DEVELOPING AN ACCURATE CLOSE-RANGE PHOTOGRAMMETRIC TECHNIQUE FOR EXTRACTING 3D INFORMATION FROM SPHERICAL PANORAMIC IMAGES

By
Naa Dedei Tagoe
TGXNAA01

Submitted to the University of Cape Town in fulfilment of the requirements for the Degree of Doctor of Philosophy in Engineering (Geomatics)

<u>Supervisors:</u>

Emeritus Prof H. Rüther

Assoc Prof J. Smit

Department of Architecture, Planning and Geomatics
Faculty of Engineering and Built Environment
University of Cape Town

November, 2016

DECLARATION

I, Naa Dedei Tagoe, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: Signed by candidate

Signature removed

Date: November 23, 2016

ACKNOWLEDGEMENT

I would like to thank my supervisors Professors Heinz Rüther and Julian Smit for their excellent guidance throughout the research period, for their continuous encouragement, and for always being approachable and taking time to discuss problems, despite other equally important responsibilities. Working with them was a real pleasure and I benefited immensely, both on a professional and a personal level.

Special thanks to Professor Heinz Rüther for not just being a supervisor but a mentor and counsellor who believes in his students. Without your support, I would have never got this far. I am also grateful to you for giving me the opportunity to conduct my research under the Zamani Research Group of which you are the lead investigator. I am grateful to the other team members of the Research Group especially, Roshan Bhurtha, Ralph Schröder and Stephen Wessels for their technical advice and support at various stages of the research. It was a great opportunity to experience the team spirit as I joined you for field campaigns especially in Ghana in collaboration with the Ghana Museums and Monuments Board (GMMB).

Further thanks are due to Professors Alfred Duker and Newton Amegbey for reading through the thesis and for their positive criticism.

I am very thankful to my family especially my husband, parents and brother for their moral support and daily encouragement, especially when difficulties seemed overwhelming.

My sincerest gratitude goes to the Schlumberger Foundation and the South African National Research Foundation for sponsoring my PhD research.

ABSTRACT

Panoramic images (panoramas) are wide-angle images that provide fields of view of up to 360°. They are acquired with a specialised panoramic camera or by stitching a series of images captured with a conventional digital camera. Panoramas have widely been used to texture 3D models generated from laser scanning, for creating virtual reality tour applications, documenting landscape and cultural heritage sites, advertising real estates and recording crime scenes.

The goal of this research was to develop an accurate close-range photogrammetric technique for the semi-automatic extraction of 3D information from spherical panoramas. This was achieved by developing a non-parametric method for the removal of distortions from images acquired from fisheye lenses as well as an algorithm, here referred to as the Minimum Ray Distance (MRD), for the fully automated approximate relative orientation of spherical panoramic images. The bundle adjustment algorithm was then applied to refine the orientation parameters of the panoramas; thus enabling accurate 3D point measurement. Finally, epipolar geometry theory was applied to the oriented panoramas to guide the interactive extraction of additional conjugate points.

The MRD algorithm has been extended to laser scanning technology for the first approximations of laser scan setup positions and scan orientation prior to a least-squares based registration. The determination of approximate scanner orientation and position parameters were accomplished using panoramic intensity images derived from full dome laser scans.

Thus, a technique for the semi-automatic extraction of 3D measurements from panoramic images has been developed in this research. The technique is most appropriate for applications which do not require dense point clouds and in situations with limited access to funds or as a quick field method to document many features in a short time. This is because a single image orientation is required for several overlapping images as compared to the normal stereo or multi-image photogrammetric approach. It is not suggested

that 3D reconstruction from spherical panoramic images should replace traditional close-range photogrammetry or laser scanning; rather, that the user of panoramic images will be offered supplementary information to the conventional and modern cultural heritage documentation approaches.

DEDICATION

I dedicate this work to my husband Dr Saviour Mantey, my daughters Jessica Nana Akua Mantey and Victoria Nana Yaa Mantey, my mum, the Late Yacoba Tagoe, my dad Mr Eric Teiko Tagoe, my brother Nii Armah Tagoe and all my family members.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1. Motivation and Problem Statement

Rapid advancements in imaging technology and relatively affordable, conventional digital cameras have contributed to the growing interest in panoramic imaging and photogrammetry.

Panoramic images[1] are wide-angle images that can provide horizontal fields of view of up to 360°. They can be acquired using low-cost techniques such as stitching a series of images captured from a single position with a conventional digital camera or with expensive specialised panoramic cameras such as the rotating line cameras. Panoramic images have a wide range of applications. Common examples include; texturing of 3D models generated from laser scanning (Sequeira and Goncalves, 2002; Klette and Scheibe, 2005) and creating virtual reality tours (Chen, 1995). Virtual reality tours generated from panoramic images have been used to record crime scenes (Chan, 2005; Toet and van Schaik, 2012), document landscape and cultural heritage sites (Addison, 2000; Koehl *et al*., 2013; Wessels *et al*., 2014), advertise real estates (Jongerius, 2003), *etc*. Beside the visualisation and navigational capabilities, virtual reality tour applications can be enhanced with the ability to make 3D measurements if such functionalities are incorporated. Photogrammetry provides such additional 3D measurement capability.

“Photogrammetry is defined as the art, science and technology of obtaining reliable information about physical objects and environment through

---

[1] The term panoramic image was coined out of a general term panorama. Panorama originates from two Greek words, “πᾶν” (pan) meaning "all" and “ὄραμα” (horama) meaning "view. In this thesis the two terms are used interchangeably.

processes of recording, measuring and interpreting photographic images and patterns of electromagnetic radiant energy and other phenomena" (Slama, Theurer and Henriksen, 1980). Close-range photogrammetry is a specialised branch of photogrammetry where the object-to-camera distance is not more than 300 m (Kemp, 1996). However, the value of 300 m for defining distance might have to be revised in view of the increasing use of drones or Unmanned Aerial Vehicle (UAV) which capture aerial images at lower flying heights. Close-range photogrammetric techniques have widely been used in the fields of archaeology, architecture, accident and crime investigations, engineering, medicine and mining.

Over the past two decades, panoramic images have been used in close-range photogrammetry. Early researchers mostly relied on cylindrical panoramic images generated from rotating line panoramic cameras (Luhmann and Tecklenburg, 2004; Maas and Schneider, 2004) or spherical panoramic images generated by stitching images captured with conventional cameras and rectilinear lenses (Fangi, 2007; 2010). The photogrammetric processing of panoramic images is generally referred to as Panorama photogrammetry. The term Spherical photogrammetry was however introduced by G. Fangi when processing spherical panoramic images (Fangi, 2007).

In Panorama (or Spherical) photogrammetry, the bundle adjustment algorithm is used to simultaneously determine the exterior orientation parameters (EOPs) of the panoramic images and the 3D coordinates of the object points (Luhmann and Tecklenburg, 2004; Maas and Schneider, 2004; Fangi, 2007; 2010). The mathematical model of the bundle adjustment algorithm, however, is non-linear, thus requiring good initial approximations. In Maas and Schneider (2004), the initial approximations were obtained by applying sequential Cassini-resection and spatial intersection. This approach of generating initial approximations for the exterior orientation parameters rely on known coordinates of control points which makes the orientation phase cumbersome and time consuming, especially for non-technical operators. Fangi (2007, 2010) generated initial approximations by applying relative orientation based on the coplanarity equations, followed by the transformation of independent models into a global coordinate system.

Like the collinearity equations, the coplanarity equations are non-linear, therefore they require good initial approximation values to solve the unknown parameters. Fangi (2007, 2010) resolved this challenge by defining arbitrary values for unknowns (orientation angles). The approximation values are subsequently refined until the iterative least squares computation converges. This process of defining the approximation values for the unknown parameters makes the orientation process cumbersome.

The research work described in this thesis therefore sought to develop a simple-to-use close-range photogrammetric technique for extracting 3D information from spherical panoramic images. Special emphasis was placed on the development of a new orientation procedure for generating initial approximation values for exterior orientation parameters of the panoramic images. The initial approximation values required to solve the exterior orientation parameters in the research are neither arbitrarily defined nor does the orientation procedure rely on coordinates of known control points.

As in the case of Fangi (2007, 2010), this research focussed on relatively low-cost techniques, thus conventional as opposed to panoramic cameras were chosen for the panorama photography. Fisheye lenses, however, were used in this research because the need to stitch several images captured from rectilinear lenses was not only considered time consuming but could possibly introduce errors in the image stitching process. A literature search, however, showed that very few researchers in the photogrammetric community have worked with fisheye lenses (Schwalbe, 2005; Van Den Heuvel *et al.*, 2006; Schneider, Schwalbe and Maas, 2009). The reason could be attributed to the fact that fisheye lenses are incompatible with already existing conventional photogrammetric software packages. The incompatibility is because the projection model for transforming 3D rays to 2D image in conventional photogrammetric software packages is central perspective while that for fisheye lenses are non-perspective. Since there are various non-perspective models (stereographic, equidistant, equisolid-angle and orthogonal), a non-parametric approach to lens distortion model can serve as a universal solution. Thus, the research work described in this thesis will explore the development of an image rectification method for the corrections of images acquired with fisheye lens systems prior to stitching the final panoramic images.

1.2. Research Background

This research work is a contribution to the activities of the Zamani project (The African Cultural Heritage Site and Landscapes Database), based at the Geomatics Division of the University of Cape Town. The aim of the Zamani project is to create a permanent metrically accurate record of important African heritage sites, to provide data for restoration and conservation purposes and to create awareness of African Heritage (Rüther, 2002). The project supports the work of researchers, archaeologists, architects and conservationists by using precise and well-established conventional survey techniques, as well as laser scanning and photogrammetry to document heritage sites (Rüther *et al.*, 2009). Interactive databases and GISs are created, supported by 3D terrain models of the immediate vicinity of the site and contextual information such as scientific publications, 3D photorealistic models of heritage structures and monuments, historical maps and general photography of current activities in and around the site (Rüther, 2002, 2011; Rüther *et al.*, 2012).

Panoramic images are an important component of the Zamani database with their use presently being restricted to visualisation of heritage sites. While laser scanning is widely used in the Zamani project for capturing dense point cloud data, it has substantial financial implications and can thus not be seen as a generally applicable tool in an environment with limited economic resources. Therefore, there is the need to develop a low-cost and accurate solution for applications which do not require dense point clouds, in situations with limited access to funds or as a quick field method to document many features in a short time. As a possible solution for such cases, this research investigates the use of panoramic images as a tool for extracting 3D information of heritage structures. The developed approach is suitable for quick 3D measurement from existing panoramic images but can also serve as a full documentation tool where laser scanning is not available.

1.3. Research Aim and Objectives

The aim of this research is to develop the concept for a low-cost and accurate close-range photogrammetric technique for extracting 3D information from spherical panoramic images.

The specific objectives which contribute to the final solution are:

(i) to develop a method for the removal of distortions from images acquired from fisheye lenses for the subsequent use of the thus corrected images in the creation of spherical panoramas;

(ii) to develop an algorithm for the approximate automated relative orientation of multiple spherical panoramas and extract 3D coordinates from the oriented images; and

(iii) to extend the relative orientation algorithm, as proposed in (ii) above, to laser scanning technology for the first approximations of laser scan setup positions and scan orientation prior to a least-squares based registration.

1.4. Hypothesis and Research Questions

Hypothesis statement: An automated process for the generation of 3D object coordinate information from spherical panoramic images will provide effective, alternative and supplementary means to standard photogrammetric 3D reconstruction. The same model can also be employed for the first approximation of laser scan setups.

In accordance with the aims and hypothesis statements, the research seeks to answer the following specific questions;

(i) Can a non-parametric model be used to accurately rectify images captured with fisheye lenses?

(ii) Can spherical panoramic images be generated from the images rectified as proposed in (i) above?

(iii) Can the initial approximation values for exterior orientation parameters of the panoramic images and 3D object point coordinates be generated automatically?

(iv) In an attempt to recover 3D coordinates from oriented spherical panoramic images, how can the epipolar geometry be applied to panoramic images and used to reduce the search space?

(v) Can the relative orientation algorithm developed for spherical panoramic images be used in laser scan registration?

1.5. Contribution of the Research

The following constitute novel contribution to knowledge:

(i) A non-parametric image rectification method to correct images acquired with both rectilinear and fisheye lens systems.

(ii) A new algorithm here referred to as the Minimum Ray Distance (MRD) for the fully automated approximate relative orientation of multiple spherical panoramic images. The MRD algorithm generates good initial approximation for all the unknown parameters for refinement by the bundle adjustment algorithm.

(iii) Extension of the MRD algorithm to laser scan technology for the first approximations of laser scans setups prior to a least squares-based registration.

In addition to these contributions, the epipolar geometry for spherical panoramic images has been implemented in the developed software to guide the interactive and automatic extraction of 3D information from two or more panoramas.

1.6. Structure of the Thesis

This thesis is organised into seven chapters. The first chapter covers the motivation and problem statement, research background, objectives, hypothesis, research questions and the main contributions of the research to the scientific community.

Chapter 2 reviews relevant literature and provides a brief theoretical background to seminal aspects of the research. Topics discussed in this chapter include the acquisition of panoramic images, panoramic image stitching, feature-based matching techniques, lens distortion modelling and the photogrammetric bundle adjustment algorithm.

Chapter 3 discusses the development of a new technique for rectifying images acquired from fisheye lenses for the subsequent use in the creation of spherical panoramic images. Experimental results and analysis of the developed approach are also presented.

Chapter 4 focuses on procedures for the orientation and extraction of 3D object coordinates from spherical panoramas. A feature-based matching approach to automatically extract conjugate feature points from spherical panoramas is presented. The MRD algorithm is then presented while the well-known bundle adjustment algorithm and the epipolar geometry are extended to spherical panoramic images.

Experimental tests and results on the orientation and extraction of 3D object coordinates from spherical panoramas discussed in the previous chapter are presented in Chapter 5.

Chapter 6 describes the extension of the MRD algorithm to laser scan technology for the purpose of scan registration.

Finally, Chapter 7 draws conclusions about the research and recommends future work.

CHAPTER 2

LITERATURE REVIEW AND THEORETICAL BACKGROUND

2.1. Acquisition of Panoramic Images

The Irish painter, Robert Barker in the 1780s used the term “panorama’’ to describe his wide-angle view paintings of Edinburgh (Oettermann, 1997; Lam, 2007; Huang, Klette and Scheibe, 2008). Thus, prior to the invention of early panoramic cameras such as J. Puchberger’s hand crank swing lens camera in 1843 and M. Grarella’s rotating lens camera in 1857, panoramas were already in existence (McBride, 1994; Oettermann, 1997; Huang, Klette and Scheibe, 2008; Faugeras, Benosman and Kang, 2013).

Today, the term panorama is collectively used to refer to a wide-angle pictorial representation of a scene visible from a single point of view. This pictorial representation could be in the form of a drawing, film, painting or a photograph. For in-depth historical background on panoramas, an interested reader is referred to McBride (1994), Oettermann (1997), Huang, Klette and Scheibe (2008); Faugeras, Benosman and Kang (2013).

Panoramic images are widely used in virtual reality tour applications such as the Google’s ‘World Wonders Project’ (Morisse, 2012) and the Cave Automatic Virtual Environment (Creagh, 2003). The ‘World Wonders Project’ allows Internet users to explore World Heritage Sites by navigating through 360° spherical panoramic views provided by the Google Street View Services. In the Cave Automatic Virtual Environment, a head-mounted, wide-view stereo display with interactive controls is used to view projected images in an immersive way.

Panoramic images are also used as complimentary data in laser scanning technology to texture 3D models (Sequeira and Goncalves, 2002; Klette and Scheibe, 2005; Held, 2012). Panoramic images have been applied in the areas of collision avoidance (Shoval, Borenstein and Koren, 1994), surveillance and security monitoring (Thibault, 2007; Wong *et al.*, 2009), medical imaging (Honey *et al.*, 2007; Estrela *et al.*, 2008) and robotic navigation (Gaspar, Winters and Santos-Victor, 2000; Briggs *et al.*, 2006).

2.1.1. Imaging Systems and Techniques for Capturing Panoramic Images

Panoramic imaging systems can be broadly grouped into two; - catadioptric and dioptric camera systems[2] (Figure 2-1). Dioptric camera systems (Figures 2-3 and 2-4) make use of lenses with field of views (FOVs) sometimes slightly exceeding 180°.

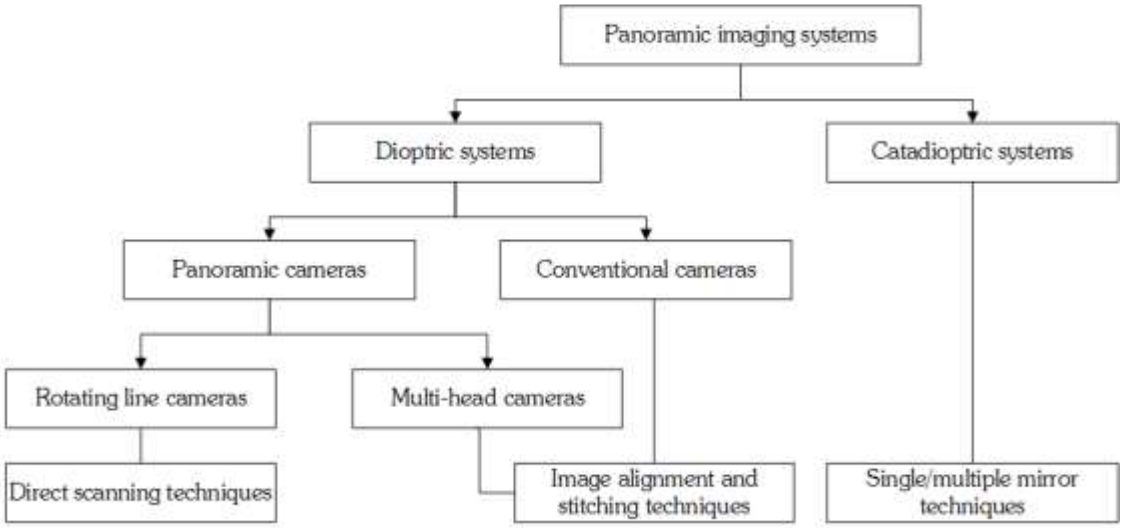


Figure 2-1 Imaging systems and techniques for capturing digital panoramic images

Catadioptric camera systems on the other hand combine lens systems with one or two curved mirrors (*e.g.* parabolic, hyperbolic mirrors) to record an

---

[2] The terms 'dioptrics' and "catoptrics" refer to the use of refracting elements (lenses) and reflecting surfaces (mirrors) respectively. Thus, 'catadioptrics' is the combination of lenses and mirror(s) in an imaging system (Hecht and Zajac, 1974) .

entire hemisphere in a single shot. An example of the catadioptric camera system is shown in Figure 2-2.

Figure 2-2 Catadioptric camera system with a hemispherical FOV (Nayar, 2001)

The drawbacks in using Catadioptric camera systems include lower image resolution than that of the conventional digital camera used in constructing them (Baker and Nayar, 1999), self-occluded areas in each image and defocuss blur resulting from the use of curved mirrors (Baker and Nayar, 1999; Huang, Klette and Scheibe, 2008). For these reasons, Catadioptric camera systems are often not used for close-range photogrammetric applications (Huang, Klette and Scheibe, 2008).

Dioptric camera systems are further discussed. These cameras are categorised into two:- panoramic and conventional digital cameras.

Panoramic Camera System

The rotating line and multi-head cameras are two categories of panoramic cameras. The rotating line camera is made up of a linear array charge-coupled device (CCD) sensor which is placed on a turntable with a direct current (DC) motor (Parian and Gruen, 2010). As the turntable rotates, the sensor captures a 360° horizontal FOV as a continuous set of vertical RBG scan lines (Huang, Klette and Scheibe, 2008; Parian and Gruen, 2010). The vertical FOV of the camera system, however, depends on the focal length of the lens in use and the size of the CCD chip (Parian and Gruen, 2010). Examples of rotating line cameras include Panoscan MK-3[3] and EYESCAN-M3/MM3[4] which are based on spherical and cylindrical projections respectively.

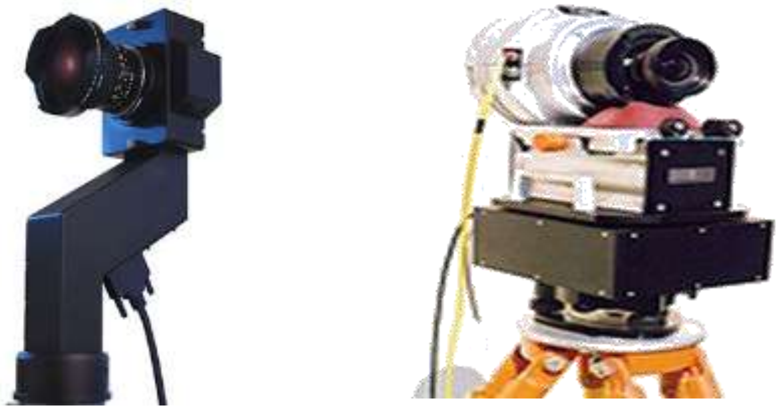

(a) (b)

Figure 2-3 Rotating line panoramic cameras: (a) Panoscan (Panoscan, 2007) (b) Eyescan-M3 (Eyescan, 2015)

The advantage of these cameras is their ability to capture high resolution seamless panoramic images without the need to stitch. However, due to the scanning nature of their image capture, the rotating line camera is suitable for recording static scenes while the camera is mounted on a steady tripod. The multi-head camera system is a group of cameras arranged on a platform. Examples of these multi-head camera systems are Ladybug 5[5] and Dodeca 2360[6] which have six and twelve cameras respectively.

---

[3] Panoscan MK-3 is produced by Panoscan Inc. - *www.panoscan.com/*

[4] EYESCAN-M3/MM3 is a joint co-operation between Kamera and System Technik, Germany and DLR, the German Aerospace Centre - www.kst-dresden.de/

[5] Ladybugs 5 is produced by Pointgrey Research Inc. - *www.ptgrey.com/*

[6] Dodeca 2360 is produced by Immersive Media. - *www.immersivemedia.com/*


(a) (b)

Figure 2-4 Multi-head panoramic cameras: (a) Ladybug 5 (Pointgrey, 2015) (b) Dodeca 2360 (Agarwal, 2007)

Multi-head camera systems have the ability to capture an entire scene with simultaneous shots from all the cameras. The individual images are subsequently stitched to form the panoramic image. However, the use of multiple cameras makes it sometimes impossible to ensure a singular perspective centre, thus, leading to parallax errors during the stitching of the panoramic images (Swaminathan and Nayar, 1999).

While panoramic cameras such as the rotating line and multi-head cameras represent a convenient means of capturing panoramic images, they are very expensive, thus making them unsuitable for low-budget projects.

Conventional Digital Camera Systems

The most practical option of generating accurate panoramic images with 360° horizontal FOV is by stitching three or more partially overlapping images captured from a single point of view as the conventional digital camera rotates about the external nodal point of its lens system. The basic hardware and software required to generate panoramic images include the camera lens and body, a panoramic head (rotator), a tripod and an image stitching software.

There are three types of conventional digital cameras: compact or ‘point and shoot’, digital single-lens reflex (DSLR) and the prosumer or ‘bridge’. The principal differences between the three cameras are the view finder mechanism, the type of lens, whether fixed or interchangeable and the size of the camera sensor. Compact cameras are generally small and light weight. They have an LCD monitor or an electronic viewfinder (EVF) or both. The operator sees a slightly displaced view compared to that of the camera’s sensor (Busch, 2000; Hirsch, 2012). The lens of the compact camera is fixed to the body of the camera. Often the camera sensors of the compact cameras are the smallest of the three camera types. DSLR cameras are relatively large and typically heavier than compact cameras. They have an optical viewfinder, made up of a mirror and a pentaprism system which enable the operator to see the same field of view as the camera sensor. The lens of a DSLR camera can be interchanged, thus one camera can be used for a wide range of photography. The size of the camera sensor is generally large (Busch, 2000; Hirsch, 2012). The bridge camera is an intermediate step between the compact and DSLR camera. It has a similar size and weight as the DSLR camera but has no interchangeable lens. The DSLR cameras are preferred over compact and bridge cameras because of their ability to interchange lenses and their relatively large sensor size which results in better image quality.

A rectilinear or fisheye lens may be attached to the DSLR camera body to capture the individual images. A major difference between these two types of lenses is the kind of image they produce. Rectilinear lenses cause straight lines in object space to be projected almost as straight lines in the image space while straight lines in object space appear to be curved in the image space when captured with a fisheye lens. Fisheye lenses can be categorised into two: - circular or true fisheye lenses and full-frame fisheye lenses. Figure 2-5 shows images captured by both rectilinear and fisheye lenses.

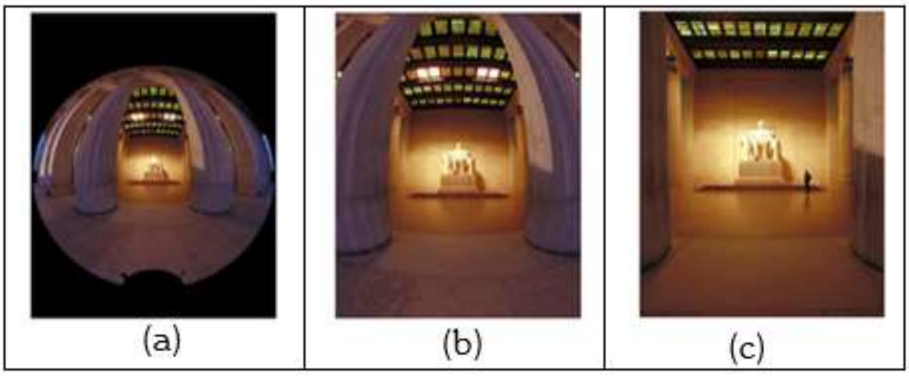


Figure 2-5 Image captured by rectilinear and fisheye lenses: (a) Circular Fisheye (b) Full-frame fisheye (c) Rectilinear (Adapted from Highton, 2008)

The circular fisheye lens captures a hemispherical view which is surrounded by the remaining unexposed sensor area with a FOV of 180° in all directions. On the other hand, the full-frame fisheye lens exposes the entire sensor area. The lens has a FOV of 180° in the diagonal direction, while the horizontal and vertical FOVs are less than 180°. Figure 2-6 shows the exposed sensor area of both circular and full-frame fisheye lenses.

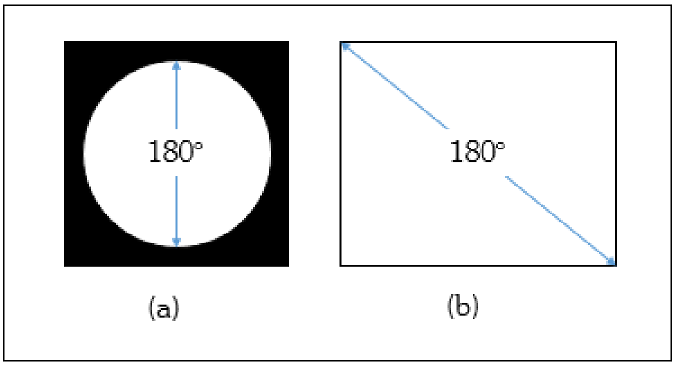


Figure 2-6 Exposed sensor area of fisheye lenses: (a) Circular fisheye (b) Full-frame (Adapted from Gløckner, 2013)

Preferably, the camera should be attached to a panoramic tripod head and placed on a firm tripod to enable it rotate about the external perspective centre of the camera's lens system. Examples of panoramic heads available on the market are Nodal ninja N3[7] and the Manfrotto 303SPH QTVR[8].

The individual images captured with the camera lens and body are rectified to remove lens distortions and subsequently stitched into a seamless panorama using a stitching software. Examples of commercial stitching software include PTGui[9] and Autopan[10].

2.1.2. Image Alignment and Stitching Techniques

Image alignment algorithms, also known as registration, are used to establish a set of geometric transformations that relates pixel coordinates of overlapping images while image stitching algorithms take these alignment estimates and blend the images seamlessly (Szeliski, 2006). These algorithms ensure that blurring or ghosting caused by parallax and scene movement as well as varying image exposures are minimised or largely eliminated (Szeliski, 2006).

In the mid-1990s, image alignment and stitching techniques for generating panoramas became common. These panoramas were used mainly for virtual reality applications such as the Apple's QuickTimeVR system (Szeliski, 2006).

The approach to aligning and stitching overlapping images to generate panoramas involve three most principal steps: - local registration, global registration and compositing.

Local registration techniques relate pixel coordinates of a pair of neighbouring images. These techniques can be categorised into two: direct or area-based techniques (Szeliski and Kang, 1995) and feature-based techniques (Zoghlami, Faugeras and Deriche, 1997; McLauchlan and Jaenicke, 2002).

---

[7] Nodal ninja N3 is produced by Bill Bailey L.L.C. – www.shop.nodalninja.com
[8] Manfrotto 303SPH QTVR is produced by Vitec Group Italia Spa - www.manfrotto.us
[9] PTGui is produced by New House Internet Services BV - www.ptgui.com
[10] Autopan is produced by Kolor SAS - www.kolor.com

For the area-based methods, pixel-to-pixel dissimilarities between neighbouring images are minimised by shifting images relatively to each other until the differences between matched pixels are minimised (Szeliski, 2006).

The area-based approach involves selecting a matching entity in one image and searching throughout the other image for its conjugate entity using for example a cross-correlation technique. Searching through the entire image is highly computation intensive, thus a hierarchical coarse-to-fine approach (based on image pyramids) or a Fourier transform can be used to speed up the computation (Szeliski, 2006). Area-based methods are generally accurate. However, they require very good initial approximation and will fail if images components to be matched are at different scales, illumination and orientations. Feature-based registration involves extracting distinctive features from the image pair and matching these features. Feature-based methods are often preferred over area-based methods because local feature operators such as Scale Invariant Feature Transformation (SIFT) (Lowe, 2004) and Speeded Up Robust Features (SURF) (Bay, Tuytelaars and Van Gool, 2006) are invariant to scale changes, illumination, robust against scene movement and are faster, if rightly implemented (Szeliski, 1996, 2006; Brown and Lowe, 2003). The feature-based registration technique using the SIFT operator is discussed in Section 2.2. The procedure involves matching features and estimating a geometric transformation matrix that maps a greater number of the feature points between the two images.

After the pairwise registration stage, a global consistent set of alignment parameters is applied to minimise the mis-registration between all pairs of images (Sawhney, 1999; Coorg and Teller, 2000; Szeliski, 2006). This is referred to as global registration. Global registration is achieved by applying a bundle adjustment algorithm based on Levenberg-Marquardt (Triggs *et al*., 1999) to calibrate the focal length and the EOPs of the camera (Szeliski, 2006).

A compositing surface or a cartographic projection is then chosen to project and stitch the aligned images. Examples of cartographic projections widely used are the cylindrical, the equirectangular and the cubic projections. During the compositing phase, blending algorithms are also applied to ensure smooth transitions between neighbouring images in the presence of illumination differences (Brown and Lowe, 2003; Szeliski, 2006). Examples of blending algorithms include Laplacian pyramid blending (Burt and Adelson, 1983a; Shum and Szeliski, 2002) and multi-band blending (Burt and Adelson, 1983b; Brown and Lowe, 2003; Luo *et al.*, 2011). Figure 2-7 shows the workflow for stitching panoramic images using feature-based image alignment techniques.

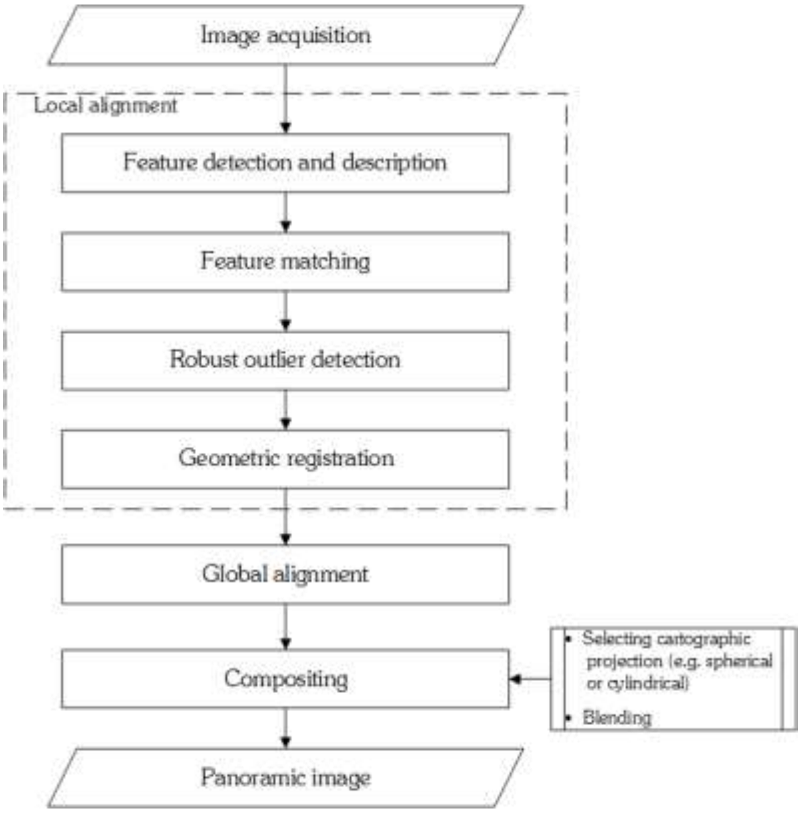


Figure 2-7 Workflow for feature-based image alignment and stitching

2.2. Overview of the SIFT Feature Matching Procedure

The procedure for extracting conjugate feature points from a pair of images using the SIFT operator involves feature detection and description, feature matching and outlier detection.

2.2.1. Feature Detection and Description

The SIFT operator plays the role of a detector and descriptor. A feature detector is used to extract distinctive features that are geometrically stable under various transformations while the descriptor is a vector of variable length that is used to uniquely describe a patch around the detected feature (Remondino, 2006). Potential features that are invariant to scale and orientation are identified on the entire image, using a difference-of-Gaussian function. Features with low contrast are filtered out. This is followed by calculating a descriptor for the local patch around the detected feature that is distinctive and invariant to properties such as change in illumination and perspectives. For detailed information about the SIFT operator an interested reader is referred to Lowe (1999, 2004).

2.2.2. Feature Matching

Once distinctive features have been detected and described in the image pair, conjugate points may be found by comparing descriptors. The comparison is often based on a nearest-neighbour search, without having any prior information about the geometry of the images. The search for conjugate point can be performed exhaustively by evaluating all possible matches between the two descriptor vectors d_a and d_b, extracted from images I_a and I_b respectively. This approach may lead to descriptors having more than one possible match as well as false matches.

Alternatively, a less computationally intense approximate approach involves using a k-d tree algorithm (Friedman, Bentley and Finkel, 1977). An example is the Best-Bin-First (BBF) method as proposed by Beis and Lowe (1997). For each descriptor in d_a, the BBF method finds the nearest descriptor in d_b by calculating the Euclidean distance between the descriptors given a threshold. This matching approach leads a one-to-one mapping of descriptors with a high possibility of mismatches especially for repetitive and less distinctive features. The resulting matches can further be filtered for uniqueness by specifying a threshold value which is measured as the ratio of the distance between the best matching feature, D_{b1} and the distance to the second best matching feature, D_{b2}.

2.2.3. Outlier Detection

For conjugate feature points that are extracted automatically, a number of outliers (false matches) exist. These outliers are often removed using iterative robust estimators such as the Random Sample Consensus (RANSAC) (Fischler and Bolles, 1981), Least Median of Squares (LMedS) (Leroy and Rousseeuw, 1987), Maximum A Posteriori Sample Consensus (MAPSAC) (Torr, 2002; Torr and Davidson, 2003) and PROgressive SAmple Consensus PROSAC (Chum and Matas, 2005).

The RANSAC algorithm is widely used as it is able to cope with a large proportion of outliers (Hartley and Zisserman, 2003). The RANSAC algorithm may be summarised as in Table 2-1.

Table 2-1 RANSAC algorithm (Fischler and Bolles, 1981; Hartley and Zisserman, 2003)

(i) From a set of data points, randomly select a sample to determine the parameters of the model.

(ii) Estimate the parameters of the model using the sample data.

(iii) Determine how many points from the sample data fits the model. (these are the inliers.)

(iv) If the number of inliers exceeds the given threshold value, re-estimate the model parameters using all the inliers and terminate.

(v) If the number of inliers is less than the given threshold value, select a new sample and repeat (i) through (iv) for a given maximum number of trials.

2.3. Lens Distortion Modelling

The most widely used parametric lens distortion model adopted by the photogrammetric community was proposed by D. C. Brown and dates back to the era of Analytical Photogrammetry.

In 1966, Brown published a paper on decentering lens distortion in which he reviewed and reconciled two earlier, conflicting theories: the thin prism and Conrady’s model. He demonstrated in his paper that with regards to the tangential component of decentering distortion, Conrady's model was in agreement with the thin prism model; however, both models were at variance with regards to the radial component. Based on Conrady's model which was formulated using the rigorous analytical ray tracing through a decentered lens, Brown proposed an alternative model for decentering distortion as follows (Equations 2-1 and 2-2):

$$\Delta x = [P_1(r^2 + 2\bar{x}^2) + 2P_2\bar{x}\bar{y}][1 + P_3 r^2 + P_4 r^4 \dots] \qquad \text{Equation 2-1}$$

$$\Delta y = [2P_1\bar{x}\bar{y} + P_2(r^2 + 2\bar{y}^2)][1 + P_3r^2 + P_4r^4 + \cdots] \qquad \text{Equation 2-2}$$

in which, $\bar{\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{x_o}$, $\bar{\boldsymbol{y}} = \boldsymbol{y} - \boldsymbol{y_o}$. In Equations 2-1 and 2-2, **x** and **y** are the measured image coordinates, $\boldsymbol{x_o}$ and $\boldsymbol{y_o}$ are the image coordinates of the principal point, **P1**, **P2** , **P3** and **P4** are the coefficients of decentring distortion and $\boldsymbol{r}$ is the radial distance from the principal point $\left(\boldsymbol{r} = [(\boldsymbol{x} - \boldsymbol{x_o})^2 + (\boldsymbol{y} - \boldsymbol{y_o})^2]^{1/2}\right)$.

Brown (1971) showed that the image of a photographed plumbline would have been a perfect straight light in the absence of radial and decentering lens distortion and therefore proposed the 8-parameter lens distortion model for the correction of image coordinates ($\Delta\boldsymbol{x}$,$\Delta\boldsymbol{y}$) as follows (Equations 2-3 and 2-4):

$$\Delta x = \bar{x}(K_1r^2 + K_2r^4 + K_3r^6 + \cdots) + [P_1(r^2 + 2\bar{x}^2) + 2P_2\bar{x}\bar{y}][1 + P_3r^2 + \cdots]$$

Equation 2-3

$$\Delta y = \bar{y}(K_1r^2 + K_2r^4 + K_3r^6 + \cdots) + [2P_1\bar{x}\bar{y} + P_2(r^2 + 2\bar{y}^2)][1 + P_3r^2 + \cdots]$$

Equation 2-4

where **K1, K2** and **K3** are the coefficients of radial distortion.

In Fraser, Shortis and Ganci (1995) and Fraser (1997), a 10-parameter lens distortion model was proposed. Fraser’s model departs from Brown's ‘standard’ 8-parameter model by ignoring **P3** and including a term to estimate camera constant and two terms to account for first-order in-plane image distortion. Equations 2-5 and 2-6 outlines Fraser’s 10-parameter lens distortion model:

$$\Delta x = -\mathrm{x_o} - \frac{\bar{x}}{c}\Delta c + \bar{x}r^2K_1 + \bar{x}r^4K_2 + \bar{x}r^6K_3 + (2\bar{x}^2 + r^2)P_1 + 2P_2\bar{x}\bar{y} + b_1\bar{x} + b_2\bar{y}$$

Equation 2-5

$$\Delta y = -y_o - \frac{\bar{y}}{c}\Delta c + \bar{y}r^2K_1 + \bar{y}r^4K_2 + \bar{y}r^6K_3 + 2P_1\bar{x}\bar{y} + (2\bar{y}^2 + r^2)P_2$$

Equation 2-6

where, **c** is the principal distance (camera constant), **Δc** represents a correction to the initial principal distance while $\boldsymbol{b_1}$ and $\boldsymbol{b_2}$ are differential scaling and non-orthogonality between the x- and y-axes.

While the above models (Equations 2-3 to 2-6) have proved efficient for rectilinear lenses, they failed when applied to fisheye lenses. This is because the projection models for transforming 3D rays to 2D image for a rectilinear lens is central perspective while that for fisheye lenses are non-perspective. Examples of the non-perspective projections are include stereographic, equidistant, equisolid-angle and orthogonal projections (Miyamoto, 1964).

With the exception of a few publications on fisheye lens distortion modelling (Schwalbe, 2005; Van Den Heuvel *et al.*, 2006; Schneider, Schwalbe and Maas, 2009), the majority of literature in this area of research are found outside the photogrammetric community (Basu and Licardie, 1995; Devernay and Faugeras, 2001; Kannala and Brandt, 2004, 2006; Hughes *et al*., 2010).

In this research, a non-parametric alternative to lens distortion modelling for rectifying fisheye images is presented. The developed approach estimates the distortion characteristics of a lens-camera combination by comparing near-error-free simulated image point positions of targets of a calibration testfield with the real image of these targets (Tagoe, Rüther and Smit, 2014).

2.4. Bundle Adjustment Algorithm

Bundle adjustment is the most widely used algorithm in photogrammetry for estimating the EOPs of a camera and the 3D coordinates of object points. The algorithm was first developed by (Brown, 1956, 1976; Brown, Davis and Johnson, 1964) and Schmid (1955) and has since undergone numerous improvements and additions (Wester-Ebbinghaus, 1980; Kruck, 1998; Fraser, 1997). The bundle adjustment algorithm is based on the collinearity equations for perspective images. The standard form of the collinearity equations for perspective images is given as in Equations 2-7 and 2-8 (McGlone, 1989; Mikhail, Bethel and McGlone, 2001):

$$x_i - x_o = c\,\frac{m_{11}(X_j - X_{\mathbf{o}}) + m_{12}(Y_j - Y_{\mathbf{o}}) + m_{13}(Z_j - Z_{\mathbf{o}})}{m_{31}(X_j - X_{\mathbf{o}}) + m_{32}(Y_j - Y_{\mathbf{o}}) + m_{33}(Z_j - Z_{\mathbf{o}})}$$ Equation 2-7

$$y_i - y_o = c\,\frac{m_{21}(X_j - X_{\mathbf{o}}) + m_{22}(Y_j - Y_{\mathbf{o}}) + m_{23}(Z_j - Z_{\mathbf{o}})}{m_{31}(X_j - X_{\mathbf{o}}) + m_{32}(Y_j - Y_{\mathbf{o}}) + m_{33}(Z_j - Z_{\mathbf{o}})}$$ Equation 2-8

where, $\boldsymbol{x_i}, \boldsymbol{y_i}$ are the measured image coordinates, $\mathbf{X}_{\boldsymbol{j}}, \mathbf{Y}_{\boldsymbol{j}}, \mathbf{Z}_{\boldsymbol{j}}$ are the object space coordinates of the measured points, $\mathbf{X_o}, \mathbf{Y_o}, \mathbf{Z_o}$ are the object space coordinates of the perspective centre of the camera, and m_{11} to m_{33} are the individual elements of the 3 x 3 orthogonal rotation matrix, $\mathbf{M}\ \left(\mathrm{M} = \mathrm{M}_{XYZ} = \mathrm{M}_{\omega}\mathrm{M}_{\phi}\mathrm{M}_{\kappa}\right)$ between image and object coordinate systems and are given by, for example, as follows: The orthogonal rotation matrix, $\mathbf{M}$ is given as:

$$\mathrm{M} = \begin{bmatrix} \mathrm{m}_{11} & \mathrm{m}_{12} & \mathrm{m}_{13} \\ \mathrm{m}_{21} & \mathrm{m}_{22} & \mathrm{m}_{23} \\ \mathrm{m}_{31} & \mathrm{m}_{32} & \mathrm{m}_{33} \end{bmatrix}$$ Equation 2-9

where

$$m_{11} = cos\,\phi\,cos\,\kappa$$

$$m_{12} = -\,cos\,\phi\,sin\,\kappa$$

$$m_{12} = -\,cos\,\phi\,sin\,\kappa$$

$$m_{13} = sin\,\phi$$

$$m_{21} = cos\,\omega\,sin\,\kappa + sin\,\omega\,sin\,\phi\,cos\,\kappa$$

$$m_{22} = \quad cos\,\omega\,cos\,\kappa - sin\,\omega\,sin\phi\,sin\,\kappa$$

$$m_{23} = -\,sin\,\omega\,cos\,\phi$$

$$m_{31} = \sin\omega \sin\kappa - \cos\omega \sin\phi \cos\kappa$$

$$m_{32} = \sin\omega \cos\kappa + \cos\omega \sin\phi \sin\kappa$$

$$m_{33} = \cos\omega \cos\phi$$

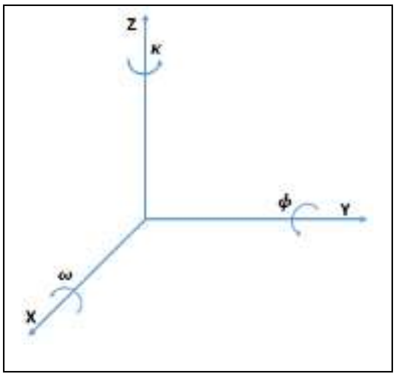


Figure 2-8 The 3D axis and directions of three rotation angles

All other terms have previously been defined. Any deviation from the collinearity condition[11] is as a result of systematic and random errors, mainly lens distortions. These distortions $(\Delta\boldsymbol{x}, \Delta\boldsymbol{y})$, often referred to as the additional parameters (APs), can be modelled as discussed in Section 2.2. Equations 2-10 and 2-11 represent the collinearity equations for perspective images extended with APs.

$$x_i - x_o + \Delta x = c\,\frac{m_{11}(X_j - X_{\mathbf{o}}) + m_{12}(Y_j - Y_{\mathbf{o}}) + m_{13}(Z_j - Z_{\mathbf{o}})}{m_{31}(X_j - X_{\mathbf{o}}) + m_{32}(Y_j - Y_{\mathbf{o}}) + m_{33}(Z_j - Z_{\mathbf{o}})} \qquad \text{Equation 2-10}$$

$$y_i - y_o + \Delta y = c\,\frac{m_{21}(X_j - X_{\mathbf{o}}) + m_{22}(Y_j - Y_{\mathbf{o}}) + m_{23}(Z_j - Z_{\mathbf{o}})}{m_{31}(X_j - X_{\mathbf{o}}) + m_{32}(Y_j - Y_{\mathbf{o}}) + m_{33}(Z_j - Z_{\mathbf{o}})} \qquad \text{Equation 2-11}$$

---

[11] The collinearity condition states that a point in object space, its corresponding point in the image plane and the perspective centre of the camera should, in the absence of lens distortions, form a straight line.

Two collinearity equations[12] per image can be written for each image point, visible in multiple images. All the equations are combined into a simultaneous solution for the unknown parameters using the Gauss Markov model of the least squares adjustment method. Since the collinearity equations are non-linear, they are linearised with respect to the unknown parameters using the first order terms of the Taylor series expansion. The linearised collinearity equations are expressed in least squares format as:

$$v = A\hat{x} - l \qquad \text{Equation 2-12}$$

where, $\boldsymbol{A}$ is the design matrix containing the coefficients of the unknown parameters and has a size of n x u (**n** is the number of observations, **u** is the number of unknowns) $\boldsymbol{l}$ is the discrepancy vector, $\boldsymbol{v}$ is the residual vector and $\hat{\boldsymbol{x}}$ is the solution vector. The solution vector ($\hat{\boldsymbol{x}}$) is computed as:

$$\hat{x} = N^{-1}A^{T}Pl \qquad \text{Equation 2-13}$$

where, P is the weight matrix of the observations, $N = A^{T}PA$ is the normal equations matrix. After the solution has converged, the variance-covariance matrix can be computed as:

$$\textstyle\sum_{x} = \hat{\sigma}_{o}^{2}Q_{x} \qquad \text{Equation 2-14}$$

where $\hat{\sigma}_{o}^{2}$ is the a posteriori variance factor estimated from the residuals as:

$$\hat{\sigma}_{o}^{2} = \frac{v^{T}Pv}{n-u} \qquad \text{Equation 2-15}$$

The discrepancy, residual and solution vectors all have a size of n x 1 while the weight matrix have a size of u x u.

---

[12] The collinearity equation here refers to either the standard collinearity equation (Equation 2-7 and 2-8) or the extended version (Equations 2-10 and 2-11).

2.4.1. Datum Definition

Without defining a datum, a photogrammetric network composed of only image coordinate observations has a datum defect (d) resulting in a non-invertible singular normal equations matrix $(N = A^T PA)$. Datum deficiency is also referred a Zero Order Design (ZOD) problem (Grafarend, 1974). In case of the bundle adjustment, the datum deficiency can be resolved by holding fixed a minimum of seven parameters or ordinates from three control points (*e.g*. two XYZ and one height). This is known as the minimum constraint solution. Alternatively, one can apply a matrix decomposition algorithm for the creation of a pseudo-inverse, referred to as free-network adjustment or inner constraint network (Blaha, 1982). If more than seven parameters are held fixed then the network is ‘over constrained’.

Parameters that are held fixed have no associated variance–covariance matrix element. In the case of the free network adjustment, error estimates for all parameters are available. The software produced for this thesis allow for all three techniques.

2.4.2. Generation of Approximate Values for the Unknown Parameters

For a bundle adjustment solution to converge to a global minimum, reliable initial approximation values are required for all unknown parameters. The unknown parameters in a standard case include the interior orientation parameters (IOPs), the EOPs of the camera(s) and the 3D coordinates of object points.

In the case of the IOPs, the principal distance and principal point coordinates may be approximated using techniques based on vanishing points *(*Van Den Heuvel, 1998, 1999; Grammatikopoulos, Karras and Petsa, 2007; Douskos *et al*., 2008). In practice, initial approximation values for the principal point coordinates and APs can sufficiently be approximated to zero while the focal length of the camera as provided by the lens manufacturer can be used as initial approximation value for the principal distance.

Initial approximation values for the EOPs of the camera(s) and 3D coordinates of the object points may be generated through the transformation of independent models after relative orientation of image pairs. Another approach involves a combination of resection and intersection techniques. Example of resection techniques often used include the closed-form space resection (Zeng and Wang, 1992; Guan *et al.*, 2008) and Cassini-resection (Maas and Schneider, 2004). If the 3D coordinates of at least six object points are known, the DLT algorithm (Abdel-Aziz and Karara, 1971) can be applied to approximate the EOPs of the camera(s). In practice, the initial approximation values of the EOPs of the camera(s) can also be generated by attaching GNSS/INS units to the camera during photography.

CHAPTER 3

IMAGE RECTIFICATION USING FICTITIOUS IMAGE COORDINATES

3.1. Overview

In this Chapter, a new technique for rectifying images acquired from fisheye lenses for the subsequent use in the creation of spherical panoramic images is presented.

The technique involves quantifying the distortion characteristics of a lens-camera combination from a single reference image. This is done by creating a distortion matrix through the comparison of near-error-free simulated image point positions of targets of a testfield with the positions of these targets on the calibration image. While the distorted image of the testfield is captured with a fisheye lens and a conventional digital camera, the virtual image coordinates of the target points of the testfield are simulated based on a fictitious image generation algorithm proposed by Rüther (1982). The difference between the distorted and fictitious image coordinates is the amount of distortion to be applied to the specific positions on the camera's sensor where the targets were imaged. Images captured with the same lens and camera are then rectified using a backward pixel mapping strategy. The rectified images are subsequently used to create spherical panoramic images.

The remaining sections of this Chapter are structured as follows: Section 3.2 presents the basic requirement for the development of the method. The procedure for the image rectification technique developed in this research is then elaborated in Section 3.3. Experimental results and analysis are presented in Section 3.4 while Section 3.5, presents the concluding remarks on the image rectification approach.

3.2. Pre-requisite for the Image Rectification Technique

The basic requirements for developing the image rectification technique are as follows:

(i) Imaging equipment to capture the image of the calibration testfield and;

(ii) Calibration testfield with known 3D object coordinates.

3.2.1. Imaging Equipment

The conventional digital camera and lens used for the image acquisition were a Nikon D200 DSLR camera and a Nikkor 10.5 mm fisheye lens respectively. The specifications of the camera and lens are provided in Tables 3-1 and 3-2 respectively.

Table 3-1 Specifications of Nikon D200 DSLR camera

Camera features	
Sensor	CCD
Maximum image size	3872 x 2592 pixels
Sensor size	23.6 x 15.8 mm
Pixel size	6.4 µm
Approx. dimensions	132 x 103 x 77 mm

Table 3-2 Specifications of the Nikkor 10.5 mm full-frame fisheye lens

Lens features	
Focal length	10.5 mm
Maximum aperture	f/ 2.8
Minimum aperture	f/ 22

3.2.2. Calibration Testfield

The calibration testfield used in this research for establishing the distortion characteristics of the fisheye lens is a near-planar point array of 91 targeted calibration points with known 3D object coordinates. The testfield is located in the Geomatics teaching laboratory at the University of Cape Town. The laboratory covers a floor area of 8 m by 8 m and is about 3 m in height. The target points of the testfield are well distributed over a wall area of approximately 6 m by 2.5 m on one side of the room and are represented by circular targets in the form of white retro-reflective disks of 14 mm diameter on a black background (Figure 3-1). Using a total station, the 3D object coordinates of these targets were determined with sub-millimetre accuracy by means of theodolite and distance measurements from a three-point base triangle. Appendix 1 shows the 3D object coordinates of these targets.

Figure 3-1 Calibration testfield at University of Cape Town

3.3. Image Rectification Procedure

The steps adopted for rectifying images acquired from fisheye lenses are:

(i) Calibration image acquisition;

(ii) Preliminary camera calibration and fictitious image generation;

(iii) Determination of geometric lens distortion characteristics and creation of distortion matrix; and

(iv) Image rectification.

3.3.1. Calibration Image Acquisition

A calibration image of the calibration testfield is captured with a fisheye lens and a camera from a central location such that the target points of the testfield fill the entire area of the camera's FOV (or ideally fill the image format). The positions of the targets on this calibration image represent the distorted image coordinates.

3.3.2. Preliminary Camera Calibration and Fictitious Image Generation

The next step is to simulate the near error-free fictitious image equivalent of the calibration image. The mathematical model of the fictitious image is based on the collinearity equations without APs (Equations 2-7 and 2-8).

To create the image coordinates of the fictitious image which simulates an error free version of the actual calibration image, the camera constant $(\boldsymbol{c})$, principal point $(\boldsymbol{x_o}, \boldsymbol{y_o})$ as well as the position $(\boldsymbol{X_o}, \boldsymbol{Y_o}, \boldsymbol{Z_o})$ and angular orientation $(\boldsymbol{\omega}, \boldsymbol{\phi}, \boldsymbol{\kappa})$ of the camera used in capturing the calibration image are required. As the true values of these are not known, a conventional camera calibration was carried out to approximate them. If the testfield is used frequently for camera calibrations a frame, similar to a panoramic head, can be permanently installed in front of the testfield, and surveyed, thus guaranteeing that the external pupil of the calibrated camera is in the same known location. In this research, the determination of the parameters for the fictitious image was carried out using the Australis software based on the standard multi-image bundle adjustment algorithm. The alternative of single image space resection was considered to be of inferior accuracy in this regard and the bundle adjustment option was adopted as a more reliable solution.

After the preliminary camera calibration step, the derived parameters $(\boldsymbol{c}, \boldsymbol{x_o}, \boldsymbol{y_o}, \boldsymbol{X_o}, \boldsymbol{Y_o}, \boldsymbol{Z_o}, \boldsymbol{\omega}, \boldsymbol{\phi}, \boldsymbol{\kappa})$ were substituted in the collinearity equation (Equations 2-1 and 2-2) to obtain the fictitious image coordinates.

3.3.3. Determination of Geometric Lens Distortion Characteristics and Creation of Distortion Matrix

After the generation of fictitious image coordinates, the geometric distortion characteristics of the lens are quantified. This is done by calculating the differences between the distorted and the near error-free fictitious image coordinates of the target points of the testfield (Figure 3.2).

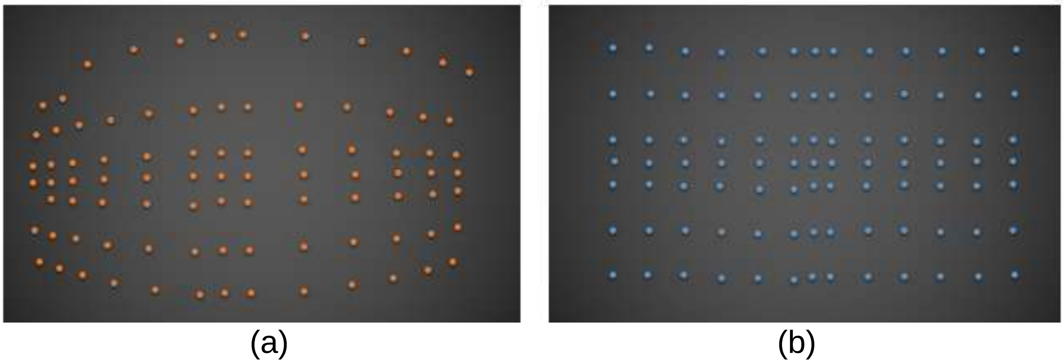

(a) (b)

Figure 3-2 Image coordinates of the calibration testfield: (a) distorted image points and (b) fictitious undistorted image points

Figure 3-3 shows the positions of the distorted and fictitious image coordinates in green and red respectively, superimposed on the calibration image.

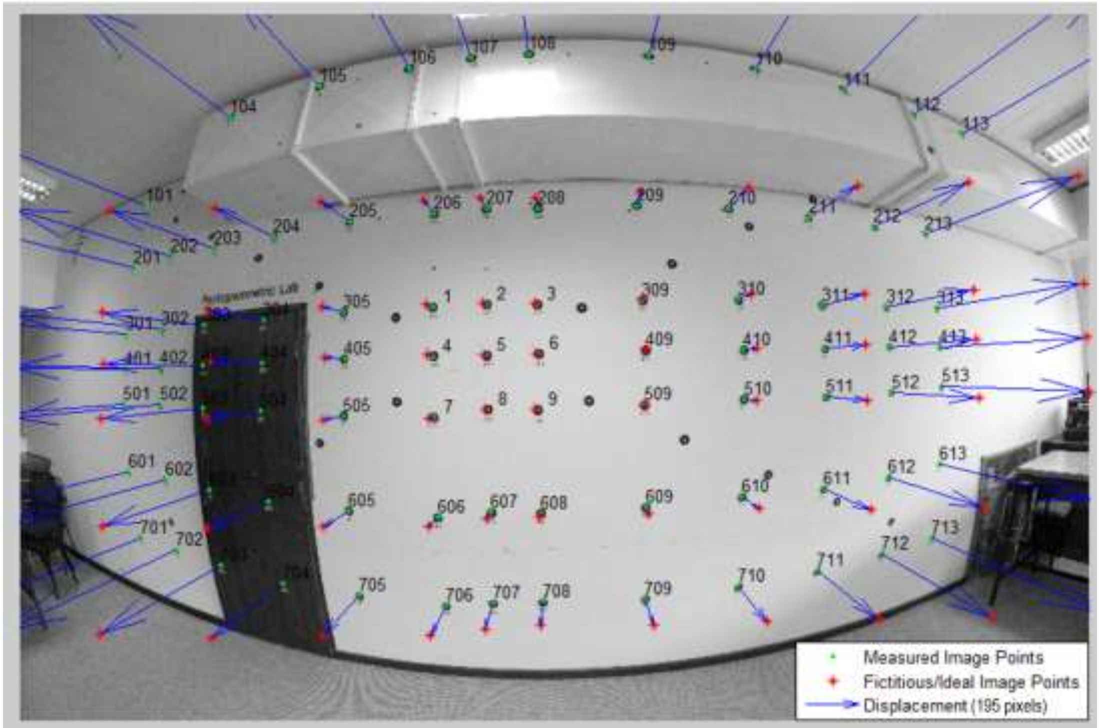


Figure 3-3 Distorted (green) and fictitious (red) image coordinates superimposed on the calibration image

The difference between the two coordinates of an image is the amount of correction to be applied to the specific positions on the camera's sensor where the targets were imaged. The corrections are expressed as:

$$\begin{bmatrix}\Delta x_i\\ \Delta y_i\end{bmatrix}_{1,2n} = \begin{bmatrix}x_i\\ y_i\end{bmatrix}_{1,2n} - \begin{bmatrix}\bar{x}_i\\ \bar{y}_i\end{bmatrix}_{1,2n} \qquad \text{Equation 3-1}$$

where $\boldsymbol{x_i}, \boldsymbol{y_i}$ are the near error-free fictitious image coordinates, $\boldsymbol{\bar{x}_i}, \boldsymbol{\bar{y}_i}$ are the distorted image coordinates of the reference image, $\Delta\boldsymbol{x_i}, \Delta\boldsymbol{y_i}$ are the departures from collinearity, representing the distortion amounts in x and y, $\boldsymbol{n}$ is the number of imaged target points involved and $\boldsymbol{i}$ is the unique identifier of the target points of the calibration testfield.

In the next step, a distortion matrix equivalent to a look-up table is created in which each pixel in the distorted image is allocated a distortion value derived by interpolating between the target points. Distortion values for pixels at the edges of the calibration image (outside the imaged calibration targets) have to be extrapolated. This emphasises the need for the calibration image to be captured close to the edge of the calibration testfield.

3.3.4. Image Rectification

In the final step, the positions of the distorted pixels are corrected to their positions in the output image where they would have appeared if there was no distortion by applying a backward pixel mapping strategy. The backward pixel mapping was chosen over the forward pixel mapping to avoid gaps or the creation of multiple values for a pixel in the resulting output image.

The procedure involves taking a pixel location on the corrected but still blank image as input (Figure 3-4) and determining which original pixel on the distorted image maps to this location after the application of correction values $(\Delta \boldsymbol{x}, \Delta \boldsymbol{y})$ from the distortion matrix. RGB values are then extracted from the distorted image to the pixel location on the corrected image. The resulting pixel location are floating point values and do not correspond to integer pixels. For this reason, their RGB values are interpolated from the RGB values of surrounding pixels. A bilinear interpolation was chosen over nearest neighbourhood and bi-cubic interpolations as it provides a compromise between computational efficiency and image quality (Gonzalez and Woods, 2011).

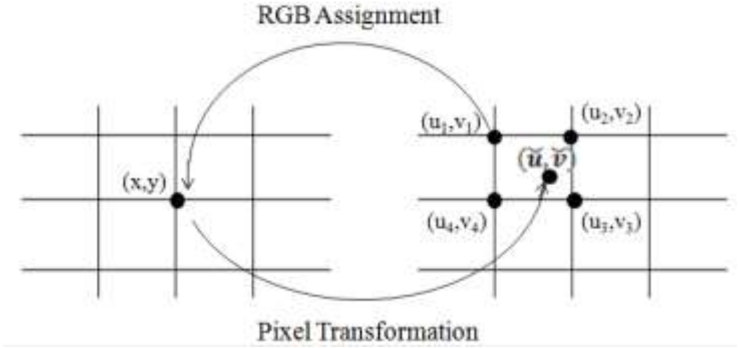
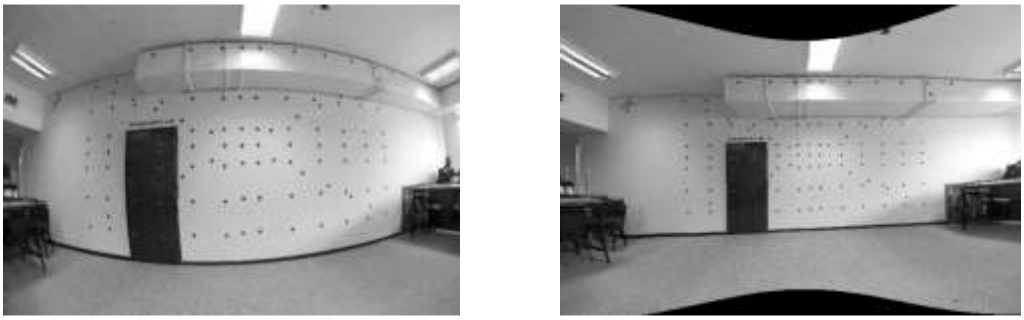


Figure 3-4 Backward pixel mapping method

3.4. Experimental Results and Analysis

Figures 3-5a and 3-5b are examples of images of the testfield before and after rectification respectively. As can be observed, the barrel effect in Figure 3-5a has been removed through the rectification process, thus converting fisheye images into rectilinear images (Figure 3-5b).



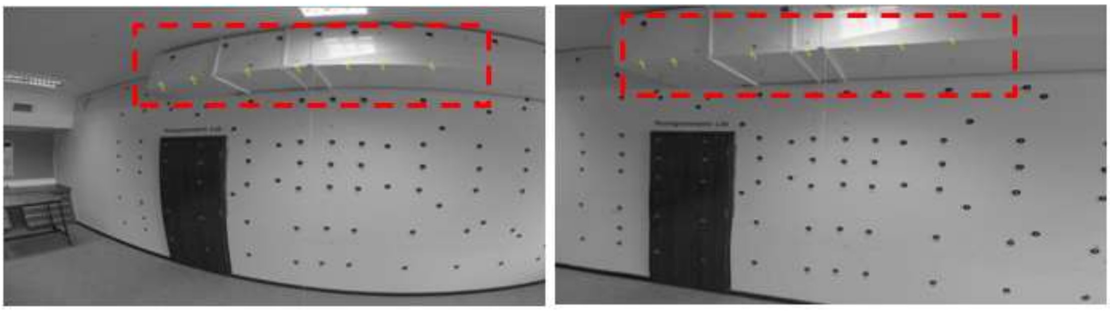
(a) (b)

Figure 3-5 Image of the calibration testfield before and after image rectification: (a) before rectification (b) after rectification

Table 3-3 shows an extract of five minimum and five maximum displacements of measured image coordinates of control points from their rectified positions. As expected, the maximum displacements in control point positions (1344.1 pixels) occur towards the edges while the and minimum displacements (1.5 pixels) occur in the image centre (Table 3-3). An obvious check of the effectiveness of the rectification technique is to verify that a straight line in object space appears as a straight line on the image after rectification. This was tested using straight line features in object space. The coordinates of these features were measured on both the distorted and rectified images. The standard deviations of the least squares fit before and after rectification were used as indicators for the straightness of lines and as a measure of the success of the rectification process. For example, the standard deviation from seven points (marked in yellow) along a straight line feature in Figure 3-6 before (180.6 µm equivalent to 28 pixels) and after (17.5 µm or approximately 3 pixels) rectification indicate that the effect of lens distortion was successfully minimised using the developed method.

Table 3-3 Minimum and maximum image coordinate displacements

ID	Fictitious Image Coordinates (mm)		Distorted Image Coordinates (mm)		Displacement mm (pixels)	
	$\bar{x}$	$\bar{y}$	x	y	d_{xy}	
509	2.10	-0.90	2.10	-0.89	0.01	(1.51)
409	2.12	0.42	2.12	0.40	0.02	(2.84)
9	-0.43	-1.01	-0.38	-1.00	0.04	(6.90)
6	-0.40	0.32	-0.35	0.31	0.05	(7.09)
3	-0.45	1.51	-0.40	1.48	0.06	(8.92)
601	-15.17	-3.73	-9.86	-2.46	5.46	(853.41)
101	-14.96	6.01	-9.48	3.78	5.92	(924.72)
701	-15.15	-6.27	9.60	-4.01	6.00	(936.81)
112	13.35	9.64	8.31	5.94	6.25	(976.00)
113	16.84	9.88	9.45	5.48	8.60	(1344.1)



(a) (b)

Figure 3-6 Verification of straight line in object space: (a) before rectification (b) after rectification

A further numerical test of the proposed method used a subset of the control points to form the distortion matrix (86 out of 91). The known locations of the five remaining randomly selected control points were compared with their interpolated values. From Table 3-4 and Figure 3-7, it is observed that Control points 610 and 5 recorded the maximum and minimum displacements respectively. The denser the control points the more precise the interpolated value. This is an indication that the proposed method is dependent on the distribution and density of control points.

Table 3-4 Displacement of interpolated control point coordinates from their measured values

ID	Measured Image Coordinates (mm)		Interpolated Image Coordinates (mm)		Displacement mm (pixels)	
	$\bar{x}$	$\bar{y}$	x	y	d_{xy}	
5	-1.63	0.29	-1.63	0.28	0.01	(1.80)
304	-8.08	1.37	-8.01	1.33	0.08	(11.77)
312	9.81	1.85	9.73	1.81	0.09	(14.01)
704	-7.96	-6.36	-7.94	-6.36	0.02	(2.84)
610	4.98	-3.39	4.75	-3.33	0.24	(37.04)

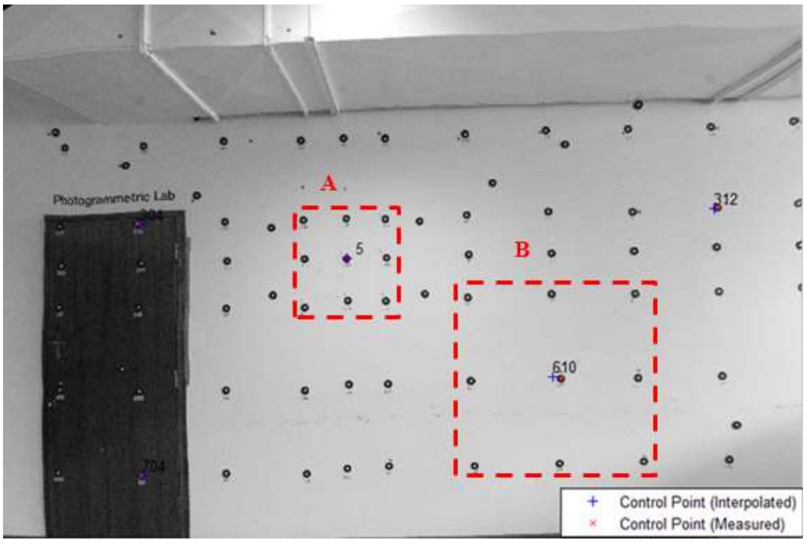
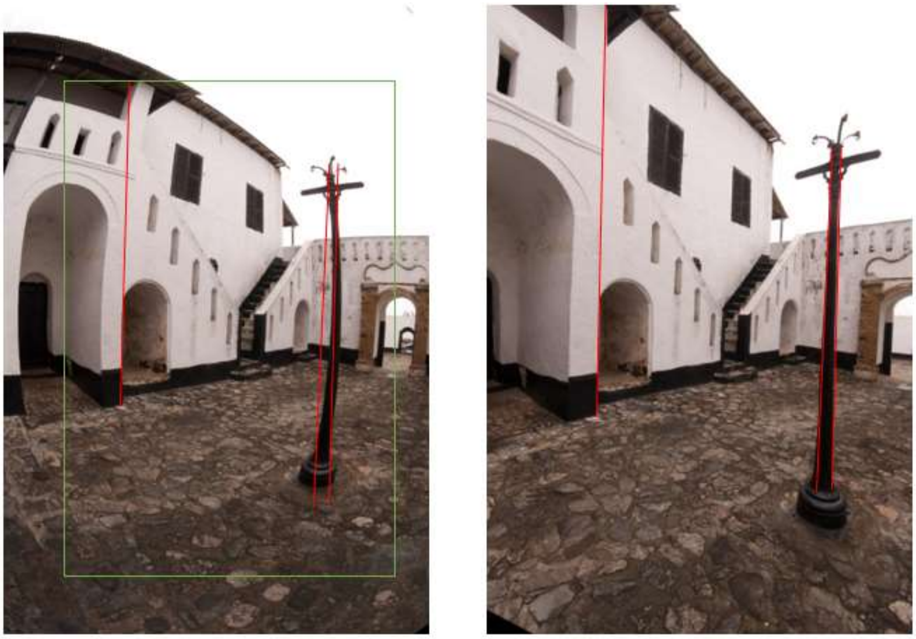


Figure 3-7 Effect of control points distribution on interpolation

Figure 3-8a shows an image of a section of a building that was captured with the camera-lens combination (Nikon D200 DSLR camera and the Nikkor 10.5 mm fisheye lens). As can be observed in Figure 3-8b, the barrel effect in the image has been removed through the rectification process.



(a) (b)

Figure 3-8 Image rectification prior to stitching panoramic images: (a) before rectification (b) after rectification

Figure 3-9 shows an example of spherical panoramic image created (with PTGui) after the original fisheye images were rectified with the proposed technique.

Figure 3-9 Spherical panorama generated from rectified fisheye images

3.5. Chapter Summary

The objective of developing an image rectification technique for images acquired from fisheye lenses has been achieved in this research. The technique is capable of removing the barrel effects produced by fisheye lens, thus transforming the fisheye into rectilinear images.

A shortcoming of the developed technique is that its efficiency relies on the density and distribution of control points of a calibration testfield. Thus, it is recommended that for higher accuracy, a dense and well-distributed set of control points should be adopted.

It has also been shown that panoramas can be created from images rectified with the proposed method. This makes the technique particularly useful for image stitching software that does not support fisheye lenses.

CHAPTER 4

ORIENTATION AND EXTRACTION OF 3D COORDINATES FROM SPHERICAL PANORAMIC IMAGES

4.1. Overview

This Chapter focuses on the algorithms and methods developed and adopted in this research for the orientation and extraction of 3D coordinates of object points from spherical panoramas.

A new approach to automatically orient and position multiple spherical panoramas with respect to each other is presented. The determination of the final 3D object coordinates is based on the bundle adjustment algorithm, which requires initial approximation values for the unknown parameters. These provisional values are provided in an automated process by the Minimum Ray Distance (MRD) algorithm developed in this research. The unknown parameters comprise of the EOPs of each panorama $(\boldsymbol{X_{oi}}, \boldsymbol{Y_{oi}}, \boldsymbol{Z_{oi}}, \boldsymbol{\omega_i}, \boldsymbol{\phi_i}, \boldsymbol{\kappa_i})$ and the 3D object coordinates of the measured image points $\left(\boldsymbol{X_j}, \boldsymbol{Y_j}, \boldsymbol{Z_j}\right)$, where the subscript $\boldsymbol{i}$ refers to the individual panorama while $\boldsymbol{j}$ refers to the object points.

Prior to orientation, image coordinates of conjugate points on the panoramas are measured either manually or automatically. Automatic conjugate point extraction was accomplished by a feature-based matching technique using the SIFT operator. When the conjugate points are extracted in automatic mode (Section 4.3), only a few of these are chosen for the MRD algorithm to avoid excessive computation time. Typically, some six to ten conjugate points suffice to provide the provisional values required for the bundle adjustment.

The workflow for the panorama orientation and extraction of 3D coordinates of object points involves five principal routines (Figures 4-1 and 4-2). These are the MRD and bundle adjustment algorithms, scaling of object coordinates, interactive extraction of conjugate points guided by the epipolar geometry and 3D similarity transformation.

First, the MRD algorithm is sequentially applied to pairs of spherical panoramas to automatically estimate the initial approximation values for the unknown parameters. This is followed by scaling all derived object coordinates to a common uniform scale. The bundle adjustment algorithm is then carried out to jointly refine the initial approximation values for the unknown parameters for all panoramas within a panorama configuration. A panoramic configuration in this thesis refers to a group of panoramas defined within a local coordinate system, where a single panorama is selected as reference and all others, known as the free panoramas are sequentially oriented with respect to the reference panorama. The initial approximation values for 3D object coordinates of any additional image points which were not included in the orientation phase as well as points interactively or automatically measured through the guidance of the epipolar geometry can then be calculated. The final 3D object coordinates of all image points are then determined by the bundle adjustment algorithm.

If all possible panorama connections have been formed and there are not sufficient points in the remaining panoramas, a new configuration is created. In a final step, the configurations can be linked with 3D similarity transformation which requires fewer points than the MRD orientation process. Given three or more control points with known object coordinates within a project area, the local coordinate system can then be georeferenced into the desired national or local coordinate system.

These processes described in the workflow are discussed in details in the following sections after introducing the spherical panoramic imaging model.

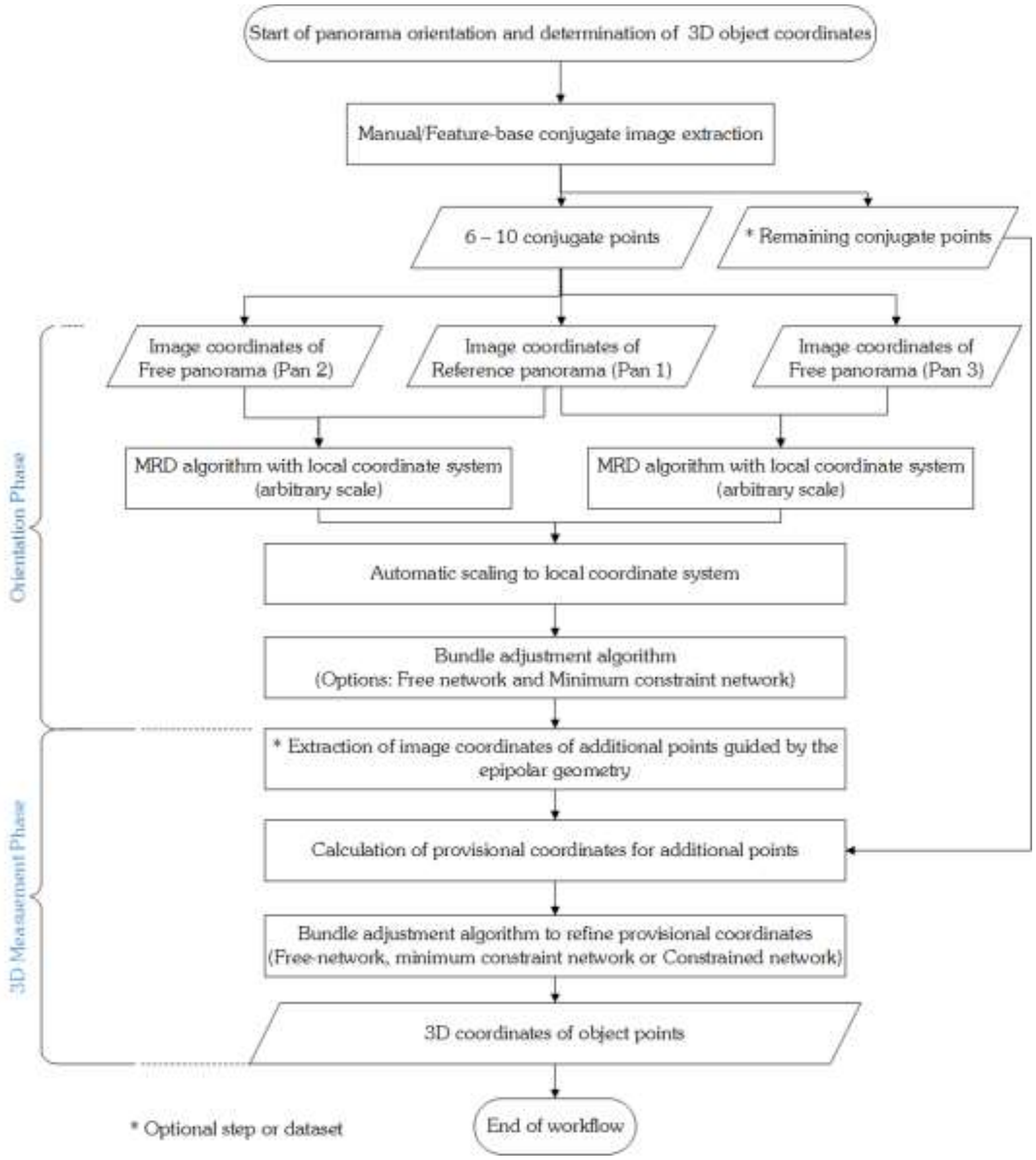


Figure 4-1 Workflow for orientation and extraction of 3D coordinates of object points

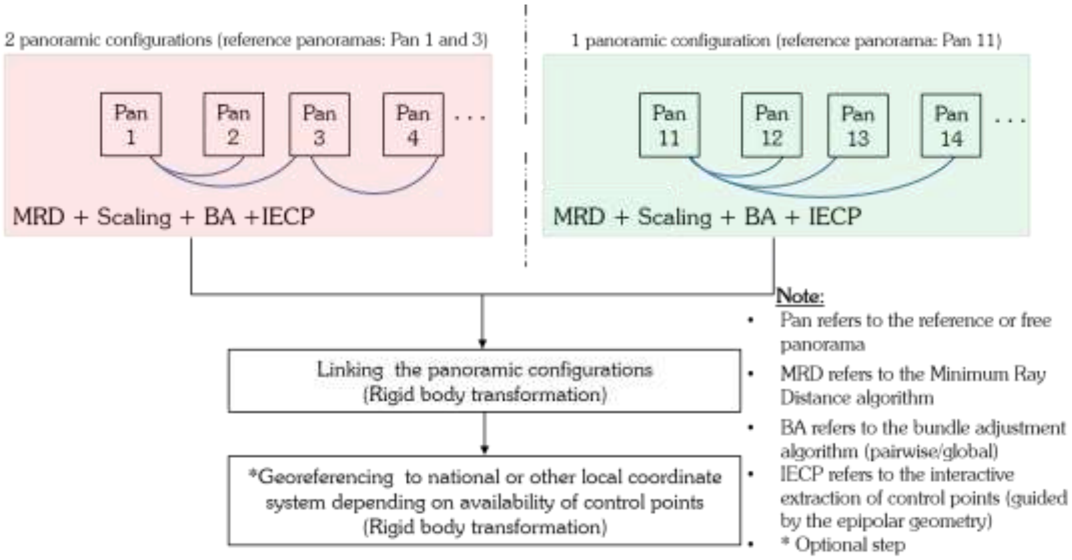


Figure 4-2 Overall workflow for panorama orientation and extraction of 3D coordinates of object points

4.2. Spherical Panoramic Imaging Model

The imaging process for transforming 3D object points onto a 2D image plane involves four different coordinate systems. These are:

(i) object coordinate system $(\boldsymbol{X_j}, \boldsymbol{Y_j}, \boldsymbol{Z_j})$,

(ii) Cartesian panoramic coordinate system $(\boldsymbol{x_i}, \boldsymbol{y_i}, \boldsymbol{z_i})$,

(iii) spherical coordinate system $(\boldsymbol{\lambda_i}, \boldsymbol{\phi_i})$, and

(iv) image coordinate system $(\boldsymbol{u_i}, \boldsymbol{v_i})$.

First, the 3D object points are mapped onto the panorama sphere in the intersection of the vector from the perspective centre of the panorama sphere $(\boldsymbol{X_{oi}}, \boldsymbol{Y_{oi}}, \boldsymbol{Z_{oi}})$ to the object with the sphere. The relationship between object $(\boldsymbol{X_j}, \boldsymbol{Y_j}, \boldsymbol{Z_j})$ and Cartesian panoramic $(\boldsymbol{x_i}, \boldsymbol{y_i}, \boldsymbol{z_i})$ coordinate system (Figure 4-3) is given as:

$$\begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = \psi \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} * \begin{bmatrix} X_j - X_{oi} \\ Y_j - Y_{oi} \\ Z_j - Z_{oi} \end{bmatrix} \qquad \text{Equation 4-1}$$

where the three separate equations describing the relationship between two coordinate systems are:

$$x_i = \psi \left[m_{11}(X_j - X_{oi}) + m_{21}(Y_j - Y_{oi}) + m_{31}(Z_j - Z_{oi}) \right] \qquad \text{Equation 4-2}$$

$$y_i = \psi \left[m_{12}(X_j - X_{oi}) + m_{22}(Y_j - Y_{oi}) + m_{32}(Z_j - Z_{oi}) \right] \qquad \text{Equation 4-3}$$

$$z_i = \psi \left[m_{13}(X_j - X_{oi}) + m_{23}(Y_j - Y_{oi}) + m_{33}(Z_j - Z_{oi}) \right] \qquad \text{Equation 4-4}$$

The $\boldsymbol{m's}$ in Equation 4-1 to 4-4 are the individual elements of the rotation matrix as defined in Equation 2-9 and ψ is the individual scale factor for each point on the panorama sphere in relation to the object.

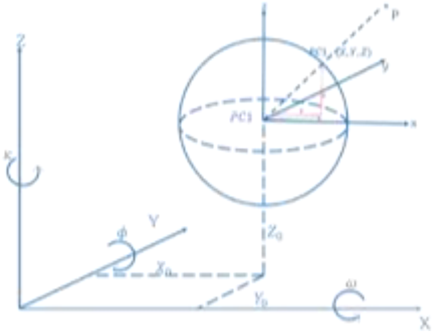


Figure 4-3 Relationship between Cartesian panoramic and object coordinate systems

On the panorama sphere, two coordinates systems can be defined. These are the Cartesian panoramic $(\boldsymbol{x_i}, \boldsymbol{y_i}, \boldsymbol{z_i})$ and spherical $(\boldsymbol{\lambda_i}, \boldsymbol{\phi_i})$ coordinate systems (Figure 4-4). The relationships between the two systems are as follows:

$$x_i = r\ \sin\lambda_i \sin\phi_i \qquad \text{Equation 4-5}$$

$$y_i = r\ \cos\lambda_i \sin\phi_i \qquad \text{Equation 4-6}$$

$$z_i = r\ \cos\phi_i \qquad \text{Equation 4-7}$$

The inverse solutions are expressed as:

$$\lambda_i = \arctan\left(\frac{x_i}{y_i}\right) \qquad \text{Equation 4-8}$$

$$\phi_i = \arccos\left(\frac{z_i}{r}\right) \qquad \text{Equation 4-9}$$

where $\boldsymbol{\lambda_i}$ represent the latitude, $\boldsymbol{\phi_i}$ is the longitude and $\boldsymbol{r}$ is the radius of the sphere.

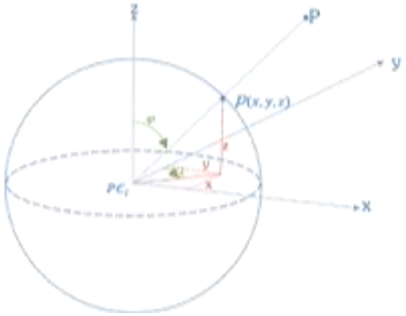


Figure 4-4 Relationship between Cartesian panoramic and spherical coordinate systems

In the final stage, the spherical surface is mapped onto the image plane. This involves the transformation from spherical $(\boldsymbol{\lambda_i}, \boldsymbol{\phi_i})$ to the image coordinate $(\boldsymbol{u_i}, \boldsymbol{v_i})$ systems based on the Equirectangular projection expressed as (Snyder, 1987; Fangi, 2007):

$$u_i = r\,\lambda_i \qquad \text{Equation 4-10}$$

$$v_i = r\,\phi_i \qquad \text{Equation 4-11}$$

The origin of the image coordinate system is at $(0,0)$ for the centre of the top-left pixel (Figure 4-5). The spherical panoramic image covers a horizontal FOV of 360° and a vertical FOV of 180°, thus the image has a width (**W**) to height (**H**) ratio of 2:1. The radius of the panorama sphere is ideally equal to the focal length of the camera. However, the actual radius of a final panorama output after stitching can be calculated using the expression:

$$r = \frac{W}{2\pi} \qquad \text{Equation 4-12}$$

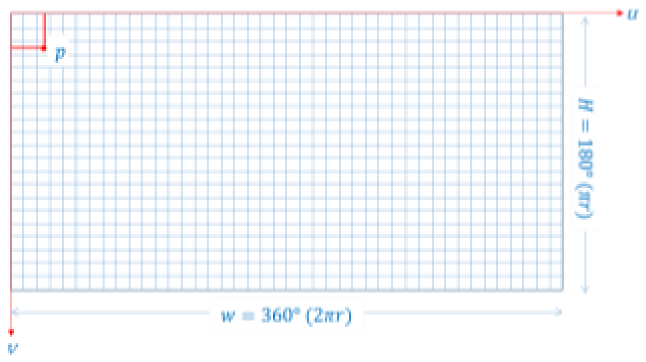


Figure 4-5 Image coordinate system

The expression for scale factor $\boldsymbol{\psi_i}$ can be derived through the relationship between the radius of the sphere ($\boldsymbol{r}$), and the distance between an object point in space ($\boldsymbol{X_j, Y_j, Z_j}$) and the perspective centre ($\boldsymbol{X_{oi}, Y_{oi}, Z_{oi}}$) as:

$$\psi_i = \frac{r}{\sqrt{(X_j - X_{oi})^2 + (Y_j - Y_{oi})^2 + (Z_j - Z_{oi})^2}}$$

Equation 4-13

Substituting, Equations 4-2 and 4-3 into 4-8 and also Equations 4-3 and 4-13 into 4-9, the collinearity equations relating the image coordinates $(\boldsymbol{u_i, v_i})$ and object points $(\boldsymbol{X_j, Y_j, Z_j})$ for spherical panoramas are expressed as follows:

$$u_i = r \, arctan\left(\frac{m_{11}(X_j - X_{oi}) + m_{21}(Y_j - Y_{oi}) + m_{31}(Z_j - Z_{oi})}{m_{12}(X_j - X_{oi}) + m_{22}(Y_j - Y_{oi}) + m_{32}(Z_j - Z_{oi})}\right)$$

Equation 4-14

$$v_i = r \, acos\left(\frac{m_{13}(X_j - X_{oi}) + m_{23}(Y_j - Y_{oi}) + m_{33}(Z_j - Z_{oi})}{\sqrt{(X_j - X_{oi})^2 + (Y_j - Y_{oi})^2 + (Z_j - Z_{oi})^2}}\right)$$

Equation 4-15

4.3. Automatic Extraction of Conjugate Points

In order to automatically extract conjugate points for the orientation of spherical panoramas, a feature-based matching technique using the SIFT operator was adopted. Initial attempts to match the descriptors after the feature detection and description stages between a pair of spherical panoramas yielded many false matches. This is because most feature-based matching algorithms like the Best-Bin-First method are designed for perspective images and not for spherical panoramas. Features in spherical panoramas often appear to be deformed especially towards the upper and lower sections of the image with straight line appearing as curved lines. Thus, making the matching algorithm not globally valid for the entire panoramas and restricting it to limited regions within the panoramas. For this reason, the matching procedure was modified by creating multiple image tiles in one panorama and matching each of these tiles with the other panorama (Table 4-1). The number and size of the image tiles are always decided by the user prior to the matching process.

Table 4-1 Procedure for extracting conjugate feature points from a pair of panoramas

For each panorama pair (*e.g.* Panorama 1 and 2)

(i) Apply a SIFT detector to extract distinctive features on both panoramas.

(ii) Represent the extracted features with the SIFT descriptors and store them with a k-d tree structure.

(iii) Define two or more small image tiles in Panorama 1 or divide the entire Panorama 1 into several image tiles.

(iv) For Panorama 2 and each image tile in Panorama 1

  - Find conjugate feature points using a nearest neighbour search procedure based on the Best-Bin-First method.
  - Reject false conjugate feature pairs by applying the RANSAC algorithm and store the best candidates.

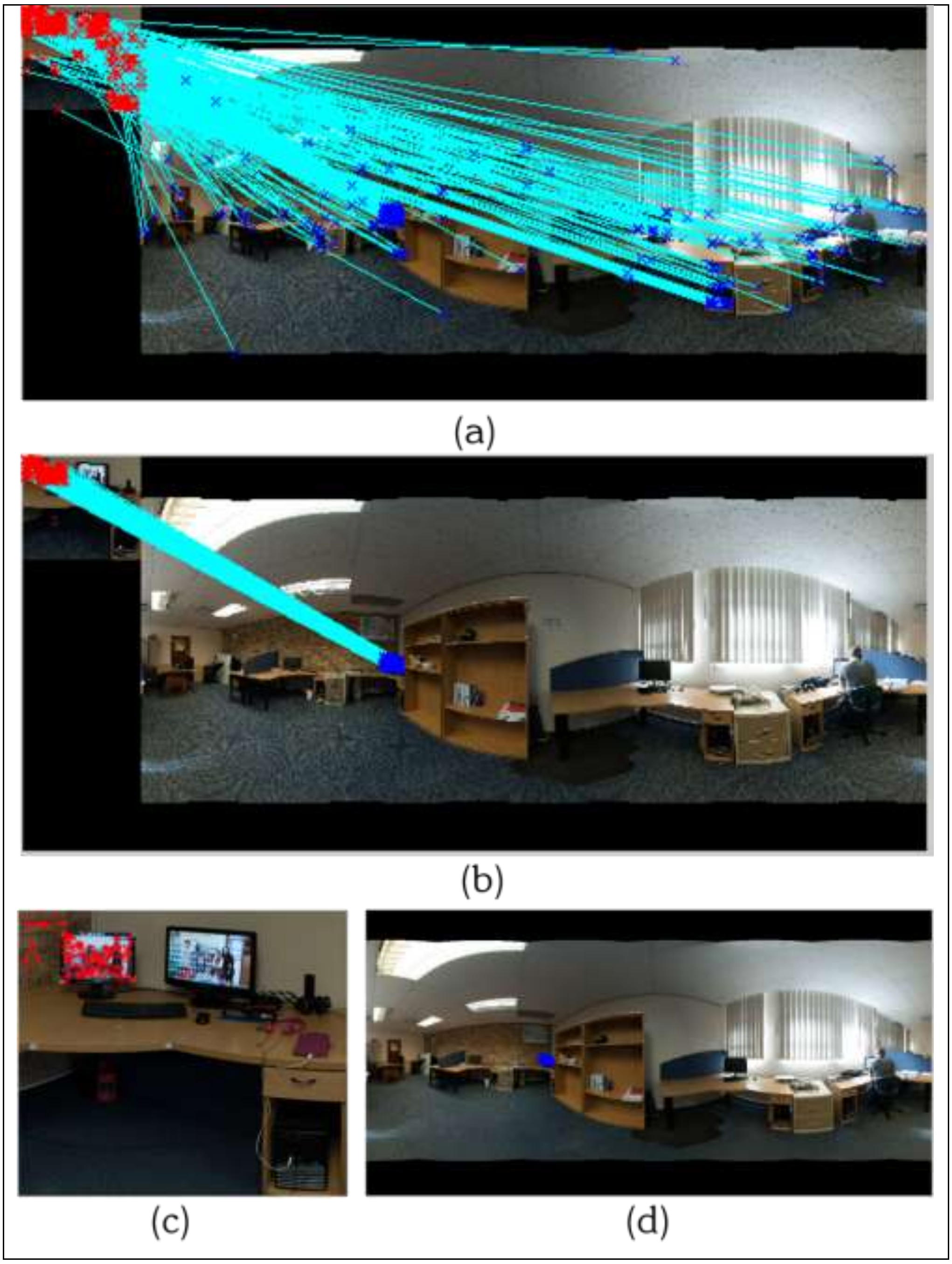
(v) Combine all valid conjugate feature points for each matching step in (iv)

The developed approach has been successfully tested on a wide number of spherical panoramas in both indoor and outdoor environments. The approach, however fails where the perspectives of the panoramas change significantly or in the presence of repetitive and indistinctive features. In such cases, the manual extraction of conjugate points is still the best option. Figure 4-6 shows two panoramas 1 and 2, with four image tiles defined in Panorama 1. Figures 4-7 to 4-8 show the nearest neighbour matching and RANSAC outlier detection results between features in Panorama 2 and three of the four image tiles of Panorama 1.

Figure 4-6 Panorama pair with four image tiles defined in one panorama

After matching features between Panorama 2 and the first image tile of Panorama 1, 516 tentative matches were obtained (Figure 4-7a). Out of these 516 matches, 252 (48.48%) outliers were removed (Figure 4-7b). Similarly, after matching features between Panorama 2 and the second image tile of Panorama 1, 137 tentative matches were obtained (Figure 4-8a). Out of these 137 matches, 93 (67.88%) outliers were removed (Figure 4-8b).

A total of 421 correct matches were obtained after the matching and outlier detection steps between Panorama 2 and the four image tiles of Panorama 1 (Figure 4-9). These can then be reduced if necessary prior to the application of the MRD orientation algorithm.



(a)

(b)

(c)

(d)

Figure 4-7 Output of feature-based matching and outlier detection procedures between Panorama 2 and image tile 1 of Panorama 1: (a) Feature matching (b) After outlier detection (c) Matched features in image tile 1 of Panorama 1 (in red) (d) Matched features in Panorama 2 (in blue)

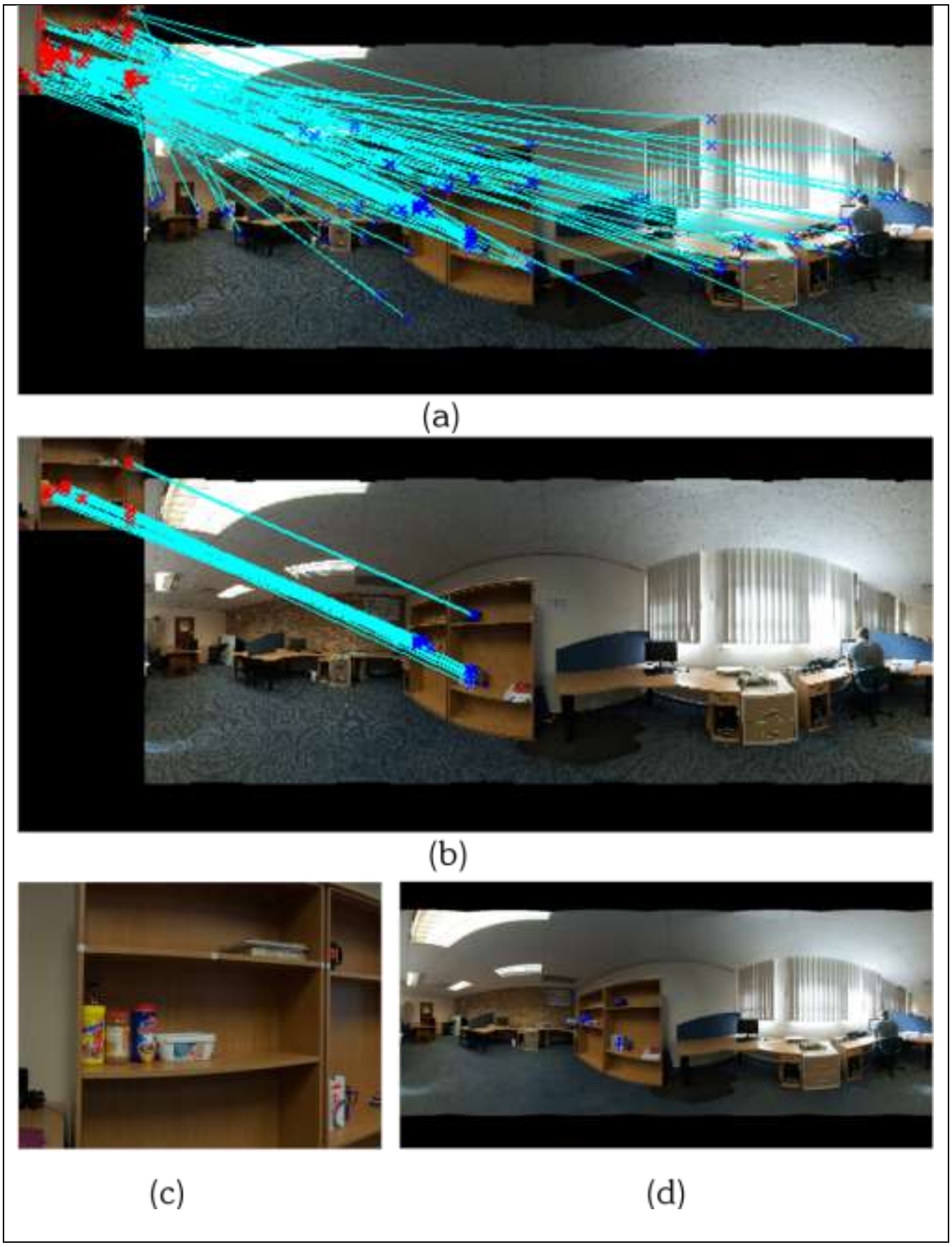


Figure 4-8 Output of feature-based matching and outlier detection procedures between Panorama 2 and image tile 2 of Panorama 1: (a) Feature matching (b) After outlier detection (c) Matched features in image tile 2 of Panorama 1 (red) (d) Matched features in Panorama 2 (blue)

Figure 4-9 Conjugate feature points after outlier detection between Panoramas 1(in red) and 2 (in blue)

4.4. Minimum Ray Distance Algorithm

The underlying concept of the MRD algorithm is the minimisation of distances between two conjugate space vectors or skewed rays. This approach is based on the premise that the vectors from the centre of separate panoramas ($\boldsymbol{PC1}$ and $\boldsymbol{PC2}$) to the same object point (P) intersect in this point, albeit not perfectly as a result of observation and other errors. The unknown orientation of the panoramas and their relative positions and heights can be determined iteratively by stepwise rotating the un-oriented panoramas with respect to each other and by iteratively changing the height displacements between the panoramas until a global minimum distance between conjugate rays are achieved. Rotations are iterated around the three coordinate axes by angles omega ($\boldsymbol{\omega_i}$), phi ($\boldsymbol{\phi_i}$) and kappa ($\boldsymbol{\kappa_i}$). If $\mathbf{L1}$ and $\mathbf{L2}$ are conjugate rays defined by the vectors $\vec{\boldsymbol{u}}$ and $\vec{\boldsymbol{v}}$ respectively, then the minimum distance between the two rays after each rotation is found between points $\boldsymbol{Pa}_{(\mathbf{s})}$ and $\boldsymbol{Pb}_{(\mathbf{t})}$ where the two vectors $\vec{\boldsymbol{u}}$ and $\vec{\boldsymbol{v}}$ have a common normal (Figure 4-10). The minimum distance between the rays ($\mathbf{L1}, \mathbf{L2}$) is mathematically expressed as follows:

$$\boldsymbol{d}(\boldsymbol{L_1}, \boldsymbol{L_2}) = \min_{\boldsymbol{Pa}_{(\mathbf{s})} \in \boldsymbol{L_1}\, \boldsymbol{Pb}_{(\mathbf{t})} \in \boldsymbol{L_2}} \left\{ \boldsymbol{d}\left(\boldsymbol{Pa}_{(\mathbf{s})}, \boldsymbol{Pb}_{(\mathbf{t})}\right) \right\} \qquad \text{Equation 4-16}$$

such that, $\boldsymbol{L1} = \boldsymbol{PC1} + \boldsymbol{s}\vec{\boldsymbol{u}}$, and $\boldsymbol{L2} = \boldsymbol{PC2} + \boldsymbol{t}\vec{\boldsymbol{v}}$. The mid-point between the two point, $\boldsymbol{Pa}_{(\mathbf{s})}$ and $\boldsymbol{Pb}_{(\mathbf{t})}$ defines the 3D position of the object point $\left(\boldsymbol{X_j}, \boldsymbol{Y_j}, \boldsymbol{Z_j}\right)$.

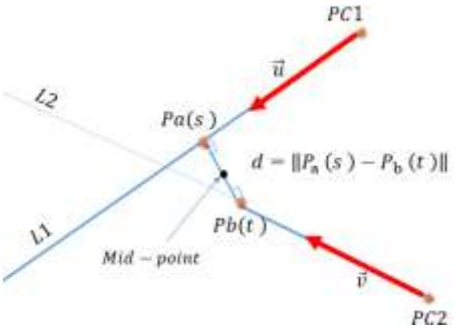


Figure 4-10 Minimum distance between two skew rays

Figure 4-11 shows a typical setup of two panoramas before and after relative orientation by MRD algorithm. The minimum distance is not only evaluated for one ray, but for a bundle of rays from each panorama centre.

(i) $\mathbf{PC1}$ and $\mathbf{PC2}$ are the perspective centres of the reference and free panoramas respectively;

(ii) $\mathbf{PC1}_i$ and $\mathbf{PC2}_i$ represent the 3D coordinates of object points on the reference and free panorama spheres; the coordinates of $\mathbf{PC2}_i$ are modified in the iteration process;

(iii) Provisional base ($\boldsymbol{B}$) is the arbitrarily chosen distance between the reference and free panoramas. The actual scale of the derived model will be determined later by measuring, or calculating an object distance from known coordinates.

Prior to the implementation of the MRD algorithm, no information about the relative position and height of the panoramas is known. EOPs $(X_{o1}, Y_{o1}, Z_{o1}, \omega_1, \phi_1, \kappa_1)$ and $(X_{o2}, Y_{o2}, Z_{o2}, \omega_2, \phi_2, \kappa_2)$, are therefore assumed as (0,0,0,0,0,0) and (0,B,0,0,0,0) for the reference and free panoramas respectively. It is also assumed that the panoramas are “quasi-horizontal” *i.e.* within 10º horizontal.

The task of the MRD algorithm is to estimate the EOPs for the free panorama(s) and object coordinate values for all image points (Figures 4-12 and 4-13).

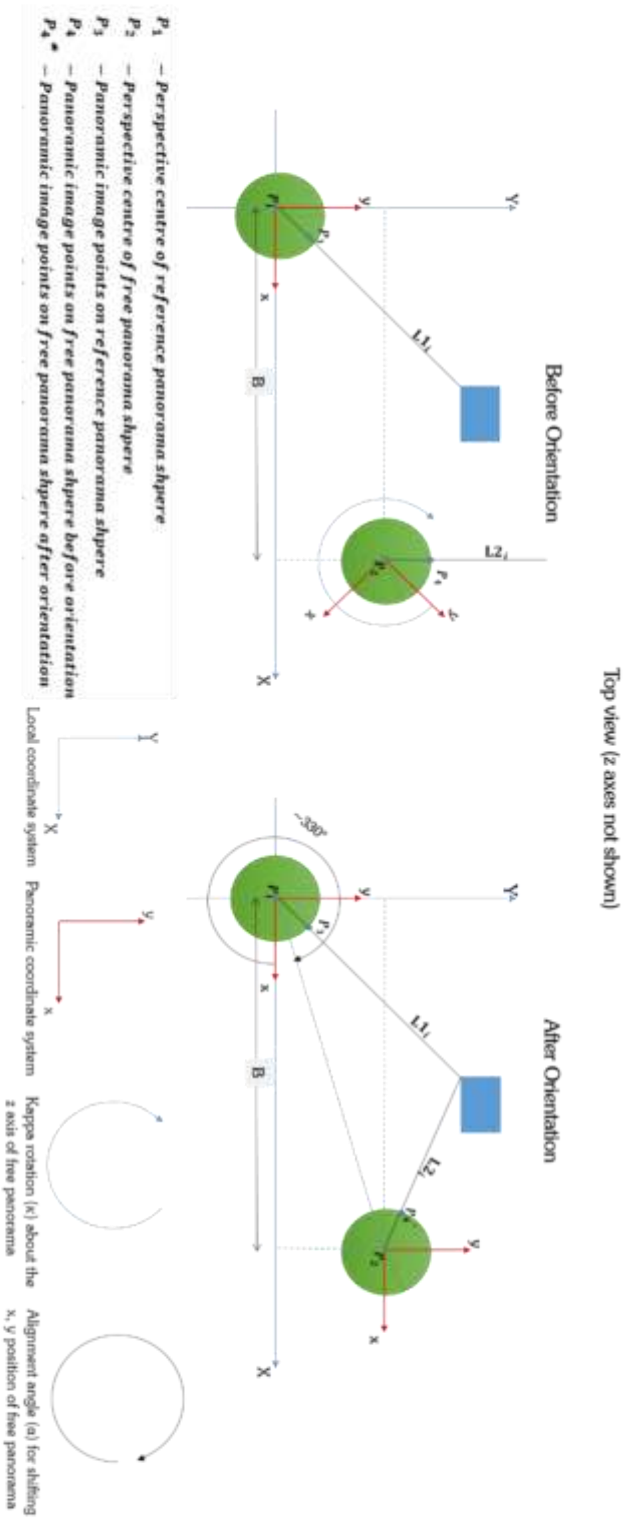


Figure 4-11 MRD algorithm configuration

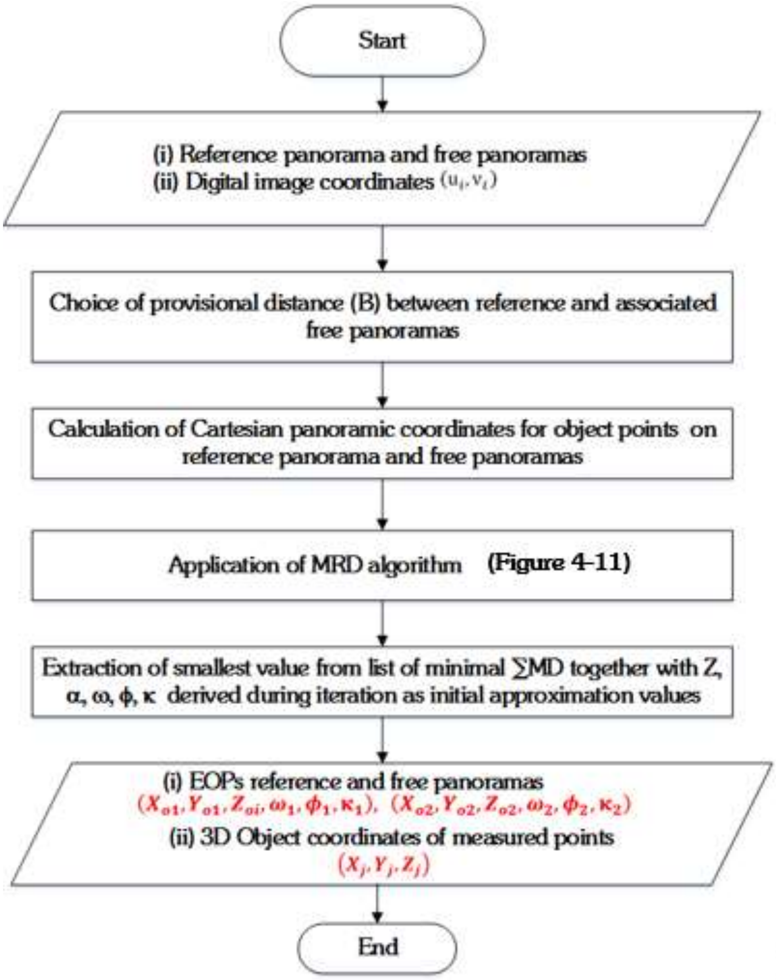


Figure 4-12 Overview of automatic estimation algorithm for the creation of initial approximation values for bundle adjustment

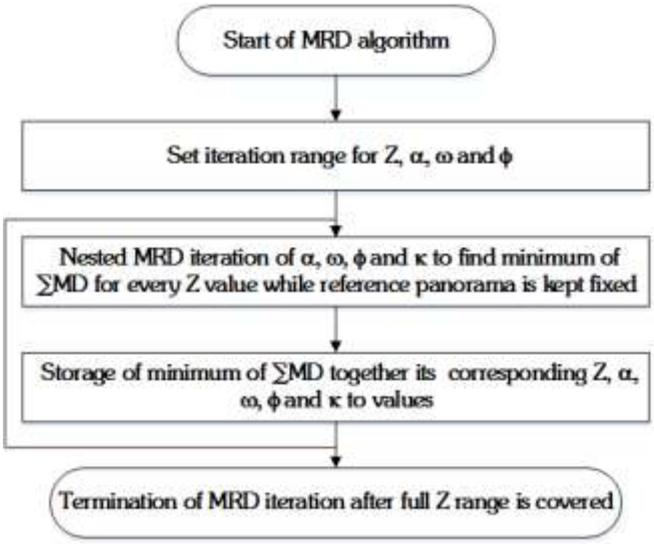


Figure 4-13 MRD algorithm

This involves keeping the EOPs of the reference panorama fixed while the 2D position $(\boldsymbol{X}, \boldsymbol{Y})$, height $(\mathbf{Z})$ and the three rotational angles $(\boldsymbol{\omega}, \boldsymbol{\phi}, \boldsymbol{\kappa})$ of the free panorama are varied in steps. To initiate the iterative orientation process, the default value for the provisional base is chosen as 1 or as a rough realistic estimation of the actual base, the height range for the iteration is then set equal to the base, B allowing iteration for the height to step through values from –B to +B in steps of B/10. Figure 4-14 shows the hill-climbing approach (Russell *et al.*, 2003) to the determination of the height $(\mathbf{Z})$ of the panoramas.

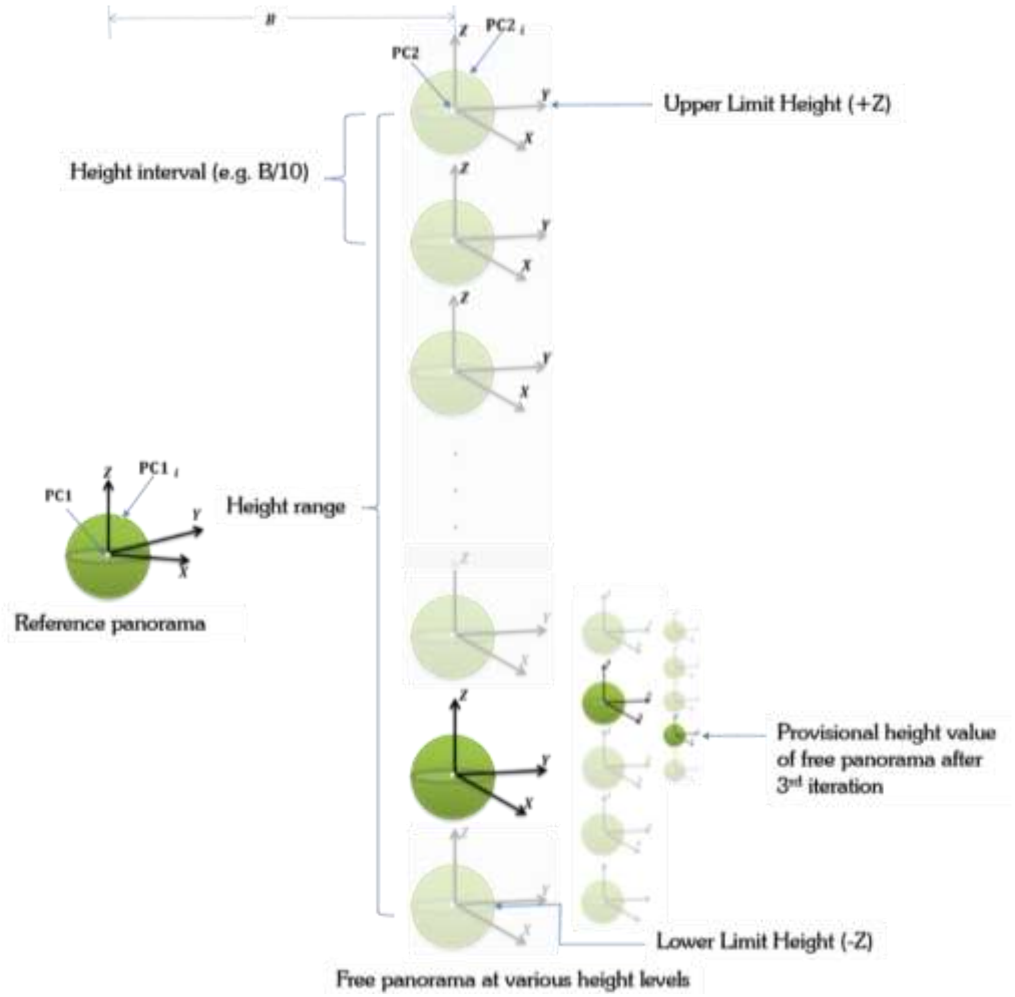


Figure 4-14 Height determination for the free panorama using the MRD algorithm

The 2D position of the free panorama $(\boldsymbol{X_2}, \boldsymbol{Y_2})$ is obtained by varying the 'alignment angle' (α) (Figure 4-15). 'Alignment angle' (α) refers here to the angle between the X-axis of the local coordinate system and the reference base in cases where the axis of the free panorama does not coincide with the axis of the local coordinate system. The iteration for $\boldsymbol{\alpha}$ ranges from 0° to 360° in intervals of 1° or 5°. In practice, one can reduce the range if prior information such as a sketch of the panorama survey is available.

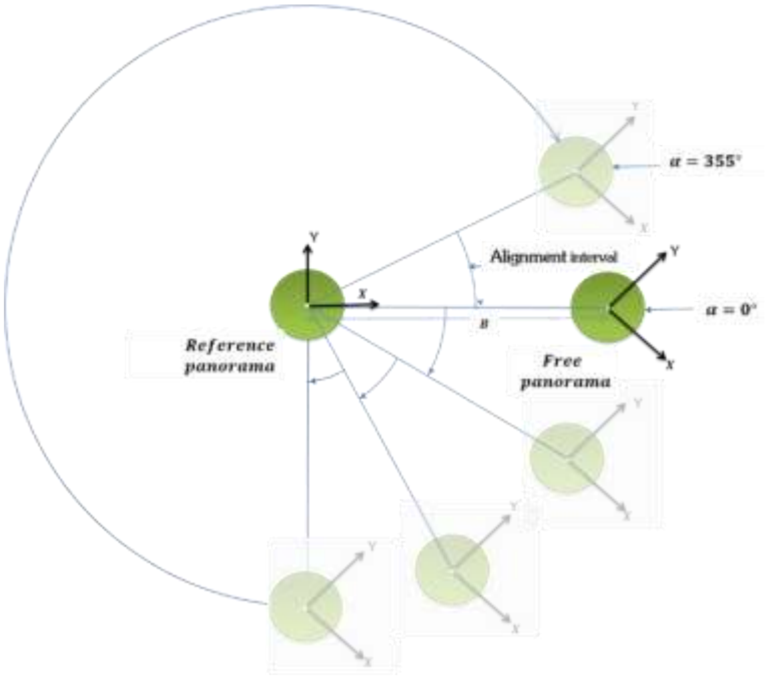


Figure 4-15 Determining suitable alignment angle for the 2D position of free panorama

It should be noted that there are two rotational angles, $\boldsymbol{\alpha}$ is the rotation of the base from its starting position, and κ is the rotation of the free panorama about its centre. The iteration for rotational angle **κ** ranges from 0° to 360° in intervals of 1° or 5°, and the iteration for rotational angles **ω** and $\boldsymbol{\phi}$ ranges from -10° to +10° in intervals of 1°. The three angles and the height difference can then be refined using smaller steps within a limited range (narrower search space) around the previous approximation.

As mentioned above, for each iteration step the minimum distances ($\boldsymbol{d}$) between the conjugate rays ($\mathbf{L1}_{\boldsymbol{i}}$ and $\mathbf{L2}_{\boldsymbol{i}}$) are determined for all image points (Figure 4-10 and Equation 4-16). The $\boldsymbol{Z}$, $\boldsymbol{\omega}$, $\boldsymbol{\phi}$ and $\boldsymbol{\kappa}$ values associated with the global minimum of the sum of all minimum distances indicate the optimal position and orientation between the two panoramas. The mid-point between the point vectors $\boldsymbol{Pa}_{(\mathbf{s})}$ and $\boldsymbol{Pb}_{(\mathbf{t})}$ then define the provisional values for 3D coordinates of the object points $(\boldsymbol{X_j}, \boldsymbol{Y_j}, \boldsymbol{Z_j})$.

4.5. Scaling of Object Coordinates

Prior to the application of the bundle adjustment for the panorama configuration, a uniform scale for all the panoramas is established automatically by determining the distance between two objects which are selected on the basis of maximum separation. If the distance between two points in object space is known or measured, then an absolute scale can be adopted.

4.6. Bundle Adjustment Algorithm

The bundle adjustment algorithm is applied twice. Once, to refine the coordinates of all object points used in the MRD iteration process and then to determine the coordinates of the additional points extracted guided by the epipolar geometry or extracted automatically via feature-based matching.

The mathematical model for the bundle adjustment algorithm is based on the collinearity equations for spherical panoramic images derived in Section 4.2 (Equations 4-14 and 4-15). Subscripts are omitted for reasons of simplicity.

$$u = r\ arctan\left(\frac{X^*}{Y^*}\right) = F_u \qquad \text{Equation 4-17}$$

$$v = r\ arccos\left(\frac{Z^*}{d}\right) = F_v \qquad \text{Equation 4-18}$$

where,

$$X^* = m_{11}(X - X_o) + \ m_{21}(Y - Y_o) + \ m_{31}(Z - Z_o)$$

$$Y^* = m_{12}(X - X_o) + \ m_{22}(Y - Y_o) + m_{32}(Z - Z_o)$$

$$Z^* = m_{13}(X - X_o) + \ m_{23}(Y - Y_o) + m_{33}(Z - Z_o)$$

$$d = \sqrt{(X - X_o)^2 + (Y - Y_o)^2 + (Z - Z_o)^2}$$

No APs were introduced into the collinearity equations as the individual images captured by the camera were rectified prior to stitching. Similar to the standard collinearity equations for rectilinear images, collinearity equations for spherical panoramic images are non-linear. These equations were linearised via Taylor series expansion as shown in Equations 4-19 and 4-20.

$$v_u = \left(\frac{\partial F_u}{\partial X_o}\right)\partial X_o + \left(\frac{\partial F_u}{\partial Y_o}\right)\partial Y_o + \left(\frac{\partial F_u}{\partial Z_o}\right)\partial Z_o + \left(\frac{\partial F_u}{\partial \omega}\right)\partial \omega + \left(\frac{\partial F_u}{\partial \phi}\right)\partial \phi + \left(\frac{\partial F_u}{\partial \kappa}\right)\partial \kappa \left(\frac{\partial F_u}{\partial X}\right)\partial X + \left(\frac{\partial F_u}{\partial Y}\right)\partial Y + \left(\frac{\partial F_u}{\partial Z}\right)\partial Z - \left(u - {F_u}^0\right) \qquad \text{Equation 4-19}$$

$$v_v = \left(\frac{\partial F_v}{\partial X_o}\right)\partial X_o + \left(\frac{\partial F_v}{\partial Y_o}\right)\partial Y_o + \left(\frac{\partial F_v}{\partial Z_o}\right)\partial Z_o + \left(\frac{\partial F_v}{\partial \omega}\right)\partial \omega + \left(\frac{\partial F_v}{\partial \phi}\right)\partial \phi + \left(\frac{\partial F_v}{\partial \kappa}\right)\partial \kappa + \left(\frac{\partial F_v}{\partial X}\right)\partial X + \left(\frac{\partial F_v}{\partial Y}\right)\partial Y + \left(\frac{\partial F_v}{\partial Z}\right)\partial Z - \left(v - {F_v}^0\right) \qquad \text{Equation 4-20}$$

where u and v are the image coordinates of the object points in the 2D format of the panorama image. ${F_u}^0$ and ${F_v}^0$ are Equations 4-17 and 4-18 respectively, evaluated at the approximate values of their parameters. The individual partial derivatives of the linearised collinearity equations with respect to the unknown parameters are given in Appendix 2.

4.7. Interactive Extraction of 3D Measurements from Spherical Panoramic Images

After the relative orientation, epipolar geometry theory was employed to guide the interactive extraction of additional feature points. The epipolar geometry of a panorama pair reduces the search space from a full 2D image space image to 1D (epipolar curve). Figure 4-16 shows the epipolar geometry between two oriented panorama spheres where P an object point in space is. All other terms used in the figures have previously been defined (Section 4.4).

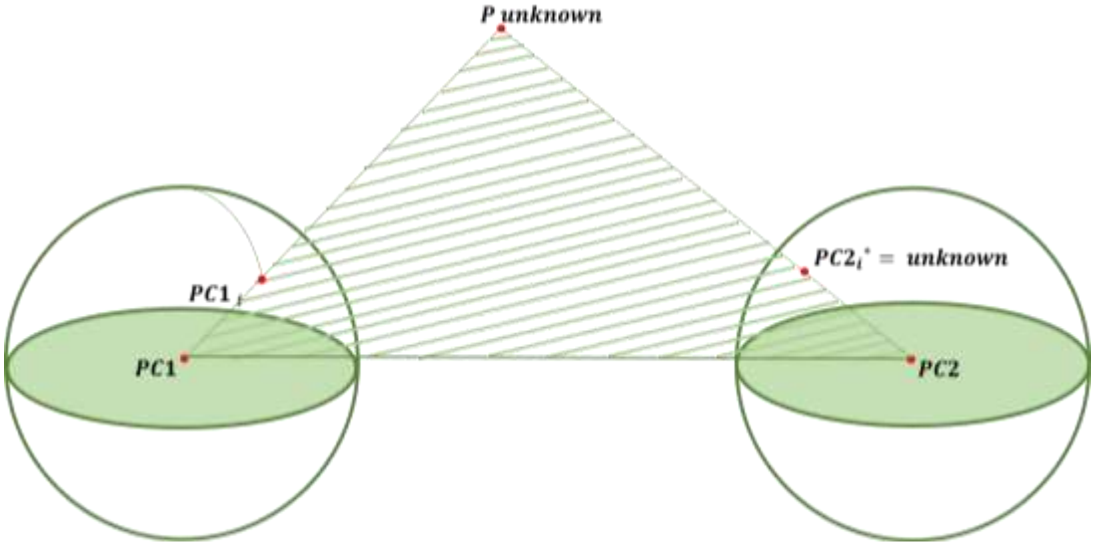


Figure 4-16 Epipolar geometry between two oriented spherical panoramas

The plane defined by $\mathrm{PC1}, \mathrm{PC2}$ and P is known as the epipolar plane. The intersection of the epipolar plane with the two oriented panorama spheres defines great circles on each of these two spheres (Figure 4-17), which appears on the panoramic image plane as curves in a sinusoidal shape known as the epipolar curve (Figure 4-18).


(a) (b)

Figure 4-17 Intersection of epipolar plane with a panorama sphere defining a great circle (a) Front view ((b) Back view

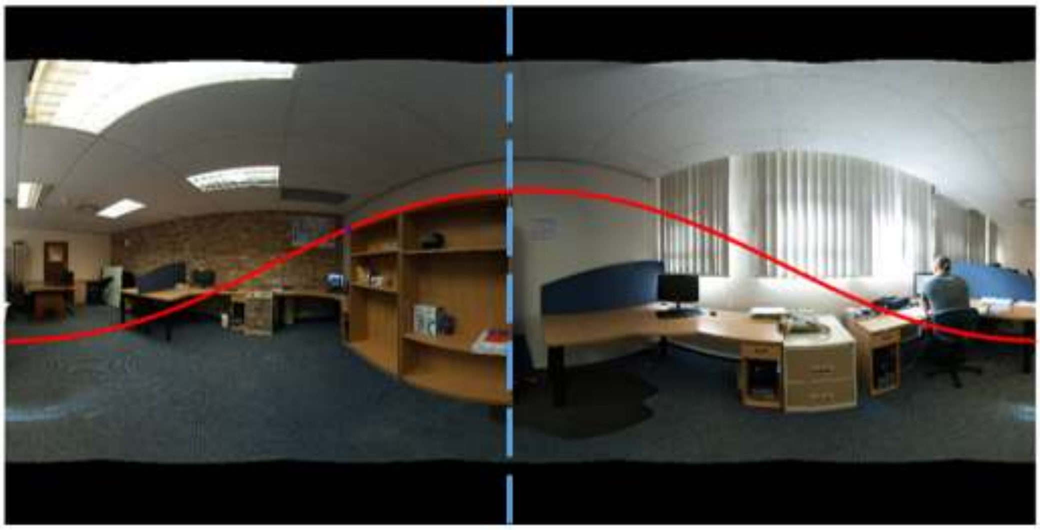


Figure 4-18 Epipolar curve on the panoramic image plane

To derive the mathematical formula for the epipolar curve on the panoramic image plane, the general equation of the plane is used. This is given as:

$$Ax + By + Cz = \mathrm{n} \cdot \mathrm{x} = 0 \qquad \text{Equation 4-21}$$

where $\begin{bmatrix} A \\ B \\ C \end{bmatrix}_{1,3}$ is normal vector to the plane and $\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{1,3}$ represent the Cartesian panoramic coordinates of any point on the sphere.

By substituting Equation 4-5 to 4-7 into Equation 4-21, the epipolar plane expressed in spherical coordinates is given as:

$$A * r \sin\lambda \sin\phi + B * r \ \cos\lambda \sin\phi + C * r \ \cos\phi \ = 0 \qquad \text{Equation 4-22}$$

From Equation 4-22, the expression for the great circle, expressed in spherical coordinates as (Fangi and Nardinocchi, 2013):

$$\phi = arctan\left(\frac{-C}{A * \sin\lambda + B * \cos\lambda}\right) \qquad \text{Equation 4-23}$$

where $\boldsymbol{A}, \boldsymbol{B}$ and $\boldsymbol{C}$ are the elements of the normal vector and $\boldsymbol{\lambda}$ is the horizontal angle of the points, which ranges from 0° to 360°.

Finally, the expression for the epipolar curve on the panoramic image plane is obtained by substituting the equation for $\boldsymbol{\lambda}$ (4-10) into Equation 4-24.

$$v = r * arctan\left(\frac{-C}{A * \sin\left(\frac{u}{r}\right) + B * \cos\left(\frac{u}{r}\right)}\right) \qquad \text{Equation 4-24}$$

The horizontal pixel coordinates of the panoramic image ranges from 1 to W where W, as defined in Section 4.2 is the width of the spherical panorama in the 2D image format.

While the epipolar geometry of a panorama pair reduces the search space from the complete 2D image space to a 1D epipolar curve, the search space can be further reduced to a single point if two or more oriented panoramas are available, in which case the search is fully automated. Figure 4-19 shows an example of the extraction of a feature point on a 2D panoramic image with the support of the epipolar geometry after the relative orientation of four panoramas.

Three epipolar curves are plotted on the single panorama intersect at two points, one in the upper-half of the panorama and the second in the lower-half. The point in the lower-half is automatically eliminated based on the minimum distance between two conjugate rays (Equation 4-19).

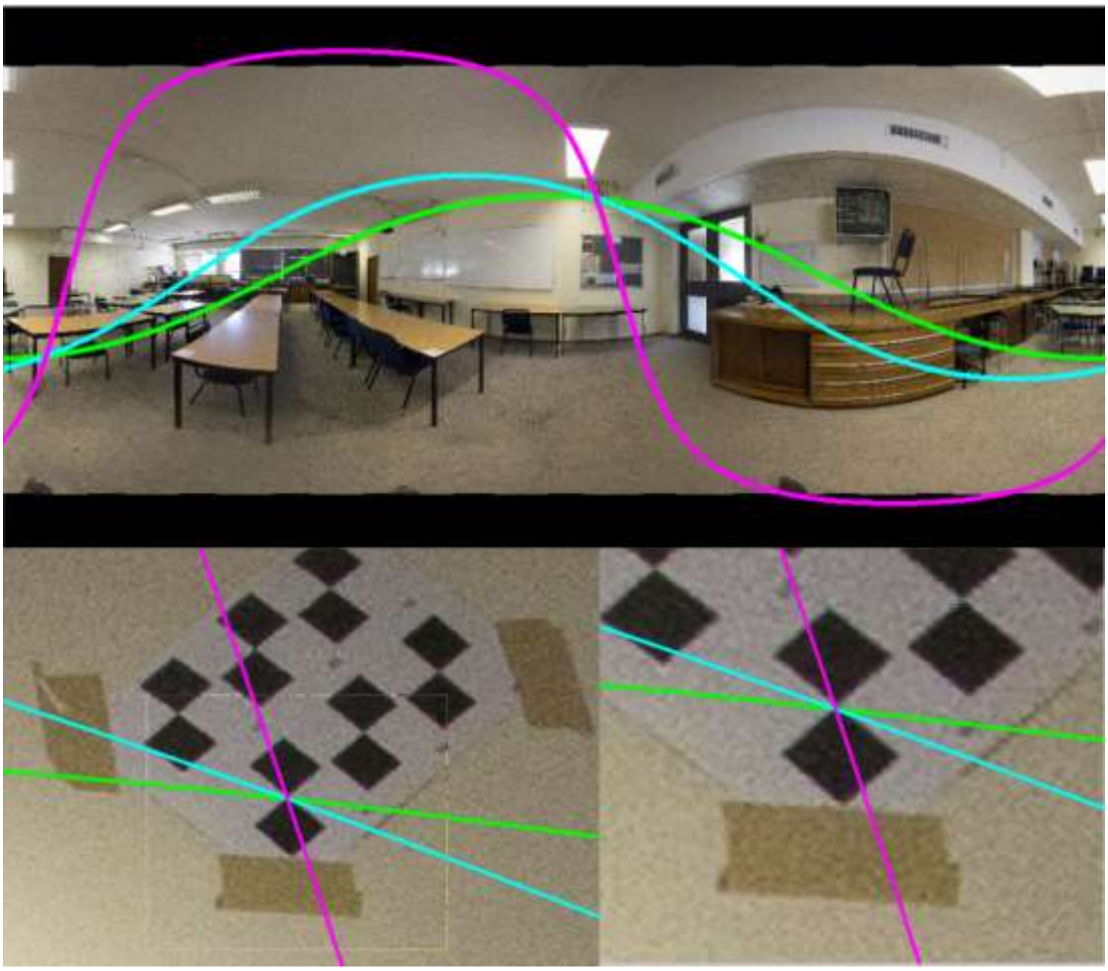


Figure 4-19 Interactive extraction of a feature point from a 2D panoramic image supported by epipolar geometry after relative orientation of four panoramas shown at different zoom levels

4.8. 3D Similarity Transformation

Two sub-routines as in the workflow for the estimation of panorama orientation and 3D object points (Figure 4-2) involve the use of common control points to:

- link individual panorama configurations into a single consistent uniform panoramic configuration; and, where possible;
- georeference panoramas to the relevant national or local coordinate system.

These tasks are carried out with the 3D similarity transformation. The 3D similarity transformation may be defined as:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_o \\ Y_o \\ Z_o \end{bmatrix} + \mu M(\omega, \phi, \kappa) \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad \text{Equation 4-25}$$

where $[X, Y, Z]'$ represents the coordinates in the target system, $[x, y, z]'$ denotes the coordinates of corresponding point in the source system. $[X_o, Y_{o,} Z_o]'$ are the translational parameters, μ is the scale factor and $M(\omega, \phi, \kappa)$ the 3 x 3 orthogonal rotation matrix whose individual elements m_{11} to m_{33} are defined in Equation 2-9.

4.9. Chapter Summary

An accurate, low-cost approach for orienting spherical panoramas and extracting 3D object coordinates from spherical panoramas has been presented in this Chapter.

The orientation procedure is based on the bundle adjustment algorithm. The main contribution of this Chapter is the development of the MRD algorithm to automatically generate initial approximate values for the bundle adjustment algorithm. Both the MRD and bundle adjustment algorithms require conjugate points as input data. These points can be measured manually or automatically. A feature-based matching approach using the SIFT operator has been presented in this Chapter to automatically measure conjugate point coordinates from the spherical panoramas.

Another important contribution of this Chapter is the implementation of the epipolar geometry between relatively oriented panorama pairs and triplets to aid interactive feature point extraction.

CHAPTER 5

EXPERIMENTAL TESTS, RESULTS AND ANALYSIS OF THE MRD ORIENTATION APPROACH

5.1. Overview

This Chapter presents and analyses the results of experimental tests carried out to verify the MRD approach of orienting panoramic images and the subsequent bundle adjustment and extraction of 3D coordinates of object points as shown in the workflows in Figures 4-1 and 4-2.

The approach was first tested at the Geomatics teaching laboratory of the University of Cape Town (indoor environment) and this is presented in Section 5.2. The laboratory was used as a testing site to establish the validity of the approach to object point measurement from panorama photography as developed in this research. The details of the testfield were presented in Section 3.1. The image coordinates of the conjugate points required to initiate the orientation process were manually measured.

In Section 5.3, another experimental test is presented using panoramas captured at Fort Jago, a heritage site in Ghana (outdoor environment). This test demonstrates the generation of sparse point cloud data, using automatically extracted conjugate feature points. The Chapter concludes with a brief summary in Section 5.4.

All experiments in this Chapter were carried out using self-developed Matlab programs including image point measurement capability, MRD and bundle adjustment algorithms as well as 3D similarity transformation. For automatic conjugate feature point extraction, the SIFT operator and matching procedure from the Vfleat open source library (Vedaldi and Fulkerson, 2010) were adopted.

5.2. Determining the 3D Object Coordinates of Target Points of the UCT Calibration Testfield

5.2.1. Data Acquisition

Four panoramas (Figure 5-1) were captured at random positions in the Geomatics teaching laboratory to determine the 3D object coordinates of the target points of the calibration testfield.

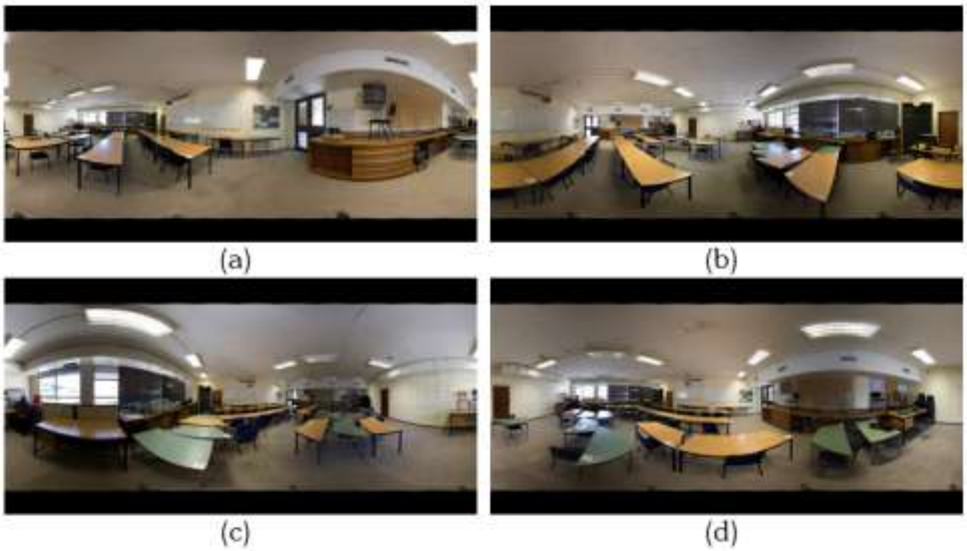


Figure 5-1 Four spherical panoramas of the calibration testfield: (a) Panorama A (b) Panorama B (c) Panorama C (d) Panorama D

The individual images of the panoramas used in this Section were captured with a Canon EOS Rebel T5 camera and the Rokinon 8 mm fisheye lens. The specifications of the camera and lens are provided in Tables 5-1 and 5-2.

Table 5-1 Specifications of Canon EOS Rebel T5 camera

Camera features	
Sensor	CMOS
Maximum image size	5184 x 3456 pixels
Sensor size	22.3 x 14.9 mm
Pixel size	4.3 µm
Approx. dimensions	129.6 x 99.7 x 77.9 mm

Table 5-2 Specifications of the Rokinon 8 mm full-frame fisheye lens

Lens features	
Focal length	8 mm
Maximum aperture	f/3.5
Minimum aperture	f/ 22

The image coordinates of fifteen points on the panoramas, eight of which were target points of the calibration testfield and seven feature points were measured manually. All measured points were common to the four panoramas. The approximate positions of the four panoramas relative to the fifteen image points are shown in Figure 5-2.

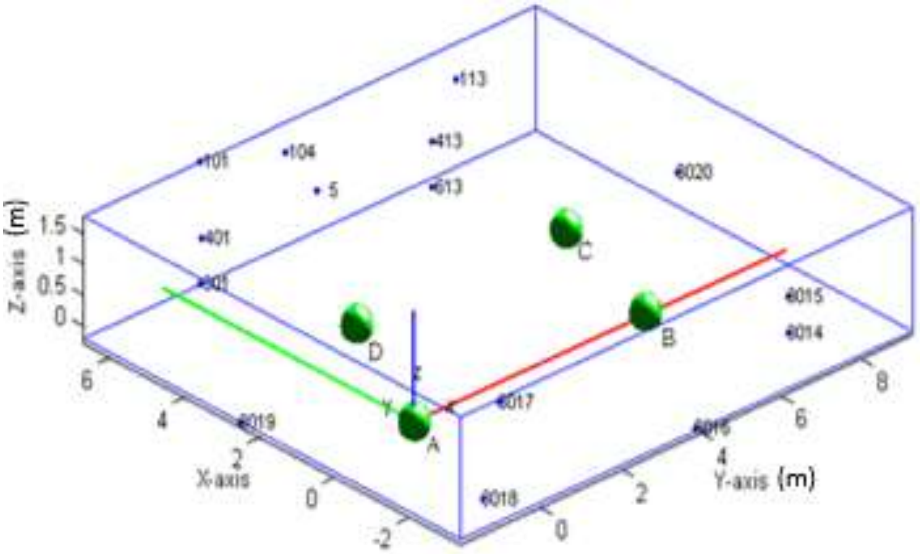


Figure 5-2 Approximate positions of the four panoramas relative to the fifteen image points in the UCT Geomatics teaching laboratory

5.2.2. Panorama Orientation and 3D Measurement Phase

Description of Procedure

Panorama A (Figures 5-1a and 5-2) was chosen as the reference and the other panoramas were oriented and positioned with respect to it. Scaling information was derived from the distance calculated from reference coordinates of two target points 601 and 613. After scaling, the initial approximation values for all

EOPs are available. The bundle adjustment algorithm was applied to refine the initial approximation values for EOP of the panorama as well as the 3D object positions of all the fifteen image points that were involved in the MRD orientation process. The epipolar geometry was then activated to guide the extraction of image coordinates of the remaining sixty-nine target points. The initial approximation values of targets were then calculated. Finally, the EOPs of the four panoramas and as well as all the 3D object coordinates of the target points of the calibration testfield were jointly refined by applying a free-network bundle adjustment.

Experimental Results

Table 5-3 shows the final EOPs of the four panoramas captured at the Geomatic teaching laboratory. The average standard deviations of the seventy-seven 3D objet coordinates of calibration target points were estimated as σ_x=0.7 mm, σ_y=1.2 mm (depth direction), σ_z = 0.6 mm and σ_{xyz} = 1.5 mm.

Table 5-3 Final EOPs of the four panoramas of the Geomatics teaching laboratory

Panorama	X (m)	Y (m)	Z (m)	ω (°)	ϕ (°)	κ (°)
A	-0.008	0.023	-0.004	0.04	-0.24	-0.10
B	5.361	-0.418	0.306	-0.49	0.58	177.37
C	6.419	2.964	0.355	0.58	-0.30	337.24
D	1.182	2.816	0.314	0.18	-0.22	353.25

5.2.3. Testing the Accuracy of the MRD Algorithm

Experimental Test

First, the EOPs of the panoramas obtained from the MRD algorithm were compared with their values after bundle adjustment. The bundle adjustment, like the MRD algorithm was also applied to a pair of panoramas at a time to provide a good basis for comparison. The bundle adjustment algorithm was based on the minimum constraint approach where the seven parameters required to define the datum are held fixed. The parameters that were held fixed in this experiment were the six EOPs of the reference panorama (A) as well as the x-ordinate of the free panorama. This allowed for direct comparison of the EOPs of the free panoramas (Tables 5-4 and 5-5).

Subsequently, the 3D object positions of the fifteen image points obtained from the MRD and bundle adjustment algorithms were also compared with their reference values obtained by survey measurements (Table 5-6). The comparison was only possible after applying a 3D similarity transformation. Four target points served as controls whilst the remaining four were used as check points to verify the accuracy of the object points.

Experimental Results and Analysis

Tables 5-4 and 5-5 compare the EOPs of the Panorama B and C respectively, after applying MRD and bundle adjustment algorithms. Table 5-6 also shows the accuracy of the 3D object coordinates of 15 points after applying MRD and bundle adjustment algorithms.

Table 5-4 Comparison of EOPs for Panorama B after applying MRD and bundle adjustment algorithms

Parameters	MRD	Bundle adjustment	Difference
X_o (m)	5.383	5.383	0.000
Y_o (m)	-0.415	-0.416	0.001
Z_o (m)	0.298	0.299	-0.001
ω (deg. decimals)	0.34	0.35	-0.01
ϕ (deg. decimals)	-0.37	-0.36	-0.01
κ (deg. decimals)	-177.72	-177.74	0.02

Table 5-5 Comparison of EOPs for Panorama C after applying MRD and bundle adjustment algorithms

Parameters	MRD	Bundle adjustment	Difference
X_o (m)	6.436	6.436	0.000
Y_o (m)	2.812	2.966	-0.155
Z_o (m)	0.310	0.338	-0.028
ω (deg. decimals)	0.42	-0.38	0.80
ϕ (deg. decimals)	-0.11	0.18	-0.29
κ (deg. decimals)	337.52	337.47	0.05

Table 5-6 Accuracy of 3D object points from MRD and bundle adjustment orientation

Parameters	Orientation of Panorama A, B		Orientation of Panorama A, C	
	MRD	Bundle adjustment	MRD	Bundle adjustment
σ_X (mm)	0.1	0.1	2.6	0.5
σ_Y (mm)	0.5	0.5	4.2	0.2
σ_Z (mm)	0.9	0.6	9.4	2.0
σ_{XYZ}(mm)	1.0	0.8	10.6	2.1

From Tables 5-4 to 5-6, it can be observed that the 3D positions of Panoramas B and C as well as the 3D object coordinates of the 15 points obtained from the MRD and bundle adjustment algorithms differed in the order of few millimetres. This is an indication that, while the accuracy of the bundle adjustment algorithm was better (Table 5-6), the MRD algorithm provided good initial approximation values. Should millimetre accuracy be desired, then the initial approximation values obtained from the MRD orientation are sufficient as final values.

However, the MRD, unlike the bundle adjustment algorithm can only be applied to a pair of panoramas at a time. It is a well-known fact, that the precision of 3D object points as well as the reliability of the photogrammetric network is greatly improved in a multi-image orientation process. Thus, including additional panorama stations will increase network redundancy and the chance of detecting gross and systematic errors. For this reason, as well as the fact that the bundle adjustment algorithm provides improvement to the

MRD results, it is advisable to always apply the bundle adjustment after the MRD algorithm.

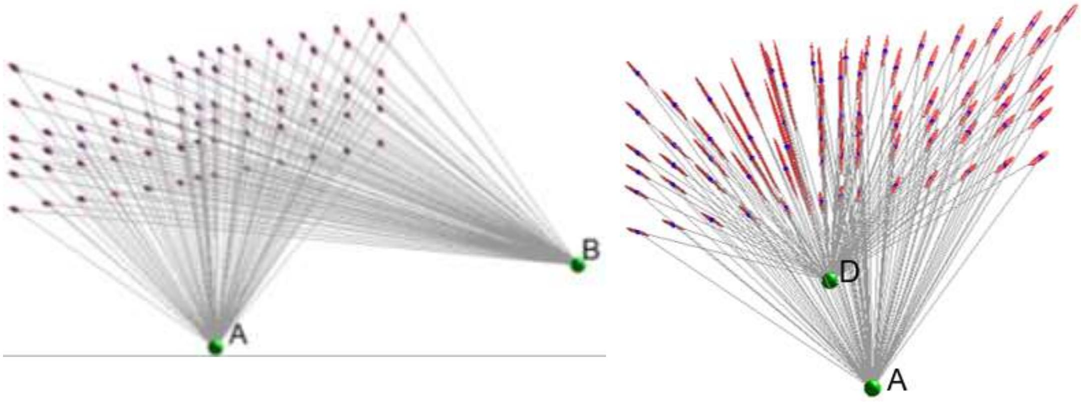
5.2.4. Influence of Imaging Configurations on the 3D Object Points

Experimental Test

A further test was carried out on two pairs of panoramas to determine (confirm) the effect of weak panorama network configurations on the determination of 3D coordinates of object points. Seventy-seven target points of the calibration testfield were used for this test. The test was performed after applying MRD algorithm and a free network adjustment to panoramas A and B (Figure 5-3a) as well as to panoramas A and D (Figure 5-3b).

Experimental Results and Analysis

Figure 5-3 and Table 5-7 show the geometry and the average standard deviation of the seventy-seven object point coordinates from network configuration 1 (Panorama A and B) and network configuration 2 (Panorama A and D) respectively after the free network adjustment.



(a) (b)

Figure 5-3 Comparison between strong and weak network configurations showing error ellipsoids for object points: (a) Network configuration 1 has large B/D ratio, (b) Network configuration 2 has small B/D ratio

Table 5-7 Comparison of average standard deviations of 77 object point coordinates obtained from two network configurations

Network configuration	σ_X (mm)	σ_Y (mm)	σ_Z (mm)	σ_{XYZ}
Network 1 (Panoramas A, B)	0.1	0.3	0.1	0.4
Network 2 (Panoramas A, D)	4.4	8.1	1.2	9.3

It is well known that large B/D ensures strong geometry in contrast to small B/D. Typical B/D of about 0.5 – 0.75 have been reported to improve precision in the depth direction while B/D ratio of 0.3 or below increases depth error (Fraser, 1984; Voltolini *et al.*, 2006; Remondino *et al*., 2013). Figures 5-3a and 5-3b demonstrates a case of strong geometry with a B/D of about 0.9 and a weak geometry with a B/D of about 0.25 respectively. It can be noted from Table 5-7 that, as the B/D decreases, the precision of the y-ordinates (depth direction) of the calibration target decreases from 0.3 mm to 8.1 mm. This justifies why the shape of the error ellipsoids in Network configuration 1 is more homogeneous and isotropic than Network configuration 2 (Figure 5-3).

The geometry of Network configuration 2 can be strengthened by including more panoramas in a multi-image orientation via bundle adjustment procedure. This is obvious from the resulting average standard deviations of the seventy-seven calibration target points obtained after the bundle adjustment orientation between all the four panoramas (Section 5.2.2).

5.3. Generation of Sparse Point Cloud from Spherical Panoramas

5.3.1. Study Area and Data Acquisition

The MRD algorithm has been tested on a number of cultural heritage sites. An example is Fort Jago, formally known Fort Sao Jago da Mina (Figures 5-4 and 5-5). The fort is located in Elmina along the coast in the “Central Region” of Ghana and was built in the 1660s by the Dutch to protect the Elmina Castle (Figure 6-2) from attacks.

Three panoramas A, B and C were captured at the central court of the Fort Jago with a Nikon D200 DSLR camera and the Nikkor 10.5 mm fisheye lens. The specifications of the camera and lens have previously been provided in Tables 3-1 and 3-2.

The goal of the experiment was to recover the position of the three panoramas and to subsequently demonstrate the generation of sparse point cloud data.

(a) (b)

Figure 5-4 Images of Fort Jago: (a) Front view (b) Central court

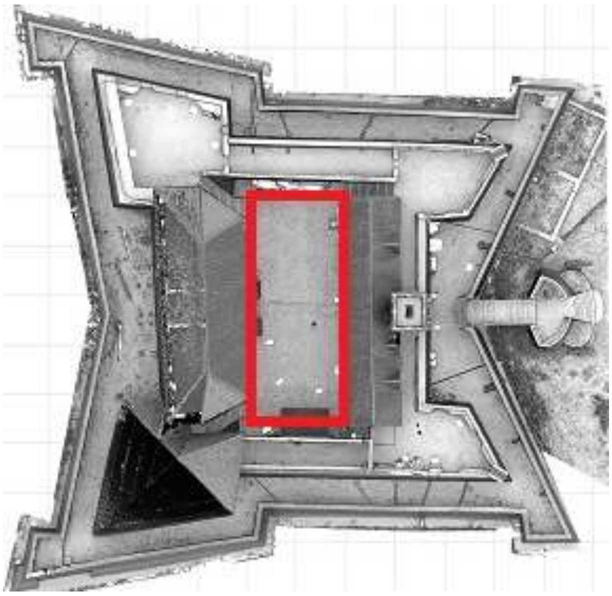


Figure 5-5 Top view of Fort Jago showing central court area (in red)

5.3.2. Procedure for the Generation of Sparse Point Clouds

Conjugate points required to initiate the orientation process were automatically extracted using the matching procedure presented in Section 4-3. First, Panorama A was chosen as the reference and sub-divided into five image tiles

(Figure 5-6). The SIFT operator was applied on Panorama A, B and C to extract feature points. Over 80,000 feature points were extracted from each panorama (Figure 5-7).

Figure 5-6 Panorama A showing the five image tiles

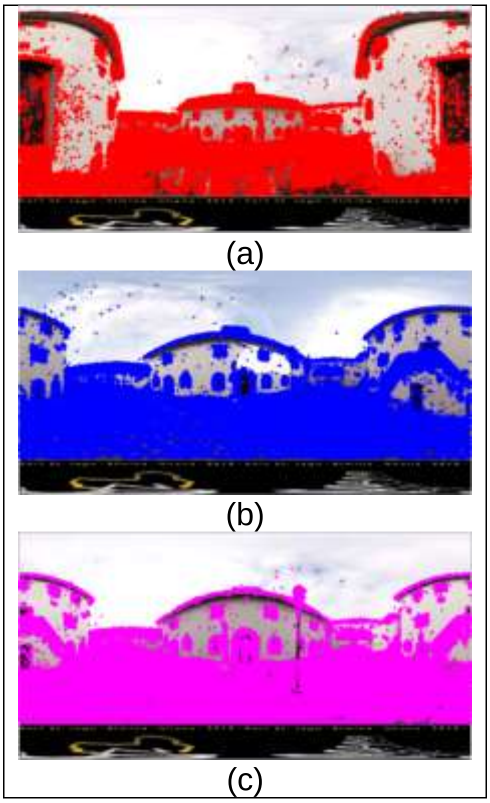

(a)
(b)
(c)

Figure 5-7 Extracted feature points using the SIFT operator: (a) 82,863 feature points on Panorama A (b) 88,590 feature points on Panorama B (c) 137,675 feature points on Panorama C

Panorama B and C were matched with each of the image tiles in Panorama A after which outliers in the matched features were removed by applying the RANSAC algorithm. The number of matched features before and after the outlier detection stages between Panorama C and each of the five image tiles of Panorama A are presented in Table 5-8. Figures 5-8 to 5-10 also show the results of the matching and outlier detection stages between Panorama C and the first three image tiles of Panorama A. After the outlier detection stages, the matched features are automatically stored on condition that the percentage of inliers are 50% and above, otherwise the results must be visually inspected before accepting or discarding. As can be noted, from Figure 5-8, the matches that were obtained after the outlier detection stages contained a number of false matches and thus were discarded.

Table 5-8 Matching and outlier detection between features in Panorama C and the five image tiles of Panorama A

Case No.	Number of points after Feature matching	Number of points after Outliers detection	% Inliers
1	308	11	3.57
2	3322	2679	80.64
3	258	81	31.40
4	564	371	65.78
5	269	10	3.72

Figures 5-8 and 5-9 show graphical representation of the matched feature points between Panorama C and the first two image tiles of Panorama A.

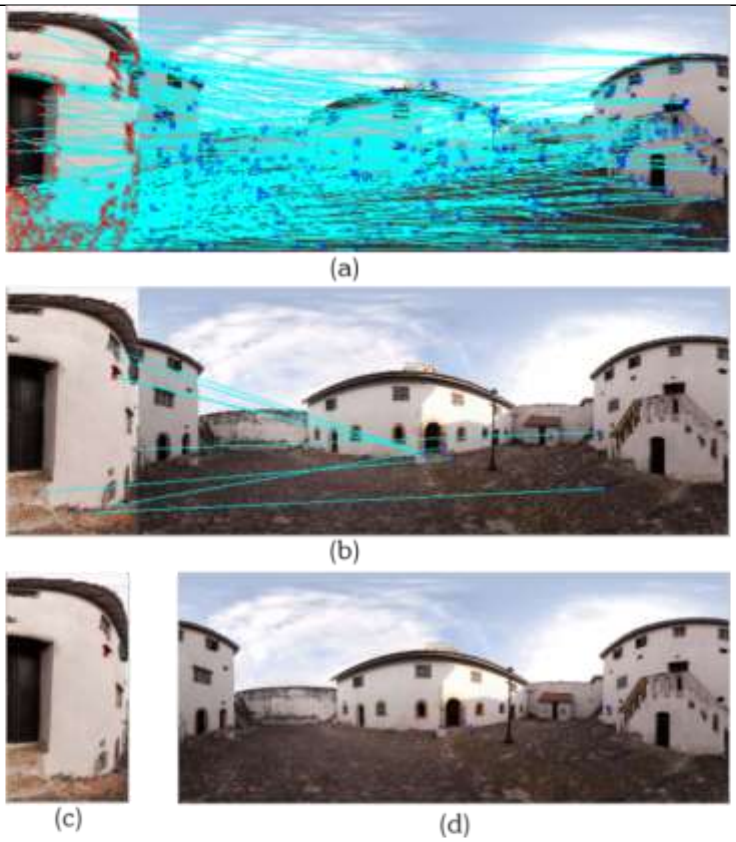


Figure 5-8 Results of feature-based matching and outlier detection procedures between Panorama C and the first image tile of Panorama A: (a) Feature matching (b) After outlier detection (c) Matched features in first image tile of Panorama A (in red) (d) Matched features in Panorama C (in blue)

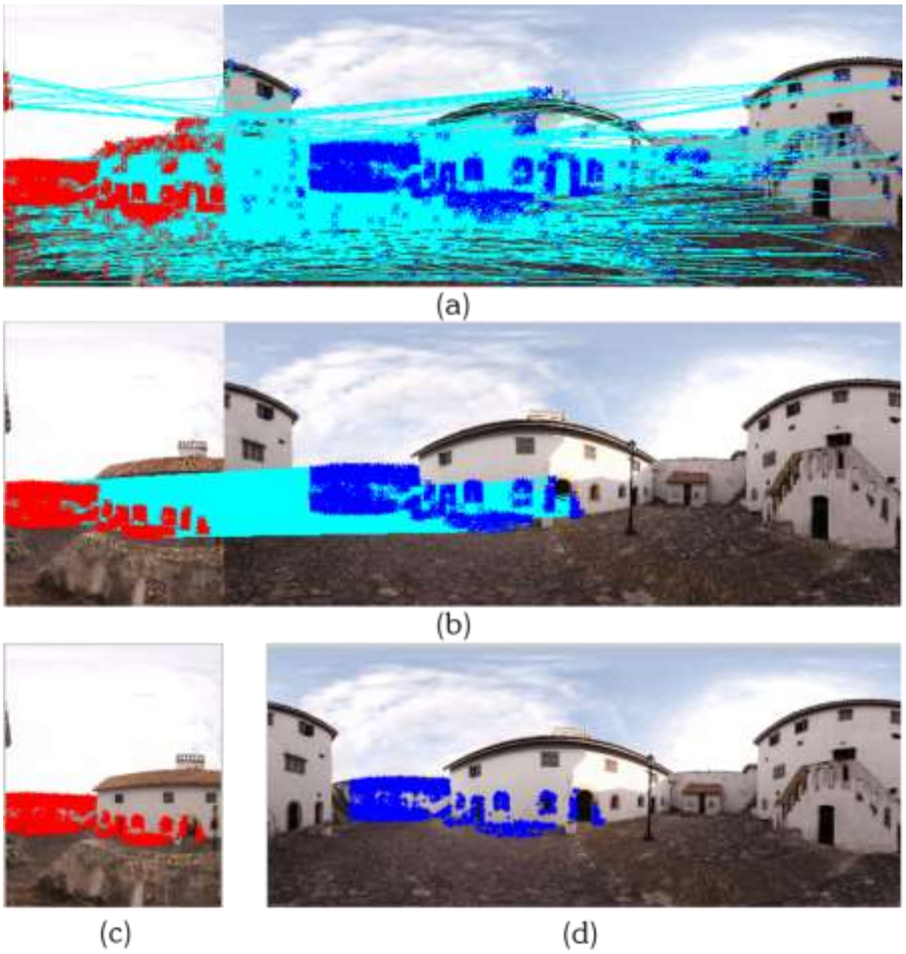


Figure 5-9 Results of feature-based matching and outlier detection procedures between Panorama C and the second image tile of Panorama A: (a) Feature matching (b) After outlier detection (c) Matched features in first image tile of Panorama A (in red) (d) Matched features in Panorama C (in blue)

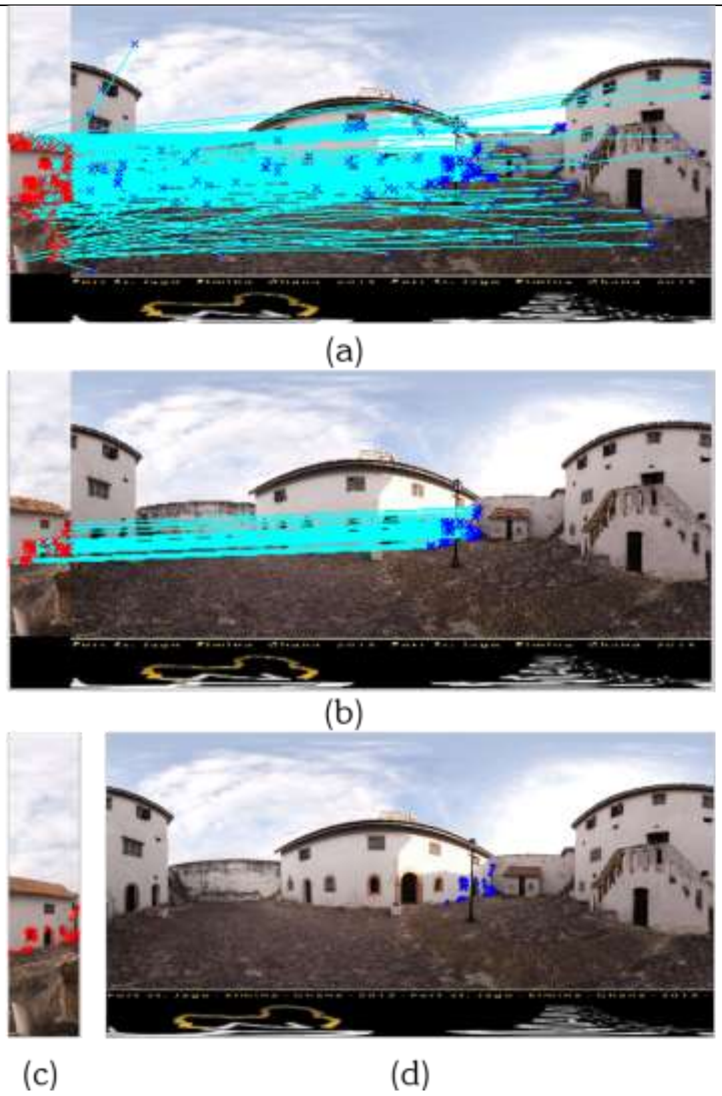


Figure 5-10 Results of feature-based matching and outlier detection procedures between Panorama C and the third image tile of Panorama A: (a) Feature matching (b) After outlier detection (c) Matched features in first image tile of Panorama A (in red) (d) Matched features in Panorama C (in blue)

A total of 3174 and 3152 correct matches were obtained from the matching procedure between panorama pair A and B and panorama pair A and C respectively. Twenty-five feature points common to the three panoramas were selected as input for the orientation phase. The MRD algorithm was then applied to orient Panoramas B and C to Panorama A after which derived objects were scaled. The initial approximation values for the position of the three panoramas as well as their orientation parameters and the 3D object coordinates of the twenty-five feature points were then refined with the bundle adjustment algorithm based on free-network adjustment. The panorama positions as well as the 3D object coordinates of all feature points were then refined with the bundle adjustment algorithm.

5.3.3. Experimental Results

The final EOPs of the three panoramas after applying the MRD and bundle adjustment algorithms are provided in Table 5-9.

Table 5-9 Final EOPs of three panoramas of the central court of Fort Jago

Panorama	X (m)	Y (m)	Z (m)	ω (°)	ϕ (°)	κ (°)
A	0.781	0.014	0.131	0.69	0.26	-1.58
B	2.642	-4.627	-2.424	1.90	0.24	342.77
C	0.461	-5.662	-2.588	0.97	0.59	0.10

The average standard deviation of the 3D coordinates of object points after applying free network adjustment was $\sigma_x = 0.027$ m, $\sigma_y = 0.013$ m, $\sigma_z = 0.008$ m. Figure 5-11 show the sparse point cloud of central court yard of Fort Jago derived from the three panoramas.

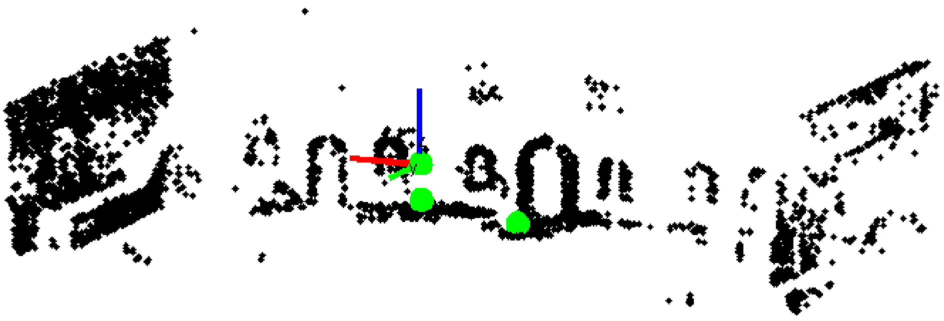


Figure 5-11 3D sparse point cloud of central court yard of Fort Jago

CHAPTER 6

APPLICATION OF MRD ALGORITHM TO LASER SCAN REGISTRATION

6.1. Background of Laser Scanning in Cultural Heritage Documentation

Due to the similarity between the relative orientation of photographic panorama orientation and laser scan registration, the MRD algorithm was tested as a potential tool for the registration of laser scans.

The terrestrial laser scanner (TLS) is an active optical sensor that captures dense 3D points in a very short time by measuring distances between scanner and object based on laser pulse travel time measurement (time-of flight) or phase-shift measurement (Vosselman and Maas, 2010). In addition to these distances, horizontal and vertical angles as well as reflectance intensity values for each point are acquired. Some scanners also capture RGB information for each point.

An important component of the laser scanning processing pipeline toward the creation of 3D models is the determination of relative orientation and position of the various scans used to create the model. This is referred to as laser scan registration. There are various approaches to the registration of laser scans. Prominent among these are the Iterative Closest Point (ICP) algorithm (Besl and McKay, 1992; Chen and Medioni, 1992) and the Least Squares 3D surface matching (LS3D) (Gruen and Akca, 2005). Both techniques require good initial approximation of the laser scan positions and orientations. Other approaches to laser scan registration derive initial approximation values from GNSS systems/IMU sensors (Alba *et al.*, 2012). Photographic or intensity images were used by Ullrich, Schwarz and Kager (2003), Al-Manasir and Fraser (2006) Dold and Brenner (2006) González-Aguilera, Rodr’iguez-Gonzálvez and Gómez-Lahoz (2009), Barazzetti (2010) and Alba *et al*. (2012)

The relative orientation of photographic panoramas and the registration of laser scans require initial approximation values for the bundle adjustment and least squares scan registration respectively. The MRD algorithm, the suitability of which has been demonstrated for the photographic panorama orientation as demostrated in this thesis, should thus also be suitable for the panorama registration. In the laser scanning enviroment, these images can either be intensity images, in which the intensity values of each point is represented by a colour scale or by RGB values, in cases where a scanner has an on-board or associated camera. The software developed in this research is designed for panoramas in equirectangular projections and if intensity or RGB images produced by laser scanner are to be used in the MRD software, then these must be in the same projection. The RGB images from the laser scanner are practically identical to photographic panoramic images and they were expected to perform in the MRD orientation process while the intensity panoramic images required testing.

6.2. Description of Study Area

Elmina castle is one of the numerous trading forts along the coasts of Ghana (formally Gold Coast) (Figure 6-1). The castle is located on 5° 05' North and 1° 21' West in Ghana’s “Central Region”. It was built in 1482 by the Portuguese in order to protect their interests in the gold trade and it was originally named São Jorge da Mina (*St. George of the Mine).* Later, the castle became the centre of the Dutch slave trade, after its capture by the Dutch in 1637. In 1872, it was seized by the British Empire. Associated with the Elmina castle is Fort St. Jago (Section 5.3 and Figure 6-2). Elmina castle remains as one of the best-preserved and most impressive castles in Africa and has been studied intensively from archaeological, architectural and historical points of view. The castle-now-turned-museum is currently being managed by the Ghana Museums and Monuments Board (GMMB). It was inscribed in 1979 by the United Nations Organisation for Education, Science and Culture (UNESCO) as a World Heritage Site (UNESCO, 2016). The 3D model of the Elmina castle is shown in Figure 6-3.

(a) (b)

Figure 6-1 Elmina Castle: Front view (a) Inner court (b)

Figure 6-2 GIS of Elmina castle and its environs

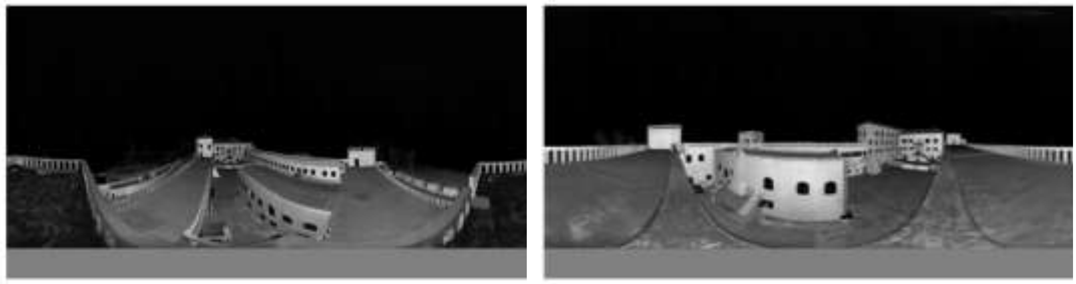
Figure 6-3 3D model of the Elmina castle

6.3. Laser Scan Registration

6.3.1. Data Acquisition

The laser scanning of the Elmina castle was part of the of the Zamani project and as earlier mentioned, the research work documented in this thesis is a contribution to the activities of the Zamani project.

For the test of the MRD algorithm for laser scan registration, seven laser scans acquired during a field campaign with Zamani research team were used. The scans were taken in June 2013 with a Z+F Imager 5010. This scanner has no inbuilt camera and only full dome panoramic intensity images are generated. Figures 6-4a and 6-4b show two examples of panoramic intensity images that were used for the registration process.



(a) (b)

Figure 6-4 Panoramic intensity images of Elmina castle used for the laser scan registration

6.3.2. Test of MRD Procedure for Laser Scan Registration

The application of the MRD algorithm (Figure 4.8) to laser scanning was tested with seven panoramas in two panoramic configurations, one of three and the other with four.

Conjugate points to be used as input for the orientation process were automatically extracted using the SIFT operator based on the procedure described in Section 4.3. The second panoramic configuration become necessary as enough feature points could not automatically be extracted from the chosen reference panorama of the first configuration. To avoid excessive computation time, these features were reduced to fourteen and sixteen common points for the two panorama configurations respectively. Of these, six feature points were common to both configurations.

The MRD algorithm was then applied to estimate the initial approximation values of laser scan setups and orientation. Derived object points were then scaled after which a free-network adjustment was applied to each setup to refine the positions of the scans and their orientation parameters. Finally, the object points of both panoramas configurations were transformed into a unique coordinate system.

6.3.3. Experimental results

Table 6-1 shows the standard deviations of the object points from the two panorama configurations after the bundle adjustment and 3D similarity transformation procedures.

Table 6-1 Results of laser scan registration

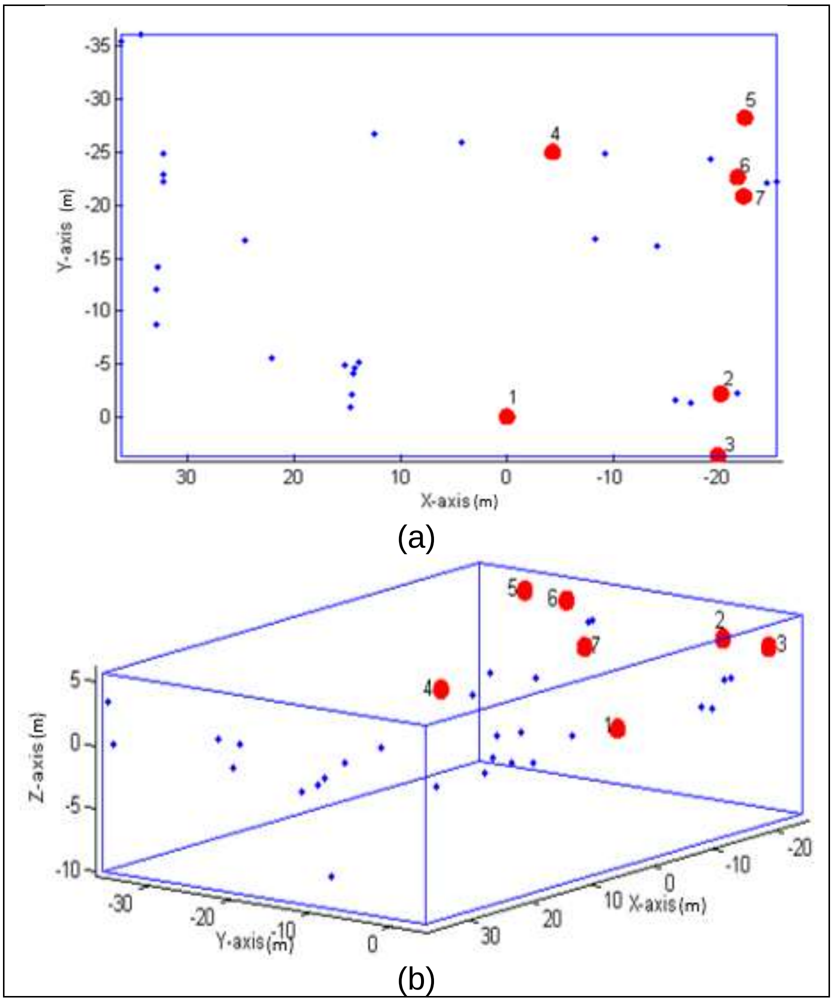
	σ_X (m)	σ_Y (m)	σ_Z (m)	σ_{XYZ} (m)
Bundle Adjustment for Setup 1	0.034	0.020	0.006	0.040
Bundle Adjustment for Setup 2	0.038	0.015	0.006	0.041
3D transformation (4 common points)	0.039	0.014	0.025	0.048

The final EOPs of the seven panoramas after the 3D similarity transformation of the two panorama configurations into a unique coordinate system are provided in Table 6-2.

Table 6-2 Final EOPs of three panoramas of the central court of Fort Jago

Panorama	X (m)	Y (m)	Z (m)	ω (°)	φ (°)	κ(°)
1	0.022	0.012	-0.008	-0.02	-0.01	0.07
2	-20.078	-2.164	4.062	0.02	0.02	42.33
3	-19.809	3.672	4.063	0.06	-0.04	89.58
4	-4.296	-24.922	-0.526	-0.95	0.69	-171.54
5	-22.406	-28.226	4.193	-0.94	0.73	106.91
6	-21.719	-22.589	4.177	-0.92	0.70	129.64
7	-22.267	-20.777	0.620	-0.94	0.78	103.45

Figures 6-5a and 6-5b show the 2D and 3D views of the positions of the seven scans (in red) and object points (in blue) after the laser scan registration process.



(a)

(b)

Figure 6-5 Scan positions and object points in 3D space after laser scan registration (a) 2D view (b) 3D view

6.4. Chapter Summary

In this Chapter, the research objective to extend the MRD algorithm to laser scanning technology has been achieved. The algorithm has been tested using intensity panoramic images of the Elmina Castle in Ghana. After conjugate points were measured automatically using the feature-based matching procedure presented in Section 4.3, the MRD algorithm was applied to determine the approximate scan positions and orientations. The free-network adjustment was finally applied to refine the initial approximation values.

CHAPTER 7

CONCLUSIONS AND RECOMMENDATION

7.1. Conclusions

This research has presented a close-range photogrammetric method for the semi-automatic extraction of 3D measurements from spherical photographic panoramas.

A component of the research was the design of a novel model for the non-parametric correction of lens distortions for images acquired with fisheye lenses. The model is based on the quantification of distortion characteristics of a lens-camera combination from a single reference image. This was achieved by creating a distortion matrix through the comparison of near-error-free simulated image point positions of targets of a calibration testfield with the positions of these targets on the reference image. The difference between the distorted and fictitious image coordinates was the amount of distortion that was applied to the specific positions on the camera's sensor where the targets were imaged. Images captured with the same lens and cameras were then rectified using a backward pixel mapping strategy. Spherical panoramas were subsequently generated from the rectified images. As a check of the effectiveness of the developed technique, a best-fit line function was used to verify that a straight line in object space appears as straight line on the image after rectification. The standard deviations before (180.6 µm equivalent to 28 pixels) and after (17.5 µm or approximately 3 pixel) fitting the best-fit line indicate that the lens distortion has been minimised. The performance of the approach was tested, by forming the distortion matrix with only a subset of the calibration points. The known locations of the remaining randomly selected control points were compared with their interpolated values. It was observed that the denser the control points the more precise the interpolated values indicating that the efficiency of the developed method is dependent on the distribution and density of control points.

An equally relevant outcome of the research is the development of the MRD algorithm for the fully automated approximate relative orientation of spherical panoramas. The MRD algorithm is based on the minimisation of distances between two conjugate space vectors or skewed rays by rotating one panorama with respect to a reference panorama. While no information about the relative position and height of the panoramas are known, the MRD algorithm assumes that the panoramas to be oriented are within 10° horizontal or “quasi-horizontal”. Six to ten conjugate points are required as input data for the MRD algorithm. These points are either extracted manually or automatically, using a feature-based matching. The initial approximation values for the EOPs of each panorama $(\boldsymbol{X_{oi}}, \boldsymbol{Y_{oi}}, \boldsymbol{Z_{oi}}, \boldsymbol{\omega_i}, \boldsymbol{\phi_i}, \boldsymbol{\kappa_i})$ and the 3D object coordinates of the measured image points $\left(\boldsymbol{X_j}, \boldsymbol{Y_j}, \boldsymbol{Z_j}\right)$ were automatically generated by the MRD algorithm are refined by the bundle adjustment algorithm.

In an attempt to recover 3D coordinates from oriented spherical panoramas, the epipolar geometry was employed to guide the interactive extraction of additional conjugate points. The epipolar geometry has been implemented for both panorama pairs and triplets. It has been shown in this research that whiles the epipolar geometry of a pair of panoramas reduces the search space from an entire 2D image space to a sinusoidal curve on the image; the search space is further reduced to a single point if multiple oriented panoramas are involved.

Following the research hypothesis that automated processes for the generation of 3D object coordinate information from spherical panoramas will provide an effective, alternative and supplementary means to standard photogrammetric 3D reconstruction. The research has investigated the possibility of extending the MRD algorithm developed in this research to laser scan registration for the first approximation of laser scan setups and scan orientations. Laser scan registration is an important component of the laser scanning processing pipeline toward the creation of 3D models. They are often accomplished using least squares technique such as Iterative Closest Point

(ICP) and Least Squares 3D surface matching (LS3D). Either techniques require good initial approximation of the laser scan position and orientation. Using intensity spherical panoramic images generated by a laser at each scan position scanner, the MRD algorithm was extended to laser scanning registration process to automatically determine the initial approximation values for scan position and orientation prior to any least squares registration.

7.2. Recommendation

As opposed to the feature-based matching procedure developed in this research for extracting conjugate points prior to panorama orientation, it is recommended that a fully-automated feature and area-based matching algorithms based on the geometry of the spherical panorama should be developed. After orientation, the epipolar geometry for spherical panoramas can be employed to cross-validate the conjugate points extracted during the fully-automated image matching stages in order to generate reliable point cloud data.

REFERENCES

Abdel-Aziz, Y.I. & Karara, H.M. 1971. Direct linear transformation into object space coordinates in close-range photogrammetry, in Proc. Symp. Close-Range Photogrammetry. *Urbana-Champaign*. 1–18.

Addison, A.C. 2000. Emerging trends in virtual heritage. *IEEE Multimedia*. 7(2):22–25. DOI: 10.1109/93.848421.

Al-Manasir, K. & Fraser, C.S. 2006. Registration of terrestrial laser scanner data using imagery. *The Photogrammetric Record*. 21(115):255–268. DOI: 10.1111/j.1477-9730.2006.00379.x.

Alba, M., Barazzetti, L., Scaioni, M. & Remondino, F. 2012. Automatic Registration of Multiple Laser Scans Using Panoramic Rgb and Intensity Images. In *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. V. XXXVIII-5/. 49–54. DOI: 10.5194/isprsarchives-XXXVIII-5-W12-49-2011.

Baker, S. & Nayar, S.K. 1999. Theory of single-viewpoint catadioptric image formation. *International Journal of Computer Vision*. 35(2):175–196. DOI: 10.1023/A:1008128724364.

Barazzetti, L. 2010. Automatic Tie Point Extraction From Markerless Image Blocks in Close-Range Photogrammetry. PhD Thesis, Politecnico di Milano, Milan.

Basu, A. & Licardie, S. 1995. Alternative models for fish-eye lenses. *Pattern Recognition Letters*. 16(4):433–441. DOI: 10.1016/0167-8655(94)00115-J.

Bay, H., Tuytelaars, T., Van Gool, L. & Gool, L. Van. 2006. SURF: Speeded up robust features. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. V. 3951 LNCS. A. Leonardis, H. Bischof, & A. Pinz, Eds. Berlin, Heidelberg: Springer. 404–417. DOI: 10.1007/11744023_32.

Beis, J.S. & Lowe, D.G. 1997. Shape indexing using approximate nearest-neighbour search in high-dimensional spaces. In *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*. 1000–1006.

Besl, P.J. & McKay, N.D. 1992. Method for registration of 3-D shapes. *Robotics-DL tentative. International Society for Optics and Photonics*. 586–606. DOI: 10.1117/12.57955.

Blaha, G. 1982. Free networks: minimum norm solution as obtained by the inner adjustment constraint method. *Bulletin G??od??sique*. 56(3):209–219. DOI: 10.1007/BF02525582.

Briggs, A., Li, Y., Scharstein, D. & Wilder, M. 2006. Robot navigation using 1D panoramic images. In *Proceedings of the 2006 International Conference on Robotics and Automation (ICRA 2006)*. 2679–2685. DOI: 10.1109/ROBOT.2006.1642106.

Brown, D.C. 1956. *The simultaneous determination of the orientation and lens distortion of a photogrammetric camera*.

Brown, D.C. 1971. Close-range camera calibration. *Photogram. Eng. Remote Sens*. 37:855–866.

Brown, D.C. 1976. The bundle adjustment—progress and prospects. *Int. Archives Photogrammetry*. 21(3):1.

Brown, M. & Lowe, D.G. 2003. Recognising panoramas. *Proceedings of the {IEEE} International Conference on Computer Vision*. 1218–1225. DOI: 10.1109/ICCV.2003.1238630.

Brown, D.C., Davis, R.G. & Johnson, F.C. 1964. The practical and rigorous adjustment of large photogrammetric nets. *Report on US Air Force, contract No AF-30 (602)-3007. DBA System*.

Burt, P.J. & Adelson, E.H. 1983a. The Laplacian Pyramid as a Compact Image

Code. *IEEE Transactions on Communications*. 31(4):532–540. DOI: 10.1109/TCOM.1983.1095851.

Burt, P.J. & Adelson, E.H. 1983b. A multiresolution spline with application to image mosaics. *ACM Transactions on Graphics*. 2(4):217–236. DOI: 10.1145/245.247.

Busch, D.D. 2000. *Digital photography for dummies quick reference*. John Wiley & Sons, Inc.

Carneiro, G. & Jepson, A.D. 2005. The distinctiveness, detectability, and robustness of local image features. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. V. 2. 296–301.

Chan, A. 2005. The Use of Low Cost Virtual Reality and Digital Technology to Aid Forensic Scene Interpretation and Recording.

Chen, S.E. 1995. Quicktime VR: An image-based approach to virtual environment navigation. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*. 29–38.

Chen, Y. & Medioni, G. 1992. Object modeling by registration of multiple range images. *International Journal of Computer Vision and Image Understanding(IJCVU)*. 10(3):145–155.

Chum, O. & Matas, J. 2005. Matching with PROSAC-progressive sample consensus. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. V. 1. 220–226. DOI: 10.1109/CVPR.2005.221.

Coorg, S. & Teller, S. 2000. Spherical mosaics with quaternions and dense correlation. *International Journal of Computer Vision*. 37(3):259–273. DOI: 10.1023/A:1008184124789.

Creagh, H. 2003. CAVE Automatic Virtual Environment. In *Proceedings of the 2003 IEEE Electrical Insulation Conference and Electrical Manufacturing*

& Coil Winding Technology Conference. 499–504. DOI: 10.1109/EICEMC.2003.1247937.

Devernay, F. & Faugeras, O. 2001. Straight lines have to be straight. *Machine Vision and Applications*. 13(1):14–24. DOI: 10.1007/PL00013269.

Devi, V.S.V., Prasad, S.M. & Bhat, H.G. 2013. An efficient algorithm for image stitching based on scale-invariant feature transform. *International Journal of Engineering*. 2(2):2305–8269.

Dold, C. & Brenner, C. 2006. Registration of terrestrial laser scanning data using planar patches and image data. *International Archives of Photogrammetry*. Vol. XXXVI(5):78–83. DOI: 10.1.1.66.1760.

Douskos, V., Kalisperakis, I., Karras, G. & Petsa, E. 2008. Fully automatic camera calibration using regular planar patterns. *International Archives of the Photogram metry, Remote Sensing and Spatial Information Sciences*. 37(5):21–26. Available: http://citeseerx.ist.psu.edu/viewdoc/download?rep=rep1&type=pdf&doi=10.1.1.144.5619.

Estrela, C., Bueno, M.R., Leles, C.R., Azevedo, B. & Azevedo, J.R. 2008. Accuracy of cone beam computed tomography and panoramic and periapical radiography for detection of apical periodontitis. *Journal of endodontics*. 34(3):273–9. DOI: 10.1016/j.joen.2007.11.023.

Fangi, G. 2007. The Multi-image spherical Panoramas as a tool for Architectural Survey. *XXI International CIPA Symposium*. (October):256--1840. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.222.4044 [2016, January 15].

Fangi, G. 2010. Multiscale Multiresolution Spherical Photogrammetry With Long Focal Lenses for Architectural Surveys. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*. XXXVIII.

Available:
http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.222.4450&rank=1
%5C [2016, January 15].

Fangi, G. & Nardinocchi, C. 2013. Photogrammetric Processing of Spherical Panoramas. *The Photogrammetric Record*. 28(143):293–311. DOI: 10.1111/phor.12031.

Faugeras, O., Benosman, R. & Kang, S.B. 2013. *Panoramic vision: sensors, theory, and applications*. Springer Science & Business Media.

Fischler, M.A. & Bolles, R.C. 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*. 24(6):381–395.

Fraser, C.S. 1984. Network Design Considerations for Non-Topographic Photogrammetry. *Photogrammetric Engineering & Remote Sensing*. 50(8):1115–1126.

Fraser, C.S. 1997. Digital camera self-calibration. *ISPRS Journal of Photogrammetry and Remote sensing*. 52(4):149–159.

Fraser, C.S., Shortis, M.R. & Ganci, G. 1995. Multi-sensor system self-calibration. In *Photonics East'95*. 2–18.

Friedman, J.H., Bentley, J.L. & Finkel, R.A. 1977. An algorithm for finding best matches in logarithmic expected time. *ACM Transactions on Mathematical Software (TOMS)*. 3(3):209–226.

Gaspar, J., Winters, N. & Santos-Victor, J. 2000. Vision-based navigation and environmental representations with an omnidirectional camera. *IEEE Transactions on Robotics and Automation*. 16(6):890–898. DOI: 10.1109/70.897802.

Gonzalez, R.C. & Woods, R.E. 2011. *Digital Image Processing*. 3rd ed. Pearson Education.

González-Aguilera, D., Rodr’iguez-Gonzálvez, P. & Gómez-Lahoz, J. 2009. An automatic procedure for co-registration of terrestrial laser scanners and digital cameras. *ISPRS J Photogramm*. 64(3):308–316. Available: http://www.sciencedirect.com/science/article/pii/S0924271608001111 [2016, January 14].

Grafarend, E.W. 1974. *Optimization of geodetic networks*. V. 33. na. DOI: 10.1029/RG020i004p00877.

Grammatikopoulos, L., Karras, G. & Petsa, E. 2007. An automatic approach for camera calibration from vanishing points. *ISPRS Journal of Photogrammetry and Remote Sensing*. 62(1):64–76. DOI: 10.1016/j.isprsjprs.2007.02.002.

Gruen, A. & Akca, D. 2005. Least squares 3D surface and curve matching. *{ISPRS} Journal of Photogrammetry and Remote Sensing*. 59(3):151–174. DOI: http://dx.doi.org/10.1016/j.isprsjprs.2005.02.006.

Guan, Y., Cheng, X., Zhan, X. & Zhou, S. 2008. Closed-form solution of space resection using unit quaternion. In *Artigo apresentado no XXI ISPRS Congress, Beijing, China*. 3–11.

Hartley, R. & Zisserman, A. 2003. *Multiple view geometry in computer vision*. Cambridge university press.

Hecht, E. & Zajac, A. 1974. Optics. *Addison-Wesley*. 301–305.

Held, C. 2012. Creating 3D models of cultural heritage sites with terrestrial laser scanning and 3D imaging.

Van Den Heuvel, F.A. 1998. Vanishing point detection for architectural photogrammetry. *International archives of photogrammetry and remote sensing*. 32(part 5):652–659. Available: http://www.lr.tudelft.nl/fileadmin/Faculteit/LR/Organisatie/Afdelingen_en_Leerstoelen/Afdeling_RS/Optical_and_Laser_Remote_Sensing/Research/Research_Fields/Heritage_and_Medical/Architectural_Photogrammetry/doc/Heuvelv anish.pdf.

Van Den Heuvel, F.A. 1999. Estimation of interior orientation parameters from constraints on line measurements in a single image. *International archives of photogrammetry …*. 32(Part 5):81–88. Available: http://delftaerospace.org/fileadmin/Faculteit/LR/Organisatie/Afdelingen_en_Leerstoelen/Afdeling_RS/Optical_and_Laser_Remote_Sensing/Publications/Papers/020-1999/doc/Heuvelinterior.pdf.

Van Den Heuvel, F.A., Verwaal, R., Beers, B., O, P. & Waardenburg, N.L.B.B. 2006. Calibration of Fisheye Camera Systems and the Reduction of Chromatic Aberration. *Technology*. 5:6. Available: http://www.isprs.org/proceedings/XXXVI/part5/paper/1267_Dresden06.pdf [2016, January 15].

Hirsch, R. 2012. *Light and lens: photography in the digital age*. CRC Press.

Honey, O.B., Scarfe, W.C., Hilgers, M.J., Klueber, K., Silveira, A.M., Haskell, B.S. & Farman, A.G. 2007. Accuracy of cone-beam computed tomography imaging of the temporomandibular joint: Comparisons with panoramic radiology and linear tomography. *American Journal of Orthodontics and Dentofacial Orthopedics*. 132(4):429–438. DOI: 10.1016/j.ajodo.2005.10.032.

Huang, F., Klette, R. & Scheibe, K. 2008. *Panoramic Imaging: Sensor-Line Cameras and Laser Range-Finders*. John Wiley & Sons.

Hughes, C., Denny, P., Jones, E. & Glavin, M. 2010. Accuracy of fish-eye lens models. *Applied optics*. 49(17):3338–3347. DOI: 10.1364/AO.49.003338.

Irani, M. & Anandan, P. 1999. About direct methods. In *International Workshop on Vision Algorithms*. 267–277.

Jongerius, J. 2003. Google Patents. Available: http://patft.uspto.gov/netacgi/nph-Parser?Sect2=PTO1&Sect2=HITOFF&p=1&u=/netahtml/PTO/search-bool.html&r=1&f=G&l=50&d=PALL&RefSrch=yes&Query=PN/6563529.

Kannala, J. & Brandt, S. 2004. A generic camera calibration method for fish-eye lenses. In *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.* V. 1. IEEE. 10–13 Vol.1. DOI: 10.1109/ICPR.2004.1333993.

Kannala, J. & Brandt, S.S. 2006. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *IEEE transactions on pattern analysis and machine intelligence*. 28(8):1335–40. DOI: 10.1109/TPAMI.2006.153.

Kemp, E.L. 1996. *Industrial Archaeology: Techniques*. (Public History Series). Krieger. Available: https://books.google.com.gh/books?id=0SXbAAAAMAAJ.

Klette, R. & Scheibe, K. 2005. Combinations of range data and panoramic images - New opportunities in 3D scene modeling. In *Proceedings of the Conference on Computer Graphics, Imaging and Vision: New Trends 2005*. V. 2005. 3–12. DOI: 10.1109/CGIV.2005.26.

Koehl, M., Schneider, A., Fritsch, E., Fritsch, F., Rachedi, A. & Guillemin, S. 2013. DOCUMENTATION OF HISTORICAL BUILDING VIA VIRTUAL TOUR: THE COMPLEX BUILDING OF BATHS IN STRASBOURG. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. XL-5/W2:385–390. DOI: 10.5194/isprsarchives-XL-5-W2-385-2013.

Kruck, E. 1998. BINGO-F: bundle adjustment for engineering applications. *User's Manual, Version*. 4.

Leroy, A.M. & Rousseeuw, P.J. 1987. Robust regression and outlier detection. *Wiley Series in Probability and Mathematical Statistics, New York: Wiley, 1987*. 1.

Lowe, D.G. 1999. Object recognition from local scale-invariant features. In *Proceedings of the Seventh IEEE International Conference on Computer Vision*. V. 2. 1150–1157. DOI: 10.1109/ICCV.1999.790410.

Lowe, D.G. 2004. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*. 60(2):91–110. DOI: 10.1023/B:VISI.0000029664.99615.94.

Luhmann, T. & Tecklenburg, W. 2004. 3-D object reconstruction from multiple-station panorama imagery. *Proceedings of ISPRS Workshop on Panorama Photogrammetry*. 34(5). Available: http://www.isprs.org/proceedings/XXXIV/5-W16/papers/PanoWS_Dresden2004_Luhmann_b.pdf [2016, January 15].

Luo, J., Shin, S.-S., Park, H.-J. & Gwun, O.-B. 2011. Stitcing for Panorama based on SURF and Multi-band Blending. *Journal of Korea Multimedia Society*. 14(2):201–209. DOI: 10.9717/kmms.2011.14.2.201.

McBride, B. 1994. *A Timeline of Panoramic Cameras*. Available: http://www.panoramicphoto.com/timeline.htm [2016, January 15].

McGlone, J.C. 1989. Analytic Data-reduction Schemes. In *Non-topographic Photogrammetry*. V. 4. 37–55.

McLauchlan, P.F. & Jaenicke, A. 2002. Image mosaicing using sequential bundle adjustment. *Image and Vision computing*. 20(9):751–759.

Mikhail, E.M., Bethel, J.S. & McGlone, J.C. 2001. *Introduction to Modern Photogrammetry*. V. 1. John Wiley & Sons Inc.

Mistry, S. & Patel, A. 2016. Image Stitching using Harris Feature Detection.

Miyamoto, K. 1964. Fish eye lens. *JOSA*. 54(8):1060–1061.

Morisse, D. 2012. *World Wonders Project : Click to explore a World Heritage site with Google*. Available: whc.unesco.org/document/122289 [2013, August 12].

Oettermann, S. 1997. *The Panorama: History of a Mass Medium*. V. 21. Zone books New York.

Parian, J.A. & Gruen, A. 2010. Sensor modeling, self-calibration and accuracy testing of panoramic cameras and laser scanners. *ISPRS Journal of Photogrammetry and Remote Sensing*. 65(1):60–76. DOI: 10.1016/j.isprsjprs.2009.08.005.

Remondino, F. 2006. Detectors and Descriptors for Photogrammetric Applications. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*. 36(3):49--54. DOI: 10.1.1.71.3114.

Remondino, F., Menna, F., Koutsoudis, A., Chamzas, C. & El-Hakim, S. 2013. Design and implement a reality-based 3D digitisation and modelling project. In *Proceedings of the DigitalHeritage 2013 - Federating the 19th Int'l VSMM, 10th Eurographics GCH, and 2nd UNESCO Memory of the World Conferences, Plus Special Sessions fromCAA, Arqueologica 2.0 et al.* V. 1. 137–144. DOI: 10.1109/DigitalHeritage.2013.6743723.

Russell, S.J., Norvig, P., Canny, J.F., Malik, J.M. & Edwards, D.D. 2003. *Artificial intelligence: a modern approach*. V. 2. Prentice hall Upper Saddle River.

Rüther, H. 1982. Relative orientation with limited control in close range Photogrammetry. University of Cape Town.

Rüther, H. 2002. An African Heritage Database the Virtual Preservation of Africa's Past Abstract : *INTERNATIONAL ARCHIVES OF PHOTOGRAMMETRY AND REMOTE SENSING*. 34(6/W6):185–192.

Rüther, H. 2011. Creating Historical Awareness in Africa. *GIM International*. 25:9. Available: http://www.gim-international.com/content/article/creating-historical-awareness-in-africa [2016, January 13].

Rüther, H., Chazan, M., Schroeder, R., Neeser, R., Held, C., Walker, S.J., Matmon, A. & Horwitz, L.K. 2009. Laser scanning for conservation and research of African cultural heritage sites: the case study of Wonderwerk Cave, South Africa. *Journal of Archaeological Science*. 36(9):1847–1856. DOI: 10.1016/j.jas.2009.04.012.

Rüther, H., Held, C., Bhurtha, R., Schroeder, R. & Wessels, S. 2012. From Point Cloud to Textured Model, the Zamani Laser Scanning Pipeline in Heritage Documentation. *South African Journal of Geomatics*. 1(1):44–59. Available: http://www.ajol.info/index.php/sajg/article/view/107036 [2016, January 13].

Sawhney, H.S. 1999. True Multi-Image Alignment and its Application to Mosaicing and Lens Distortion Correction True Multi-Image Alignment and its Application to Mosaicing and Lens Distortion Correction. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. 21(3):22. DOI: 10.1109/34.754589.

Schmid, H.H. 1955. *An analytical treatment of the problem of triangulation by stereophotogrammetry*.

Schneider, D., Schwalbe, E. & Maas, H.-G. 2009. Validation of geometric models for fisheye lenses. *ISPRS Journal of Photogrammetry and Remote Sensing*. 64(3):259–266. DOI: 10.1016/j.isprsjprs.2009.01.001.

Schwalbe, E. 2005. Geometric modelling and calibration of fisheye lens camera systems. In *Camera*. V. XXXVI. Part 5/W8. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.84.6969&rep=rep1&type=pdf.

Sequeira, V. & Goncalves, J.G.M. 2002. 3\text{D} reality modelling: photo-realistic 3\text{D} models of real world scenes. In *Proc. First International Symposium on 3D Data Processing Visualization and Transmission*. 776–783.

Sequeira, V., Ng, K., Wolfart, E., Gonçalves, J.G.M. & Hogg, D. 1999. Automated reconstruction of 3D models from real environments. *ISPRS Journal of Photogrammetry and Remote Sensing*. 54(1):1–22. DOI: http://dx.doi.org/10.1016/S0924-2716(98)00026-4.

Shoval, S., Borenstein, J. & Koren, Y. 1994. Mobile robot obstacle avoidance in a computerized travel aid for the blind. In *Proceedings of the 1994 IEEE*

International Conference on Robotics and Automation. IEEE Comput. Soc. Press. 2023–2028. DOI: 10.1109/ROBOT.1994.351166.

Shum, H.Y. & Szeliski, R.S. 2002. Google Patents. Available: https://www.google.ch/patents/US6271847.

Slama, C.C., Theurer, C. & Henriksen, S.W. 1980. *Manual of photogrammetry*. American Society of photogrammetry.

Snyder, J.P. 1987. *Map Projections: A Working Manual*. V. 1395. US Government Printing Office. DOI: 10.2307/1774978.

Swaminathan, R. & Nayar, S.K. 1999. Polycameras: Camera Clusters for Wide Angle Imaging CUCS-013-99. *Department of Computer Science, Columbia University, New York, New York*. 0. Available: citeulike-article-id [2016, January 15].

Szeliski, R. 1996. Video mosaics for virtual environments. *IEEE Computer Graphics and Applications*. 16(2):22–30. DOI: 10.1109/38.486677.

Szeliski, R. 2006. Image Alignment and Stitching: A Tutorial. *Foundations and Trend in Computer Graphics and Vision*. 2(1):1–104. DOI: 10.1561/0600000009.

Szeliski, R. & Kang, S.B. 1995. Direct Methods for Visual Scene Reconstruction. In *Workshop on Representation of Visual Scenes, Cambridge, Massachusetts, USA*. 26–33. DOI: 10.1109/WVRS.1995.476849.

Tagoe, N.D., Rüther, H. & Smit, J. 2014. A Pragmatic Approach for Lens Distortion Correction from a Fictitious Image. In *Proceedings of AfricaGEO 2014*. Cape Town. 1–12. Available: http://www.africageoproceedings.org.za/.

Thibault, S. 2007. Enhanced surveillance system based on panomorph panoramic lenses. In *Defense and Security Symposium*. T.T. Saito, D. Lehrfeld, & M.J. DeWeert, Eds. International Society for Optics and Photonics. 65400E–65400E–8. DOI: 10.1117/12.719495.

Toet, A. & van Schaik, M.G. 2012. Effects of signals of disorder on fear of crime in real and virtual environments. *Journal of Environmental Psychology*. 32(3):260–276. DOI: 10.1016/j.jenvp.2012.04.001.

Torr, P. 2002. Bayesian Model Estimation for epipolar geometry and Generic Manifold Fitting. *International Journal of Computer Vision*. 50(1):35–61. DOI: doi:10.1023/A:1020224303087.

Torr, P.H.S. & Davidson, C. 2003. IMPSAC: Synthesis of importance sampling and random sample consensus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 25(3):354–364. DOI: 10.1109/TPAMI.2003.1182098.

Triggs, B., McLauchlan, P.F., Hartley, R.I. & Fitzgibbon, A.W. 1999. Bundle adjustment—a modern synthesis. In *International workshop on vision algorithms*. 298–372.

Ullrich, A., Schwarz, R. & Kager, A. 2003. Using Hybrid Multi-Station Adjustment for an Integrated Camera Laser-Scanner System. In *Proceedings of Optical 3D Measurement Techniques*. na. 298–304.

Vedaldi, A. & Fulkerson, B. 2010. VLFeat: An open and portable library of computer vision algorithms. In *Proceedings of the international conference on Multimedia*. 1469–1472.

Voltolini, F., Remondino, F., Pontin, M. & Gonzo, L. 2006. Experiences and Considerations in Image-Based Modeling of Complex Architectures. *International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences*. XXXVI(5):309–314.

Vosselman, G. & Maas, H.G. 2010. *Airborne and Terrestrial Laser Scanning*. Whittles.

Wessels, S., Ruther, H., Bhurtha, R. & Schroeder, R. 2014. Design and creation of a 3D virtual tour of the world heritage site of Petra, Jordan. In *Proceedings of AfricaGEO 2014*. 1–12.

Wester-Ebbinghaus, W. 1980. Aerial photography by radio controlled model helicopter. *The Photogrammetric Record*. 10(55):85–92.

Wong, W.K., Liew, J.T.Y., Loo, C.K. & Society, I.C. 2009. Omnidirectional Surveillance System for Digital Home Security. In *Proceedings of the 2009 International Conference on Signal Acquisition and Processing*. IEEE. 8–12. DOI: 10.1109/icsap.2009.13.

Zeng, Z. & Wang, X. 1992. A General Solution of a Closed-Form Space Resection. *Photogrammetric Engineering and Remote Sensing*. 58(3):327–338.

Zoghlami, I., Faugeras, O. & Deriche, R. 1997. Using geometric corners to build a 2D mosaic from a set of images. In *Computer Vision and Pattern Recognition*. 420–425. DOI: 10.1109/CVPR.1997.609359.

APPENDIX

APPENDIX 1: 3D OBJECT COORDINATES OF TARGETS ON THE CALIBRATION TESTFIELD

ID	X	Y	Z
1	3.3677	0.8279	1.998
2	3.632	0.8368	2.0007
3	3.8725	0.841	1.9914
4	3.3715	0.8273	1.7597
5	3.6313	0.8359	1.7532
6	3.8763	0.8409	1.7542
7	3.3678	0.8296	1.4642
8	3.6314	0.8374	1.5
9	3.8641	0.8422	1.4915
101	0.8842	0.7701	3.0162
102	1.384	0.7849	3.0204
103	1.8682	0.7934	2.9928
104	2.3726	0.1622	2.9738
105	2.9259	0.177	2.987
106	3.3692	0.1914	2.9833
107	3.6465	0.1961	2.9917
108	3.8941	0.196	2.9856
109	4.3975	0.2112	2.9865
110	4.8971	0.2263	2.9914
111	5.4029	0.2401	2.9937
112	5.9295	0.2632	2.9966
113	6.3994	0.2606	2.9992
201	0.8885	0.7717	2.5148
202	1.3884	0.787	2.5194
203	1.8646	0.7966	2.4971
204	2.3672	0.8067	2.4915
205	2.8763	0.8174	2.5038
206	3.3617	0.8257	2.4924
207	3.6264	0.8338	2.4992
208	3.8848	0.8404	2.4887
209	4.379	0.8534	2.5004
210	4.8732	0.866	2.5071
211	5.3749	0.8778	2.4996
212	5.8771	0.8906	2.4976
213	6.3734	0.9033	2.5144
301	0.8843	0.7722	2.0135
302	1.3815	0.7887	2.0147
303	1.8616	0.7686	1.9986
304	2.3644	0.7778	1.9956
305	2.8793	0.8191	2.0033
309	4.3721	0.8543	2.0028
310	4.8694	0.867	2.0081
311	5.3797	0.8781	1.9922
312	5.8753	0.8912	1.9968
313	6.373	0.9042	2.014
401	0.8931	0.772	1.7635
402	1.3845	0.7881	1.7553
403	1.8742	0.7687	1.7545
404	2.3732	0.777	1.7581
405	2.8898	0.8186	1.7595
409	4.3766	0.8531	1.7608
410	4.8799	0.8664	1.7536
411	5.3761	0.8773	1.7545
412	5.8721	0.8909	1.7658
413	6.3679	0.9033	1.7627
501	0.8757	0.7728	1.5125
502	1.3729	0.79	1.514
503	1.8618	0.7694	1.5023
504	2.3567	0.7792	1.4908
505	2.881	0.8203	1.4704
509	4.363	0.8539	1.5008
510	4.8705	0.8674	1.5087
511	5.3717	0.8781	1.4986
512	5.872	0.8925	1.4991
513	6.366	0.9048	1.5114
601	0.8794	0.7739	1.0155
602	1.3678	0.7918	1.0129
603	1.8671	0.7714	1.0025

ID	X	Y	Z
604	2.3601	0.7804	0.9865
605	2.8758	0.8213	0.9735
606	3.3609	0.8319	0.9634
607	3.6283	0.8384	0.9949
608	3.8656	0.8441	0.993
609	4.3662	0.8554	1.0022
610	4.8645	0.8684	1.0149
611	5.365	0.8801	0.9973
612	5.8642	0.893	0.9982
613	6.3649	0.9058	1.0152
701	0.8906	0.7775	0.5119
702	1.3635	0.793	0.5117
703	1.8602	0.7726	0.5106
704	2.3743	0.7824	0.4864
705	2.8693	0.8231	0.476
706	3.3615	0.8339	0.4661
707	3.6092	0.8407	0.4939
708	3.8599	0.8465	0.5043
709	4.3732	0.8572	0.4975
710	4.8798	0.8707	0.5032
711	5.3783	0.8823	0.5063
712	5.8798	0.8956	0.5076
713	6.3754	0.9074	0.5105

APPENDIX 2: PARTIAL DERIVATIVES OF THE LINEARISED COLLINEARITY EQUATION

This section provides the partial derivatives of the unknown parameters in the collinearity equation for spherical panoramic images developed in Section 4.4. For the sake of convenience, the simplified collinearity equations (Equation 4-17 and 4-18) are repeated as follows:

$$F_u = r \, arctan\left(\frac{X^*}{Y^*}\right) = u \qquad \text{Equation A1}$$

$$F_v = r \, arccos\left(\frac{Z^*}{d}\right) = v \qquad \text{Equation A2}$$

where,

$$X^* = m_{11}(X - X_o) + \ m_{21}(Y - Y_o) + \ m_{31}(Z - Z_o)$$

$$Y^* = m_{12}(X - X_o) + \ m_{22}(Y - Y_o) + m_{32}(Z - Z_o)$$

$$Z^* = m_{13}(X - X_o) + \ m_{23}(Y - Y_o) + m_{33}(Z - Z_o)$$

$$d = \sqrt{(X - X_o)^2 + (Y - Y_o)^2 + (Z - Z_o)^2}$$

The partial derivatives of the collinearity equation for the ***u*** coordinate of a point (Equation A-10) with respect to the unknown parameters are as follows:

$$\frac{\partial F_u}{\partial \omega} = r\left[\frac{\big(m_{31}(X - X_o) + m_{32}(Y - Y_o) + m_{33}(Z - Z_o)\big)(X^*)}{{X^*}^2 + \ {Y^*}^2}\right]$$

$$\frac{\partial F_u}{\partial \phi} = r\left[\frac{\begin{gathered}\big(\beta_{11}(X - X_o) + \beta_{12}\ (Y - Y_o) + \beta_{13}\ (Z - Z_o)\big)(Y^*) - \\ \big(\beta_{21}(X - X_o) + \beta_{22}\ (Y - Y_o) + \beta_{23}\ (Z - Z_o)\big)(X^*)\end{gathered}}{{X^*}^2 + \ {Y^*}^2}\right]$$

$$\frac{\partial F_u}{\partial \kappa} = r\left[\frac{(m_{12}(X - X_o) - m_{11}\ (Y - Y_o))(Y^*) - \big(m_{22}(X - X_o) + m_{21}(Y - Y_o)\big)(X^*)}{{X^*}^2 + \ {Y^*}^2}\right]$$

$$\frac{\partial F_u}{\partial X} = r\left[\frac{m_{11}(Y^*) - m_{21}(X^*)}{X^{*2} + Y^{*2}}\right]$$

$$\frac{\partial F_u}{\partial Y} = r\left[\frac{m_{12}(Y^*) - m_{22}(X^*)}{X^{*2} + Y^{*2}}\right]$$

$$\frac{\partial F_u}{\partial Z} = r\left[\frac{m_{13}(Y^*) - m_{23}(X^*)}{X^{*2} + Y^{*2}}\right]$$

The partial derivatives of the collinearity equation for the $\boldsymbol{v}$ coordinate of a point (Equation A-2) with respect to the unknown parameters are as follows:

$$\frac{\partial F_v}{\partial \omega} = -\left[\frac{m_{21}(X - X_o) + m_{22}(Y - Y_o) + m_{23}(Z - Z_o)}{d\sqrt{1 - \frac{Z^{*2}}{d^2}}}\right]$$

$$\frac{\partial F_u}{\partial \phi} = -r\left[\frac{\beta_{31}(X - X_o) + \beta_{32}(Y - Y_o) + \beta_{33}(Z - Z_o)}{d\sqrt{1 - \frac{Z^{*2}}{d^2}}}\right]$$

$$\frac{\partial F_v}{\partial \kappa} = -r\left[\frac{m_{32}(X - X_o) + m_{31}(Y - Y_o)}{d\sqrt{1 - \frac{Z^{*2}}{d^2}}}\right]$$

$$\frac{\partial F_v}{\partial X} = -r\left[\frac{d^2\, m_{31} + (X - X_o)(Z^*)}{d^3\sqrt{1 - \frac{Z^{*2}}{d^2}}}\right]$$

$$\frac{\partial F_v}{\partial Y} = -r\left[\frac{d^2\, m_{32} + (Y - Y_o)(Z^*)}{d^3\sqrt{1 - \frac{Z^{*2}}{d^2}}}\right]$$

$$\frac{\partial F_v}{\partial Z} = r\left[\frac{d^2\, m_{33} + (Z - Z_o)(Z^*)}{d^3\sqrt{1 - \frac{Z^{*2}}{d^2}}}\right]$$

where

$$b_{11} = -\sin\phi\cos\kappa$$

$$b_{12} = \sin\phi\sin\kappa$$

$$b_{13} = \cos\phi$$

$$b_{21} = \sin\omega\cos\phi\cos\kappa$$

$$b_{22} = -\sin\omega\cos\phi\sin\kappa$$

$$b_{23} = \sin\omega\sin\phi$$

$$b_{31} = -\cos\omega\cos\phi\cos\kappa$$

$$b_{32} = \cos\omega\cos\phi\sin\kappa$$

$$b_{33} = -\cos\omega\sin\phi$$