

Mikhail, E. M. (1976). Observations and least squares. IEP- Dun-Donnelley. Harper & Row with contributions by F. Ackermann, 497 pages.

Mikolajczyk, K. & Schmid, C. (2004). Scale and affine invariant interest point detectors. International Journal of Computer Vision Vol. 60, Issue 1, pp. 63-86.

Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T. & Gool, L. V. (2005). A comparison of affine region detectors. International Journal of Computer Vision 65 (1), pp. 43-72.
http://dx.doi.org/10.1007/s11263-005-3848-x.

Mugnier, C. J., Foerstner, W., Wrobel, B., Paderes, F. & Munjy, R. (2004). Chapter 3: The mathematics of photogrammetry. In: Manual of photogrammetry, J. C. McGlone, (Ed.), V Edn., ASPRS, pp. 181-316.

Nikon (2002). The Nikon guide to digital photography with the D100 camera.

Nikon (2009). Nikon imaging products 2002 release for Nikon D100.
http://imaging.nikon.com/products/imaging/technology/archives/digitalcamera/slr/200 2/d100/index.htm (accessed April 2009).

NPL (2001). Modelling for non-contacting CMMs. Project 2.3.1/2/3 - large scale metrology. Ref. MPU 8/6.1.1 - 09121999, Technical report, University College London. National Physical Laboratory Leica UK, 48 pages.

Ono, T., Akamatsu, S. & Hattori, S. (2004). A long range photogrammetric method with orthogonal projection model. In: Proceedings of the International Society for Photogrammetry and Remote Sensing (ISPRS '04), Commission III, Vol. III, Part B3, 6 pages.

Optoengineering (2009). Telecentric lenses brochure. Technical report. Optoengineering.
http://www.opto-engineering.com/telecentric.php (downloaded August 2009).

Otepka, J. (2004). Precision target mensuration in vision metrology, PhD thesis. University of Technology, Vienna, 87 pages.

Parian, J. A. (2007). Sensor modeling, calibration and point positioning with terrestrial panoramic cameras, PhD thesis, Swiss Federal Institute of Technology, Zurich, 121 pages.

Photomodeler (2009). Photomodeler official website. measuring and modeling the real world. Eos systems inc.
http://www.photomodeler.com/index.htm?gclid=CKaJhbPKoZ0CFZ1h4wodbloJ9A (accessed October 2009).

Pollefeys, M., Gool, L. V., Vergauwen, M., Cornelis, K., Verbiest, F. & Tops, J. (2003). 3D capture of archaeology and architecture with a hand-held camera. In Proceedings of the ISPRS Workshop on Vision Techniques for Digital Architectural and Archaeological Archives, Vol. XXXIV, Part 5/W12, The International Archive of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Ancona, Italy, pp. 262-267.

Pollefeys, M., Gool, L. V., Vergauwen, M., Verbiest, F., Cornelis, K., Tops, J. & Koch, R. (2004). Visual modeling with a hand-held camera. International Journal of Computer Vision Vol. 59 (3), pp. 207-232.

Press W. H., Teukolsky S. A., Vetterling W. T. & Flannery B. P. (2005). Numerical Recipes in C++. The Art of scientific computing. Second Edition, Cambridge University Press, 1002 pages.

Ray, S.F., (1988). Chapter 23: Perspective and Projection. In: Applied Photographic Optics. Imaging Systems for Photography, Film and Video. Focal Press.

Remondino, F. (2006). Detectors and descriptors for photogrammetric applications. In: R. S. W. Foerstner (Ed.), Symposium of ISPRS Commission III Photogrammetric Computer Vision PCV 2006, Vol. XXXVI, part 3, ISPRS Commission III - Photogrammetric Computer Vision and Image Analysis, Bonn, Germany, pp. 49-54.

Remondino, F., Pelagoti, A., Mastio, A. D. & Uccheddu, F. (2009). Novel data registration techniques for art diagnostics and 3d heritage visualization. Optical 3-D Measurement Techniques IX 1, pp. 1-10.

Remondino, F. & Zhang, L. (2006). Surface reconstruction algorithms for detailed close-range object modelling, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, ISPRS Commission III Symposium, Bonn, Germany Vol. XXXVI, part 3, pp. 117-123.

Rieke-Zapp, D. H., Tecklenburg, W., Peipe, J., H, H. & Haig, C. (2009). Evaluation of the geometric stability and the accuracy potential of digital cameras - comparing mechanical stabilization versus parameterization. ISPRS Journal of Photogrammetry and Remote Sensing 64, pp. 248-258.

Rieke-Zapp, D.H., Hastedt, H., Tecklenburg, W., Peipe, J. & Luhmann, T. (2008). Performance evaluation of several high-quality digital cameras. International Society for Photogrammetry and Remote Sensing, Commission V papers, Vol. XXXVII, part B5, pp. 7-12.

Robson, S. & Kyle, S. (2004). Lecture notes on engineering metrology, Technical report, Department of Geomatic Engineering, University College London. 327 pages.

Robson, S., Rova, M., Morgan, L., Heuman, J., Heathcote, M., Lees, D., Caner, H. & Mills, C. G. (2008). Measurement of transparent plastic sculptures using photogrammetry, laser scanning and touch probing for conservation purposes. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII (Part B5), pp. 207-212.

Rova, M., Robson, S. & Cooper, M. (2008*a*). Multi-image calibration and measurement of close range objects with parallel projection cameras. In: Proceedings of the Remote Sensing and Photogrammetry society conference "Measuring change in the Earth system". University of Exeter, 4 pages.

Rova, M., Robson, S. & Cooper, M. (2008b). Multistation bundle adjustment with a machine vision parallel camera system: An alternative to the perspective case for the measurement of small objects, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences Vol. XXXVII (Part B5), pp. 45-50.

Rova, M., Robson, S. & Cooper, M. (2009). Analysis of scale invariance within an affine multi-image adjustment algorithm. In: Proceedings of the 9th conference in Optical 3-D Measurement Techniques. 10 pages.

Sakane, S., Ishii, M. & Kakikura, M. (1987). Occlusion avoidance of visual sensors based on a hand-eye action simulator system: Heaven, *Advanced Robotics* 2 (2), pp. 149-165.

Schenk, T. (1999). *Digital photogrammetry*, Vol. I, TerraScience, Ohio State University, USA, 422 pages.

Scotchlite (2009). 3M™ Scotchlite™ Reflective Material.
http://solutions.3m.com/wps/portal/3M/en_US/ScotchliteReflectiveMaterial/Scotchlite/?WT.mc_id=www.3m.com/scotchlite (accessed August 2009).

Seitz, P., Veitze, O. & Spirig, T. (1995). From pixels to answers - recent developments and trends in electronic imaging. International Archives of Photogrammetry and Remote Sensing, Vol. 30, Part 5W1, ISPRS Intercommission Workshop "From pixels to Sequences", 12 pages.

Shapiro, L. S. (1995). Affine analysis of image sequences, PhD Thesis, Cambridge University Press, 210 pages.

Shi, J. & Tomasi, C. (1994). Good features to track. In: Proceedings of the IEEE Society on Computer Vision and Pattern Recognition (CVPR '94), Seattle, pp. 593-600.

Shortis, M., Bellman, C., Robson, S., Johnston, G. & Johnson, G. (2006). Stability of zoom and fixed lenses used with digital SLR cameras. International Archives of Photogrammetry and Remote Sensing ISSN: 1682-1750, 36 (5), pp. 285-290.

Shortis, M. R. & Beyer, H. (1996). Chapter 5: Sensor technology for digital photogrammetry and machine vision, Whittles Publishing, pp. 106-155.

Shortis, M. R., Clarke, T. A. & Short, T. (1994). A comparison of some techniques for the subpixel location of discrete target images, SPIE, Videometrics III 2350, pp. 239-250.

Shortis, M. R., Seager, J. W., Robson, S. & Harvey, E. S. (2003). Automatic recognition of coded targets based on a Hough transform and segment matching, In: Proceedings of SPIE series. Videometrics Conference No7, Vol. 5013, pp. 202-208.

Smith, G. E. (2009). The invention and early history of the CCD, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 607 (1), pp. 1-6.

Sonka, M., Hlavac, V. & Boyle, R. (1999a). Chapter 4: Image pre-processing. In: Image processing, analysis and machine vision, Brooks/Cole Publishing Company, pp. 57-122.

Sonka, M., Hlavac, V. & Boyle, R. (1999b). Chapter 6: Shape representation and description. In: Image processing analysis and machine vision, Brooks/Cole Publishing Company, pp. 228-289.

Sonka, M., Hlavac, V. & Boyle, R. (1999c). Chapter 9: 3D vision, geometry and radiometry. In: Image processing, analysis and machine vision, Brooks/Cole Publishing Company, pp. 441-507.

Sonka, M., Hlavac, V. & Boyle, R. (1999d). Image processing, analysis and machine vision, second Edn, Brooks/Cole Publishing Company, Pacific Grove, 770 pages.

Sony (2001), Technical manual. Sony (Ver. 1.0).

Strecha, C., Gool, L. V. & Fua, P. (2008). A generative model for true orthorectification. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVII. Part B3a, Beijing, China, pp. 303-308.

Tecklenburg, W., Luhmann, T. & Hastedt, H. (2001). Camera modelling with image - variant parameters and finite elements, Gruen, A., Kahmen, H. (Eds.), Optical 3-D Measurement Techniques V, pp. 328-335.

Tomasi, C. & Kanade, T. (1992). Shape and motion from image streams: A factorization method. Full report on the orthographic case, Technical report, Carnegie Mellon University, Pittsburgh, PA, USA, 36 pages.

Triggs, B. (1998). Autocalibration from planar scenes. In: Proceedings of the 5th European Conference on Computer Vision (ECCV '98), Vol. I, Springer Berlin / Heidelberg, London, UK, pp. 89-105.

Triggs, B., McLauchlan, P. F., Hartley, R. I. & Fitzgibbon, A. W. (2000). Bundle adjustment - a modern synthesis. In: Proceedings of the International Workshop on Vision Algorithms: Theory and Practice (ICCV '99). Springer-Verlag, London, UK, pp. 298-372.

Trucco, E. & Verri, A. (1998). Introductory techniques for 3-D computer vision, Prentice Hall PTR, Upper Saddle River, NJ, USA, 343 pages.

Tsai, R. Y. (1987). A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses, IEEE Journal of Robotics and Automation Vol. RA-3(4), pp. 323-344.

VDI/VDE (2009). Vdi guideline: VDI/VDE - handbuch optische technologien.
 http://www.vdi.de/401.0.html?&L=1&tx_vdirili_pi2%5BshowUID%5D=89692 (accessed July 2009).

VMS (2009). Vision measurement system developed by S. Robson (UCL, UK) and M. Shortis (RMIT, Australia).
http://www.geomsoft.com/ (accessed May 2009).

Watanabe, M. & Nayar, S. K. (1997). Telecentric optics for focus analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (12), pp. 1360-1365.

Weinhaus, F. M. & Devich, R. N. (1999). Photogrammetric texture mapping onto planar polygons, In: Graphical models and image processing, Vol. 61, Academic Press, San Diego, CA, ETATS-UNIS (1995-1999) (Revue), pp. 63-83.

Wolfram, M. (2009*a*). 'Geometric centroid definition'.
http://mathworld.wolfram.com/GeometricCentroid.html (accessed October 2009).

Wolfram, M. (2009*b*). Singular value decomposition definition.
http://mathworld.wolfram.com/SingularValueDecomposition.html (accessed October 2009).

Wrobel, B. P. (2001). Chapter 2: Minimum solutions for orientation. In A. Gruen and T.S. Huang (Eds.), Calibration and orientation of cameras in computer vision, Springer - Verlag Berlin / Heidelberg, pp. 7-62.

Xu, G. & Zhang, Z. (1996). Epipolar geometry in stereo, motion and object recognition, Kluwer Academic Publishers, 313 pages.

Zeng, Z. & Wang, X. (1992). A general solution of a closed - form space resection. Photogrammetric Engineering and Remote Sensing 58 (3), pp. 327-338.

Zhang, Z. (1999). Flexible camera calibration by viewing a plane from unknown orientations. In 'ICCV '99', Vol. 1, pp. 666-673.

http://computer.org/proceedings/iccv/0164/vol%201/01640666abs.htm (accessed October 2009).

Zhang, Z. (2009). Microsoft easy camera calibration toolbox.

http://research.microsoft.com/en-us/downloads/7e9de40f-06db-452c-a0f2-4fabb4f20f52/ (accessed October 2009).

Zitova, B. & Flusser, J. (2003). Image registration methods: A survey, Image and Vision Computing 21 (11), pp. 977-1000.

http://dx.doi.org/10.1016/S0262-8856(03)00137-9.

Appendix A: Essential elements for affine model formation

This appendix gives the basic elements needed for the development of the affine sensor model. Particularly the following are listed: Coordinate system definitions (see section A.1.), rotation matrix (see section A.2.), numerical first order derivatives of the affine camera model (see section A.3.), notation for least squares adjustment (see section A.4.) as well as camera model arrangement (see section A.5.).

A.1. Coordinate system definitions

The coordinate systems applied in the image formation process are illustrated as described accordingly (see Figure 1). It is pointed that x,y is the photo - coordinate system (given in mm) and N, M is the associate digital image coordinate system (given in pixels). The digital image coordinate system is linked with the photo-coordinate system through an affine transform (see equations (A.1) and (A.2)).
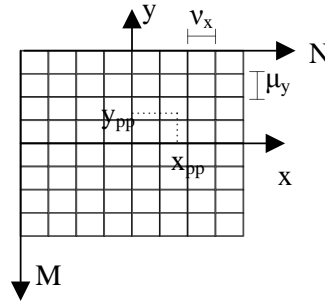
Appendix A - Figure 1: The relation between digital and photo coordinate systems.

$$x = v_x (N - \frac{N_x}{2}) \ \& \ y = -\mu_y (M - \frac{M_y}{2}) \tag{A.1}$$

$$N = \frac{x}{v_x} + \frac{N_x}{2} \ \& \ M = -\frac{y}{\mu_y} + \frac{M_y}{2} \tag{A.2}$$

Where:

x,y= photo-coordinate system (mm)

x_{pp}, y_{pp}= principal point location (mm)

N,M= image-coordinate system (pixels)

Nx, My= image size (pixels)

v_x, μ_y= pixel size in x and y (mm)

For the purpose of the affine-sensor model development, image measurements were initialized externally to the developed method within the software tool VMS 8.0. The structure of the .obs file within the system is given as: 'Photo id, Target id, x coordinate (mm), y coordinate (mm), standard deviation in x (mm), standard deviation in y (mm), residual in x (mm), residual in y (mm) and measurement flag (0 or -1)' (see Figure 2).

```
##################################################################################
# VMS Project:  Photo data output
# Sunday, March 22, 2009
#
# Project name: legonikon
#
# 2749  Photo Obs
#  Photo  Target    X(mm)    Y(mm)    sdx  sdy   res x   res y
#
#
# Used image measurements - at last adjustment
     1026     2011  2.05289  1.73421  2.2064  2.2064  -0.4873 -0.1646    0
     1025     1000  3.03020  2.51099  2.2064  2.2064  -0.2899  0.7698    0
     1025     1019  2.97068  1.91054  2.2064  2.2064  -0.0492  0.4871    0
     1049    20049 -0.72281 -1.44884  0.2167  0.2167   0.0159 -0.8476    0
     1033      400  6.81271  1.32663  0.4983  0.4983  -5.5208 -1.1356   -1
#
# Unused and rejected image measurements - at last adjustment
#
# End of file
##################################################################################
```
Appendix A - Figure 2: Sample of .obs file in VMS 8.0.

As a result, within this work system implementation (computations and modelling) was treated directly on the photo-coordinate system (defined in mm) without the requirement for any system-image transformation.

A.2. Rotation matrix

The 3*3 rotation matrix that establishes the relationship between image and object space coordinate systems with its elements defined as trigonometric functions of the three rotations angles omega (ω), phi (φ) and kappa (κ) is given as follows.

$$R = \begin{bmatrix} \cos\varphi\cos\kappa & \sin\omega\sin\varphi\cos\kappa + \cos\omega\sin\kappa & -\cos\omega\sin\varphi\cos\kappa + \sin\omega\sin\kappa \\ -\cos\varphi\sin\kappa & -\sin\omega\sin\varphi\sin\kappa + \cos\omega\cos\kappa & \cos\omega\sin\varphi\sin\kappa + \sin\omega\cos\kappa \\ \sin\varphi & -\sin\omega\cos\varphi & \cos\omega\cos\varphi \end{bmatrix} \qquad (A.3)$$

A.3. Affine camera model: numerical first order derivatives

The partial derivatives of the simplified collinearity condition function f(x) with respect to the model variables (tx, ty, ω, φ, κ, X, Y, Z, s, k_3) are calculated according to the following equations A.4 - A.23.

$$\frac{\partial fx}{\partial tx} = 1 \qquad (A.4)$$

$$\frac{\partial fx}{\partial ty} = 0 \qquad (A.5)$$

$$\frac{\partial fy}{\partial tx} = 0 \qquad (A.6)$$

$$\frac{\partial fy}{\partial ty} = 1 \qquad (A.7)$$

$$\frac{\partial fx}{\partial t\omega} = s*(\cos\omega*\sin\varphi*\cos\kappa - \sin\omega*\sin\kappa)*Y + \qquad (A.8)$$
$$+ s*(\sin\omega*\sin\varphi*\cos\kappa + \cos\omega*\sin k)*Z$$

$$\frac{\partial fx}{\partial t\varphi} = s*(-\sin\varphi*\cos\kappa)*X + s*(\sin\omega*\cos\varphi*\cos\kappa)*Y + \qquad (A.9)$$
$$+ s*(-\cos\omega*\cos\varphi*\cos\kappa)*Z$$

$$\frac{\partial fx}{\partial t\kappa} = s*(-\cos\varphi*\sin\kappa)*X + s*(-\sin\omega*\sin\varphi*\sin\kappa + \cos\omega*\cos\kappa)*Y + \qquad (A.10)$$
$$+ s*(\cos\omega*\sin\varphi*\sin\kappa + \sin\omega*\cos\kappa)*Z$$

$$\frac{\partial fy}{\partial t\omega} = s*(-\cos\omega*\sin\varphi*\sin\kappa - \sin\omega*\cos\kappa)*Y + \qquad (A.11)$$
$$+ s*(-\sin\omega*\sin\varphi*\sin\kappa + \cos\omega*\cos\kappa)*Y$$

$$\frac{\partial fy}{\partial t\varphi} = s*(\sin\varphi*\sin\kappa)*X + s*(-\sin\omega*\cos\varphi*\sin\kappa)*Y + \qquad (A.12)$$
$$+ s*(\cos\omega*\cos\varphi*\sin\kappa)*Z$$

$$\frac{\partial fy}{\partial t\kappa} = s*(-\cos\varphi*\cos\kappa)*X + s*(-\sin\omega*\sin\varphi*\cos\kappa - \cos\omega*\sin\kappa)*Y +$$
$$+s*(\cos\omega*\sin\varphi*\cos\kappa - \sin\omega*\sin\kappa)*Z$$

(A.13)

$$\frac{\partial fx}{\partial X} = s*(\cos\varphi*\cos\kappa)$$

(A.14)

$$\frac{\partial fx}{\partial Y} = s*(\sin\omega*\sin\varphi*\cos\kappa + \cos\omega*\sin\kappa)$$

(A.15)

$$\frac{\partial fx}{\partial Z} = s*(-\cos\omega*\sin\varphi*\cos\kappa + \sin\omega*\sin\kappa)$$

(A.16)

$$\frac{\partial fy}{\partial X} = s*(\cos\varphi*\sin\kappa)$$

(A.17)

$$\frac{\partial fy}{\partial Y} = s*(-\sin\omega*\sin\varphi*\sin\kappa + \cos\omega*\cos\kappa)$$

(A.18)

$$\frac{\partial fy}{\partial Z} = s*(\cos\omega*\sin\varphi*\sin\kappa + \sin\omega*\cos\kappa)$$

(A.19)

$$\frac{\partial fx}{\partial s} = (\cos\varphi*\cos\kappa)*X + (\sin\omega*\sin\varphi*\cos\kappa + \cos\omega*\sin\kappa)*Y +$$
$$+(-\cos\omega*\sin\varphi*\cos\kappa + \sin\omega*\sin\kappa)*Y$$

(A.20)

$$\frac{\partial fy}{\partial s} = (-\cos\varphi*\sin\kappa)*X + (-\sin\omega*\sin\varphi*\sin\kappa + \cos\omega*\cos\kappa)*Y +$$
$$+(\cos\omega*\sin\varphi*\sin\kappa + \sin\omega*\cos\kappa)*Y$$

(A.21)

$$\frac{\partial fx}{\partial k_3} = (x^2 + y^2)*x$$

(A.22)

$$\frac{\partial fy}{\partial k_3} = (x^2 + y^2)*y$$

(A.23)

A.4. Notation for least squares adjustment

The notation of symbols utilized within the least squares adjustment procedure is given below.

Notation of symbols used within the least squares adjustment	
$l =$	Vector of observation variables
$\bar{l} =$	True values of the vector of observation variables
$\hat{l} =$	Vector of least squares estimate of observation variables $(\bar{l})$
$v =$	True values of the vector residuals
$\hat{v} =$	Least squares estimate of v
$\bar{x} =$	True values of the vector of parameters; mean value
$x^o =$	Approximate values of $\bar{x}$
$x =$	True values of the corrections to x^o, $x = \bar{x} - x^o$
$\hat{x} =$	Least squares estimate of x
$n =$	Number of observations
$m =$	Number of parameters
$A =$	Design or Jacobian matrix (n rows, m columns)
$W =$	Weight matrix (n rows, n columns)
$C =$	Covariance matrix (m rows, m columns)
$N =$	Normal equations matrix (m rows, m columns)

Appendix A - Table 1: Notation of symbols utilized within the least squares adjustment.

A.5. Camera model arrangement

Considering a test set of j=10 views, i=10 point targets and k=1 camera the design matrix is populated with the partial derivatives given above. It is noted that the exemplary design matrix has been structured for the inner constraints datum case; hence it is augmented with seven additional datum rows (see section 4.5.1.).
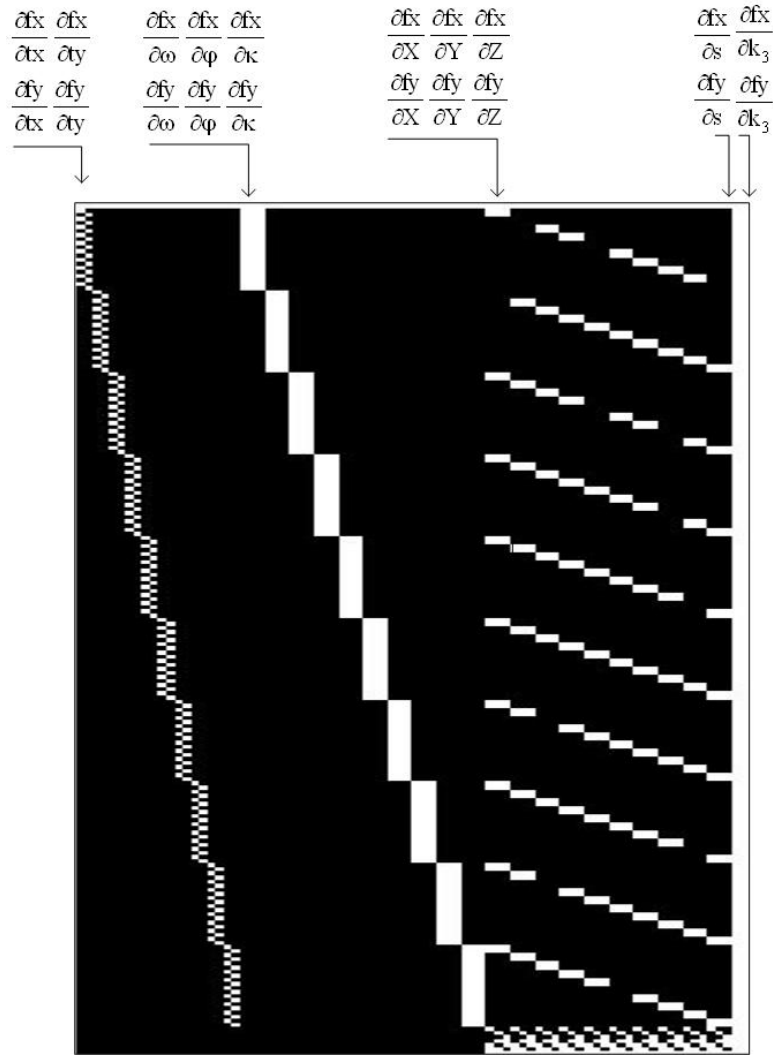
Appendix A - Figure 3: Design matrix arrangement for a sample of j=10 views, i= 10 point data and k= 1cameras.

Appendix B: Processing framework

This appendix describes the input and output data files of the implemented processes. Data measurement and initialization have been treated within the VMS 8.0 software tool (see sections B.1. and B.2.) whilst affine multi-view modelling has been implemented within the developed framework (see section B.3.).

B.1. VMS 8.0 software settings

VMS 8.0 software for network adjustment requires that the user adjusts three types of settings given here:

- Image measurement parameters: For 2D location (measurement) of point-based target data.
- Adjustment settings: For data handling within the initialization procedures (resection and intersection).
- Network adjustment settings: For data handling within the network adjustment procedure.

- Image measurement parameters:

Appendix B - Figure 1: Image measurement parameters menu in VMS 8.0.

- Adjustment settings:
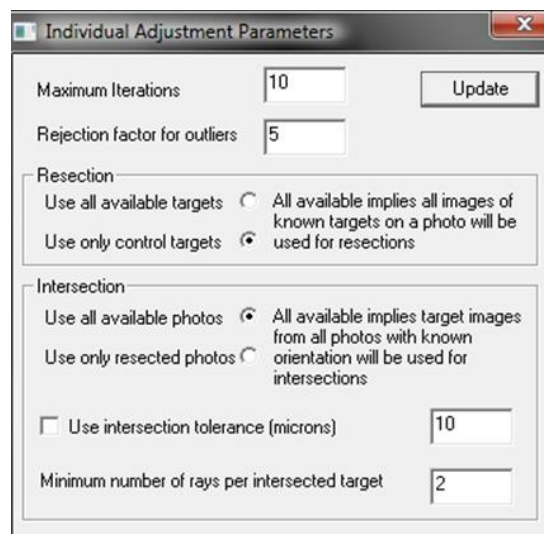
Appendix B - Figure 2: Adjustment parameters menu in VMS 8.0.

- Network adjustment settings:
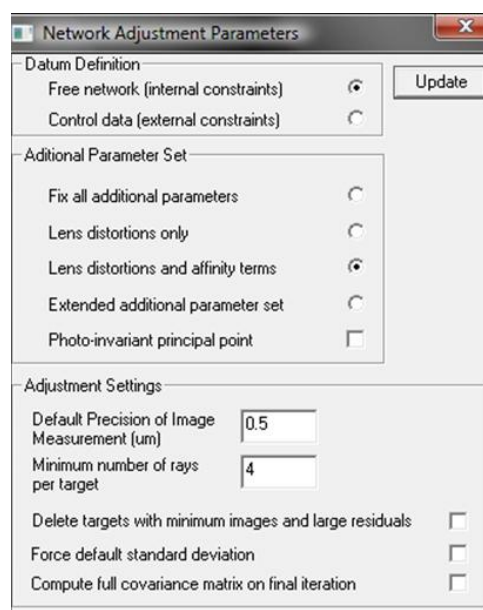
Appendix B - Figure 3: Network adjustment parameters menu in VMS 8.0.

B.2. VMS 8.0 software data files

The VMS 8.0 software's data files are given below.

B.2.1. Input VMS 8.0 data files

The input data files of the perspective sensor model follow the format identified within the VMS 8.0 software. Example data files are given below:

- Target file (.tar): Contains the 3D targets co-ordinates.
- Calibration file (.cal): Contains the internal camera geometry (interior orientation).
- Photo file (.pho): Contains the cameras positions and orientations (exterior orientation).
- Observations file (.obs): Contains the 2D image observations.

TARGET FILE INPUT FORMAT (.tar):
```
################################################################################
# VMS Project:  Target data output
# Tuesday, March 24, 2009
#
# Project name: legonikon
#
# 178  Targets
# Targets in millimetres, precisions in 1000*millimetres
#
#    Target      X       Y       Z  flag    sdx     sdy     sdz
#
    20045   65.5735   75.5262   23.4447 0    3.8866    4.6782    6.4878
      200    2.4575  100.0290    3.3735 7   10.7861    5.8654   10.1860
       …        …        …        …     …      …         …         …
      107   -0.5143    2.7531   96.7073 7   13.5413   19.2530   11.4413
#
# End of file
################################################################################
```

CALIBRATION FILE INPUT FORMAT (.cal):
```
################################################################################
# VMS Project:  Camera data
# Tuesday, March 24, 2009
#
# Project name: legonikon
#
# This file contains information on 1 cameras
# Parameters : 1=PPx,2=PPy,3=PD,4-6=radial,7-8=decentring,9=orthogonality,10=affinity
 1
#
# Calibration parameters for camera 1
```

```
  1     0.1137       0.0020
  2     0.0404       0.0073
  3    29.5784       0.0048
  4   -1.8822e-004   2.6425e-006
  5    2.2259e-006   1.1925e-007
  6   -2.0575e-008   1.4543e-009
  7    0.0000e+000   2.6565e-006
  8    7.7070e-006   2.8038e-006
  9    0.0000e+000   8.5655e-006
 10    3.8332e-005   1.0764e-005
#
#  x and y pixel size in mm and x, y image size in pixels
  0.00780  0.00780    3008     2000
#
#
#  Fiducial Mark Data
# Camera  Point  Ref X     Ref Y     Std X   Std Y    Obs X     Obs Y    Std X   Std Y
#
# No fiducial information for this camera
#
# End of file
############################################################################
```

PHOTO FILE INPUT FORMAT (.pho):

```
############################################################################
# VMS Project:  Photo data
# Tuesday, March 24, 2009
#
# Project name: legonikon
#
# This file contains information on 52 photos
# Exterior orientation parameters for photo: 1000
# Photo     X        Y        Z       Omega      Phi      Kappa     Camera
   1000    33.6101   404.4008   354.6350   -45.2831   -0.8795    91.5219  1
   1002     0.0000     0.0000     0.0000     0.0000    0.0000     0.0000  1
    …         …          …          …          …         …          …
   1051  -256.4724   401.4178    82.2784   -81.4259  -44.5892    15.8017  1
#
# End of photo orientation file
############################################################################
```

It is noted that where the orientation parameters are assigned with zero values, these indicate that that the associated photo has not been measured and will be discounted from the calculations.

OBSERVATIONS FILE INPUT FORMAT (.obs) :

```
################################################################################
# VMS Project:  Photo data output
# Tuesday, March 24, 2009
#
# Project name: legonikon
#
# 2745  Photo Obs
#  Photo  Target    X(mm)    Y(mm)    sdx   sdy   res x   res y
#
#
# Used image measurements - at last adjustment
     1026     2011  2.05289  1.73421  0.4035  0.4035 -0.2650 -0.0584    0
     1033      400  6.81271  1.32663  0.4035  0.4035 -3.1485  0.2948   -1
      …        …   …        …        …        …       …       …        …
     1032    30025  2.69271  0.71370  0.4035  0.4035 -0.2079  0.6499    0
#
# Unused and rejected image measurements - at last adjustment
#
# End of file
################################################################################
```

It is noted that where the observation flag is -1, this indicate that the associated observations has been rejected from the system and will be not used in the subsequent computations.

B.2.2. Output VMS 8.0 data files

The output data files of the VMS 8.0 software under the 'photogrammetry' processing menu are listed here:

- The initial exterior orientations file: Outputs the exterior orientations data (.log) of the initializations procedures.

- The resections file: Outputs the updated exterior orientations data (.log) of the resection procedures.

- The intersections file: Outputs the 3D targets coordinates (.log) of the intersections procedures.

- The network adjustment file: Outputs the bundle adjustment report (.log) of the bundle adjustment procedures.

INITIAL EXTERIOR ORIENTATIONS FILE (.log):

```
################################################################################
*** Vision Measurement System (VMS) Initial Orientation Solutions ***

Version 8.0 - Stuart Robson and Mark Shortis - August 2008


Project name : legonikon

VMS initial orientation log file legonikon_init_orient.log written on Thu Mar 12 20:55:14 2009


Only measured control point targets (those with a known location) used for initial orientation
computation


 Photo Camera    X       Y       Z      Omega    Phi     Kappa
            (millimetres)                 (Degrees)
 1000    1   34.7101   387.3291   338.7801  -45.0838   -0.8447   91.4651
 1002    1    0.0000    0.0000     0.0000    0.0000    0.0000    0.0000 insufficient targets imaged:
measure some more!
  …     …    …         …          …         …         …        …
 1051    1  -240.1238  383.7137    86.2845  -81.1419  -45.6671   14.4972

Summary of    52 initial orientation computations

  44 computations OK
   0 computation failures
   8 insufficient image observations
   0 non-convergent solutions

*** End of VMS initial orientation log file ***
################################################################################
```

It is noted that the software sets zero values to the orientation parameters of the non-measured and hence excluded from the calculations photos.

RESECTIONS FILE (.log):

```
################################################################################

*** Vision Measurement System (VMS) Resection Solutions ***

Version 8.0 - Stuart Robson and Mark Shortis  - August 2008


Project name : legonikon

VMS resection log file legonikon_resect.log written on Thu Jun 18 19:46:45 2009


Only measured control point targets (those with a known location) used for resection computation

Maximum iteration count of the solutions :     10

Outlier rejection factor (image residuals) (microns):     5.0
```

Photo	Camera	X	Y	Z	Omega	Phi	Kappa	RMS Image Residuals	Image Counts Valid	Rejected
		(millimetres)			(Degrees)			(microns)		
1000	1	37.5575	386.0050	339.8901	-45.1949	-0.5102	91.5658	1.31	4	1
1002	1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.00	0	0 Too

few targets or image area coverage

…	…	…	…	…	…	…	…	…	…	…
1051	1	-241.4290	386.3624	80.9816	-81.4474	-44.3574	15.7887	8.42	7	0

Mean resection image measurement residual RMS (microns):   3.7897
Mean valid target image observations: 4.10

Summary of   52 resection computations

  44 computations OK
   0 computation failures
   8 insufficient image observations
   0 non-convergent solutions

*** End of VMS resection log file ***

##

It is noted that the software sets zero values to the orientation parameters of the non-measured and hence excluded from the calculations photos.

INTERSECTIONS FILE (.log):

##

*** Vision Measurement System (VMS) Intersection Solution ***

Version 8.0 - Mark Shortis and Stuart Robson - August 2008

Project name : legonikon

VMS intersection log file legonikon_intersect.log written on Wed Mar 18 21:26:40 2009

All target coordinates are unconstrained (free network or internal datum)

All photographs with a known location/orientation (resection) will used in the solutions

Maximum iteration count of the solutions :   10

Outlier rejection factor (image residuals) :   5.0

Target	X	Y	Z	Index	sX	sY	sZ	Mean Image Residuals		Image Counts Valid	Rejected
		(millimetres)			(microns)			(microns)			
					x	y					
20045	65.3209	64.8550	28.4933	0	25.1	25.2	13.4	58.96	8.84	2	1
200	2.1541	99.8931	3.4229	7	14.6	21.9	15.0	13.00	12.23	40	0
…	…	…	…	…	…	…	…	…	…	…	…
107	-0.8992	3.1723	96.5628	7	128.5	189.3	61.7	5.97	9.04	5	0

Mean	21.5	37.3	22.3	2.56	2.15	14.7

Summary of 183 intersection computations

  183 computations OK
   0 computation failures
   0 insufficient image observations
   0 incomplete solutions
   0 solutions over tolerance

Summary of Target Images by Photo

Photo	Target Counts	RMS Residuals (microns)
1000	50	3.15
1001	58	3.95
…	…	…
1051	63	4.40
Mean	51.7	3.32

*** End of VMS intersection log file ***

###

NETWORK ADJUSTMENT FILE (.log):

###

*** Vision Measurement System (VMS) Self-calibrating Photogrammetric Network Solution ***

Version 8.0 - Mark Shortis and Stuart Robson - August 2008

Project name :   legonikon

VMS bundle adjustment log file legonikon_network.log written on Tue Mar 24 12:47:04 2009

<<< Program control variables >>>

Network datum definition type :   generalised internal constraints

Additional parameter set type :   PP, PD, lens distortion and affinity parameters, block-invariant PP

Maximum iterations for a solution :   10

Default target image precision :   0.40

Minimum images for a network target : 4

Rejection criterion for image errors : 5.0

<<< Initial Camera Calibration Sets >>>

Units : millimetres

Calibration set : 1

Parameter	Value	Precision
1	0.114	1.000
2	0.040	1.000
3	29.578	1.000
4	-1.8822e-004	1.000e+000
5	2.2259e-006	1.000e+000
6	-2.0575e-008	1.000e+000
7	0.0000e+000	0.000e+000
8	7.7070e-006	1.000e+000
9	0.0000e+000	0.000e+000
10	3.8332e-005	1.000e+000

Pixel size in mm   Format size in pixels

x	y	x	y
0.0078	0.0078	3008	2000

<<< Initial Camera Locations >>>

Photo	Cal Set	X	Y	Z	Omega	Phi	Kappa	#Images
		(millimetres)			(degrees)			
1000	1	33.610	404.401	354.635	-45.28	-0.88	91.52	51
1002	1	0.000	0.000	0.000	0.00	0.00	0.00	0 (insufficient images)
…	…	…	…	…	…	…	…	…
1051	1	-256.472	401.418	82.278	-81.43	-44.59	15.80	63

<<< Initial Targets >>>

Target	X	Y	Z	Index	sX	sY	sZ	#Images
	(millimetres)				(microns)			
20045	65.5735	75.5262	23.4447	7				8
200	2.4575	100.0290	3.3735	7				41
…	…	…	…	…				…
107	-0.5143	2.7531	96.7073	7				5

<<< Initial Survey Measurements >>>

Units (azimuths and angles) :   ddd.mmsss and seconds of arc
Units (distances and level differences) : millimetres and microns

RO Targ	At Targ	To Targ	Measurement	Type	Precision	Residual	Significance	Inst Ht	Targ Ht
	1000	1005	74.0250 Slope dist	50.0	0.000	0.000			
	1005	1010	74.3800 Slope dist	50.0	0.000	0.000			
	…	…	…	…	…	…			
	2009	2000	43.8700 Slope dist	50.0	0.000	0.000			

<<< Input Summary >>>

Number of camera calibration sets :   1
Number of target image observations :   2745

Total number of exposures :   52
Number of exposures in the network :   44

Total number of targets :   178
Number of targets in the network :   178

Total number of survey measurements :   10
Number of survey measurements in the network :   10

*** Results for the Calibration Solution ***

Solution completed successfully after 10 iteration(s)

Unit weight estimate (sigma zero) :   1.00
RMS image residual (microns) :   0.37

Number of rejected target images :   62

Number of observables in the network :   5376
Number of unknowns in the network :   812
Number of redundancies in the network :   4564

<<< Updated Camera Calibration Set : 1 >>>

Parameter	Value	Precision	Correction	Significance	Max. Effect
		(millimetres)		(microns)	
1	0.1137	0.0020	0.0000	0.00	
2	0.0404	0.0073	0.0000	0.00	
3	29.5784	0.0048	0.0000	0.00	
4	-1.8822e-004	2.643e-006	2.121e-011	0.00	-88.0
5	2.2259e-006	1.193e-007	2.101e-012	0.00	62.7
6	-2.0575e-008	1.454e-009	-5.009e-014	0.00	-34.9
7	0.0000e+000	2.657e-006	0.000e+000	0.00	0.0
8	7.7070e-006	2.804e-006	1.877e-011	0.00	0.5
9	0.0000e+000	8.566e-006	0.000e+000	0.00	0.0
10	3.8332e-005	1.076e-005	3.409e-010	0.00	0.3

Lens Distortion Profiles for Camera Calibration Set: 1

Units : microns

Radius	Radial Distortion		Decentring Distortion	
	Value	Precision	Value	Precision
0.0	-0.00	0.00	0.00	0.00
1.0	-0.19	0.00	0.01	0.00
2.0	-1.44	0.02	0.03	0.01
3.0	-4.59	0.05	0.07	0.03
4.0	-10.10	0.08	0.12	0.04
5.0	-18.18	0.11	0.19	0.07
6.0	-29.11	0.16	0.28	0.10
7.0	-44.09	0.26	0.38	0.14
8.0	-66.58	0.63	0.49	0.18

Correlation Parameters for Camera Calibration Set: 1

```
 1   1.00
 2   0.00  1.00
 3   0.03  0.59  1.00
 4  -0.03  0.04 -0.10  1.00
 5   0.12 -0.02  0.07 -0.93  1.00
 6  -0.11  0.03 -0.04  0.86 -0.97  1.00
 8  -0.01  0.91  0.40  0.13 -0.09  0.08  1.00
10  -0.07 -0.15 -0.06 -0.09  0.07 -0.07 -0.29  1.00
```

<<< Updated Camera Location Values and Precisions >>>

Photo	Cal Set	X	Y	Z	Omega	Phi	Kappa	sX	sY	sZ	sO	sP	sK
		(millimetres)			(degrees)			(microns)			(seconds of arc)		
1000	1	33.6101	404.4008	354.6350	-45.283	-0.880	91.522	26.0	91.7	85.9	18.0	51.4	6.1
1002	Indeterminate - insufficient target images												
1051	1	-256.4724	401.4178	82.2784	-81.426	-44.589	15.802	85.7	89.8	23.4	70.1	20.7	47.8

<<< Target Image Precisions and Residuals >>>

Units : microns

			Mean Precision		Mean Residual		Max Residual	
Photo	#Targets	#Rej	x	y	x	y	x	y
1000	51	0	0.40	0.40	0.30	0.36	0.67	1.11
1002	Indeterminate - insufficient image observations							
1051	63	0	0.40	0.40	0.46	0.35	1.63	1.43
Mean	61.0	1.4	0.40	0.40	0.83	0.73		
All	2683	62					19.15	20.03

<<< Updated Targets >>>

	Coordinates				Precisions			Mean Residuals		Max Residuals			
Target	X	Y	Z	Index	sX	sY	sZ	x	y	x	y	#Images	Photo
List (y=yes, n=no, R=rejected)													
	(millimetres)				(microns)			(microns)					
20045	65.5735	75.5262	23.4447	7	3.9	4.7	6.5	0.12	0.24	0.23	0.35	8	
nnxxxxxxxxnnnyyyynnnnnynynnnnnnnnnnnnynynnnnnnnnnnnnnynnnnnnnnnnnyn													
200	2.4575	100.0290	3.3735	7	10.8	5.9	10.2	0.78	0.59	1.22	1.21	25	
yRxxxxxxxxRRRRRRnRRRRRRRRyyyyyyyyyyyyynyyyyyyyyyyyyyyRny													
…	…	…	…	…	…	…	…	…	…	…	…	…	…
107	-0.5143	2.7531	96.7073	7	13.5	19.3	11.4	0.85	0.92	1.30	1.43	5	
nnxxxxxxxxxnnnnnnnnnnnnnnnnnnnnnnyyynnnnnnnnnnnynnnnnnnnny													
Mean										15.1			

<<< Target Precision Summary >>>

Units : microns

Target	sX	sY	sZ
Minimum	30029	3.96	3.38	1.58
Mean		5.98	5.08	5.92
Maximum	101	11.98	23.32	16.73

Mean precision of target coordinates :   5.66

Relative precision for the network 1 :   36000

<<< Updated Survey Measurements >>>

Units (azimuths and angles) :   ddd.mmsss and seconds of arc
Units (distances and level differences) : millimetres and microns

RO Targ	At Targ	To Targ	Measurement	Type	Precision	Residual	Significance	Inst Ht	Targ Ht
1000	1005	74.0250	Slope dist	50.0	57.6	1.15	0.000	0.000	
1005	1010	74.3800	Slope dist	50.0	-70.9	1.42	0.000	0.000	
…	…	…	…	…	…	…	…	…	
2009	2000	43.8700	Slope dist	50.0	-142.3	2.85	0.000	0.000	

 RMS Residual   103.28

*** End of VMS calibration log file ***

##

B.3. Framework space

The description of the processing framework is given below.

B.3.1. Data structure

Development and implementation of the affine algorithm required the generation of two data structures (embedded in the available melb.h header file). These two structures are given here.

- PHOTO_ORTHO data structure: identifies the photo structure (members for affine photo data implementation).

- MAT_POINTERS data structure: identifies the matrices structure (1D and 2D arrays members) for algorithm handling.

- PHOTO_ORTHO STRUCTURE:

```
typedef struct      PHOTO_ORTHO_T
{
        long    m_camera;       // Camera name
        long    m_photo;        // ID of this photo
        long    xsize;          // width of the digital image in pixels
        long    ysize;           // height of the digital image in pixels
        double  pixel_x;        // x pixel size in mm.
        double  pixel_y;        // y pixel size in mm.
      //double  ext_ori[6];     //Camera location X,Y,Z, and orientation Omega,Phi,Kappa
(meters and degrees dec)
        //double ext_std[6];    // Camera Stdev X,Y,Z and Stdev Omega,Phi,Kappa
        double ortho_ori[6];    //Parallel camera scale, orientations omega, phi, kappa and
translations tx, ty
        double ortho_std[6];    //Parallel camera std scale, omega, phi, kappa, tx, ty
        double  rot_array[9];   // Rotation matrix stored in row order.
        double  int_ori[AP_MAX];// Camera interior orientation parameters
xp,yp,pd,k1,k2,k3,p1,p2,a1,a2,extended APs
        double  int_std[AP_MAX];// Camera interior orientation standard deviations
        double  int_std_file[AP_MAX];// Camera interior orientation initial standard deviations
        double  int_res[AP_MAX];// Camera interior orientation residuals
xp,yp,pd,k1,k2,k3,p1,p2,a1,a2,extended APs
        double  int_normat[AP_MAX_Q]; // Camera interior orientation cofactor matrix
        double  quality;        // Quality value : estimate of unit weight for the photo
        double  rms;            // RMS image residual for the photo
        short   total_rays;     // Total number of rays to this photo
        short   used_rays;      // Number of used rays to this photo
        short  pho_flag;        // Flag for photo setup
        short   rot_hir;        // Flag for rotation hirarchy
        long    epoch_id;       // Epoch id
        char    image_file[CHLIM];// photo image file name
        double  min_target_depth;// Depth of the nearest target
        double  max_target_depth;// Depth of the farthest target
        DPOINT          dpFootprint[4];// footprint of the photo at the maximum target depth
        double  coef[16];       // Coefficients for IO transformations
        char    transf_type;    // IO Transformation type // p = 16, b = 8, a = 6, s = 4, t = 3
        //long  photo_count_ortho; //Parallel camera, number of photos - added 01102007
}       PHOTO_ORTHO;
```

-MAT_POINTERS STRUCTURE:

```
typedef struct MAT_POINTERS_T //matrices allocated in mrscal -added 21092007
{
        double* ans_vec;             //Vector of parameters of LSA ans_vec[max_unknowns] (x)
        double* corr_ans_vec;        ////Vector of increments to parameters of LSA
ans_vec[max_unknowns] (x)
        double* nor_vec;             //Normals array vector nor_vec[max_unknowns] (At*W*b)
        double** des_mat_full;       //Design matrix in 2D des_mat_full[equations][unknowns]
(A)//added 25092007
        double** nor_mat_full;       //Normals array matrix in 2D nor_mat_full[unknowns][unknowns]
N=(At*W*A)//added 25092007
        double* nor_vec_full;        //normals array vector atwb for the lse
        double* obs_vec_full;        //Correction vector to the observations//added 16102007
        double* weight_vec_full;     // Weight vector associated with the weight matrix full
        double** inv_nor_mat_full;   //normals matrix inverse
        long** photo_loc;            //Photo station parameter locations
        long** trans_loc;            //Translation parameter locations
        long** scale_loc;            //Scale parameter locations
        double** rot;                 //Rotation matrices for each photograph
```

```
        long** targ_loc;                //Target coordinate column locator
        //double** ab_bb;               //Intermidiate matrix for LSA solution: ab_bb[n_unk][7]
        //double*** ab_mat;             //off-diagonal sub-matrix for targets: ab_mat[n_targ][n_unk][3]
        //double** bm_mat;              //normals vector components for each target:
bm_mat[n_unk|n_targ][3]
        //bm_mat should be set to n_unk or n_targ, whichever is larger
        double* nor_mat;                //normals array matrix stored columnwise: nor_mat[] N=(At*W*b)
        double** photo0;                //initial values for camera stations: photo0[n_photo][6]
        double** targ0;                 //initial values for target coordinates: targ0[n_targ][3]
        //long** pp_para_loc;//pp parameter locations
        //double** pp_para0; //initial values for pps
        long n_eqns;                    //default number of equations at allocation time// added 25092007
        long n_unks;                    //default number of unknowns at allocation time
        long tar_count;                 //default number of targets at allocation time
        long photo_count;               //default number of targets at allocation time
        double* res_vec;                //residual vector after lse adjustment
        double** cov_mat_full;          //covariance matrix after lse adjustment
} MAT_POINTERS;
```

B.3.2. Processing menus description

The menus of the developed framework space are given below:

- Bundle adjustment menu:
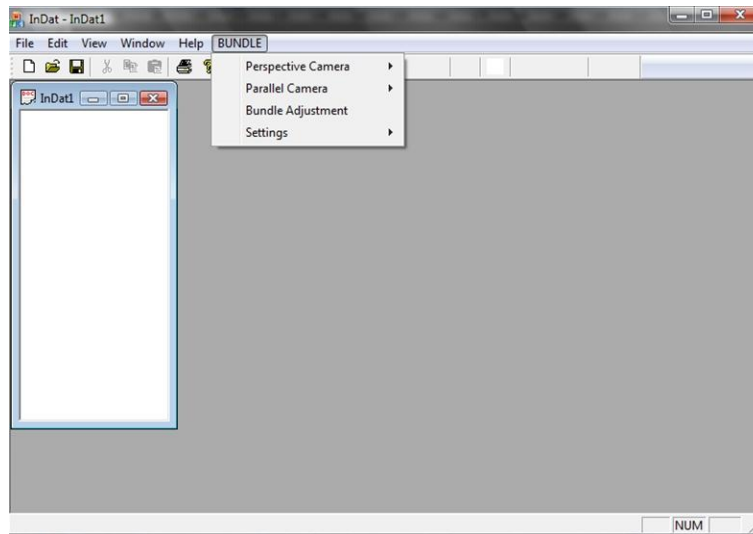
Appendix B - Figure 4: Bundle adjustment framework.

- Read perspective camera data files (VMS 8.0 format): Perspective camera menu opens and subsequently saves target, calibration, photo and observations data files.
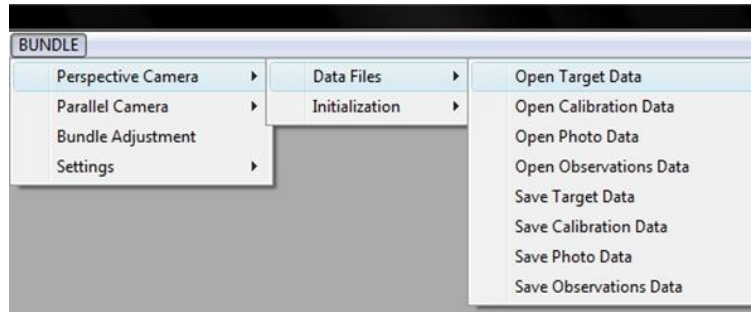
Appendix B - Figure 5: Perspective camera menu.

- Initialize perspective camera procedures: Initialization menu performs initial exterior orientation, resection and intersection procedures.
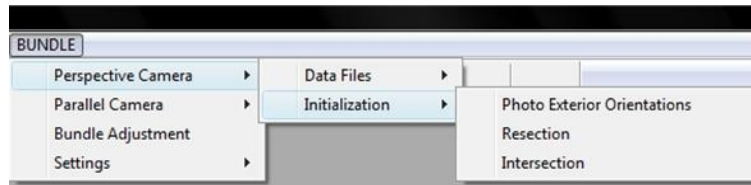
Appendix B - Figure 6: Initialization menu.

- Initialize parallel camera procedures: Orientation menu reads and saves orientations data files. Generate translations: performs 2D cameras locations calculations. Generate 3D targets points: computes 3D targets coordinates.
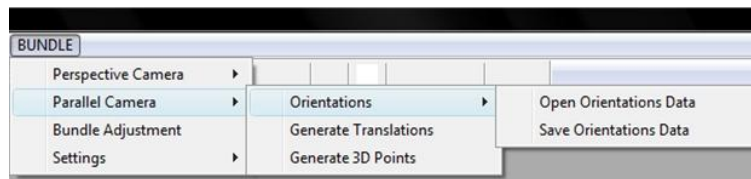
Appendix B - Figure 7: Parallel camera menu.

B.3.3. Input processing files

The input data files required for the implementation of the parallel sensor model follow the general format of the perspective sensor model as described above with modifications. The data files are listed as follows.

TARGET FILE (.tar)

```
################################################################################
# VMS Project:  Target data output
# Thursday, April 02, 2009
#
# Project name: sony_sx900
#
# 86  Targets
# Targets in millimetres, precisions in 1000*millimetres
#
#    Target      X       Y       Z  flag    sdx     sdy     sdz
#
     30016   52.0970  94.5440  48.4662  7   1.9000  5.2000  1.8000
      1009   72.8786  67.7193  86.6971  7   5.1000  2.7000  8.4000

       …        …        …        …      …    …       …       …
     20020   79.5663  71.4596  51.8730  7   7.1000  4.1000  1.9000
#
# End of file


################################################################################
```

CALIBRATION FILE (.cal)

```
################################################################################
# VMS Project:  Camera data
# Friday, May 26, 2006
#
# Project name: SonyCam-SWITAR:10mm
#
# This file contains information on 1 cameras
# Parameters : 1=PPx,2=PPy,3=PD,4-6=radial,7-8=decentring,9=orthogonality,10=affinity
 1
#
# Calibration parameters for camera 1
    1     0.0000      0.0000
    2     0.0000      0.0000
    3 100000.0000      0.0000
    4    0.0000e+000    1.0000e+000
    5    0.0000e+000    0.0000e+000
    6    0.0000e+000    0.0000e+000
    7    0.0000e+000    0.0000e+000
    8    0.0000e+000    0.0000e+000
    9    0.0000e+000    0.0000e+000
   10    0.0000e+000    0.0000e+000
#
#  x and y pixel size in mm and x, y image size in pixels
  0.00478  0.00478    1024     768
#
#
#  Fiducial Mark Data
# Camera  Point  Ref X    Ref Y   Std X   Std Y   Obs X   Obs Y   Std X   Std Y
#
# No fiducial information for this camera
#
# End of file
################################################################################
```

It is noted that indirect initialization requires that the camera constant c is set to a large value such that (c=100,000.0 mm) and the standard deviation of the third power term of the radial lens distortion polynomial (parameter 4) is set to 1.0 for inclusion in the calculations.

PHOTO FILE (.pho)

```
################################################################################
# VMS Project:  Photo data
# Thursday, April 02, 2009
#
# Project name: sony_sx900
#
# This file contains information on 85 photos
# Exterior orientation parameters for photo: 1001
# Photo    X      Y      Z     Omega      Phi     Kappa    Camera
   1001  0.0000 0.0000 0.1600  -92.1378  -13.9839  -95.3905 1
   1002  0.0000 0.0000 0.1600  -91.1046  -14.1578  -90.9038 1
    …      …      …      …        …         …         …       …
   1145  0.0000 0.0000 0.1600  -64.6832   28.5346   49.4593 1
#
# End of photo orientation file
################################################################################
```

It is noted that the first two parameters correspond to the 2D cameras locations the values of which will be updated after back-substitution from the affine camera sensor.

OBSERVATIONS FILE (.obs)

```
################################################################################
# VMS Project:  Photo data output
# Thursday, April 02, 2009
#
# Project name: sony_sx900
#
# 553  Photo Obs
#  Photo  Target    X(mm)    Y(mm)    sdx   sdy   res x   res y
#
#
# Used image measurements - at last adjustment
    1001     1019  -1.84762  0.78812  0.5000  0.5000 -0.5216  0.1714    0
    1001     1000  -1.78758 -1.40620  0.5000  0.5000 -0.3189 -0.5313    0
     …       …      …        …         …       …      …       …         …
    1092    20032   0.82131  1.00758  0.5000  0.5000 -0.7636 -0.2837    0
#
# Unused and rejected image measurements - at last adjustment
#
# End of file
```

B.3.4. Output processing files

The output data files under the 'parallel camera' and 'bundle adjustment' processing menus are listed here:

- Generate translations (.log): Outputs the updated orientations file of the back-substitution procedure.

- Generate 3D points (.log): Outputs the 3D coordinates file of the 3D targets calculation procedure.

- Bundle adjustment (.log): Outputs the solution file of the bundle adjustment solution in two modes external datum and inner datum constraints.

GENERATE TRANSLATIONS (.log)

```
##############################################################################
# VMS Project:  Photo data
# This file contains information on 85 photos
# Exterior orientation parameters for photo: 1001
# Photo    X        Y        Z       Omega      Phi      Kappa     Camera
   1001   -3.5543  -6.3773   0.1600  -92.1378  -13.9839  -95.3905 1
   1002   -6.2914  -6.3716   0.1600  -91.1046  -14.1578  -90.9038 1
    …        …        …        …         …         …         …      …

   1145    0.9669  -0.2349   0.1600  -64.6832   28.5346   49.4593 1
#
# End of photo orientation file
##############################################################################
```

GENERATE 3D POINTS (.log).

```
##############################################################################
#<< Parallel Camera Project - Intersection Solution >>
#Units: mm
#Multi-view intersection converged successfully after 0002 iterations.
#Number of redundancies in the intersection solution:  0846
#s_aposteriori: 69.8041
#<< UPDATED TARGETS LOCATIONS >>
#Units: mm
#Target  X         Y         Z         flag  sdx      sdy       sdz
#
30016   52.1595   94.4384   48.5095     7   66.9488  255.4150  66.9691
1009    72.9971   68.2253   87.1588     7   70.3455  176.4792  84.1095
  …       …         …         …         …      …        …         …
20020   79.9308   72.3601   51.9579     7  258.1666  390.0057  75.2454
#
# End of 3D targets file

##############################################################################
```

BUNDLE ADJUSTMENT (.log).

###

EXTERNAL DATUM CONSTRAINTS

###

<< Parallel Camera Project - Bundle Adjustment >>
<< Network Datum Definition Type: e >>
<< INITIAL PHOTO LOCATIONS>>
 Units: mm, degrees

Photo	Camera	tx	ty	Scale	Omega	Phi	Kappa	Rays
1001	1	-3.554	-6.3773	0.1600	-92.1378	-13.9839	-95.3905	5
1002	1	-6.2914	-6.3716	0.1600	-91.1046	-14.1578	-90.9038	6
…	…	…	…	…	…	…	…	…
1145	1	0.9669	-0.2349	0.1600	-64.6832	28.5346	49.4593	5

<< INITIAL TARGETS LOCATIONS>>
Units: mm

Target	X	Y	Z	Index	sX	sY	sZ	Images
30016	52.1595	94.4384	48.5095	0	66.9488	255.415	66.9691	11
1009	72.8786	67.7193	86.6971	7	5.1000	2.7000	8.4000	12
…	…	…	…	…	…	…	…	…
20020	79.5663	71.4596	51.873	7	7.1000	4.1000	1.9000	9

<< Number of Equations: 1338 >>
<< Number of Unknowns: 0685 >>
Iteration: 0003
Number of redundancies in the network : 0653
Unit weight estimate (sigma zero) : 2.2600

Image residuals (microns):

x	y	Mean
0.7296	0.4763	0.6029

Lens Distortion Profile for Camera Calibration
Units: microns

Radius	Value
0.0000	0.0000
0.5000	0.0289
1.0000	0.2313
1.5000	0.7805
2.0000	1.8502
2.5000	3.6136

<< UPDATED ORIENTATIONS LOCATIONS>>
 Units: Values mm, degrees Precisions microns, degrees

Photo	Camera	tx	ty	Scale	Omega	Phi	Kappa	stx	sty	ss	so	sp	sk	Rays
1001	1	-3.7052	-6.6006	0.1659	-92.2531	-13.9763	-95.4445	19.4582	18.1860	0.0000	0.1085	0.1034	0.0329	5
1002	1	-6.5044	-6.5856	0.1659	-91.1612	-14.1185	-90.9403	8.6938	10.8058	0.0000	0.0580	0.0461	0.0460	6
….	…	…	…	…	…	…	…	…	…	…	…	…	…	…
1145	1	1.0109	-0.2579	0.1659	-64.6506	28.5629	49.4359	20.6365	24.9394	0.0000	0.0703	0.0639	0.0411	5

<< UPDATED TARGETS LOCATIONS >>
Units: mm, microns

Target	X	Y	Z	Index	sX	sY	sZ	Images

30016 52.0975 94.5801 48.4683   0 3.0693 10.5697 3.0819   11
1009   72.8837 67.7244 86.6937   7 3.6725 5.3298 4.8033   12
…   …   …   …   …   …   …   …   …
20020 79.5706 71.4604 51.8676   7 6.4348 7.5573 2.5350   9

<< UPDATED ADDITIONAL PARAMETERS DATA (k1) >>
Units: mm
Value Precision Correction Significance
2.312710354e-004 2.4014e-005 1.4908e-005 0.6208

<<< TARGETS PRECISION SUMMARY >>>
Mean (microns):
sX   sY   sZ
10.1599 17.5822 7.7488

<< FULL CORRELATIONS MATRIX>>
 1.000   0.047   0.133   0.110   0.014   0.050   0.001   0.019   -0.004   0.011
 0.004   0.002   -0.001   0.000   -0.004   0.001   -0.002   0.002   -0.002   0.002
-0.002   0.001   0.001   0.003   -0.000   0.004   0.000   0.004   -0.002   0.004
-0.006   0.004   -0.003   0.019   -0.005   0.016   0.004   0.074   -0.032   -0.008
-0.005
 …   …   …   …   …   …   …   …   …   …
…
##

INNER DATUM CONSTRAINTS

##
<< Parallel Camera Project - Bundle Adjustment >>
<< Network Datum Definition Type: i >>
<< INITIAL PHOTO LOCATIONS>>
 Units: mm, degrees
Photo Camera tx   ty   Scale Omega   Phi   Kappa   Rays
 1001   1 -3.5543 -6.3773 0.1600 -92.1378 -13.9839 -95.3905   5
 1002   1 -6.2914 -6.3716 0.1600 -91.1046 -14.1578 -90.9038   6
 ….
 1145   1 0.9669 -0.2349 0.1600 -64.6832 28.5346 49.4593   5

<< INITIAL TARGETS LOCATIONS>>
Units: mm
Target X   Y   Z   Index sX   sY   sZ   Images
30016 52.1595 94.4384 48.5095   0 66.9488 255.4150 66.9691   11
 1009 72.8786 67.7193 86.6971   7 5.1000 2.700 8.4000   12
…
20020 79.5663 71.4596 51.8730   7 7.1000 4.1000 1.9000   9

<< Number of Equations: 1111 >>
<< Number of Unknowns: 0685 >>
Iteration: 0003
Number of redundancies in the network : 0426
Unit weight estimate (sigma zero) : 2.5279

Image residuals (microns):
x   y   Mean
0.6861 0.4391 0.5626

Lens Distortion Profile for Camera Calibration
Units: microns

Radius Value
0.0000 0.0000
0.5000 0.0280
1.0000 0.2239
1.5000 0.7556
2.0000 1.7911
2.5000 3.4983

<< UPDATED ORIENTATIONS LOCATIONS>>
 Units: Values mm, degrees Precisions microns, degrees

Photo	Camera	tx	ty	Scale	Omega	Phi	Kappa	stx	sty	ss	so	sp	sk	Rays
1001	1	-3.7040	-6.5922	0.1659	-92.2799	-13.9103	-95.5057	39.5662	28.9734	0.0000	0.2113	0.1751	0.0826	5
1002	1	-6.5067	-6.5635	0.1659	-91.1936	-13.9928	-90.9730	33.4803	29.5572	0.0000	0.18613	0.1544	0.0926	6
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
1145	1	0.9844	-0.2906	0.1659	-64.5100	28.5968	49.3478	138.2115	124.4356	0.0000	0.4456	0.3552	0.1874	5

<< UPDATED TARGETS LOCATIONS >>
Units: mm, microns

Target	X	Y	Z	Index	sX	sY	sZ	Images
30016	52.1110	94.5747	48.4634	0	37.6660	67.5979	42.5347	11
1009	72.8891	67.7051	86.6932	7	41.4308	85.3256	46.7768	12
…	…	…	…	…	…	…	…	…
20020	79.5739	71.4647	51.8620	7	41.5566	62.8952	37.5903	9

<< UPDATED ADDITIONAL PARAMETERS DATA (k1) >>
Units: mm
Value Precision Correction Significance
2.238894011e-004 2.9664e-005 1.7741e-005 0.5981

<<< TARGETS PRECISION SUMMARY >>>
Mean (microns):
sX   sY   sZ
57.6372 142.6282 60.5613

<< FULL CORRELATIONS MATRIX>>

1.000	-0.286	0.721	-0.307	0.099	-0.187	0.120	-0.150	0.146	-0.122
-0.129	-0.097	-0.093	-0.084	0.034	-0.101	0.050	-0.142	0.052	-0.136
0.050	-0.133	-0.120	-0.079	-0.130	-0.093	-0.135	-0.109	-0.205	-0.164
-0.228	-0.179	-0.196	0.284	-0.216	0.256	-0.200	0.468	0.063	0.529
0.360									
…	…	…	…	…	…	…	…	…	…
…									

##

Appendix C: Processing data examples

This appendix provides a supportive description of processed data required for the critical analysis while decomposing the problem within the successive thesis chapters. In particular the main sections that will be listed here cover: data implementation for simulation analysis of the affine sensor model (see section C.1.), numerical target image quality characteristics (see section C.2.), evaluation of typical camera calibration parameters (see section C.3.) as well as some comparative radial lens distortion profiles results from affine bundle adjustment processing (see section C.4.).

C.1. Computed data in simulation analysis

Prior to the implementation of the affine bundle adjustment algorithm a simulation project was implemented based upon a synthetic 3D cube which was processed for three different imaging cases (see section 5.6.).

C.1.1. Input 3D target coordinates of synthetic cube

The 3D target coordinates of the synthetic cube were designed as follows.

3D TARGET FILE FOR SIMULATION PROJECT:

```
################################################################################
# VMS Project:  Target data output
# Monday, November 24, 2008
#
# Project name: cube_simu_4_4
#
# 9  Targets
# Targets in millimetres, precisions in 1000*millimetres
#
#    Target     X       Y       Z  flag    sdx      sdy      sdz
#
     100    4.5000   29.8000  -15.7000  7   25.0000   25.0000   25.0000
     101   14.5000   29.8000  -15.7000  7   25.0000   25.0000   25.0000
     102   14.5000   39.8000  -15.7000  7   25.0000   25.0000   25.0000
     103    4.5000   39.8000  -15.7000  7   25.0000   25.0000   25.0000
     104    4.5000   29.8000   -5.7000  7   25.0000   25.0000   25.0000
     105   14.5000   29.8000   -5.7000  7   25.0000   25.0000   25.0000
     106   14.5000   39.8000   -5.7000  7   25.0000   25.0000   25.0000
     107    4.5000   39.8000   -5.7000  7   25.0000   25.0000   25.0000
     108    9.5000   34.5000  -10.7000  7   25.0000   25.0000   25.0000
#
# End of file
################################################################################
```

C.1.2. Histograms of bundle adjustment residuals

Bundle adjustment runs for the simulated geometric cases produce the following residual histograms separated in x and y directions. Here the bundle adjustment residuals are given for the two-view geometry case (see section 5.4.2.) (red framed figures), three-view geometry case (see section 5.4.3.) (green framed figures) and seven-view geometry case (see section 5.4.4.) (blue framed figures).
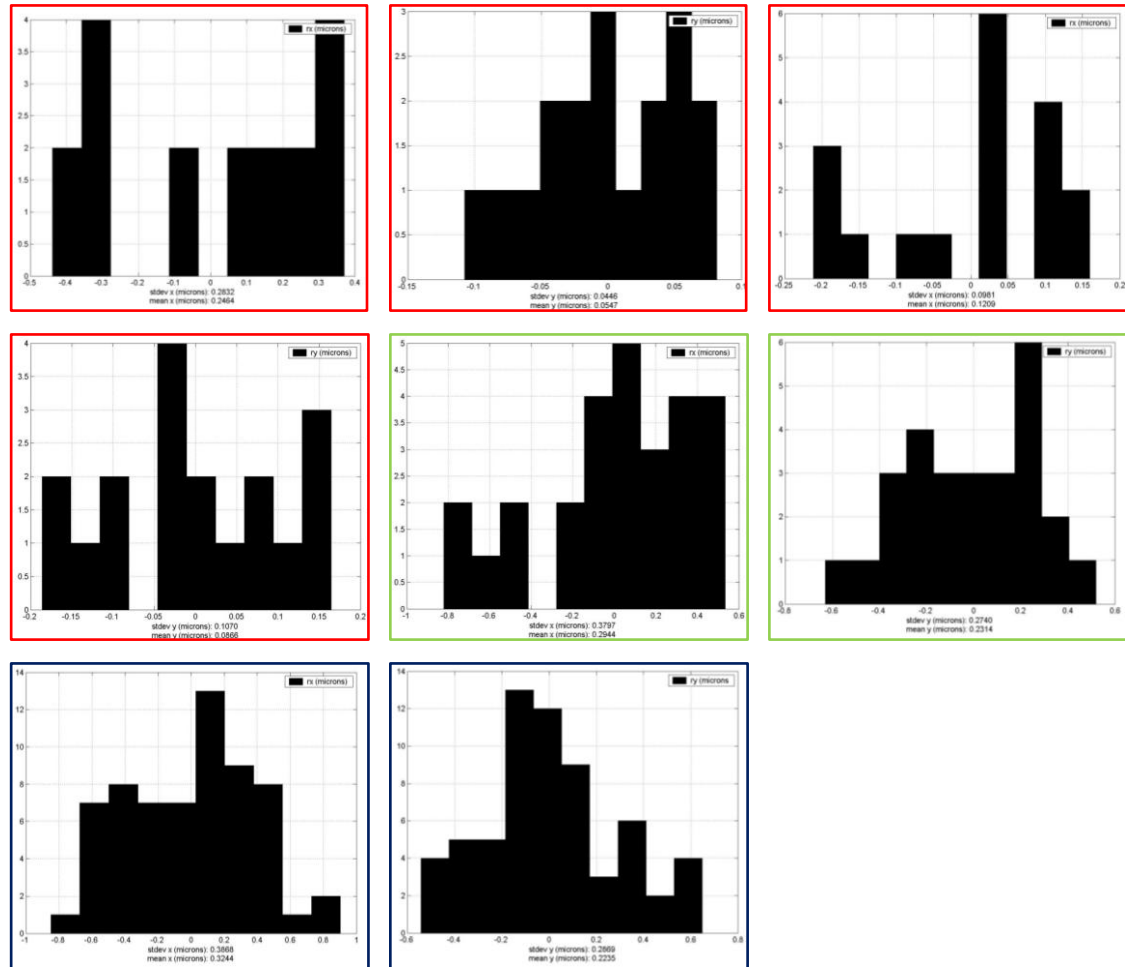

Appendix C - Figure 1: Histograms of residuals - BA solution, simulation project.

C.2. Numerical target image quality characteristics

For analysis of the image measurement quality of the measured targets within the image network a set of image targets are selected. The first sample illustrates a range of different targets for both perspective and affine imaging geometries acquired with the Kodak and Sony camera systems (see Figure 2). Some extracted numerical properties of the imaged targets are given in Table 1 to draw the characteristics of the measured image features. The following target images are tabulated as: (a.)= 2mm

diameter retro-reflective target (Kodak system), (b.)= 1mm diameter white marker (Kodak system), (c.)= 0.5mm diameter white marker (Sony system), (d.)= 1mm diameter (Kodak system), (e.)= 1mm diameter white marker (Kodak system), (f.)= 2mm diameter retro-reflective target (Kodak system), (g.)= 1mm diameter white marker (Kodak system), (h.)= 0.5mm diameter white marker (Sony system), (i.)= 1mm diameter white marker (Kodak system) and (j.) 1mm diameter white marker (Sony system).
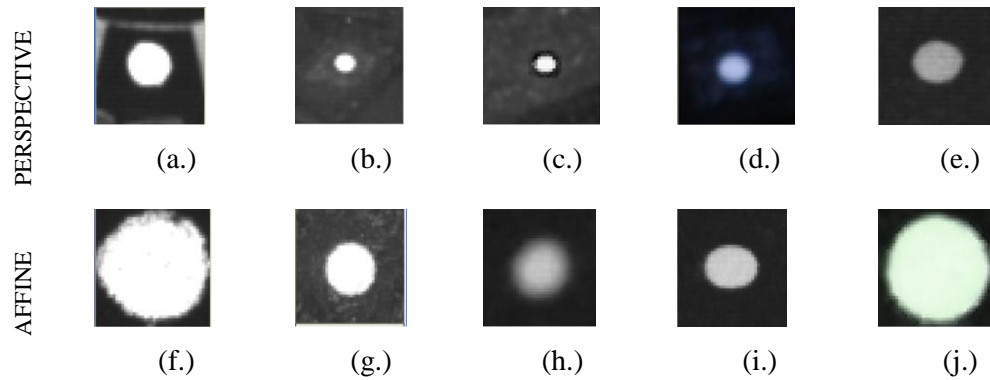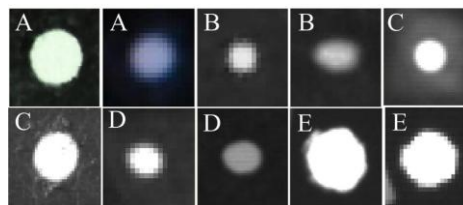

Appendix C - Figure 2: Signalized point targets (magnification window 40x40 pixels).

Camera	PP	AP	PP	AP	PP	AP	PP	AP	PP	AP
Target	a	f	b	g	c_{RGB}	h_{RGB}	d_{RGB}	i	e	j_{RGB}
Ø (pixels)	16	36	4	18	6	17	10	8	8	34
FB	255	255	255	255	248	193	175	207	166	236
					248	193	188			251
					248	193	232			232
BB	37	37	60	66	61	28	17	33	37	23
					61	28	23			25
					61	28	35			22

Appendix C - Table 1: Signalized targets image characteristics. Table notation: PP= perspective projection, AP= affine projection, Ø= target diameter, FB= foreground brightness, BB= background brightness, a - i= coded image targets.

An additional set of target images are displayed as a set of perspective and affine views (see Figure 3) characterized by their histograms and measurement parameters (see Figure 4 and Table 2 accordingly).

Appendix C - Figure 3: Signalized image points - perspective (top) and affine (bottom) views.
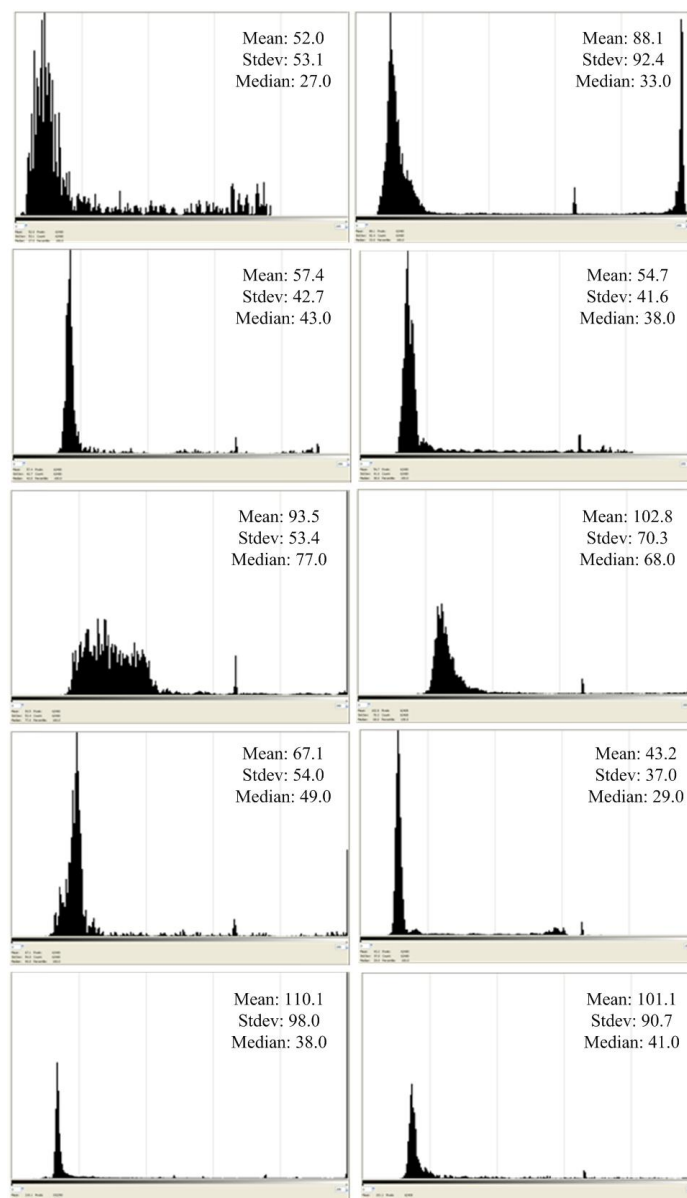
Appendix C - Figure 4: Brightness histograms of point pairs (A-E) perspective and affine views.

Target	Ø (pixels)	IW (pixels)	LM	TT	S	GT
A [PP]	10	20 x 20	w_centroid	histogram	circular	on
A [AP]	33	60 x 60	w_centroid	histogram	circular	off
B [PP]	6	20 x 20	w_centroid	histogram	circular	on
B [AP]	17	40 x 40	ellipse_fit	histogram	circular	on
C [PP]	8	36 x 36	w_centroid	histogram	circular	on
C [AP]	18	38 x 38	w_centroid	histogram	circular	on
D [PP]	8	20 x 20	w_centroid	histogram	circular	on
D [AP]	18	40 x 40	w_centroid	histogram	circular	on
E [PP]	16	24 x 24	w_centroid	histogram	circular	on
E [AP]	36	60 x 60	w_centroid	histogram	circular	off

Appendix C - Table 2: Parameters of measurement method within VMS 8.0. Table notation: Ø = target diameter, IW= image window, LM= location method, TM= threshold type, S= shape, GT= geometric tests.

C.3. Computed camera calibration parameters

In the course of the experimental analysis data initialization was implemented through conventional perspective bundle adjustment estimation (VMS 8.0). The software embeds an extended additional parameters model to accommodate internal geometric errors ordered as 1: principal point x component (PPx), 2: principal point y component (PPy), 3: principal distance (PD), 4-6: radial distortions, 7-8: decentring distortions, 9: orthogonality, 10: affinity terms. Bundle adjustment processing dictates parameters inclusion judged by their extracted significance values. The example here illustrates the output calibration data for the utilized C1 dataset, camera system CAMP2 (Kodak Megaplus ES1.0; Fujinon TV / f=12.5 mm lens) (see section 6.4.1.1.). The bundle adjustment was processed based on an inner constraints datum with appropriate 1-5 up to the k_5 radial lens distortion term internal model parameters. The derived camera parameters include the updated calibration terms and the associated correlation coefficients together with the illustrated radial lens distortion profile (see Figure 5 and Figure 6).

```
            <<< Updated Camera Calibration Set : 1 >>>

    Parameter      Value      Precision     Correction   Significance    Max. Effect
                              (millimetres)                              (microns)

        1         -0.0626       0.0015       -0.0001        0.04
        2         -0.0534       0.0035       -0.0000        0.01
        3         12.6806       0.0042        0.0003        0.08
        4       -1.1408e-003   1.424e-005  -1.129e-006      0.08          -160.4
        5        8.5689e-006   4.537e-007   3.201e-008      0.07            32.6
        6       -9.4123e-008   5.302e-008   0.000e+000      0.00            -9.7
        7       -3.8429e-005   1.461e-005   0.000e+000      0.00            -1.0
        8       -2.2393e-005   1.448e-005   0.000e+000      0.00            -0.6

            Correlation Parameters for Camera Calibration Set: 1

        1    1.00
        2   -0.04  1.00
        3    0.12  0.70  1.00
        4   -0.25 -0.02 -0.29  1.00
        5    0.11  0.18  0.31 -0.93  1.00
```

Appendix C - Figure 5: Camera calibration parameters and correlation coefficients - system CAM_P2, dataset C1.
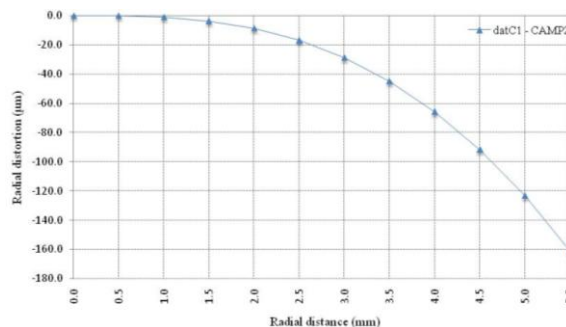
Appendix C - Figure 6: Radial les distortion profile - system CAM_P2, dataset C1.

C.4. Comparative radial lens distortion profiles

In chapter 6 (see section 6.5.2.) the affine bundle adjustment was processed in order to evaluate method behaviour in object space. This was done by application of the method in datasets B2, D2 and E2. The bundle adjustment was processed with the external constraints datum and the implemented calibration model (third power term of the radial lens distortion polynomial). Results from this test have already been assessed in detail. This section provides a display of the radial lens distortion profiles for the three implemented datasets B2, D2 and E2 (see Figure 7). It is clearly evident that the target occupancy varies within the image frame for radial lens distortion calculation reaching a maximum of 43.3μm at a radial distance of 4.5mm.
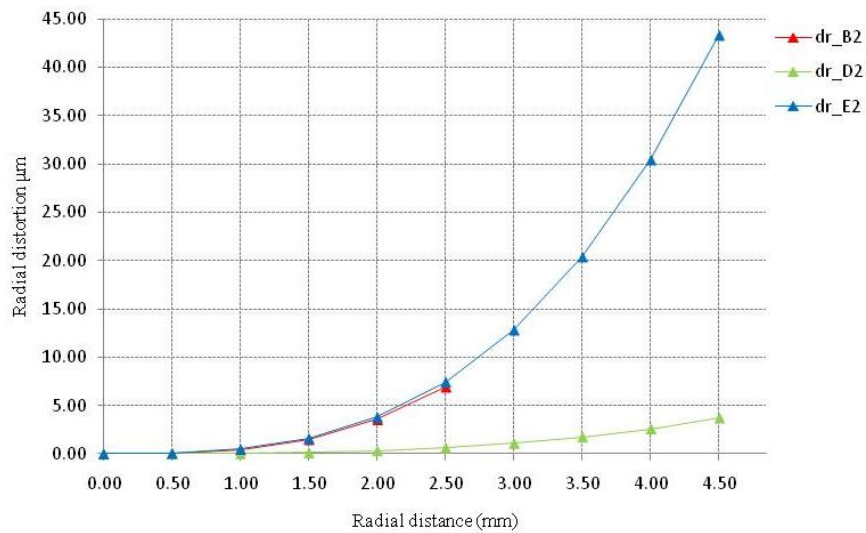
Appendix C - Figure 7: Comparative radial lens distortions – datasets B2, D2 and E2.