

MARIANA BATISTA CAMPOS

**A PORTABLE MOBILE TERRESTRIAL SYSTEM WITH
OMNIDIRECTIONAL CAMERA FOR CLOSE RANGE APPLICATIONS**



PRESIDENTE PRUDENTE
2019



UNIVERSIDADE ESTADUAL PAULISTA
CAMPUS DE PRESIDENTE PRUDENTE
FACULDADE DE CIÊNCIAS E TECNOLOGIA
Programa de Pós-Graduação em Ciências Cartográficas

MARIANA BATISTA CAMPOS

**A PORTABLE MOBILE TERRESTRIAL SYSTEM WITH
OMNIDIRECTIONAL CAMERA FOR CLOSE RANGE APPLICATIONS**

Thesis presented to “*Programa de Pós
Graduação em Ciências Cartográficas*” at
São Paulo State University – UNESP -
School of Technology and Sciences.
Supervisor: Dr. Antonio M. G. Tommaselli
Co-Supervisor: Dr.^a Eija Honkavaara

PRESIDENTE PRUDENTE
2019

C198p

Campos, Mariana Batista

A portable mobile terrestrial system with omnidirectional camera for close range applications / Mariana Batista Campos. -- Presidente Prudente, 2019
146 p.

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências e Tecnologia, Presidente Prudente

Orientador: Antonio Maria Garcia Tommaselli

Coorientadora: Eija Honkavaara

1. Personal mobile terrestrial system. 2. Mobile mapping. 3. Polydioptric system calibration. 4. Omnidirectional image matching. 5. Incremental bundle adjustment. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências e Tecnologia, Presidente Prudente. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

CERTIFICADO DE APROVAÇÃO

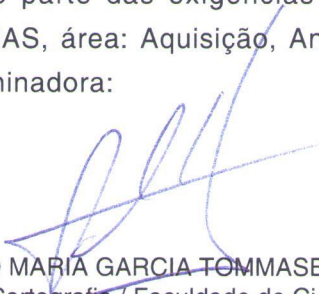
TÍTULO DA TESE: A PORTABLE MOBILE TERRESTRIAL SYSTEM WITH OMNIDIRECTIONAL CAMERA
FOR CLOSE RANGE APPLICATIONS

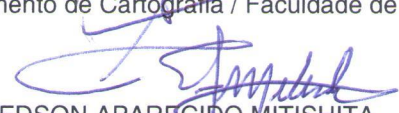
AUTORA: MARIANA BATISTA CAMPOS

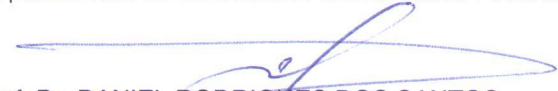
ORIENTADOR: ANTONIO MARIA GARCIA TOMMASELLI

COORIENTADORA: EIJA HONKAVAARA

Aprovada como parte das exigências para obtenção do Título de Doutora em CIÊNCIAS
CARTOGRÁFICAS, área: Aquisição, Análise e Representação de Informações Espaciais pela
Comissão Examinadora:


Prof. Dr. ANTONIO MARIA GARCIA TOMMASELLI
Departamento de Cartografia / Faculdade de Ciencias e Tecnologia de Presidente Prudente


Prof. Dr. EDSON APARECIDO MITISHITA
Departamento de Geomática / Universidade Federal do Paraná


Prof. Dr. DANIEL RODRIGUES DOS SANTOS
Departamento de Geomática / Universidade Federal do Paraná


Prof. Dr. MAURICIO GALO
Departamento de Cartografia / Faculdade de Ciencias e Tecnologia de Presidente Prudente

Prof. Dr. ANTERO KUKKO
FGI / Finnish Geospatial Research Institute

VIDEOCONFERÊNCIA



Presidente Prudente, 13 de março de 2019

ACKNOWLEDGMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. Standard PhD scholarship from 03/2015 to 03/2017 and from 08/2017 to 02/2019 (Grant: 1481339). Internship abroad at Finnish Geospatial Research Institute (*Programa de Doutorado Sanduiche no Exterior* - PDSE) from 04/2017 to 07/2017 (Grant: 88881.135114/2016-01);

This work was also supported by:

Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP—Grant: 2013/50426-4) and Academy of Finland (Grant: 273806), which jointly funded the international project Unmanned Airborne Vehicle - based 4D Remote Sensing for Mapping Rain Forest Biodiversity and Its Change in Brazil (UAV_4D_Bio) conducted at São Paulo State University (UNESP) and at the Finnish Geospatial Research Institute in the National Land Survey of Finland (FGI);

Postgraduate program in Cartographic Sciences (Photogrammetry group), São Paulo State University (UNESP) - School of Technology and Sciences;

Finnish Geospatial Research Institute (FGI) in the National Land Survey of Finland.

ABSTRACT

This research proposes a new technique for close-range mobile data acquisition and processing, consisting of a backpacked light-weight low-cost system. This system integrates an omnidirectional camera and a GPS/IMU system (Global Positioning Systems/Inertial Measurement Unit System) with a tailored photogrammetric processing chain to obtain sensor location and 3D points coordinates using the fisheye images. Omnidirectional systems, based on multiple cameras covering a full-spherical field of view, have been used in close range photogrammetry applications. The use of omnidirectional systems is especially motivated by their 360° coverage around the sensor, which allows more features to be tracked in a single image shot, and by the light weight and low cost of some off-the-shelf omnidirectional cameras. This kind of systems have been named as Personal Mobile Terrestrial System (PMTS). There are only few studies focusing on PMTS using omnidirectional systems. This research assessed the performance of an omnidirectional PMTS based exclusively on low-cost technologies to indirectly estimate forest and outdoor urban features. An accuracy evaluation of GPS and IMU sensors and the development of rigorous photogrammetric processing considering fisheye geometry were performed. PMTS data, i.e. fisheye images and navigation data, are input information for the photogrammetric process. The proposed photogrammetric process focused on omnidirectional camera modelling, feature-based matching and bundle adjustment considering fisheye geometry. Experimental assessments showed that the integrated sensor orientation approach using navigation data as the initial information and a rigorous photogrammetric process can increase the trajectory accuracy, especially in obstructed areas, such as dense tropical forests and some urban environments. Overall, using a postprocessing approach with PMTS data and ground control points, the EOPs (Exterior Orientation Parameters) were estimated with standard deviations of 0.1° for sensor attitude and centimetric accuracy for sensor position. The point cloud generated had an accuracy consistent with the range of the pixel size in the object space units in the central part of omnidirectional images (3.5–8 cm). A PMTS real-time application was simulated in laboratory. The PMTS trajectory was estimated with a planialtimetric accuracy of 0.7 m, while, the 3D map was simultaneously computed with an average accuracy ranging between 0.5 m and 2 m.

Keywords: Personal mobile terrestrial system, Mobile mapping, Polydioptric system calibration, Omnidirectional image matching, Incremental bundle adjustment.

CONTENTS

CHAPTER I INTRODUCTION

1. OVERVIEW AND MOTIVATION	9
1.1 OBJECTIVES	10
1.2 CONTENT AND CONTRIBUTIONS	11
1.3 INTERNATIONAL COOPERATION.....	12

CHAPTER II A PERSONAL MOBILE TERRESTRIAL SYSTEM: DEVELOPMENT AND APPLICATIONS

2. SCOPE	13
2.1 BACKPACK OMNIDIRECTIONAL SYSTEMS: RELATED WORKS.....	13
2.2 PMTS SETUP AND DATA INTEGRATION	16
2.2.1 Technical information of the Omnidirectional system.....	17
2.2.2 Navigation system	20
2.2.3. PMTS data integration	22
2.3. PMTS ASSESSMENTS	25
2.3.1 Test areas.....	25
2.3.2 Assessment of the positional and attitude trajectory accuracies	26
2.3.2.1 Positional Accuracy Experiments	26
2.3.2.2 Attitude Accuracy Experiments	28
2.3.3 Assessment of object reconstruction	28
2.3.3.1. General data set and processing	31
2.3.3.2 PMTS point cloud accuracy assessment	32
2.4. PMTS RESULTS	33
2.4.1 Positional and attitude trajectory accuracies	34
2.4.2 Accuracy of object reconstruction.....	38
2.4.2.1 Results - Experiment II.A.....	38
2.4.2.2 Results - Experiment II.B.....	39
2.4.2.3 Results - Experiment II.C.....	41
2.5. DISCUSSION	42
2.6. CONCLUSIONS	45

CHAPTER III GEOMETRIC MODEL OF A DUAL-FISHEYE 360° IMAGING SYSTEM

3. SCOPE	47
3.1POLYDIOPTRIC SYSTEM CALIBRATION: OVERVIEW AND CHALLENGES	47
3.2 DUAL 360° IMAGING SYSTEM.....	49
3.2.1 Ricoh Theta S Data – Still image	49
3.2.2 Ricoh Theta S Data – Dual-fisheye video	52
3.3 360° CAMERA CALIBRATION FIELD	55

3.4 EXPERIMENTAL ASSESSMENT FOR RICOH THETA S.....	56
3.4.1 Data set.....	56
3.4.2 Mathematical Model	57
3.4.2.1 Experiments A and B - Single and Simultaneous camera calibration without relative orientation constraints	59
3.4.2.2 Experiment C - Simultaneous camera calibration with PCs' distance and relative rotation matrix as stability constraint.	61
3.4.2.3 Experiment D - Simultaneous camera calibration with base elements and relative rotation matrix as stability constraint.	62
3.5 RESULTS AND DISCUSSION.....	63
3.5.1 Preliminary analysis	63
3.5.2 Calibration with Bundle Adjustment.....	64
3.5.3 Dual camera system relative orientation	67
3.5.4 Accuracy assessment.....	68
3.6 CONCLUSION.....	70

CHAPTER IV

A RECURSIVE FEATURE SEARCH METHOD FOR FISHEYE IMAGE MATCHING

4. SCOPE	72
4.1 OMNIDIRECTIONAL IMAGE MATCHING APPROACHES.....	72
4.2 SIFT_{RFS}: A PROPOSAL TO IMPROVE FBM IN FISHEYE IMAGES	74
4.2.1 Epipolar geometry on the sphere domain.....	76
4.2.2 Search space window and bundle adjustment	79
4.2.3 Image Matching	81
4.3 MATERIALS AND EXPERIMENTS.....	82
4.3.1 Data sets	82
4.3.2 Preliminary assessment: Why SIFT?	84
4.3.3 Experiment I - A comparison of omnidirectional matching approaches based on SIFT	85
4.3.4 Experiment II – Sensor Position and attitude estimation	86
4.3.5 Performance Assessments.....	87
4.4 RESULTS AND DISCUSSION.....	87
4.4.1 Comparative analysis	88
4.4.1.1 Detection rate and location correctness.....	88
4.4.1.2 Repeatability	90
4.4.1.3 Geometric distribution.....	91
4.4.2 Bundle adjustment results	93
4.5 CONCLUSIONS	94

CHAPTER V

AN INCREMENTAL BUNDLE ADJUSTMENT BASED ON WIDE ANGLE IMAGE GEOMETRY

5. OVERVIEW	96
5.1 OPEN PROBLEMS ON REAL-TIME CLOSE RANGE PHOTOGRAMMETRY WITH OMNIDIRECTIONAL SENSORS.....	98

5.1.1 Omnidirectional match network and bundle adjustment geometry	98
5.1.2 Real-time strategies	101
5.2 MATHEMATICAL MODEL AND SENSOR STOCHASTIC PROPERTIES	103
5.2.1 Functional model.....	103
5.2.2 Stochastic treatment	105
5.2.2.1 Image coordinates in the image space	106
5.2.2.2 3D coordinates in the object space	110
5.2.2.3 Initial Exterior Orientation Parameters	111
5.3 INCREMENTAL BUNDLE ADJUSTMENT METHODOLOGY	111
5.3.1 PMTS sequential data acquisition	112
5.3.2 Image Matching	113
5.3.3 Initial map (approximated values for the ground coordinates of the tie points).....	114
5.3.4 Incremental bundle adjustment estimation.....	116
5.4 EXPERIMENTAL ASSESSMENTS.....	117
5.4.1 Dataset	117
5.4.2 Photogrammetric process setup.....	119
5.4.3 Quality Control: EOPs and 3D Map accuracies	120
5.5 RESULTS.....	121
5.5.1 Performance assessment of PMTS initial values	121
5.5.2 Incremental bundle adjustment results	122
5.5.3 PMTS trajectory and the 3D ground coordinates accuracy.....	126
5.5 DISCUSSION AND CONCLUSIONS	128

CHAPTER VI FINAL REMARKS

FINAL REMARKS.....	130
REFERENCES.....	133
APPENDIX A	141

CHAPTER I

INTRODUCTION

1. OVERVIEW AND MOTIVATION

Mobile Mapping Technologies have begun to be developed and earn attention in the last few years. They can provide fast, cost-effective and complete data collection, considering the dynamic and continuous displacement of the sensors. Mobile Terrestrial Systems (MTS) are composed of navigation and remote sensors, such as cameras and LASER scanners, integrated in a common moving platform (SCHUWARZ and EL-SHEIMY, 2004). Different types of platforms can be used in terrestrial mapping, such as cars (KUKKO et al., 2012), small Unmanned Aerial Vehicles (SHEN et al., 2012, ROCA et al., 2013), boats (MORAES et al., 2016), ground robots (KLINGENSMITH et al., 2016), and a person as operator (KUKKO et al., 2012; CORSO and ZAKHOR, 2013; LIANG et al., 2014; WEN et al., 2016; FORSMAN et al., 2016 and OVELAND et al., 2017). Therefore, the motivation of this Ph.D. research is driven on how to acquire indirect data with a ground perspective in difficult access places, such as forest and urban environments, considering an accurate, economic and fast methodology.

Personal Mobile Terrestrial System (PMTS) consists of a set of sensors that are carried by a human operator, for instance in a backpack, to acquire measurements of the environment while walking. The use of PMTS is increasing in mobile mapping applications, because they are offering dynamic data acquisition on ground perspective in places where the use of wheeled platforms is unfeasible, such as forest and some city mapping application (sidewalks and indoor buildings). PMTS have become more popular with emerging technologies, for instance, miniaturized navigation sensors and off-the-shelf omnidirectional cameras, enabling flexible and low-cost mobile mapping approaches.

Emerging sensor technologies are changing the way of mapping the environment. Sensors with nonconventional internal geometry, i.e., omnidirectional, have been increasingly used in Close Range Photogrammetry (CRP) applications. However, most of these sensors have not been developed for high accuracy metric purposes, and therefore rigorous methods are required in data acquisition and processing chain to obtain satisfactory results. Thus, in order to contribute to the development of a lite PMTS and potential applications of these off-the-shelf sensors, this research presents a low-cost PMTS approach composed by an

omnidirectional camera with two fisheye lenses and off-the-shelf inertial navigation system, specially arranged in a backpack, aiming close-range applications. Our study was the first to assess the performance of a low-cost backpack-mounted system with omnidirectional camera in the estimation of forest and urban parameters, focusing on the accuracy evaluation of low-cost GPS (Global Positioning Systems) receiver and MEMS-IMU (Microelectromechanical systems - Inertial Measurement Unit) sensors and on the development of photogrammetric processing methodologies using an omnidirectional system.

Omnidirectional systems, based on multiple cameras that cover a full-spherical field of view, are an attractive alternative due to the imaging of larger areas, which allows more features to be tracked in a single image, and the low-weight and low-cost of these cameras that have recently entered the market (e.g. Giroptic 360cam, Go Pro 360°, Kodak PixPro SP360, Google Jump system, Ricoh Theta S, Samsung Gear 360 and LG 360). Therefore, the proposed PMTS will enable a dynamic image acquisition of a set of features, simultaneously, in different perspectives, for a fast and systematic terrestrial surveying. However, the usability of omnidirectional sensors requires adaptations in the classic photogrammetric process, considering its sensor geometry particularities. The challenges of fisheye images include nonperspective inner geometry, huge scale and illumination variations between scenes, large radial distortion and nonuniformity of spatial resolution. In this regard, this research aims to extend the well-known photogrammetric process from the perspective to omnidirectional images improving PMTS results, which involves omnidirectional system modelling, fisheye image matching and a sequential bundle adjustment using fisheye images. PMTS developments and data processing are the main subjects assessed and discussed in the following sections.

1.1 OBJECTIVES

The main goal of this research is to develop a methodology for close-range data acquisition with a portable mobile terrestrial system and a photogrammetric processing chain to obtain 3D points coordinates with fisheye images. In this regard, the Ph.D. research has the following specific objectives:

- (1) To evaluate the feasibility of a low-cost PMTS with omnidirectional system for forest and city mapping;

- (2) To model the Ricoh Theta S omnidirectional system inner geometry, studying a suitable method for the calibration of a Poly-dioptric spherical imaging systems, based on fisheye lenses models;
- (3) To investigate the performance of state-of-art interest operators for omnidirectional image, identifying the main limitations and proposing improvements on Feature-Based Matching (FBM) techniques for original fisheye images;
- (4) To study mathematical models and stochastic treatment of image observations obtained with omnidirectional cameras based on fisheye lenses to be applied in an incremental bundle adjustment with the PMTS data, aiming the estimation of Exterior Orientation Parameter (EOP) and a 3D sparse point cloud in real-time applications.

1.2 CONTENT AND CONTRIBUTIONS

Thesis content can be summarized in two main purposes: PMTS development and feasibility study (Chapter 2) and photogrammetric processes (Chapter 3 to Chapter 5) considering fisheye image geometry (omnidirectional system calibration, fisheye image matching and incremental bundle adjustment). Chapter 2 summarizes the studies developed by Campos et al. (2017), Campos et al. (2018a) and Castanheiro et al. (2018). A new PMTS approach with a low-cost backpack-mounted omnidirectional camera system is introduced. A technical description of the proposed PMTS, which includes sensors setup, integration and synchronization, is presented. Then, a PMTS feasibility study for forest and outdoor city mapping applications is performed. The results obtained are compared with other PMTS approaches.

Chapter 3 is based on the work presented by Campos et al. (2018b), and, it presents a process for the calibration of a poly-dioptric spherical imaging systems (for instance, Ricoh Theta S), in order to enable their metric use in close-range photogrammetry applications. The calibration methodology proposed is based on fisheye lenses models. Experiments were performed to verify the set of interior and relative orientation parameters that better reproduces the bundle that generated the 360° image of Ricoh Theta S. The main contribution presented in this chapter is a simultaneous bundle adjustment approach with stability constraints on base elements

and rotation matrix angles, which achieved accuracies comparable with conventional photogrammetric calibration with perspective cameras.

Chapter 4 presents an approach to improve FBM techniques on fisheye images with a recursive reduction of the search space based on epipolar geometry. The epipolar restriction is computed with the equidistant mathematical model and initial EOPs determined from the PMTS navigation sensors. The interest points are detected with the SIFT (Scale-Invariant Feature Transform) operator and the candidate matches are searched in a reduced search space. Experimental assessments were performed with data sets collected by the proposed PMTS, aiming outdoor applications. This chapter extends the study presented by Castanheiro et al. (2017).

Chapter 5 proposes an incremental bundle adjustment to estimate omnidirectional sensor position and attitude (sensor pose) and the 3D tie points coordinates (map points) simultaneously during PMTS acquisition. A methodology for PMTS data acquisition and processing based on Chapters 2 to 4 are applied. A mathematical model consistent with the omnidirectional sensor geometry and a stochastics modeling of the PMTS data are assessed. The summary of thesis results and conclusions are presented in Chapter 6.

1.3 INTERNATIONAL COOPERATION

This Ph.D. research is linked to an international project jointly conducted by researchers from São Paulo State University (UNESP) and the Finnish Geospatial Research Institute in the National Land Survey of Finland (FGI) funded by FAPESP (Fundação de Amparo a Pesquisa do Estado de São Paulo - Grant: 2013/50426-4) and AKA (Academy of Finland - Grant: 273806). The project “Unmanned Airborne Vehicle - based 4D Remote Sensing for Mapping Rain Forest Biodiversity and Its Change in Brazil (UAV_4D_Bio)” aims to develop methodologies and technologies for biodiversity mapping and for indicators of environmental change in Evo test forest in Finland and in regeneration areas of tropical forest in the interior of São Paulo State.

CHAPTER II

A PERSONAL MOBILE TERRESTRIAL SYSTEM: DEVELOPMENT AND APPLICATIONS

2. SCOPE

Chapter 2 introduces a technical description of the proposed Personal Mobile Terrestrial System (PMTS) and a PMTS feasibility study for forest and outdoor city mapping applications. The results and discussions of this chapter are based on the works developed by Campos et al. (2017), Campos et al. (2018a) and Castanheiro et al. (2018).

2.1 BACKPACK OMNIDIRECTIONAL SYSTEMS: RELATED WORKS

Indoor and outdoor city mapping aimed at developing a Building Information Model (BIM) was one of the main motivations for the development of embedded systems in a backpack, commonly named as PMTS. Studies involving PMTS for these applications have increased in recent years.

Corso and Zakhor (2013) used an arrangement of five LASER scanners in a backpack platform combined with two fisheye cameras and an IMU for indoor mapping purposes. The authors achieved centimetric accuracy in three-dimensional (3D) reconstruction. Another backpack platform using a mobile LASER scanning system for indoor and outdoor urban environments was evaluated by Lauterbach et al. (2015). These authors developed a GNSS (Global Navigation Satellite Systems)-free approach based on LASER scanners (SICK LMS 100 and Riegl VZ-400) and a low-cost IMU (Phidgets 1044), with the aim of a simultaneous location and mapping solution (SLAM). A GNSS-free approach using an Extended Kalman Filter (EKF)-based method for indoor environments was also discussed by Wen et al. (2016). The system proposed in Wen et al. (2016) has three 2-D LASER scanners (UTM-30LX) and an IMU (Xsens MTi-10) specially arranged in a backpack. Indoor data obtained with this backpacked system was integrated to outdoor Terrestrial LASER Scanner (TLS) data in Wen et al. (2018), obtaining a 3D point cloud model with centimetric accuracy. A portable system with multi digital cameras was also presented by Tseng et al. (2016) aiming outdoor city mapping applications, in which eight single-lens cameras (Sony NEX-5N) and a GPS receiver were used obtaining an accuracy higher than 0.5 m.

Another potential application of PMTS is forest mapping. Forest preservation and management requires information about trees and underlying vegetation, such as location, height, volume, stem straightness and Diameter at Breast Height (DBH), to ensure proper management of the forest system (WEST, 2009). This type of information is used to understand how forests grow and develop, which requires detailed close-range measurements. Kukko et al. (2012) introduced a backpack personal system composed of a Mobile LASER Scanning system (MLS - FARO Photon 120) and a NovAtel SPAN, which integrates a GPS receiver and an IMU, for forest monitoring in Finland. The authors obtained a 3D point cloud with accuracy in the range of 2-5 cm. However, this PMTS had some disadvantages, such as the need for two operators and the high system weight (approximately 30 kg), which limited its operability. These limitations were reduced by Liang et al. (2014), with the development of a 10-kg PMTS with a GNSS receiver (NovAtel Flexpak6), an IMU (NovAtel UIMU-LCI) and a FARO Focus3D 120 LASER scanner. This system was applied for forest purposes and estimated DBH with centimetric accuracy. A backpack approach with multi digital cameras was presented by Forsman et al. (2016). In this work, a multi-camera system in a backpack structure was developed to estimate tree attributes in low-density forest plots based on terrestrial photogrammetric techniques. This system comprised five synchronized cameras: three Canon EOS 7D with Sigma fixed lenses and two Canon EOS 40D with Canon zoom lenses, weighing a total of 13 kg. The cameras were specially positioned to improve redundancy and to reduce the risk of false matching and occlusions. The process of indirectly estimating DBH with this system contains some sequential steps: cameras and system calibration, image acquisition, image processing, matching using SIFT and RANSAC (RANDOM Sample Consensus) methods, and circle fitting to estimate tree diameter. DBH was obtained with a Root Mean Square Error (RMSE) of 3.9 cm. More accurate results for forest mapping were presented recently by Oveland et al. (2017). In this work, a Velodyne VLP16 LASER scanner and an Applanix APX-15 UAV system (GNSS and IMU data) were assembled in a backpack to capture forest data. After data processing, DBH was estimated with an RMSE of 1.5 cm.

It is worth mentioning that other optical sensors have been recently assessed for forest applications, such as RGB-D (TOMAŠTÍK et al., 2017 and HYYPPÄ et al., 2017). RGB-D sensors combine RGB data with depth (D) data acquired using structured light with an IR (InfraRed) sensor. For instance, Tomaščík et al. (2017) estimated DBH using GoogleTango with an average accuracy (RMSE) of 1.9 cm. Similar results were presented by Hyyppä et al. (2017),

estimating DBH of individual trees with a RMSE of 1.9 cm and 0.73 cm using Kinect and GoogleTango systems, respectively. However, RGB-D sensors have a limited useful range.

The works mentioned above show the relevance and increasing number of studies involving PMTS in backpack platforms in the last years. Furthermore, commercial solutions with a LASER scanner and omnidirectional systems were recently presented by Leica (LEICA GEOSYSTEMS, 2017) and Google (FISHER, 2017). The Leica Pegasus Backpack is equipped with a Dual Velodyne VLP-16 LASER scanner and five cameras for a 360° field of view, whereas the Google Street View backpack (Trekker) is composed of an omnidirectional system with 15 cameras, also enabling full 360° imaging. Therefore, some research trends can be highlighted: (1) the use of off-the-shelf navigation sensors (LAUTERBACH et al., 2015; WEN et al., 2016 and OVELAND et al., 2017), which can be combined in PMTS with (2) optical sensors (FISHER, 2017; FORSMAN et al., 2016 and TSENG et al., 2016) and (3) the employment of mobile LASER scanner solutions in the PMTS (KUKKO et al., 2012; CORSO and ZAKHOR, 2013; LIANG et al., 2014; WEN et al., 2016 and OVELAND et al., 2017), particularly motivated by recent light-weight MLS (Mobile LASER Scanning), such as Velodyne Puck LITE and Riegl miniVUX-1UAV. Furthermore, commercial low-cost MLS, such as YdLiDAR X4 360°, are potential possibility for PMTS systems, but they were not assessed yet, especially aiming forest applications. These low-cost MLS systems are limited by point cloud density and high beam divergence.

Considering these trends, this chapter aims to contribute to the following issues. First, despite the diversity of off-the-shelf navigation sensors that have entered the market and the increasing use of these sensors in MTS and personal navigation, uncertainties about the data accuracy of these sensors remain. In this regard, this chapter presents an overview, experiments and discussions of the data accuracy of the low-cost system used and highlights its potential and limitations. Second, although MLSs provide fast data acquisition, dense surface coverage, high relative accuracy and real-time point cloud visualization, some limitations remain. MLS is still a high-cost system and produces a huge amount of data, which could be a further problem for data processing and long-term storage. Furthermore, MLSs require an accurate navigation system, and real-time data production and visualization requires accurate Real Time Kinematic (RTK) positioning with base stations nearby. The use of omnidirectional systems (MAAS, 2008), especially with multi cameras, has also gained attention due to significant improvements in recent

years, including lighter, smaller, lower cost and more diverse commercial models (Giroptic 360cam, Go Pro 360°, Kodak PixPro SP360, Lady bug5, Google Jump system, Ricoh Theta S, Samsung Gear 360, LG 360 and Nikon KeyMission 360). Omnidirectional sensor solutions can outperform other sensors, such as LASER scanners, in terms of field of view, the amount of usable information from the environment for mapping, simple operating principles and intuitive data interpretation (VALIENTE et al., 2014). However, research focused on PMTS using optical systems, especially those based on omnidirectional cameras, are still underexplored. In this regard, this chapter presents the description and assessment of a PMTS with a backpacked omnidirectional system and off-the-shelf navigation sensors for forest and outdoor city mapping.

2.2 PMTS SETUP AND DATA INTEGRATION

Figure 1 shows the proposed PMTS. The system components cost approximately US\$ 450 and weight 3 kg (Appendix A). The backpack was equipped with a rigid structure in a “T” design, thereby the omnidirectional system was positioned in the top of this structure with the GPS receiver and antenna (Figure 1.b), whereas the IMU, Arduino microprocessor and SD (Secured Digital) card (Figure 1.c) were placed below to avoid image occlusions.

The PMTS can be divided into two sets of sensors: an optical imaging sensor (Ricoh Theta S camera) and navigation sensors (IMU and GPS). Section 2.2.1 presents the imaging sensor configuration, whereas Section 2.2.2 shows the navigation sensor settings. System integration and data synchronization is summarized in Section 2.2.3.

Figure 1. Personal Mobile Terrestrial System (PMTS): (a) Backpack platform structure and PMTS sensors; (b) Top view of platform with the omnidirectional camera and GPS receiver; (c) Navigation sensors; (d) Example of an operator carrying the PMTS inside a boreal forest site and (e) urban area.



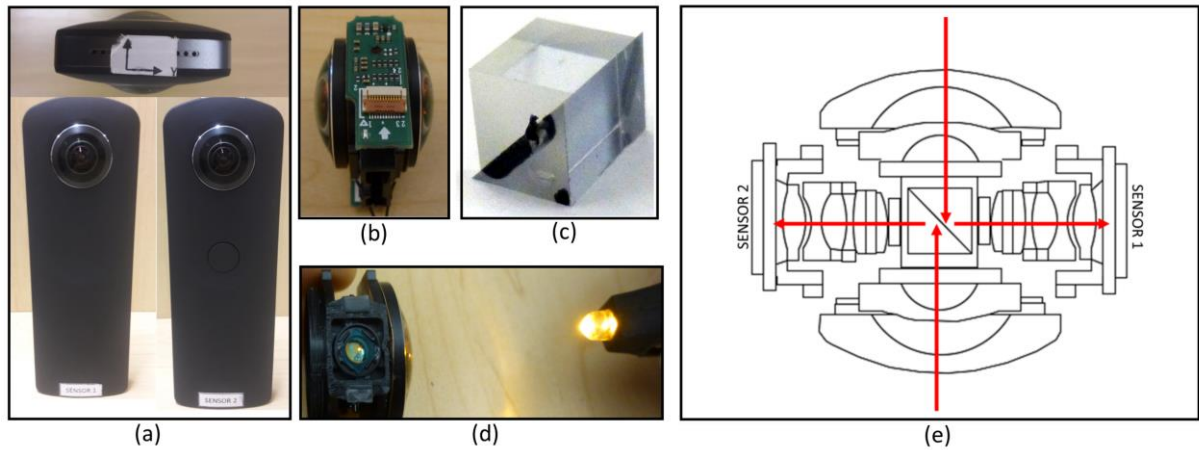
Source: Adapted from Campos et al. (2018a) and Castanheiro et al. (2018).

2.2.1 Technical information of the Omnidirectional system

Omnidirectional images can be obtained using panoramic cameras, fisheye lenses, poly-dioptric solutions and catadioptric systems (MAAS, 2008). The omnidirectional system used in the PMTS is a poly-dioptric camera named Ricoh Theta S. The Ricoh Theta S optical configuration is composed of two fisheye lenses, providing a 360° field of view (each with more than a 180° field of view, in a back-to-back position) and two complementary metal oxide semiconductor (CMOS) sensors (1/2.3 inches) with 7.2 megapixels on each side of the camera. The bundle of rays that generate the images are internally redirected to the sensors (which are rotated 90°) using a cube composed of two prisms, as can be seen in Figure 2.d. The optical system of the Ricoh Theta S is shown in Figure 2 (CAMPOS et al., 2018b and XIUMING, 2015): in (a) both sides of the Ricoh Theta S are displayed; in (b) one lateral side of the camera is presented, showing the Sensor 1 position; (c) shows the internal cube that enables the redirection of the bundle of rays that generates the image to the sensor, as illustrated in (d); and finally (e) depicts the set of lenses,

the internal cube and both sensors of the Ricoh Theta S. Table 1 shows the Ricoh Theta S technical specifications (THETA DEVELOPERS, 2017). More details about the omnidirectional system used and the camera calibration methodology proposed can be found in Chapter 3 (CAMPOS et al. 2018b).

Figure 2 . Ricoh Theta S optical system: (a) both sides of the camera; (b) Sensor 1 position on one side of the Ricoh Theta S; (c) internal cube; (d) redirection of the bundle of rays that generates the image to the sensor through the internal cube; (e) set of lenses, internal cube and both sensors of the Ricoh Theta S



Source: Campos et al, (2018b) and XiuMing, (2015)

Table 1. Ricoh THETA S digital camera and fisheye lens settings.

SETTINGS	CAMERA RICOH THETA S
Sensor Size	Two 1/2.3" CMOS Sensors (7.2 Mpx each)
Still Image/ Dual Fisheye Video	2688x2688 pixels in each sensor/ 960x1080 pixels in each sensor
Principal Distance	1.43 mm
Camera Dimensions	44 mmx130 mm x 22.9 mm – weight 125 g
Fittings	Remote control, RSWC201 wireless, HDMI and USB.

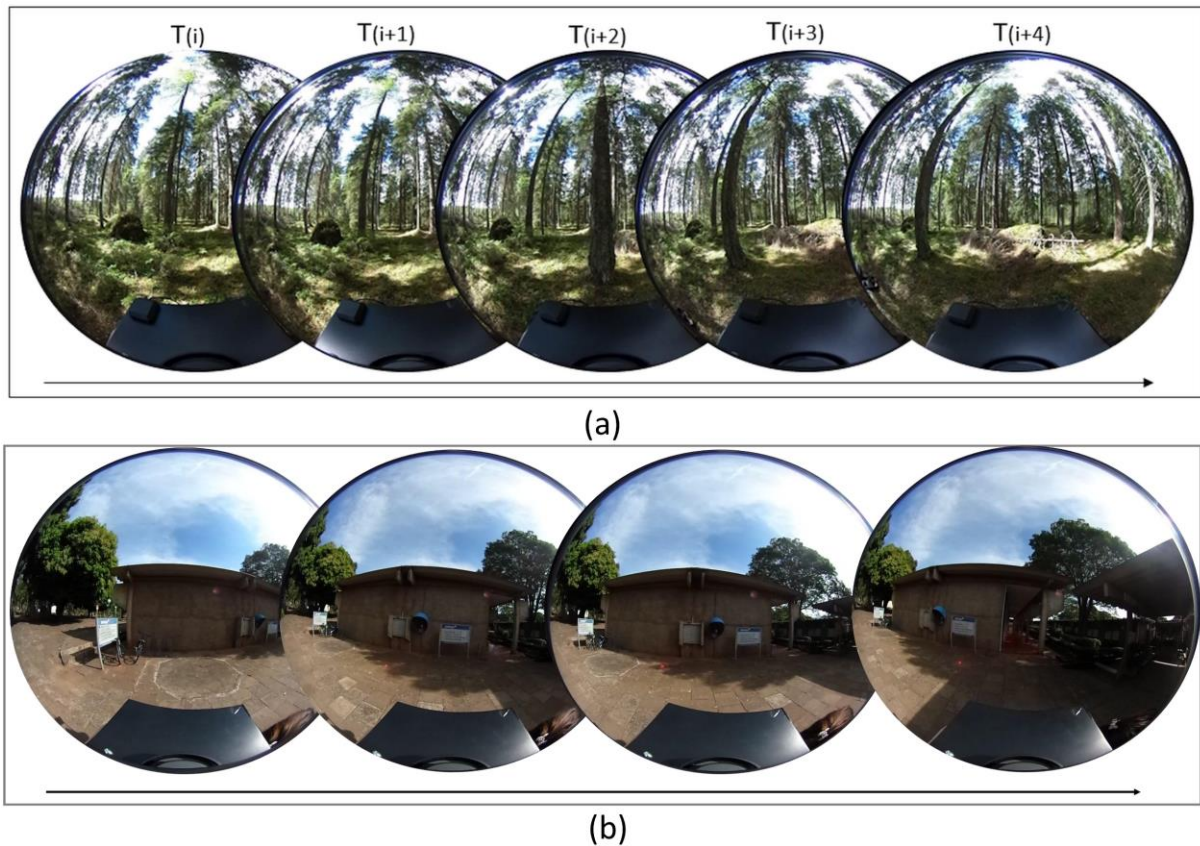
The Ricoh Theta S data can provide either still images (5376 x 2688 pixels) or dual-fisheye video (1920 x 1080 pixels). The still image mode requires an 8-second processing time between successive image shots. This time delay hinders an appropriate image base geometry considering the PMTS displacement during a survey. Furthermore, uncertainties in the accuracy of image stitching remain that are not easily modelled. Since the stitching parameters are unknown, further errors may accumulate in the photogrammetric processing. Considering this drawback, the dual-fisheye video mode is used in the PMTS to acquire omnidirectional data, providing a dense

and continuous frame acquisition. Ricoh Theta S limitations are discussed with more details in Chapter 3.

PMTS operation time is directly associated to the camera internal battery and storage. The Ricoh Theta S can record 65 minutes of dual-fisheye video in a 1920 x 1080 resolution mode, which is enough to collect data, for instance, from a standard forest plot with 32 m x 32 m dimension or a city mapping trajectory of 3 km, approximately. The main disadvantage of the dual-fisheye video, compared to the still images, is its low resolution. Ricoh Theta S has an Instantaneous Field Of View (IFOV) of 0.2° (3.5 mrad). Therefore, the 3D reconstruction of objects that are located farther than 5 m from the camera in the forest case (pixel size in object space units > 1.7 cm) and farther than 20 m in the urban case (pixel size in object space units > 7 cm) - at the optical axis direction – can present inaccurate results, especially due to the low level of detail of these objects. Improving the camera resolution will also augment this distance limit.

The dual fisheye frames extracted from the Ricoh Theta S video can be converted automatically (with a MATLAB Script) into two individual frames (one for each camera that composes the Ricoh Theta S) with dimensions of 960 x 1080 pixels and a pixel size of 0.005 mm. More details about this conversion is further presented in Chapter 3. These individual frames obtained from video mode will be called omnidirectional images in the chapters. Figure 3 shows examples of an omnidirectional image sequence (frame videos) used as input data in the subsequent photogrammetric processes for forest (a) and city mapping (b).

Figure 3. Example of sequential omnidirectional images acquired in a forest (a) and urban (b) areas using the Personal Mobile Terrestrial System.



Source: Adapted from Campos et al, (2018a)

2.2.2 Navigation system

The navigation system is composed of a GPS receiver (Ublox NEO-6M) and an IMU (MPU 6050) integrated with an Arduino microprocessor (Blackboard, produced by RoboCore), and the sensor data are recorded with an SD card adapter. The electronic schema of the navigation system is presented in (Appendix A), as well as the Arduino scripts.

The Ublox NEO-6M is a low-cost single frequency GPS receiver with a compact architecture and low power consumption, which is ideal for battery-operated systems. Like most low-cost GPS receivers, this device has an internal GPS active antenna designed for tracking L-band signals in the frequency of 1575 MHz. The processed data are delivered in the NMEA (National Marine Electronics Association) protocol, which provides the position, velocity and time (PVT) data computed by the receiver and detailed information on the satellites tracked. In addition, the NEO-6M supports Differential GPS (DGPS) operation using RTCM (Radio Technical

Commission for Maritime Services) correction messages from a local reference station or aiding network. Alternatively, raw GPS observations can be accessed by sending specific messages from RTKlib (TAKASU and YASUDA, 2009), which provides the user with many options for obtaining coordinates, with the goal of activating a low-cost mobile mapping solution. Other receivers from Ublox, such as NEO-6T and NEO-6P, provide GPS raw data output and may be better options depending on the user application.

The positions retrieved with the low-cost GNSS receivers are mainly affected by the atmospheric refraction, errors on the broadcast ephemeris, multipaths and measurement noise. Such errors and the use of pseudoranges create a difficult scenario for achieving sub-metric accuracy. However, its affordable cost, small size, and low power consumption have inspired many efforts in the development of techniques and algorithms to improve its positional solution. For instance, Takasu and Yasuda (2009) developed a low-cost unit linked to a high-precision algorithm for performing RTK positioning and obtained a centimetric error with a reasonable fixing ratio of 50-60% using a baseline length of 7 km. Hedgecock et al. (2013) also developed a differential algorithm for short baselines (< 3.5 km) to reach a relative positional error of approximately 10 to 35 cm, depending on the speed of the receiver. These studies were performed in areas with unobstructed sky conditions. For forests and city paths, challenging conditions must be considered, such as signal distortion (positive or negative interferences) and attenuation due to multipath or obstruction of the GPS signals by vegetation (branches, trunks and leaves) and urban features (buildings and roofs). In this regard, some studies have been performed to estimate low-cost GPS accuracy in obstructed sky environments, such as forest (RODRÍGUEZ-PÉREZ et al., 2007; KAARTINEN et al., 2015). These studies have provided meter-level positional accuracy driven by forest canopy characteristics, such as the stand density. These good results inspired an evaluation of Ublox NEO-6M positional accuracy in both open-sky and forest environments, which results are presented in Section 2.4.1.

The inertial measurement unit used, named MPU 6050, is a (*InvenSense*) combination of three axis MEMS accelerometers and MEMS gyroscopes (six degrees of freedom - 6DoF) integrated in a GY-521 electronic board (digital motion processor). MPU 6050 provides accelerations and gravity data for the X, Y and Z axis using a 16-bit digital analogue converter. Then, the angles (yaw, pitch and roll) are estimated in real time using an EKF approach. With the aim of acquiring more suitable data from MPU6050, a previous calibration to estimate the three

gyroscopes and three accelerometer offsets was performed using a I2Cdev device library code (motion apps library, ROWBERG, 2017). The calibration process aims to compute the leveling errors (zero error) of the three gyroscopes and three accelerometers (six parameters) related to the PMTS platform in an initial static position. These six parameters estimated in the IMU calibration were used in the main Arduino script for data integration (Section 2.2.3).

Some works have focused on the feasibility analysis of MPU 6050 applications in mobile mapping. Alam et al. (2014) presented a comparison of MPU6050 performance using complementary filters, for instance, the explicit complementary filter (ECF) and gradient descent-based orientation filter (GDOF) and showed improvements in terms of time processing and angle estimation. Marino et al. (2016) presented an application using the MPU 6050 in a vehicle platform. These authors assessed the accuracy of the acceleration measurement with this IMU over time of trajectory and obtained a standard deviation of 0.1921 m/s^2 . The performance of a low-cost MEMS-based orientation sensor associated with a Kalman filter and a digital imaging sensor has been also analyzed for the estimation of a two-wheel vehicle (motorcycle) trajectory, showing meter-level accuracy in position estimation (GUARNIERI et al., 2013). The behavior of motorcycles resembles that of a PMTS system since in both cases a significant variation in roll angle can be observed that is not observed for four-wheel vehicles. These variations affect the estimation of the yaw and pitch angles (GUARNIERI et al., 2013). In addition to these studies, the accuracy of the trajectory obtained with MPU 6050 in the PMTS is discussed in Section 2.4.1. The navigation sensors' properties are summarized in Table 2.

Table 2. Main technical specifications of MPU 6050 and Ublox NEO-6M.

MPU6050	Sensor	Full scale Range	Resolution	Linearity	Sensitivity
	Accelerometers	$\pm 4g$	0.001198 m/s^2	0.5%	8192 LSB*/g
	Gyroscopes	$\pm 250^\circ/\text{s}$	$0.0076294^\circ/\text{s}$	0.2%	131 LSB*/($^\circ/\text{s}$)
UBLOX NEO 6M	Sensor	Time pulse frequency range	Time pulse resolution	Velocity resolution	Heading resolution
	GPS L1 frequency	0.25 Hz to 1 kHz	30 ns	0.1m/s	0.5°

*LSB (Least Significant Bit per unit).

2.2.3. PMTS data integration

PMTS provides omnidirectional images, GPS positions and orientation angles from the IMU. Time synchronization and estimation of position and angle offsets between each

measurement unit are therefore required to extract the corresponding attitude and position of the camera for each instant in which images were acquired. This section describes the procedures from acquisition through postprocessing for data integration.

An Arduino script was developed to synchronize the navigation sensors. The algorithm starts by defining the sensor configuration. MPU 6050 and Ublox NEO-6M were set to a baud rate of 9600 bps (bits per second). The GPS receiver had a data collection rate of 1 Hz, and the IMU data collection rate was 10 Hz. After the configuration settings, both sensors are initialized. In case of signal or connection loss, the sensor initialization function is restarted. The main loop was developed to collect NMEA GPS and IMU raw data, compute the yaw, pitch and roll angles (γ , ρ , r) from the IMU raw data using EKF, and record both on the SD card. The Arduino microcontroller communicates with the SD card using a Serial Peripheral Interface (SPI). Navigation data were recorded on the SD card considering a GPS data frequency with a rate of 1 Hz.

One difficulty in PMTS data integration is the communication between the omnidirectional system and the navigation system, since Ricoh Theta S does not provide a serial port for this purpose. In this regard, the Coordinated Universal Time (UTC) obtained from Ublox NEO-6M was used for sensor data integration. The Ricoh Theta S records 29 dual-fisheye video frames per second, and the camera clock was set in Local Time (LT). Therefore, frame videos could be associated with the navigation data considering the transformation between LT and UTC. Visual synchronization was implemented by means of a small LED (Light Emitting Diode). The light pulse was synchronized with the GPS record and appeared in the dual-fisheye video frames. This synchronization was performed with an average error of 0.04 seconds. A typical person walks at an average speed of 1.4 m/s, resulting in, at most, a 5-cm synchronization error in the object space, which is less than the expected positioning error using low-cost GPS receivers (TAKASU and YASUDA, 2009 and HEDGECK et al., 2013).

Another concern is the estimation of linear and angular misalignments between sensors to fully calibrate the system. PMTS calibration consists of two main steps: first, Relative Orientation Parameters (ROP) is estimated between cameras in the omnidirectional system, which has been performed in the camera calibration process considering the Exterior Orientation Parameters (EOP) estimated in the bundle adjustment (Chapter 3). Then, the parameters between the cameras and navigation system sensors are estimated, which consist of lever-arms (offsets)

between the camera projection centers and GPS antenna phase center and boresight misalignments between the IMU and cameras.

Since ROP between cameras are known (Chapter 3), the lever-arm and boresight can be calculated for camera 1 and extended to camera 2. Several methods have been proposed to estimate system calibration parameters with direct and indirect measurement approaches. Regardless of the method used to estimate the lever-arms and boresight angles, the accuracy of these parameters should be superior to the system noise level. Therefore, considering an expected centimetric accuracy for PMTS data, the lever-arms (offsets) between each camera projection center and GPS antenna phase center (ΔX , ΔY , Δh) were directly measured in the laboratory using a caliper with 0.05-mm accuracy. Table 3 presents the angular and positional offsets estimated between the Ricoh Theta S (camera 1) and PMTS platform (CAM1/Platform) and between camera 1 and camera 2 (ROP), estimated as presented in Chapter 3.

Table 3. Angular and positional offsets in the Personal Mobile Terrestrial System.

OFFSETS	ΔX (mm)	ΔY (mm)	Δh (mm)	Δr (°)	$\Delta \rho$ (°)	$\Delta \gamma$ (°)
CAM1/Platform	101.4	85.05	138.0	-78.75	29.45	174.57
ROP	-0.79	0.41	1.904	0.0149	0.0526	179.66

Boresight misalignment angles are difficult to directly measure in the laboratory with suitable accuracy, and thus they are usually estimated in an indirect process. With this aim, a bundle adjustment was previously performed using the Structure from Motion (SfM)-based software Agisoft PhotoScan to estimate the camera attitude indirectly (reference matrix), and the platform attitude was directly measured using MPU6050. Because the camera and platform orientation angles are known, the 3×3 rotation matrixes between both systems (R_{CAM} and R_{IMU} , respectively) can be computed. Then, the boresight misalignment angles (Δr , $\Delta \rho$, $\Delta \gamma$) were computed as presented in Equation 1 and Equation 2, in which R_b is the rotation matrix with boresight angles (SCHWARZ et al., 1993). Other methods to estimate boresight misalignment angles between the IMU and camera systems can be found in Mostafa (2001).

$$R_b = R_{cam} * R_{IMU}^{-1}; \quad (1)$$

$$\Delta \rho = \text{asin}[R_{b\ 3,1}]; \Delta r = \text{asin}\left[\frac{-R_{b\ 3,2}}{\cos \Delta \phi}\right]; \Delta \gamma = \text{acos}\left[\frac{R_{b\ 1,1}}{\cos \Delta \phi}\right]. \quad (2)$$

2.3. PMTS ASSESSMENTS

Section 2.3 summarizes the PMTS feasibility study to estimate system accuracy for close range applications, focusing on forest and outdoor building areas. Experiment I aim to assess the positional and attitude trajectory accuracy of the low-cost navigation system (Section 2.3.2). Navigation data will be used as initial information in Experiment II (Section 2.3.3), which supports a more reliable photogrammetric process, for instance, in image matching and bundle adjustment procedures. Experiment II evaluates the object reconstruction performance using the PMTS data, considering three different environments (Experiment II - trials A to C).

2.3.1 Test areas

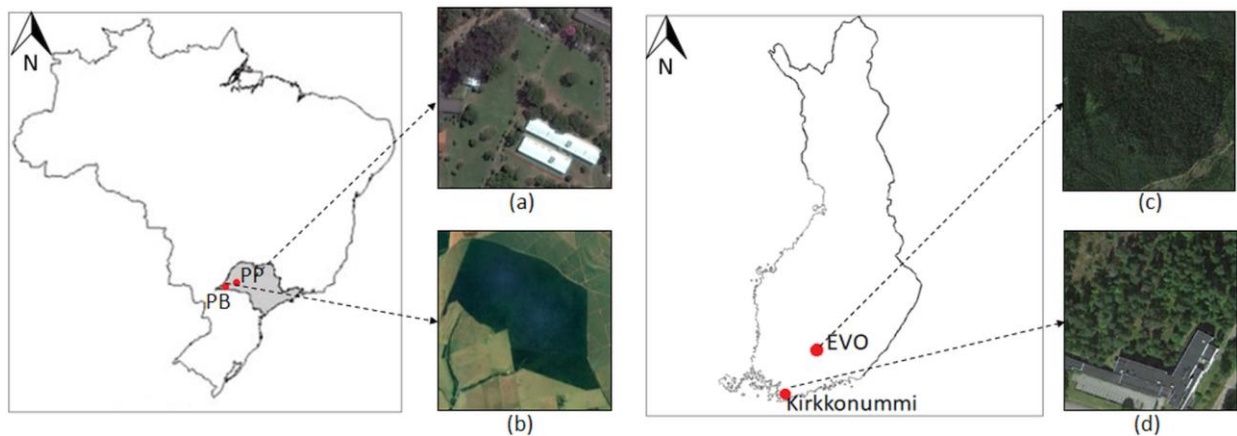
Four different test areas located in Brazil and Finland were imaged to assess the positional and attitude trajectory and the object reconstruction performance using the PMTS data. Figure 4 illustrates these areas.

Test areas 1 and 2 are located in Brazil. Test area 1 ($22^{\circ} 07' S$, $51^{\circ} 24' W$) is located inside the São Paulo State University (UNESP) campus at Presidente Prudente (PP; Figure 4.a). Two sites had been mapped in Test area 1 (a and b). Test site 1.a is a narrow path covered with sparse vegetation areas (*Eucalyptus* trees and ground vegetation - grass) and urban features (small building, light poles and traffic signs), enabling to analyses PMTS performance under a diversity of features to be tracked. Test site 1.b is composed of small buildings (height lower than 5 m) located at a sidewalk, where wheeled vehicles are not allowed (Figure 3.b). Test area 2 is an Atlantic tropical forest reserve named *Ponte Branca* (PB), which is located in the West region of São Paulo State ($22^{\circ} 25' S$, $52^{\circ} 30' W$; Figure 4.b). The phytophysiology of the Atlantic Forest biome in test area 2 is a seasonal semi-deciduous forest, typical of southeast Brazil. This type of vegetation is a transitional forest between coastal forest and non-forest formations comprising trees with a high canopy (e.g., *Ceiba speciosa*, *Cedrela odorata* and *Handroanthus ochraceus*) and a large number of small trees (e.g., *Syagrus romanzoffiana* and *Plinia cauliflora*). Both areas were used to assess the performance (positional accuracy) of the Ublox NEO-6 GPS receiver by allowing experiments with distinct canopies density. Furthermore, test area 1 was also used to evaluate object reconstruction accuracy.

Test areas 3 and 4 are located in Finland. These areas were used to assess the attitude measurement accuracy and the object reconstruction performance, respectively. Test area 3

(Figure 4.c) is part of a deciduous boreal forest region at Evo in Southern Finland ($61^{\circ} 19' \text{ N}$, $25^{\circ} 11' \text{ E}$), with a tree density of 625 stems/ha. Test area 4 is located in Kirkkonummi ($60^{\circ} 04' \text{ N}$, $24^{\circ} 32' \text{ E}$) in the backyard of the Finnish Geospatial Research Institute (FGI) office building (Figure 4.d). This area has sparsely distributed mature trees with a tree density of less than 300 stems/ha. The forest composition in both test areas (3 and 4) includes six tree species: Scots pine (*Pinus sylvestris*), Norway spruce (*Picea abies*), Birch (*Betula sp.*), Junipers (*Juniperus communis*), Eurasian aspen (*Populus tremula*) and rowans (*Sorbus aucuparia*), illustrated in Figure 3.a.

Figure 4. Test areas: (a) Test area 1 – UNESP campus at Presidente Prudente, São Paulo, Brazil; (b) Test area 2 – Ponte Branca at Euclides da Cunha, São Paulo, Brazil; (c) Test area 3 – boreal forest at Evo, Finland; (d) Test area 4 – FGI at Kirkkonummi, Finland.



Source: Campos et al, (2018a)

2.3.2 Assessment of the positional and attitude trajectory accuracies

2.3.2.1 Positional Accuracy Experiments

Ublox NEO-6M experiments were performed considering the concepts of absolute and relative positional accuracy (ISO, 2011). Absolute positional accuracy can be assessed considering the closeness of the measured value to a reference value (ISO, 2011). To acquire reference positioning data, a Topcon HIPER SR (L1/L2) GNSS receiver was installed in the backpack platform to acquire a reference trajectory to assess the Ublox receiver. Topcon SR data were acquired in kinematic mode, and relative positioning in a post-processed mode was computed with RTKlib (TAKASU and YASUDA, 2009). The reference data were obtained with 87% of the points with the ambiguity-fixed solution. The standard deviation of the positioning

solution with Topcon HIPER SR receiver data (reference) was 8.8 mm in N, 9.2 mm in E, and 21.3 mm in height. Ublox NEO-6M data were acquired in NMEA format.

With the aim of comparative analysis, Ublox NEO-6M and Topcon HIPER SR observations can be synchronized by the GPS time, setting the same collection rate (1 Hz), and determining the physical offset between them in the platform. Therefore, observations were measured by Ublox NEO-6M and Topcon HIPER SR receivers at the same seconds (GPS time) of survey associated to the trajectory. The distance between the coordinates obtained by Ublox NEO-6M and Topcon HIPER SR for the same second of survey, with the subtraction of the offset, can be considered as errors in the trajectory positioning obtained with Ublox NEO-6M. The offsets between receivers were measured directly using a caliper with 0.05 mm accuracy. Planimetric (D_{plan}) and planialtimetric ($D_{planialt}$) errors were analyzed by comparing the distance (D) between point coordinates (East, North and height – E, N, h respectively) measured by the Ublox NEO-6M (P_{ublox}) and the reference –Topcon HIPER SR (P_{ref})– corrected from offsets at a frequency of 1 Hz (Equations 3 to 5).

$$D = P_{ublox} - P_{ref} + offsets, \quad (3)$$

$$D_{plan} = \sqrt{(E_{ublox} - E_{ref})^2 + (N_{ublox} - N_{ref})^2}, \quad (4)$$

$$D_{planialt} = \sqrt{(E_{ublox} - E_{ref})^2 + (N_{ublox} - N_{ref})^2 + (h_{ublox} - h_{ref})^2}. \quad (5)$$

The relative positional accuracy is the closeness of the relative positions of the features in a data set to their respective positions accepted as a reference (ISO, 2011) and provides different insights on the positional accuracy compared to absolute assessments. For the relative positional accuracy assessment, a set of distances between consecutive points acquired with 1 Hz frequency was computed for Ublox NEO-6M ($R_{D1(i)}$) and for Topcon HIPER SR ($R_{D2(i)}$). Then, the differences between the relative distances ($\Delta R_{D(i)}$) estimated for both sensors in the same period were compared (Equation 6). Experiments were performed considering trajectories in open and vegetated areas inside test areas 1 (a) and 2, with the aim of assessing both the positional accuracy and loss of signal in covered areas under simulated forest conditions.

$$\Delta R_{D(i)} = R_{D1(i)} - R_{D2(i)}. \quad (6)$$

2.3.2.2 Attitude Accuracy Experiments

Data acquisition for the attitude accuracy experiment was performed in test area 4. The absolute accuracy of the PMTS attitude measurement with MPU6050 ($\gamma_{MPU6050}$; $\rho_{MPU6050}$; $r_{MPU6050}$) was evaluated by comparing MPU6050 data to the reference data (γ_{ref} ; ρ_{ref} ; r_{ref}) acquired simultaneously with a more accurate device (VN-100 Rugged IMU), which was also installed in the backpack platform ($D\gamma$; $D\rho$; Dr).

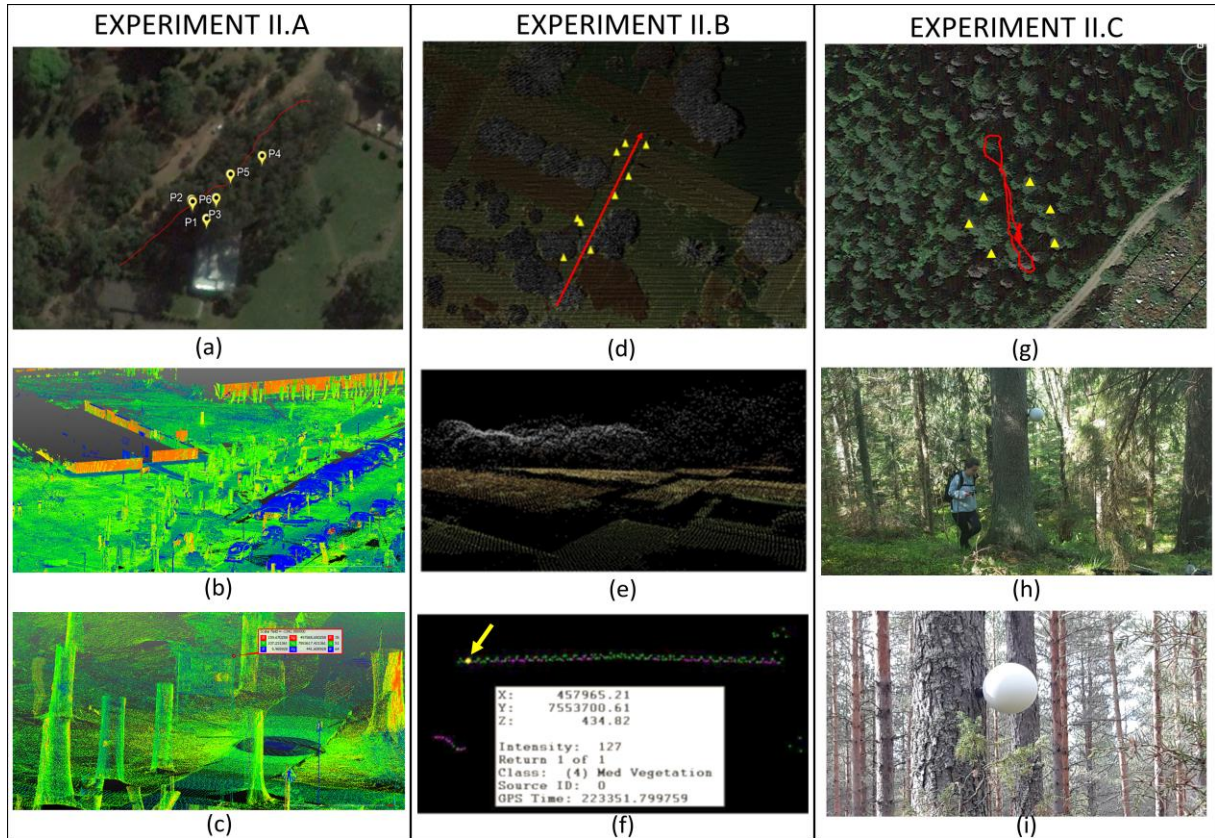
The estimated boresights misalignment angles between IMUs ($\Delta\gamma=-36^\circ$; $\Delta\rho=0.26^\circ$; $\Delta r=-3.47^\circ$) were accounted for in the analysis (Equation 7). The VN-100 Rugged is a high-performance IMU and Attitude Heading Reference System (AHRS) comprising 3-axis accelerometers, 3-axis gyros, 3-axis magnetometers and a barometric sensor (VECTORNAV, 2017). This IMU has a nominal angular resolution and alignment error of 0.05° degree, which provides good reference data.

$$D\gamma = \gamma_{MPU6050} - \gamma_{ref} + \Delta\gamma; D\rho = \rho_{MPU6050} - \rho_{ref} + \Delta\rho; Dr = r_{MPU6050} - r_{ref} + \Delta r. \quad (7)$$

2.3.3 Assessment of object reconstruction

3D object reconstructions in three different environments (Experiment II, trials A to C) with the SfM-based software Agisoft PhotoScan (AGISOFT LLC, 2017) using PMTS data were performed. PhotoScan projects were built for all experiments setting the mathematical model (fisheye), existing Interior Orientation Parameters (IOP, Chapter 3), initial EOP from the off-the-shelf navigation system (Section 2.2.2) and system offsets (Section 2.2.3 – Table 3). The standard data processing performed for all experiments is detailed in Section 2.3.3.1. Figure 5 summarize Experiment II, trials A to C.

Figure 5. Experiment II.A to II.C: **a**, **d** and **g** illustrates PMTS trajectory and GCP distribution for experiments II.A, II.B and II.C, respectively; **b** and **e** show a 3D visualization of the terrestrial and aerial point cloud used as reference data to measure GCP in experiments II.A and II.B, respectively; **c** and **f** show examples of CGP measurement for experiment II.A and II.B in the reference point cloud using CloudCompare and Fugro Viewer software, respectively; **h** and **i** exemplify the positioning of the white spheres in the tree used as GCP in experiment II.C.



Experiment II.A (CAMPOS et al., 2017) was developed in test area 1 (a). A set of 57 omnidirectional images were acquired to reconstruct the environment surrounding a path of 40.6 m (Figure 5.a), considering only the data from Sensor 1. The pixel size in object space units of the images in this dataset ranges between 2.5 to 5 cm considering the central part of the image. The pixel size in object space units is higher than 5 cm in the image periphery. Omnidirectional images have a non-uniform spatial resolution, which explains the considerable variation in pixel size in object space units along the image. Consecutive images were acquired with an average base length of 0.7 m. Ground Control Points (GCP) were obtained from a LASER point cloud collected with a TLS (Leica ScanStation P40) in the same area, with an average accuracy of 5 mm (Figure 5.b). The 3D ground coordinates of seven (7) well-defined features in the LASER point cloud (GCP) were

obtained using the CloudCompare software (GIRARDEAU-MONTAUT, 2017). A sample of TLS point cloud used as reference data in Experiment II.A is shown in Figure 5.c.

Experiment II.B (CASTANHEIRO et al., 2018) was performed in test area 1 (b). The test area 1 (b) is composed by 5 buildings with different anthropic features (Section 2.4.2.2, Figure 10). The PMTS is an alternative to obtain complementary urban data to the conventional aerial photogrammetric surveys or Aerial LASER Scanning (ALS) mapping. Despite of aerial Photogrammetry and ALS survey had been a standard technique for urban data acquisition, some relevant features for city mapping applications, such as building's façade and land use, are only available from a terrestrial perspective. Therefore, the main goal of this experiment is to analyze the feasibility of the PMTS for the generation of complementary 3D urban models without the need for direct measurements. Since ALS is the most used system for urban data acquisition, we proposed in this experiment the use of an aerial point cloud as source of control data (GCP), aiming a complementary approach using ALS and PMTS data for a complete urban mapping without direct measurements. In this regard, a trajectory of approximately 70 m (Figure 5.d) was covered with a set of 80 dual fisheye video frames. The dual fisheye frames were converted to 160 omnidirectional images with pixel size in object space units ranging from 2 to 5 cm, considering the data from both sensors that composes Ricoh Theta S system. The nine GCP used (four to Sensor 1 and five to Sensor 2) were obtained from an aerial LASER point cloud (RIEGL LMS-Q680i) that covers the same test area (Figure 5.e), with a density of 5.8 points/m² and a planialtimetric accuracy of approximately 30 cm (TOMMASELLI et al., 2018). Figure 5.f shows an example of GCP measurement (building edge) in the aerial LASER point cloud using FugroViewer software. A profile of a building (roof) was obtained using the 3D visualization tools of FugroViewer software.

Experiment II.C (CAMPOS et al., 2018) was performed in the Evo forest. A set of 66 dual fisheye video frames from test area 3 was selected to reconstruct a path in a 16 x 16 m plot with PMTS (Figure 5.g), resulting in 132 omnidirectional images (Section 2.2.1), with a pixel size in object space units ranging from 3.5 to 7 cm. Six targets (white spheres) to be used as GCPs were installed in some tree stems inside the plot (Figure 5.h and Figure 5.i). GCPs were accurately positioned using a tachymeter (total station) with an average accuracy of 3 mm. Just a few GCPs were available. The number of spheres was originally optimized to provide sufficient georeferencing for the TLS scans. However, the distribution of the spheres aiming TLS processing

did not provide an optimal geometry for photogrammetric process. Furthermore, the installation of GCPs inside the forest plot is challenging due to occlusions and accessibility.

2.3.3.1. General data set and processing

The PMTS data were integrated as described in Section 2.2.3. Each omnidirectional image has its position and attitude collected by the navigation system. The camera position was obtained in 3D geographic coordinates with Ublox NEO-6M. These coordinates were converted to UTM coordinate system (Universal Transverse Mercator – 22 zone) for Brazilian test areas (Experiments A and B) and to the ETRS -TM35FIN system (E, N, h) for Finnish test areas (Experiments C), which is compatible with the coordinate system that is used for GCPs. For those positions in which a loss of signal occurred, camera coordinates and attitude were estimated by an interpolation process (linear). The accuracy values obtained in the assessment of Ublox NEO-6M and MPU6050 (Section 2.4.1) were used as standard deviations for the initial EOP data in PhotoScan processing. PhotoScan project also requires camera calibration information. Especially using fisheye lenses, it is recommended to provide IOP estimated from a previous camera calibration process using a mathematical model implemented for wide angle lens geometry (AGISOFT LLC, 2017). In the PhotoScan processing, initial EOPs and IOPs are re-estimated in a bundle block adjustment. Therefore, pixel size and IOP (principal distance and principal point coordinates) of the omnidirectional images estimated from a previous camera calibration process (Chapter 3) were included and set in pre-calibrated mode. Lens distortions using the fisheye model option were not estimated by PhotoScan in the version (1.2) used.

After project configuration, the PhotoScan workflow consists of GCP measurement in the object and image spaces (1); tie point generation and EOP estimation with a SfM algorithm (2); and dense image matching (3) for point cloud generation.

- (1) Experiment II, trials A and B, uses GCPs obtained from terrestrial (Figure 5.b) and aerial (Figure 5.e) LASER point cloud, respectively. The 3D ground coordinates were measured in LASER point cloud using CloudCompare and FugroViewer software, as shown in Figure 5.c and Figure 5.f, respectively. The GCP from Experiment II.C were targeted with white spheres and directly measured by tachymeter (Figure 5.i). All GCPs were identified and measured manually in the image.

- (2) Tie points were generated using the PhotoScan accuracy option “high”, which uses the full resolution images in the processing (AGISOFT LLC, 2017). Gradual automatic filtering was performed to exclude outliers; the reprojection accuracy and reconstruction uncertainty methods were used. The errors in the image observations were less than 1 pixel. Finally, a bundle block adjustment was performed to refine the camera positions, camera attitudes and camera calibration parameters, resulting in a final set of oriented images. Table 4 presents the RMSE for the GCPs. The final RMSE of the GCPs for experiments A and C were approximately 9 and 6 cm, respectively, which is comparable to the range of the pixel size in the object space units in these datasets. Experiment II.B presented a total GCP error of 37 cm, which can be explained considering the ALS point cloud accuracy (± 30 cm) used as reference (TOMMASELLI et al., 2018) and the manual GCP measurements performed using the Fugro Viewer software (Figure 5.f).
- (3) A dense point cloud was generated using the multi-view stereo-reconstruction method (available in PhotoScan software) and the oriented image set. An ultra-high accuracy mode was used, producing a point cloud density, on average, with a distance of 1.7 cm between points (DEWEZ et al., 2016 and AGISOFT LLC, 2017). A depth filtering approach (aggressive mode) was selected to reduce outliers, particularly those resulted from the movement of tree leaves and low vegetation.

Table 4. RMSE of GCP estimation after bundle adjustment for experiment II, trials A to C.

RMSE	$\Delta E(m)$	$\Delta N(m)$	$\Delta H(m)$	Total Error (m)
EXPERIMENT II.A	0.064	0.059	0.042	0.096
EXPERIMENT II.B	0.153	0.221	0.265	0.377
EXPERIMENT II.C	0.060	0.021	0.005	0.064

2.3.3.2 PMTS point cloud accuracy assessment

Different approaches can be used to evaluate point cloud accuracy (LEHTOLA et al., 2017), for instance, considering another point cloud as a reference (absolute accuracy) or using control distances (relative accuracy). Comparisons between forest point clouds are challenging because the correspondences and similarity cannot be guaranteed since the environment mapped is not static (e.g building façades) and the number of outliers can be high. Therefore, one option is to extract a subset of well-defined and stable features in the point clouds,

such as trees stems, and then compute the Euclidian distance among them (LEHTOLA et al., 2017). In other words, the absolute positional accuracy of the PMTS point cloud was evaluated considering how close the features of this cloud were to those of a point cloud used as a reference (ISO, 2011).

The absolute positional accuracy was estimated for experiments A and C. The reference data for both experiments were measured using the Leica ScanStation P40 (TLS) in a static survey in the same forest area and reference system and processed in Leica Cyclone 9.0 software, with a 3D positional accuracy of 3 mm at a 50-m range. This was considered an adequate reference since centimetric accuracy is expected for the PMTS point cloud. Therefore, the PMTS point cloud was compared to the TLS point cloud using CloudCompare software in both experiments. PMTS and TLS point clouds were segmented to ensure that only the compatible features, such as tree stems, were compared. A cross section (clipping box tool) was defined to extract a delimited section from the point clouds. Automatic filtering using CloudCompare was performed in the delimited section to exclude isolated points (points farther than 1 meter of the nearest neighbor) and remaining outliers. Then, the Euclidian distances between the features from the PMTS and TLS point clouds were estimated using the least squares plane local modelling strategy available in CloudCompare. The least squares local method computes a set of small planes that fits a point and its neighbors (six nearest points) in the reference point cloud, enabling a point to plane comparison (GIRARDEAU-MONTAUT, 2017).

The relative positional accuracy (ISO, 2011) of the PMTS point cloud was evaluated considering the discrepancy between distances measured directly in the test area using a caliper and indirectly estimated in the PMTS point cloud, considering well-defined features, such as tree stems and building façades. This analysis is performed for experiments A and B.

2.4. PMTS RESULTS

This section details the results from Experiments I (Section 2.4.1) and II (Section 2.4.2) obtained with the PMTS proposed. In summary, results show that PMTS approaches can be very advantageous for forest and city mapping and enable fast and systematic surveying. These systems can provide detailed and accurate measurements of their surrounding areas from a ground perspective, enabling environmental analysis that cannot be achieved with manual techniques or non-terrestrial sensors. This represents an advance in close-range surveying.

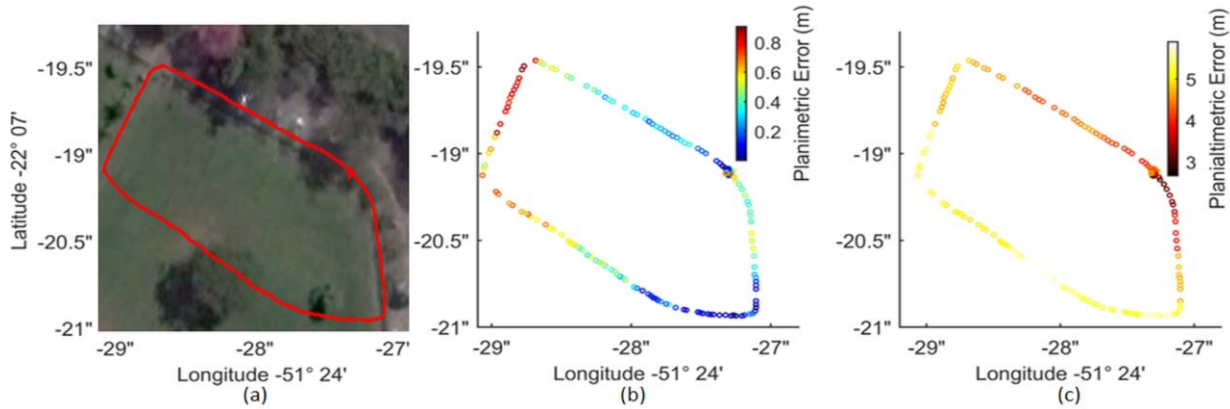
2.4.1 Positional and attitude trajectory accuracies

Figure 6 shows the discrepancies between the Ublox NEO-6M and Topcon HIPER SR trajectories (distances between points in both trajectories collected with 1 Hz frequency) obtained in the open-area experiment (test area 1.a - Figure 6.a).

The planimetric discrepancies (D_{plan}) varied from 7 cm to 1 m (Figure 6.b). Ublox NEO-6M presented good results for the horizontal coordinates, with an average discrepancy of 34 cm with respect to the Topcon SR positions, a standard deviation of 21 cm and a RMSE of 40 cm. The major problem observed in the Ublox NEO-6M positioning was related to height estimation. When vertical differences were incorporated in the calculation of the 3D point errors (Equation 5), an average discrepancy of 4.33 m with respect to the Topcon SR positional coordinates was obtained with a 1-m standard deviation and RMSE of 4.45 m. For instance, Ublox NEO-6M heights are, on average, 4 m systematically higher than Topcon SR height measurements along the PMTS trajectory. The planialtimetric discrepancies (D_{planalt}) are illustrated in Figure 6.c; most of the errors were related to the weak altimetric geometry.

Height estimation is a critical issue in GPS positioning because the positioning accuracy significantly decreases due to loss of satellites, multipath errors and poor satellite geometry. For instance, Figure 6.c illustrates an error increase in the trajectory estimation near to the vegetated area, causing signal obstruction and consequently weak observation geometry. It is reasonable to say that such vegetation degraded significantly the GNSS signals tracked from that direction with more impact in the altitude determination. Since height estimation is more affected by the GNSS geometry degradation than the horizontal coordinates, an absolute accuracy varying between 1 and 5 m is expected in height measurements using the Ublox NEO-6M receiver. This problem can be minimized by using the height differences between observations instead of the absolute values. The relative positions between observations can be used as constraints in a bundle adjustment (BLÁZQUEZ and COLOMINA, 2012). The relative accuracy was estimated comparing the distance between consecutive observations obtained with Ublox NEO-6M and Topcon HIPER SR (reference), see Section 2.3.2.1. The relative accuracy results showed that for displacements corresponding to 1 s (1 Hz), the differences in the distances (ΔR_D – Equation 6) had an average value ($\bar{x}$) of 2.6 cm, with a standard deviation (σ) of 7.8 cm and a RMSE of 8.2 cm. These results are summarized in Table 5.

Figure 6. Comparative analyses between the Ublox NEO-6M single-frequency GPS receiver and the dual-frequency GNSS receiver Topcon SR in the open-area trajectory: (a) the open-area trajectory obtained by Ublox (red), (b) planimetric error (D_{plan}) and (c) planialtimetric error ($D_{planalt}$) in the Ublox trajectory with a one-second interval.



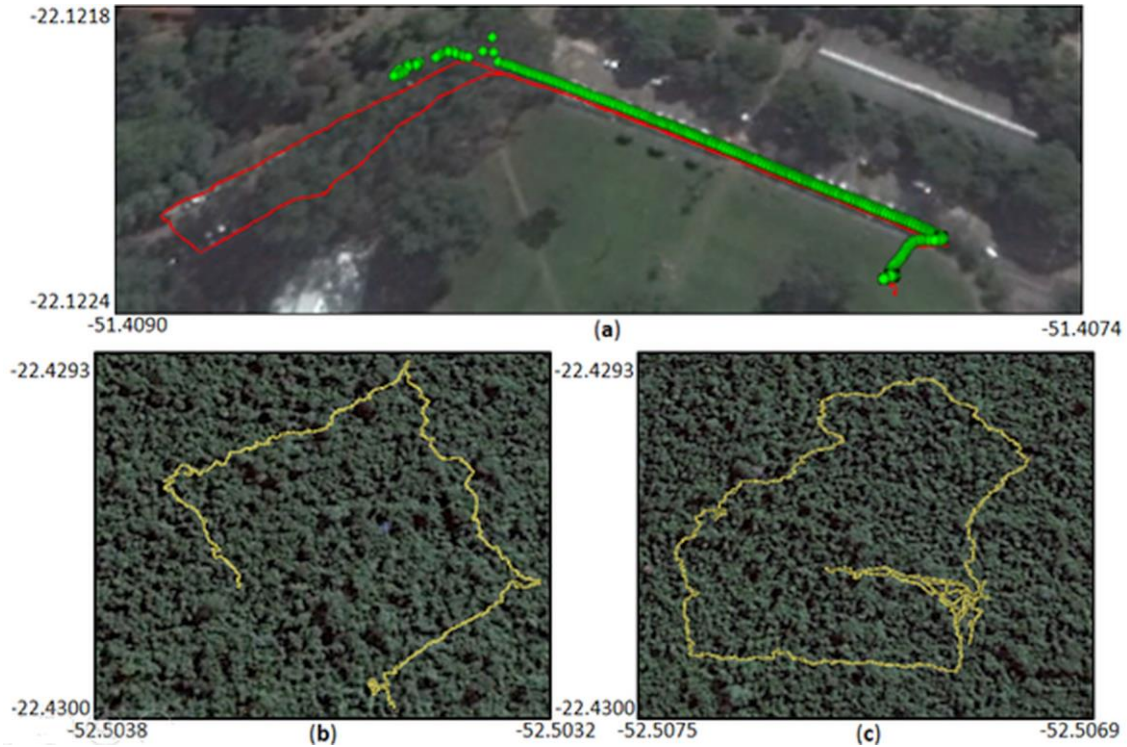
Source: Campos et al, (2018a)

Considering the experiments in the vegetation part of test area 1 (Figure 7.a), major problems have been observed earlier in places under forest canopies due to the number of available satellites (see e.g. NAESSET, 2001 and KAARTINEN et al., 2015). Figure 7.a shows the trajectory obtained with the Ublox NEO-6M receiver (in red) and the points collected by the Topcon SR receiver (in green). A set of seven data collections was performed in the same area on different days and times. A huge loss of signal occurred in all acquisitions with the Topcon SR receiver in covered areas in which less than the minimum of four satellites were tracked. However, with the Ublox NEO-6M receiver, it was possible to reproduce the trajectory inside the vegetated area.

This result can be explained by the Topcon SR filter components, which are more rigorous in rejecting low-power signals and non-RHCP (Right Hand Circular Polarization) signals than the Ublox NEO-6M, which acquires more signals, even those of low quality. Kaartinen et al., (2015) achieved a similar conclusion. In this work, the positioning accuracies of single-frequency (NVS NV08C-CSM) and double-frequency (NovAtel FlexPak6) GNSS receivers were assessed under forest canopies and in motion. The single-frequency devices also presented better results than dual-frequency equipment under forest canopy due to the high-sensitivity of the single-frequency antennas to track weaker signals. In addition, the relative positioning, as applied for Topcon SR data, requires simultaneous observations from the base and rover, which is more likely to be performed with fewer observations than the point positioning used by the Ublox.

Figures 7.b and 7.c present examples of data collection with the Ublox NEO-6M receiver in test area 2 and show that Ublox NEO-6M reproduced the trajectory with meter-level accuracy even under a high-density canopy.

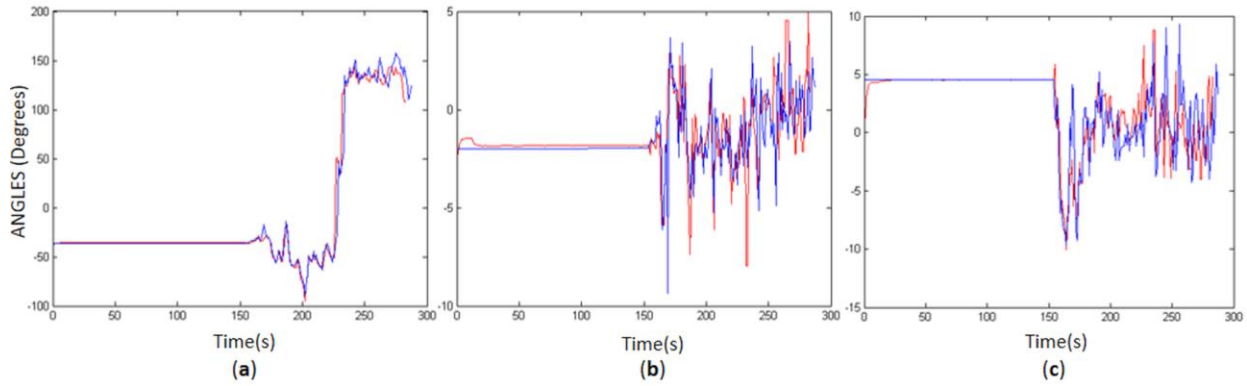
Figure 7. Ublox NEO-6M positioning accuracy. (a) Personal Mobile Terrestrial System trajectory in test area 1 with Ublox NEO-6M (red) and Topcon HiPer SR (green). (b) and (c) are examples of the trajectory with Ublox NEO-6M (yellow) inside test area 2.



Source: Campos et al, (2018a)

The results of the attitude trajectory assessment are presented in Figure 8, in which the yaw (a), pitch (b) and roll (c) angles measured with the VN-100 Rugged are depicted by the blue lines (reference) and MPU6050 by the red lines. Despite some outliers, the platform attitude was well reproduced by the MPU6050 data. Since MPU6050 is noise-sensitive, most of the outliers were due to sudden variations in PMTS attitude and loss of GPS signal, when the yaw error drift was not corrected. Since the GNSS receiver data has some gaps in covered areas, such as test area 4, the accumulated errors in position and attitude calculated from the IMU data increase without a complementary navigation sensor to correct the trajectory drift.

Figure 8. MPU 6050 attitude accuracy, comparing VN-100 Rugged IMU (blue) and MPU6050 (red): yaw (a), pitch (b) and roll (c).



Source: Campos et al, (2018a)

Table 5 presents the statistics of mean ($\bar{x}$), standard deviation (σ) and RMSE of the discrepancies between attitude angles estimated with VN-100 Rugged and MPU6050 in the same instants of time (1 Hz). In a first assessment, outliers were considered in the statistics, which revealed the influence of noise in the attitude estimation (Figure 8). The accuracy decreases mainly in the yaw angle, in which discrepancies between MPU6050 and reference can reach 7° , whereas the pitch and roll errors varies between 1° and 3° . However, these results can be improved by combining the disturbance rejection filter with the EKF estimation. In this regard, a simple noise reduction was performed using a moving-average filter, considering as a threshold the RMSE obtained for each angle in the approach without the noise filter. The MPU6050 results are better when outliers are not considered. The RMSE of the discrepancies between MPU6050 and the reference data are approximately 1° for pitch and roll and 3° for yaw.

Table 5. Summary of the results for the navigation sensors: (1) Mean ($\bar{x}$), Standard deviation (σ) and RMSE of discrepancies between MPU6050 and the reference data with a noise filter for yaw, pitch and roll angles (γ , ρ , r) and (2) $\bar{x}$, σ and RMSE of discrepancies between UBLOX NEO 6M and the reference data in the absolute and relative positional accuracy assessments.

Angles		$\bar{x}$ ($^\circ$)	σ ($^\circ$)	RMSE ($^\circ$)
MPU6050	Yaw (γ)	-0.0888	3.0652	3.0588
	Pitch (ρ)	0.0190	0.5474	0.5463
	Roll (r)	0.0999	1.0690	1.0710
Accuracy		$\bar{x}$ (m)	σ (m)	RMSE (m)
UBLOX NEO 6M	Absolute (Planimetric)	0.34	0.21	0.40
	Absolute (Planialtimetric)	4.34	1.0	4.45
	Relative (Planialtimetric)	0.026	0.078	0.082

The Ublox NEO-6M coordinates and MPU6050 data can be used as input data for many photogrammetric processes, such as constraints in the traditional bundle block adjustment. The approach considering navigation data is well known in photogrammetry as the Integrated Sensor Orientation (ISO). For practical purposes, an ISO approach can be used to improve image matching by reducing the search space and the number of wrong matches, reducing the number of GCP, reducing the drift in parameter convergence, and mitigating uncorrected systematic errors (CRAMER, 2001 and IP et al., 2007).

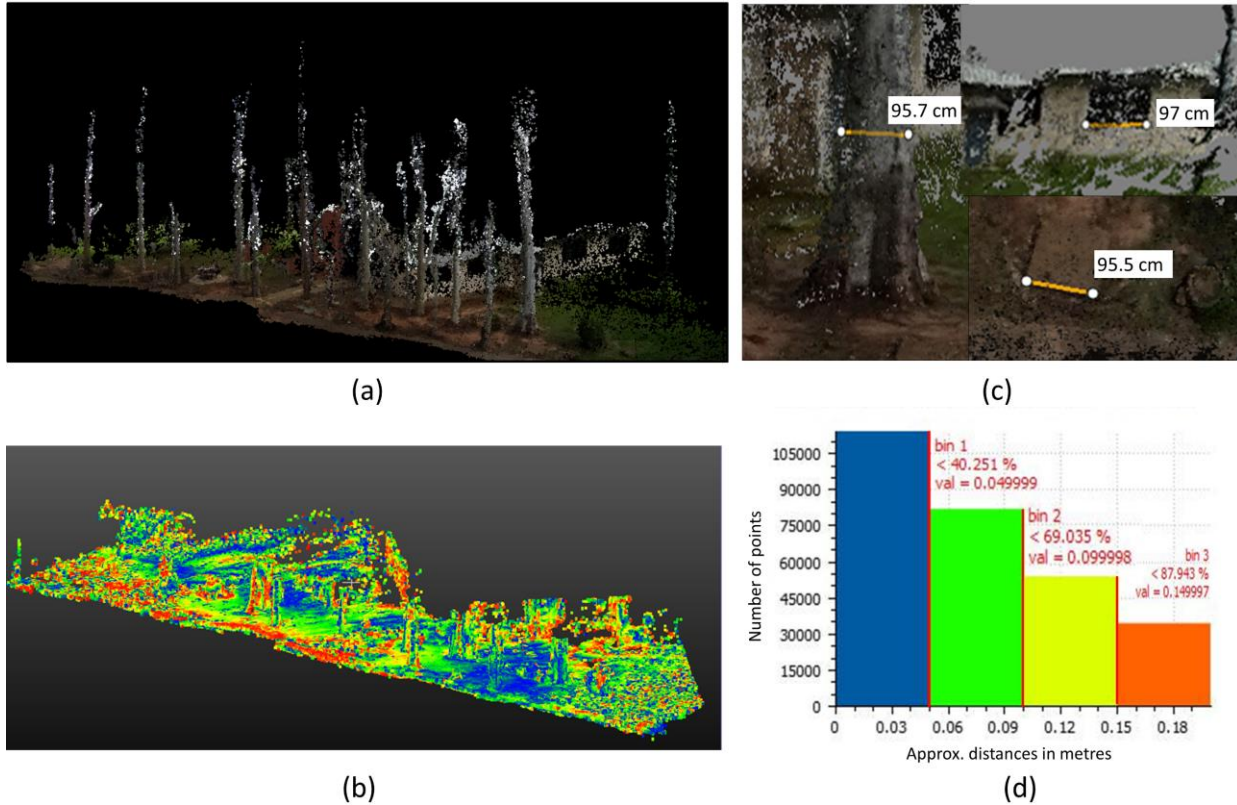
The accuracy values presented in this Section were used as standard deviations for the input data (EOP) in the object reconstruction process with Photoscan. Therefore, the camera positions were imported in PhotoScan with standard deviations of 4 m and attitude angles with standard deviations of 3° . It is important to point out that in the PhotoScan version (1.2) used, it is not clear how the initial position and attitude data obtained from the PMTS are applied in the PhotoScan bundle adjustment.

2.4.2 Accuracy of object reconstruction

2.4.2.1 Results - Experiment II.A

Figure 9 presents the results related to PMTS absolute accuracy for Experiment II.A, which show (a) the filtered point cloud generated with PMTS data, (b) the PMTS point cloud classified according to the Euclidian distance from the TLS reference and (d) the histogram of the classified PMTS point cloud in four classes. Preliminary clipping and filtering processes were performed in the CloudCompare software to compare the PMTS and the TLS point clouds (Section 2.3.3.2), mainly because the TLS reference mapped only a height range of 2 m, as illustrated in Figure 5.b. Tree stems, ground and buildings features were compared (least squares plane local modelling strategy). Considering a set of 283423 points that composes the filtered PMTS point cloud generated, 40.3% of the points (114219) have a Euclidian distance within 5 cm from the reference, while 69% (195561) is within 10 cm. These points were reconstructed with an average accuracy compatible with the pixel size in object space units of the omnidirectional images in this dataset. Less accurate results were obtained for 31% of the points with error ranging between 10 and 20 cm, especially in the PMTS point cloud extremities.

Figure 9. Accuracy assessment of object reconstruction (Experiment II.A results): (a) filtered Personal Mobile Terrestrial System point cloud, (b) classified Personal Mobile Terrestrial System point cloud according to the Euclidian distance from TLS reference, (c) example of feature measurements with PhotoScan tools and (d) histogram of the estimated errors in feature positions.



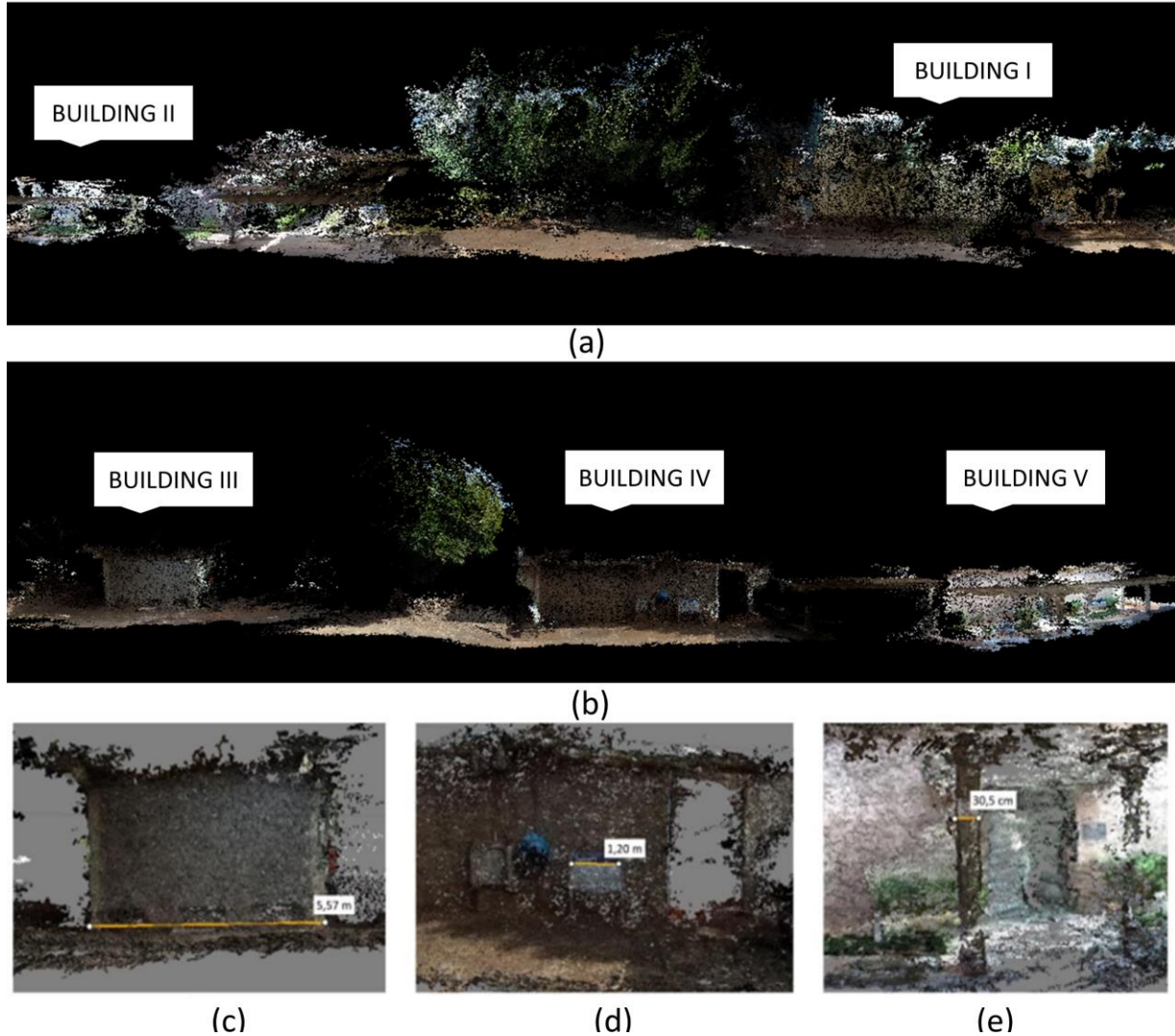
Relative accuracy was estimated according to 12 measurements performed directly in the test area using a caliper and considering well-defined features, for instance, 6 measurements are related to tree stems (DBH, 1.30 m), 3 to building façades and 3 to anthropic objects (traffic signs and power boxes) on the PMTS point cloud, as illustrated in Figure 9.c. The statistics of mean ($\bar{x}$), standard deviation (σ) and RMSE of the discrepancies between the distances measured directly and indirectly in the point cloud (Figure 9.c) were estimated, resulting in 0.002 m, 0.003 m and 0.005 m, respectively. The relative accuracy result does not consider systematic errors.

2.4.2.2 Results - Experiment II.B

The accuracy of object reconstruction on Experiment II.B was performed considering the relative accuracy and the completeness of the PMTS point cloud (ISO, 2011). These quality criteria were considered, aiming to verify the applicability in multipurpose cadastre,

in which relative measurements such as the build area is required. Figure 10.a and Figure 10.b show the PMTS point cloud generated with Sensor 1 and Sensor 2, respectively. Figure 10.c to Figure 10.e show examples of measurements performed in the PMTS point cloud.

Figure 10. Accuracy assessment of object reconstruction (Experiment II.B): (a) Personal Mobile Terrestrial System point cloud from Sensor 1 and (b) from Sensor 2, while, c to e are example of distance measurements for relative accuracy assessment.



A set of 13 measurements was performed. The discrepancies between the distances measured directly in the test area and indirectly in the PMTS point cloud were shown in Table 6. The statistics of mean ($\bar{x}$), standard deviation (σ) and RMSE of these discrepancies were 4.3 cm, 5.5 cm and 6.6 cm, respectively. An average relative accuracy of 6 cm was obtained in this experiment, which can be considered adequate for urban multipurpose cadastre (*Portaria n°*

511/2009). Some outliers can be noticed in Table 6, for instance in the window of building I, with a discrepancy of 20 cm from the reference. These outliers demonstrate the limitations of optical systems in the 3D reconstruction of homogeneous and repetitive pattern areas. Therefore, repetitive pattern features were more prone to errors in the PMTS point cloud.

Completeness analysis consists of identifying the excess and the absence of data (features, attributes and relationships) in a dataset compared with its specification. A set of 33 features were previously specified as interest features to be reconstructed. From this total, 79% of the interest features had been reconstructed. Seven interest features were not reconstructed, which can be associated to a limitation of the omnidirectional system due to the low resolution of the images, since these objects were farther than 20 m from the PMTS

Table 6. Relative positional accuracy assessments: The discrepancies values and $\bar{x}$, σ and RMSE of the discrepancies between the distances measured directly in the test area and indirectly in the Personal Mobile Terrestrial System point cloud.

Building	Features	Reference (m)	PMTS point cloud (m)	Discrepancy (m)
I	Façade	16.650	16.600	0.050
	Window	2.000	1.800	0.200
	Traffic sign	1.220	1.200	0.020
II	Façade	12.900	12.910	0.010
	Traffic sign	1.220	1.180	0.040
	Pillar	0.300	0.315	0.015
	Bench	1.400	1.390	0.010
III	Façade	5.560	5.570	0.010
IV	Façade	10.840	10.800	0.040
	Window	1.530	1.500	0.030
	Traffic sign	1.220	1.200	0.020
V	Façade	12.910	12.800	0.110
	Pillar	0.300	0.305	0.005
$\bar{x}$				0.043
σ				0.055
RMSE				0.066

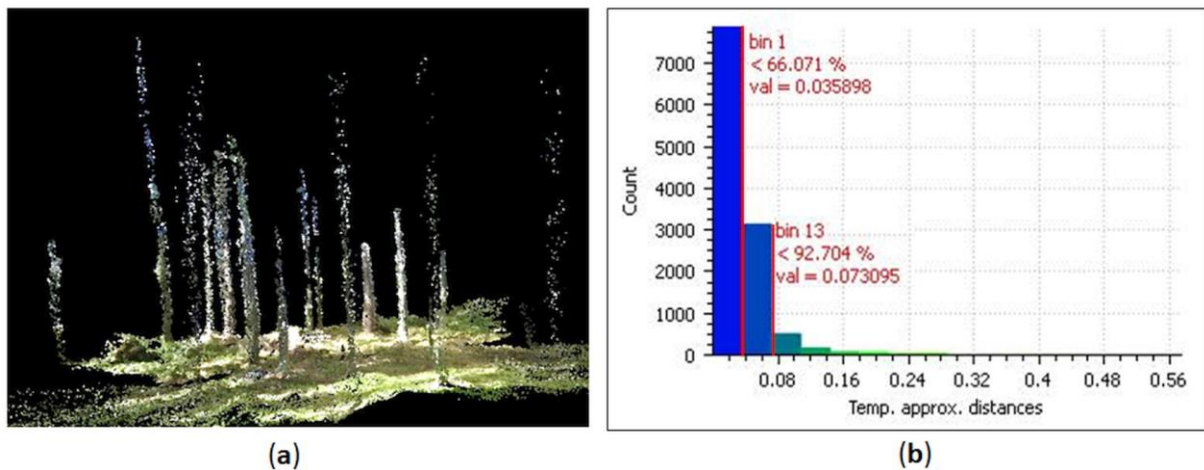
2.4.2.3 Results - Experiment II.C

Figure 11 illustrates the object reconstruction results from Experiment II.C, aiming forest applications, which show (a) the filtered point cloud generated with PMTS data and (b) the histogram of the PMTS point cloud classified in four classes according to the Euclidian distance from the TLS reference.

The accuracy of object reconstruction of the tree stems (11910 points) was assessed by comparing the PMTS cross section to the TLS reference. The Euclidian distances

between points of the tree stems in the PMTS point cloud and planes in the reference (least squares plane modelling – 6 points) were calculated (Figure 11.b) in CloudCompare software. From this set, 66% of the points had errors less than 3.5 cm (7861 points), consistent with the range of the pixel size in object space units of the images (3 to 7 cm in the central part); 15.6% of the errors (1853) were within the range of 3.5 to 5 cm, and 10.8% of the errors were within the range of 5 to 7 cm (1293). Therefore, the errors in the stem points acquired by the PMTS were close to the range of the pixel size in object space units of the images in 92.4% of the points, when considering the normal distribution; 7.6% of the points (904) had errors larger than 7 cm. The statistics of the average and standard deviation of the discrepancies between the PMTS and reference were 3.4 cm and 3.7 cm, respectively, comparable to the values for other PMTS approaches for forest mapping (KUKKO et al., 2012; LIANG et al., 2014 and FORSMAN et al., 2016)

Figure 11. Accuracy assessment of object reconstruction. (a) Personal Mobile Terrestrial System point cloud after tree leaf filtering and (b) histogram of the estimated errors in tree positions.



Source: Campos et al, (2018a)

2.5. DISCUSSION

The aim of the experimental assessments was to analyze the feasibility of the developed PMTS for forest and outdoor city mapping. Table 7 summarizes the results obtained for experiment II, trials A to C. This section presents a discussion of PMTS performance compared to other backpacked approaches for these applications that use optical systems (FORSMAN et al., 2016; TSENG et al., 2016) or LASER scanners (KUKKO et al., 2012; LIANG et al., 2014; LAUTERBACH et al., 2016; OVELAND et al., 2017)

Table 7. Absolute and relative accuracies of the PMTS point clouds from experiment II.

Accuracy	EXPERIMENT II.A	EXPERIMENT II.B	EXPERIMENT II.C
<i>Absolute (cm)</i>	10.0	----	7.0
<i>Relative(cm)</i>	0.5	6.6	----

In optical systems, the level of accuracy that can be achieved is directly related to the pixel size in object space units and intersection geometry, which depends on the sensor size, focal length, sensor-object distance and image scale. Omnidirectional images have a reduced angular resolution and a larger image scale variation when compared to perspective images; however, the increased field of view allows more features to be matched between images, thereby reducing the number of images required to map the environment and enabling a fast and systematic survey. For instance, the optical approach with five perspective cameras in a side-by-side configuration proposed by Forsman et al. (2016) is limited by system weight (13 kg) and the implementation of a single-camera method considering differences in perspective between cameras. Only a limited area is acquired from the center of the plot. Therefore, a PMTS with omnidirectional system represents an alternative to improve the limitations pointed out in Forsman et al. (2016) while achieving comparable accuracy. With this motivation, Tseng et al. (2016) presented an omnidirectional system composed of eight single-lens cameras in back-to-back position. This system provided a panoramic image resulting from a stitching process. Therefore, techniques for perspective images, such as the traditional bundle adjustment, can be applied. However, the stitching process can include errors, such as blurring, and further computational cost. Our PMTS preserves the original geometry of the Ricoh Theta S images to avoid additional errors in the photogrammetric process.

The main advantages of the Ricoh Theta S in relation to the systems developed in Forsman et al. (2016) and Tseng et al. (2016) are its low weight and cost and the internal synchronization. However, this compact solution, with the current configuration, has low image resolution. The results showed that due to this low resolution, the reconstruction accuracy is higher only for features near the trajectory. The imaging range of the system with acceptable accuracy is 5 m. Objects farther than 5 m from the imaging system were not reconstructed or were inaccurate, especially in forest environments. This is the main limitation of the PMTS compared to LASER approaches (KUKKO et al., 2012; LIANG et al., 2014; LAUTERBACH et al., 2016 and OVELAND et al., 2017). This issue can be solved by changing the Ricoh Theta S to another omnidirectional camera with higher resolution and trigger control for image exposure, such as

LadyBug5. The methodology presented in Chapter 2 can be extended to any other omnidirectional camera model.

Advantages with respect to LASER systems can also be mentioned, especially for forest mapping. The use of initial information (IOP, EOP and ROP) combined with a minimum number of GCPs enables an accurate trajectory estimation (centimetric level), even in covered areas. This is the main advantage of optical systems, compared to LASER PMTS approaches. Most LASER approaches have been assessed in boreal forests (KUKKO et al., 2012 and LIANG et al., 2014), which have sparse tree distributions. Therefore, dense forests, such as test area 2, create new challenges for PMTS application, especially in terms of trajectory estimation. Indirect methods for the determination of EOPs, such as image bundle adjustment, are a good alternative for forest mapping, especially in dense areas. Direct methods based on GNSS and IMU may not be suitable to the required accuracy, especially in cases of signal loss (RODRÍGUEZ-PÉREZ et al., 2007). Furthermore, a direct determination requires robust and high-accuracy equipment, which, in general, means high costs. However, there are also critical challenges in the use of passive optical systems in forest environments. Forests, especially the dense ones, have a huge illumination variation and shadows occurrence in the images. The presence of shadows can negatively affect image quality and photogrammetric processes, including image matching, image classification, automatic detection of objects and loss of information, which requires an image pre-processing step (FREITAS et al., 2017). Therefore, it is also important to highlight the possibility of combining an omnidirectional system with independent georeferencing capability with other complementary technologies, such as a LASER scanner. For instance, an optical system can complement the LASER scanner in position and attitude trajectory estimation, planimetric accuracy and assigning RGB colours for LASER point clouds. In addition, the LASER scanner can increase the mapping range area, penetrate vegetation with good accuracy in depth, and is subject to fewer problems in shadow areas and homogeneous areas, such as indoor environments.

With respect to PMTS accuracy, most of the PMTS mentioned in Section 2.1 achieved a centimetric-level of accuracy in 3D object reconstruction ranging between 2 and 5 cm. The object reconstruction results showed that the PMTS point cloud has an absolute accuracy ranging from 5 cm to 10 cm (Table 7). Therefore, the system presented in this section can be considered consistent with other PMTS approaches, especially when compared to other optical systems.

2.6. CONCLUSIONS

A low-cost and lightweight PMTS comprising an omnidirectional camera with off-the-shelf navigation systems was presented and assessed in this study. Experiments explored the feasibility of the proposed PMTS for forests and outdoor city mapping, which can be considered the largest applications of backpacked systems. Our results provided new insights on the suitability of these novel sensors, which have recently entered the market.

First, despite positioning challenges in forest areas, for instance, due to multipath, signal obstruction and measurement noise, the GPS point positioning using UBLOX NEO-6M reproduced the PMTS trajectory with centimetric accuracy in planimetry and meter-level accuracy in altimetry. These data can be used, for instance, as a source of initial information in a traditional bundle adjustment. Another interesting possibility for the application of low-cost GPS data is to estimate the relative displacements between observations along the trajectory, which can be used as relative constraints in the bundle adjustment. However, the use of MEMS-IMU is more challenging, mainly because the measurement accuracy decreases quickly as a function of time due to deterministic and random errors. Particularly in forests, where GPS signal are constantly obstructed, these errors increase without a trajectory drift correction. More rigorous methods for modelling these errors and filtering noises are required.

Second, our study was the first to assess the performance of a low-cost backpack-mounted omnidirectional camera system in the estimation of forest and urban features. Centimetric accuracy was obtained for object reconstruction using the PMTS data processed with PhotoScan software. These results are comparable to those for other PMTS approaches for forest and city mapping, which are currently ten to hundreds of times more expensive. Modelling of individual tree stems is a fundamental process for extracting DBH measurements and other forest parameters along the stem, while, the indirect measurement of building's façade and land use information are some goals of the multipurpose cadastre. Both measurements are commonly performed manually, which would be extremely laborious, thus highlighting a potential application of this system. Furthermore, the acquired images combined to the 3D models generated can be considered a complete documentation of the environments with geometric and radiometric information.

Based on our results, we can identify some drawbacks related to the PMTS that provide scope for future research. (1) The 8-s time gap between successive still images with the

Ricoh Theta S prevents the dynamic sequential acquisition of images, leading to the decision to use frame videos with lower resolution in our system. Therefore, this technique can be extended for other dual systems, such as the Samsung Gear 360, LG 360 and Nikon KeyMission 360. (2) Due to the image resolution, the accuracy of the 3D reconstruction is limited for an imaging range less than 5 m, especially in forest areas, and thus more strips are required to cover the area than for LASER scanner PMTS approaches. (3) Further investigations should be performed to evaluate the effects on PMTS navigation system data of using an integrated sensor orientation approach, with rigorous analysis of the stochastic properties of the data. (4) Further experiments considering larger test areas are also suggested. (5) Omnidirectional images do not follow a perspective geometry, which requires modifications of classic photogrammetric models, such as camera calibration, image matching and bundle adjustments. Therefore, the following sections presented some methodologies to adapt the classic photogrammetric models to omnidirectional geometry.

CHAPTER III

GEOMETRIC MODEL OF A DUAL-FISHEYE 360° IMAGING SYSTEM

3. SCOPE

Chapter 3 presents a rigorous technique for the calibration of the 360° dual imaging system (Ricoh Theta S) used in the PMTS presented in Chapter 2. The experiments are based on calibration with a bundle adjustment using the equidistant fisheye lens model and relative orientation stability constraints, such as distance, base elements and rotation matrix between individual cameras' coordinate systems. This chapter summarizes the geometric model and assessment of Ricoh Theta S presented by Campos et al. (2018b).

3.1 POLYDIOPTRIC SYSTEM CALIBRATION: OVERVIEW AND CHALLENGES

Polydioptric systems are composed of multiple cameras specially arranged in a structure to cover a full spherical field of view, aiming at 360° image generation from multiple frames (Maas, 2008). The capabilities of these systems, in terms of device developments and data processing technology, have been improved significantly in recent years, with lighter, smaller and low-cost approaches such as the Giroptic 360cam, Go Pro 360°, Kodak PixPro SP360, Ladybug5, Google Jump system and Ricoh Theta S. However, the use of 360° images in close range photogrammetry still requires some research in camera modelling and calibration, since most of these systems were not developed for metric purposes.

Many methods have been proposed to estimate interior orientation parameters (IOPs), varying according to image acquisition configuration, mathematical model (such as standard, extended and generic), statistical approach and level of automation. Luhmann et al. (2016) have discussed the emerging diversity of imaging systems with different internal geometry characteristics (especially non-traditional ones, such as polydioptric systems) that have required new camera calibration assessments to support applications in close range photogrammetry.

Two challenges can be pointed out in the calibration of polydioptric systems. First, cameras used in photogrammetry and computer vision usually follow the central perspective (pinhole) geometric model, which cannot be applied to cameras with fisheye lenses, which are extensively used in polydioptric systems (ABRAHAM and FORSTNER, 2005). Consequently, image acquisition does not follow the collinearity condition, requiring a suitable calibration model reflecting the behaviour of the fisheye lenses (SCHNEIDER et al., 2009; HUGHES et al., 2010).

Second, considering that these systems are composed of multiple cameras, relative orientation parameters (ROPs) need to be estimated to make feasible the use of a single set of exterior orientation parameters (EOPs) for the component images and to generate composite images. In these new commercial systems, cameras are arranged in the same compact and closed structure, which hinders direct ROP measurements and requires indirect estimation. Usually, a simultaneous camera calibration for IOP and ROP estimation, considering polydioptric systems' challenges, has not been considered in existing MATLAB toolboxes for fisheye lens calibration, such as those proposed by Scaramuzza et al. (2006), Kannala and Brandt (2006) and Bouguet (2002), or even in commercial software such as PhotoScan (AGISOFT, 2016) and PhotoModeler (EOS Systems, 2014).

The problem of simultaneous multiple-camera calibration has been already discussed in the literature (HE et al., 1993; KING et al., 1994; KING et al., 1995; ESPINHOSA et al., 2008; BAZAN et al., 2009; TOMMASELLI et al., 2010; TOMMASELLI et al., 2013 and LICHTI et al., 2015). He et al. (1993) developed a technique for stereo camera calibration by introducing a set of relative orientation constraints with equations for rotation stability and weighted constraints for the base distance, based on directly measured values. King (1994, 1995) also worked with techniques for a stereo camera calibration, either by modifying the collinearity equations or by introducing relative orientation constraints. Tommaselli et al. (2010) introduced a technique for dual oblique camera calibration that was based on the concept of stability of the relative orientation elements, but considering random variations in the relationships between the ROPs. This technique was used to accurately calibrate a pair of synchronized digital cameras to produce virtual images from two oblique frames (TOMMASELLI et al., 2013). A similar approach was used by Lichti et al. (2015), who presented a single-step calibration of a dual fluoroscopic imaging system using a self-calibrating bundle adjustment with geometric constraints, based on the stability of the relative orientation. The largest part of these works deals with stereo cameras and the generation of virtual images from oblique frames, which have a different geometry than polydioptric systems, especially those with fisheye lenses.

A study of the calibration of a polydioptric spherical imaging system (Ricoh Theta S) was performed by Aghayari et al. (2017). In their approach, IOPs and EOPs for each camera that comprises the system were independently estimated with a purposely designed calibration network. The ROPs were subsequently calculated from the EOP values of each image

pair. However, the stability of the relative orientation was not considered in their approach as an additional constraint. In this regard, this chapter presents a technique and experimental assessments for the calibration of the Ricoh Theta S, based on fisheye lens models and focusing on the use of stability constraints of ROPs between cameras. The set of IOPs and ROPs that best reproduces the bundle that generated the 360° image, together with a proper camera calibration field developed to simulate a 360° environment, were evaluated considering four different calibration approaches.

3.2 DUAL 360° IMAGING SYSTEM

Sensor modelling initially requires an understanding of the physical imaging system configuration. Therefore, this section presents the geometric configuration of the Ricoh Theta S and the fisheye lens mathematical model that best fits this system calibration. Technical information of the Ricoh Theta S was presented in Section 2.2.1. As initially mentioned in section 2.2.1, the Ricoh Theta S can provide a still image (in JPEG format) or a dual-fisheye video (as an MP4 file), which is deeply detailed in the Sections 3.2.1 and 3.2.2.

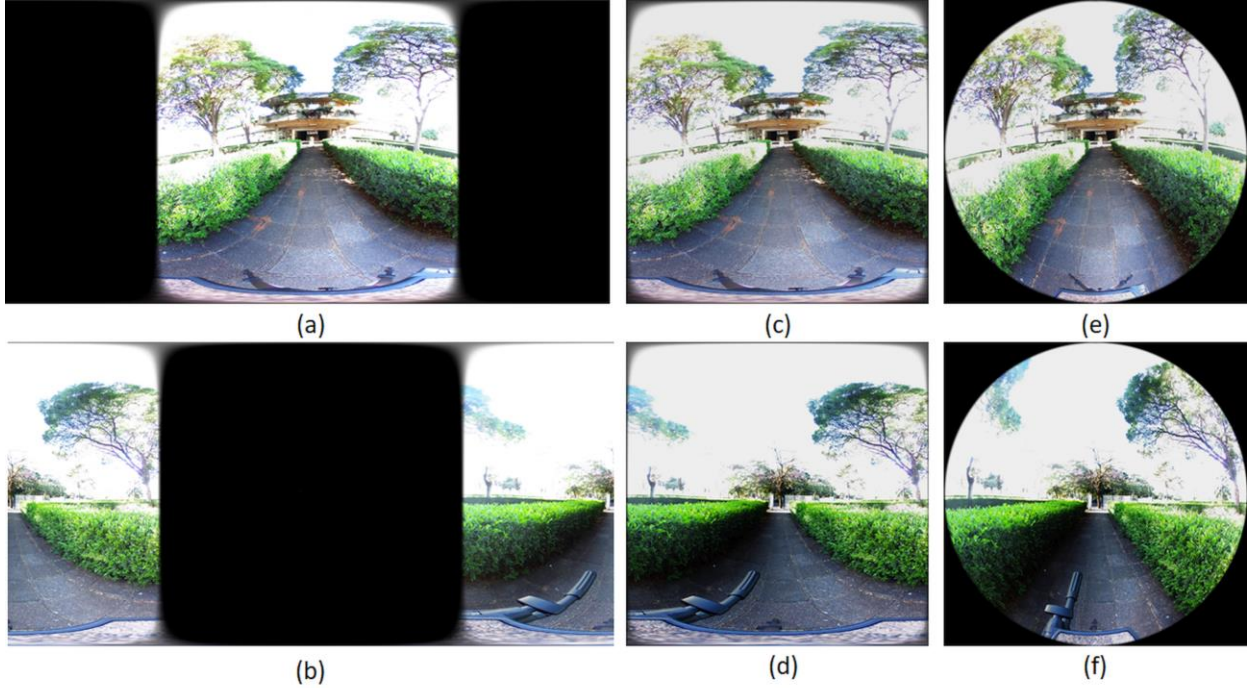
3.2.1 Ricoh Theta S Data – Still image

The raw image of the Ricoh Theta S (still image) is a panoramic image with dimensions of 5376 x 2688 pixels (2688 x 2688 pixels each image from sensor 1 and 2). The original dual-fisheye image is converted internally (in a process that takes 8 seconds) to a panoramic image in an equirectangular projection. The equirectangular image is a composite of the images from both sensors with a further tilt correction, which is achieved using the attitude data (image metadata) obtained from an internal gyroscope sensor. The image from Sensor 1 is positioned in the center of the post-processed image, while the image from Sensor 2 forms the left- and right-hand outer parts of the processed image (Figure 12).

In Figure 12.a light penetration to the lens of Sensor 2 was blocked, showing the position of the image coming from Sensor 1 in the processed image. Similarly, Figure 12.b shows the two halves of the image coming from Sensor 2 in the final equirectangular image, when light penetration to the lens of Sensor 1 was blocked. Considering this arrangement, the images coming from each sensor can be recovered. Figures 12.c and Figures 12.d show equirectangular images from Sensors 1 and 2, respectively (2688 x 2688 pixels), and thus the Ricoh Theta S can be treated as a two-head digital camera. A reverse transformation from an equirectangular projection image to

a fisheye image can be done for each sensor individually, as shown in Figures 12.e for Sensor 1 and Figures 12.f for Sensor 2.

Figure 12. Ricoh Theta S image composition for an outdoor scene: (a) and (b) illustrate the equirectangular image composition when one of the sensors is blocked from light; separate equirectangular images from Sensor 1 (c) and Sensor 2 (d); (e) and (f) show the recovered fisheye image for Sensors 1 and 2, respectively.



Source: Campos et al, (2018b)

Equirectangular images recovered from each sensor (Figure 12.c and Figure 12.d), with dimensions N_x and N_y , store pixels whose image coordinate system (column c and row l) have the upper-left corner as the origin. This image can be normalised to x_r and y_r coordinates, ranging between -1 and +1, considering the origin as the image center (Equation 8). Since an equirectangular image was created as a function of the fisheye image (hemispherical image), it is possible to assume polar coordinates (ρ, λ) for this equirectangular image, ranging from $-\pi$ to $+\pi$ for ρ and $-\pi/2$ to $+\pi/2$ for λ (Figure 13.a), obtained with Equation 9. These polar coordinates can be converted to 3D spherical coordinates (p_x, p_y, p_z) with the origin in the center of a hemisphere, using Equation 10 (BOURKE, 2016).

$$x_r = -\left(-1 + 2 * \frac{c}{N_x}\right); \quad y_r = \left(-1 + 2 * \frac{l}{N_y}\right) \quad (8)$$

$$\rho = \frac{\pi}{2} + \left(x_r * \frac{\pi}{2}\right); \quad \lambda = \left(y_r * \frac{\pi}{2}\right) \quad (9)$$

$$p_x = \cos(\lambda) * \cos(\rho); \quad p_y = \cos(\lambda) * \sin(\rho); \quad p_z = \sin(\lambda) \quad (10)$$

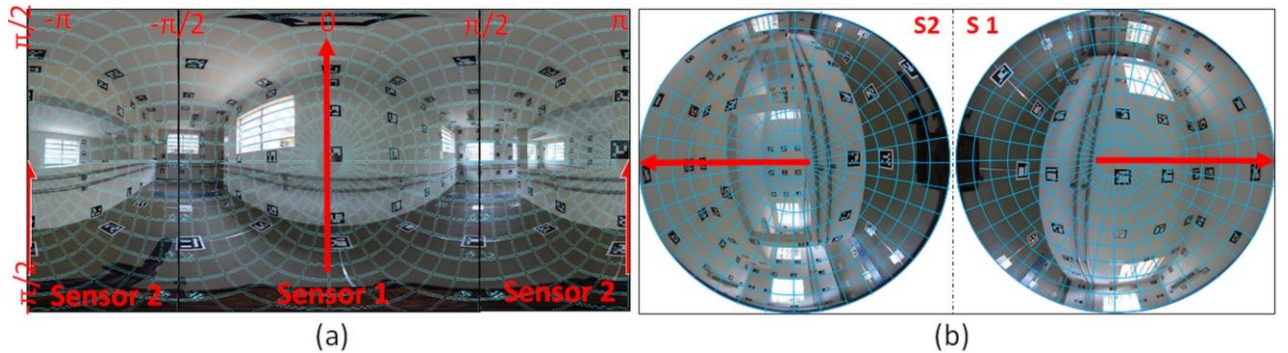
Since 3D coordinates over the sphere are known (p_x, p_y, p_z), it is possible to estimate the spherical coordinates (θ, r_s), as presented in Equation 11, in which r_s is the spatial distance between the sphere center and point position, considering the Field of View (FOV) of the camera in radians, as stated by Bourke (2016). Then, new coordinates ($x_{\text{fish}}, y_{\text{fish}}$) of the points from 3D semi-sphere to 2D space can be calculated with Equation 12. These new point positions are not integer values, therefore pixels resampling is required. A fisheye image can be obtained using interpolation methods such as nearest neighbour or bilinear interpolation.

$$\theta = \tan^{-1} \frac{p_z}{p_x}; \quad r_s = \frac{2}{\text{FOV}} * \tan^{-1} \frac{\sqrt{p_x^2 + p_z^2}}{p_y} \quad (11)$$

$$x_{\text{fish}} = r_s * \cos \theta; \quad y_{\text{fish}} = r_s * \sin \theta \quad (12)$$

Finally, the dual-fisheye image composition can be restored in a reverse process, including tilt effect that have been removed and considering the 90° sensors rotation, as shown in Figure 13.b. Tilt parameters are available in the image metadata, however the Ricoh Theta S internal correction process is unknown as well as the accuracy of the available tilt angles values.

Figure 13. (a) Ricoh Theta S still-image projection, and (b) transformed to a dual-fisheye configuration. Blue lines indicate the projection of rays in each case.



Source: Campos et al, (2018b)

3.2.2 Ricoh Theta S Data – Dual-fisheye video

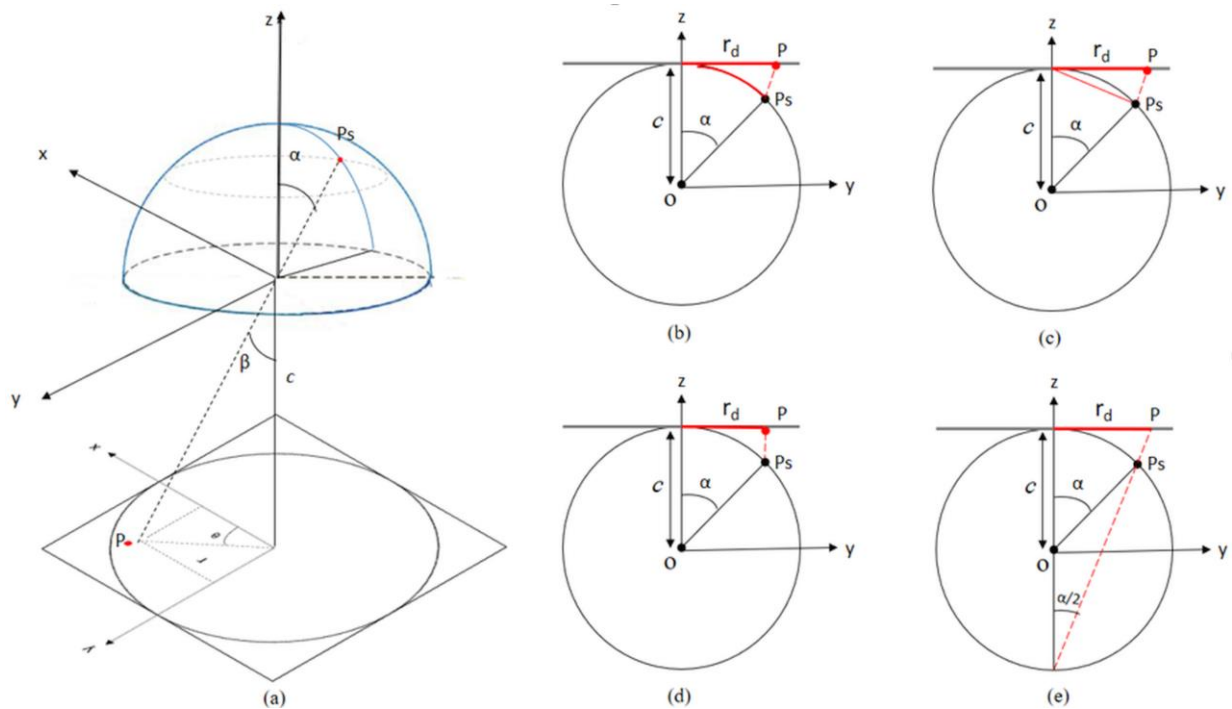
Despite the fact that the original fisheye image projection might be recoverable, uncertainties concerning the accuracy of the tilt correction and image stitching still exist. As recently discussed by Aghayari et al. (2017), these errors are not easily modelled since stitching parameters are unknown, resulting in further errors (for instance, in radial distortion estimation). Therefore, considering that the main objective of the calibration process is to accurately determine the internal characteristics of the camera, these uncertainties can affect the IOP estimation and consequently the successful analytical reconstruction of the bundle that generated the image. Furthermore, another disadvantage of the dual-fisheye image composition (still image) is the 8-second time gap between successive image shots due to the Ricoh Theta S internal stitching process. This time delay for image acquisition might not be suitable for dynamic photogrammetric applications, such as in mobile mapping, hindering appropriate image location and reconstruction by a bundle adjustment.

Considering these problems with still-image capture, the Ricoh Theta S video mode can be a practical alternative for many applications, since dual-fisheye video is stored in the camera without any internal processing. Furthermore, data acquisition has a frame rate of 29 dual frames per second. Video can be converted to individual frames, similar to those in Figure 13.b, but with dimensions of 1920 x 1080 pixels and a 0.005 mm pixel size. The frames from each sensor (960 x 1080 pixels) can then be recovered, as described for still images. The drawback is the low resolution of the frames from video when compared to still images, as presented in Table 1 (Section 2.2.1); but the advantage is the high frequency of image capture. Therefore, the Ricoh Theta S frame video of each sensor separately was used as input data in all experimental processes and it will henceforth be called “images”.

Concerning the fisheye lens mathematical model that best describes the relationship between image and ground coordinates for the Ricoh Theta S, the four well-known fisheye lens projections (stereographic, equidistant, orthogonal and equi-solid-angle) (HUGHES et al., 2010) were tested in a preliminary study. The radial distances (r_d) of a set of points in the 360° calibration field (see the following section) were calculated for each projection and compared to the radial distance measured in the Ricoh Theta S image. The value of r_d is basically a function of the principal distance and the incident angle α . The nominal focal length (Table 1; Section 2.2.1) was used for radial distance estimation, since the calibrated principal distance was not available

before the calibration process. The incident angle was calculated considering the targets with known coordinates in the calibration field and the camera position and orientation. Figure 14 depicts the principles of four fisheye projection models: (b) equidistant; (c) equi-solid-angle; (d) orthogonal; and (e) stereographic. Figure 14.a show the principal distance (c), incidence angle (α), emergent angle (β) and radial distance (r_d).

Figure 14. (a) Geometry of image acquisition with fisheye lens cameras. (a) The projection of a point P_S (X_c, Y_c, Z_c) to the hemisphere (in black lines). The reprojection of the point P_S to the image plane (in red) with: (b) equidistant, (c) equi-solid-angle, (d) orthogonal, and (e) stereographic projections (adapted from RAY, 2001 and HUGHES et al., 2010).

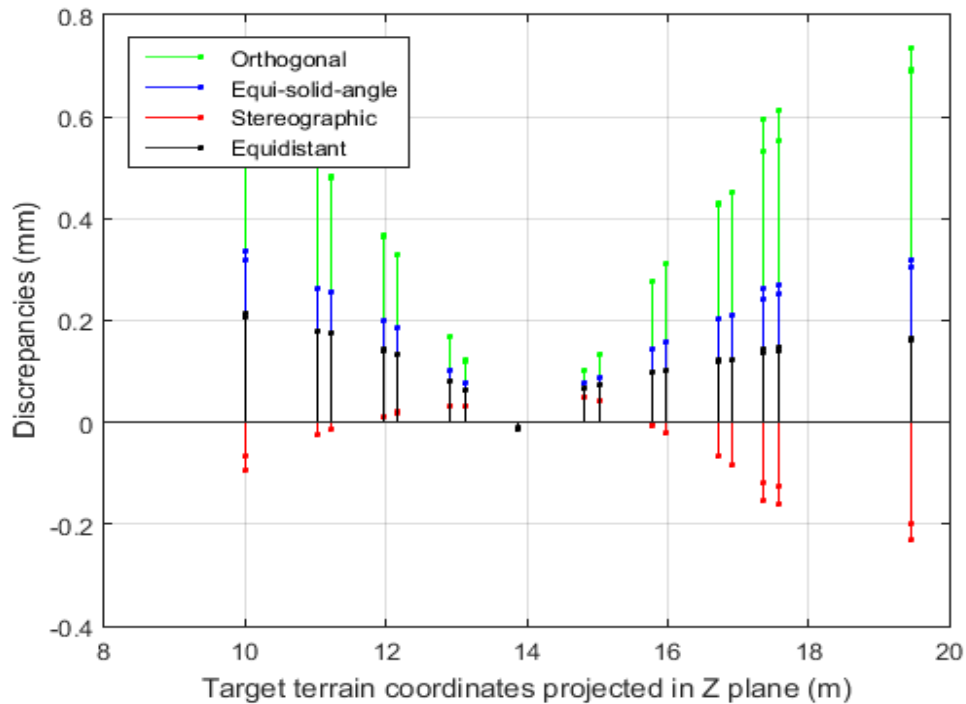


Source: Campos et al, (2018b)

The equidistant projection presented better results for the Ricoh Theta S with a standard deviation (r) of radial distance discrepancies of 0.05 mm and a symmetrical distribution, as illustrated by Figure 15, where Z equal to 14 m corresponds to the position of the camera in the calibration field local system, during image acquisition for this experiment, as better discussed in Section 3.3. This outcome is in line with the developer's specification (THETA DEVELOPERS, 2017) and other works, such as Abraham and Forstner (2005), which indicates that an equidistant projection is followed by most fisheye lenses. Less satisfactory results were obtained with the orthogonal ($r = 0.22$ mm), equi-solid-angle ($r = 0.10$ mm) and stereographic ($r = 0.06$ mm)

projections. Figure 15 presents the discrepancies between the radial distance calculated with stereographic (red), equidistant (black), orthogonal (green) and equi-solid-angle (blue) projection and the radial distance extracted from Ricoh Theta S image, considering a set of points from the 360° calibration field (Section 3.3).

Figure 15 - Discrepancies between the radial distance computed with the fisheye lens projection and in Ricoh Theta S image.

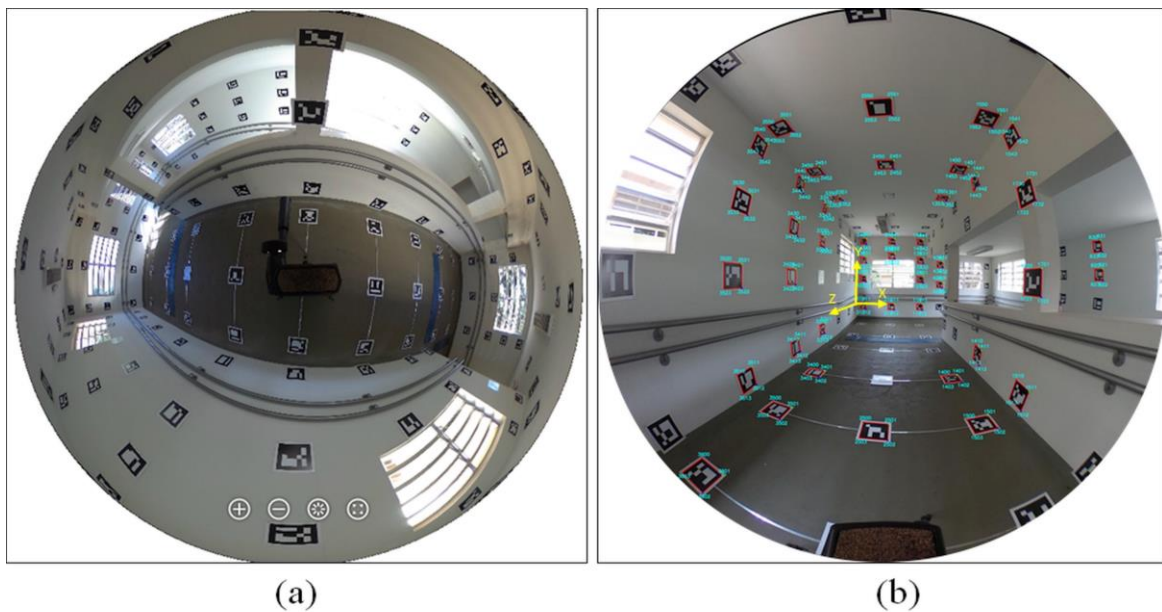


Observing the Ricoh Theta S images carefully (for example, Figure 13.b), it is possible to notice that the field of view of each camera covers more than 180° (the same targets appear in images of both sensors). According to Ray (2001), this model can be named as hyper-hemispherical, when $\alpha \sim 110^\circ$. When $|\alpha| > 90^\circ$, Z has a negative value and a double mapping problem occurs, which cannot be modelled with the geometry presented in Figure 14.a, considering a spherical image surface. This limitation in the mathematical model proposed (see the “Experimental Assessment”- Section 3.4) can be circumvented by considering a hyperbolic image surface. Therefore, the image measurements of these points ($|\alpha| > 90^\circ$) were not considered in the calibration process, and the study of a suitable treatment of this geometry is left for future work.

3.3 360° CAMERA CALIBRATION FIELD

The calibration of spherical cameras requires that object points are evenly distributed around the camera (to produce suitable geometry) and also that they are accurately measured in the images. Therefore, a proper 360° camera calibration field is required. A critical process in camera calibration is the extraction of accurate image point coordinates. Automatic measurements enable image coordinates to be determined with sub-pixel precision, as was shown by Reiss and Tommaselli (2011), Galo et al. (2012); Fraser (2013) and Silva et al. (2014). In this regard, this 360° calibration field was composed of ARUCO coded targets (GARRIDO-JURADO et al., 2014).

Figure 16. (a) 360° camera calibration field. (b) ARUCO target detection with a Ricoh Theta S image.



Source: Campos et al, (2018b)

A sample of 160 ARUCO targets, with dimensions of 21 cm x 21 cm, was selected to populate the calibration field. The distribution and the choice of targets were planned considering a right-handed system (Figure 16.b) and target numeration related to its ground positions (X , Z , Y). The detection process returned the identification (number) and the position of the four corners of each ARUCO target in the image (MUÑOZ-SALINAS, 2012), as shown in Figure 16.b, providing image coordinates with sub-pixel accuracy (SILVA et al., 2014). Another option to avoid systematic errors caused by edge detection is to use the centroid of each ARUCO

target computed from the measured corners. Figure 16 shows spherical images of the 360° calibration field (a) and the ARUCO targets detected in the Ricoh Theta S fisheye image (b). This field is about 8 m long, 4 m wide and 3 m high, providing a significant range variation to reduce correlations between EOPs and IOPs (c and Z_0) in the calibration with the bundle adjustment.

Ground coordinates of the corners of each ARUCO target were previously estimated. However, they could also be determined as unknowns in a self-calibration. To achieve an acceptable level of accuracy, these ground coordinates were determined using photogrammetric methods with a high resolution camera, a Sony Nex (6000 x 4000 pixels). This camera, previously calibrated with a 15.86 mm principal distance, was used, generating a set of 35 images from the calibration field. With this set, the 3D coordinates of target corners were determined with a bundle adjustment in an arbitrary local reference system. Ground control points were measured with a total station, and some distances were also measured using a precision calliper with 0.1 mm precision. The bundle adjustment resulted in ground coordinates with 3 mm accuracy, which could be used as ground control in the Ricoh Theta S calibration.

3.4 EXPERIMENTAL ASSESSMENT FOR RICOH THETA S

This section summarizes the data set used for experimental assessments (section 3.4.1) and the mathematical model that best fits for Ricoh Theta S (section 3.4.2), detailing the experimental assessments performed (section 3.4.2.1 to section 3.4.2.3) to determine a methodology for polydioptric systems calibration. The results following the calibration process describes in this section are presented in Section 3.5.

3.4.1 Data set

Ricoh Theta S data (still image or frame video) needs to be prepared for camera calibration, using the process described in Chapter 2. Because of the uncertainties in the still image processing, as previously described, dual-fisheye video was utilized as input data in the camera calibration process. The methodology described using frame video can be extended to be used with still images, as long as raw data is recovered (Chapter 2), and also for another dual-cameras model such as the Samsung Gear 360.

A total of twenty (20) Ricoh Theta frame video clips, with an approximate pixel size in object space units ranging between 4 mm and 6 mm, were acquired in the 360° calibration

field by remote shooting (Wi-Fi access), using a levelled tripod specially positioned over targets with known ground coordinates. Convergent images and rotated images around the z-axis, with kappa varying between 0° , -90° , 90° and 180° , were taken to avoid correlations between interior and exterior orientation parameters (KENEFICK et al., 1972).

After data acquisition, fisheye frame video sequences for each sensor were recovered (images) for image measurements. Twenty (20) images were used for camera calibration of Sensor 1 and sixteen (16) images for Sensor 2. As the frame video sequences were taken simultaneously for both sensors, some images had a non-uniform target distribution, and for that reason they were excluded, which explains the use of different sets of images for the calibration of sensor 1 and 2. From this set of images, thirteen (13) are image pairs acquired at the same instance, which will be considered for the ROP estimation.

As presented in Chapter 3, ARUCO targets were located and automatically measured in the images with subpixel accuracy; thus, standard deviations of 0.5 pixel for image observations in both components (c , l) were assumed. In this case, all image points have known ground coordinates (X , Y , Z) with a 0.003 m accuracy. Therefore, a standard deviation of 0.005 m was set in the bundle adjustment for the ground coordinates. A self-calibration could also be performed, which requires at least seven constraints to define the object reference frame, without any control points. However, since the camera's field of view and scale variation in omnidirectional images are huge, the best solution was obtained considering the calibration field ground coordinates as weighted constraints. Finally, initial values for EOP were estimated based on camera position in the calibration field with a standard deviation of 0.5 m for camera position (X_0 , Y_0 , Z_0) and 10° for attitude parameters (ω , ϕ , κ).

3.4.2 Mathematical Model

The calibration processes were implemented based on the unified approach to least squares (combined adjustment with additional constraints, as presented by Mikhail and Ackerman, 1976), considering as mathematical model the equidistant fisheye lens equations combined with the well-known Conrady-Brown distortion model (BROWN, 1971). In equidistant projection, the radial distance (r_d) is given by the product of the principal distance (c) and the incident angle α ($r_d = c \cdot \alpha$). The incident angle α is computed from the ratio of the planimetric distance and the Z coordinate in the object space. The image coordinates of a point can be

calculated decomposing r_d from the equidistant projection function in the components x' and y' , as expressed in Equation 13. Equidistant equations are composed by adjusted observations (L_a) and unknowns (X_a) in the same model, represented by $F(L_a, X_a) = 0$. Parameters estimation is done with Least Squares combined model with constraints (WELLS and KRAKIWSKY, 1971; MIKHAIL and ACKERMAN, 1976). This approach was implemented in C/C++ language on the CMC (Calibration of Multiple Cameras) program, based on Tommaselli et al., (2013) and Marcato-Junior et al., (2015).

$$\begin{aligned} x' - x_0 + \Delta_x - c \cdot \frac{X_c}{\sqrt{X_c^2 + Y_c^2}} \cdot \arctan\left(\frac{\sqrt{X_c^2 + Y_c^2}}{Z_c}\right) &= 0 \\ y' - y_0 + \Delta_y - c \cdot \frac{Y_c}{\sqrt{X_c^2 + Y_c^2}} \cdot \arctan\left(\frac{\sqrt{X_c^2 + Y_c^2}}{Z_c}\right) &= 0 \end{aligned} \quad (13)$$

Since the purpose of camera calibration is to estimate the IOP, such as principal distance (c), principal point coordinates (x_0, y_0) and coefficients for systematic errors correction (mainly lens distortion), equidistant projections can be associated with a distortion model ($\Delta x, \Delta y$) that is composed of polynomials describing symmetric radial distortion as a function of coefficients, k_1, k_2 and k_3 (BROWN, 1971), decentring distortion, whose coefficients are p_1 and p_2 (CONRADY, 1919; BROWN, 1971) and affinity distortion parameters A and B (HABIB and MORGAN, 2003). Equation 14 presents the distortion model, in which x_f ($x' - x_0$) and y_f ($y' - y_0$) are the measured image point coordinates expressed in a photogrammetric reference system whose origin is the camera perspective center and r is the planimetric radial distance between the principal point and the point in the image (Equation 15).

$$\begin{aligned} \Delta x &= x_f (k_1 r^2 + k_2 r^4 + k_3 r^6) + P_1 (r^2 + 2x_f^2) + 2P_2 x_f y_f + Ax_f + By_f \\ \Delta y &= y_f (k_1 r^2 + k_2 r^4 + k_3 r^6) + P_2 (r^2 + 2y_f^2) + 2P_1 x_f y_f + Ay_f \end{aligned} \quad (14)$$

$$r = \sqrt{x_f^2 + y_f^2} \quad (15)$$

In a first trial, the significance of each IOP described in the distortion model presented in Equation 14 was analyzed. A simple assessment of IOP's significance can be done

comparing the parameter magnitude with its standard deviation. As briefly presented by Chandler, Fryer and Jack (2005), parameters can be removed if the estimated value is less than the respective standard deviation. The decision whether the estimated IOP are significantly or not can also be done considering other methods based on statistic tests, such as t-student (HABIB and MORGAN, 2005). In preliminary calibration tests, this behaviour was noticed with affinity distortion parameters (A and B) for Sensor 1 and Sensor 2. The effects of these parameters in the image coordinates were also compared with the image measurement error. Affinity distortion effects are smaller than the image measurement error, which is observed in most modern digital cameras. Consequently, A and B parameters were removed from the set of IOP.

Considering this well-known mathematical model, four calibration approaches (experiments A to D) were performed, aiming at a Ricoh Theta S geometric assessment. These approaches were implemented in the in-house developed software CMC. An independent IOP determination was done for each camera (single camera calibration), then the relative orientation between them was computed as a function of two sets of EOP estimated in the bundle adjustments (Exp. A). In a second approach, IOP were estimated in a simultaneous camera calibration process, but ROP were still determined in a further step (Exp. B). Since cameras are arranged in a stable and rigid structure, it is possible to assume that a stability between cameras exists during image acquisition; therefore, two trials with simultaneous bundle adjustment with relative orientation stability constraints were performed. In the first one, the stability of the relative rotation matrix and the distance between the camera's perspective center were introduced as constraints (Exp. C), and then the relative rotation matrix and base elements were considered as stability constraints (Exp. D). The results are presented and discussed in Section 3.5.

3.4.2.1 Experiments A and B - Single and Simultaneous camera calibration without relative orientation constraints

In Experiment A, a set of IOP was obtained for each camera separately in a single calibration process, considering the mathematical model described in Section 3.4.2, without relative orientation constraints. The estimation of relative orientation parameters between sensors was done later, as a second step in this system calibration process.

The angles of relative orientation ($\Delta\omega$, $\Delta\phi$, $\Delta\kappa$), the distance (D_{PC}) between both cameras' perspective centers PC (X_0 , Y_0 , Z_0) and base elements (B_x , B_y , B_z) were computed from

the estimated values of the EOP of each camera station (Equation 16 to Equation 19), which were estimated considering the individual bundle adjustment for each camera. The relative orientation matrix (R^R), composed of $\Delta\kappa$, $\Delta\phi$, $\Delta\omega$, rotations, was obtained considering the rotation matrix of Sensor 1 (R^1 ; κ^1 , ϕ^1 , ω^1) and the rotation matrix of Sensor 2 (R^2 ; κ^2 , ϕ^2 , ω^2) for images acquired in the same instance (i), as expressed in Equation 16 (BAZAN et al., 2009). The relative orientation angles, $\Delta\omega$, $\Delta\phi$ and $\Delta\kappa$ can then be estimated individually from R^R with Equation 19.

$$R^R = R^1 (R^2)^{-1} \quad (16)$$

$$D_{PC_i} = \sqrt{(X_0^1 - X_0^2)^2 + (Y_0^1 - Y_0^2)^2 + (Z_0^1 - Z_0^2)^2} \quad (17)$$

$$[B_X \ B_Y \ B_Z]^T = R^1 [X_0^1 - X_0^2 \ Y_0^1 - Y_0^2 \ Z_0^1 - Z_0^2]^T \quad (18)$$

$$\Delta\phi = \text{asin}[R_{(3,1)}^R]; \quad \Delta\omega = \text{asin}\left[\frac{-R_{(3,2)}^R}{\cos \Delta\phi}\right]; \quad \Delta\kappa = \text{acos}\left[\frac{R_{(1,1)}^R}{\cos \Delta\phi}\right] \quad (19)$$

In Experiment B, IOP, EOP for both cameras and object coordinates of photogrammetric points were simultaneously estimated in the same least squares combined adjustment process but, still without additional relative orientation constraints. ROP were derived from the set of EOP as described in Experiment A with Equation 16 to Equation 19. The main advantage of this second approach is that all bundles of rays that formed the 360° image are considered in the same process, but no constraints on the relative orientation between cameras were introduced. To introduce such constraints, several options are available.

The ROP could be introduced either as unknowns or as absolute constraints applied based on measured values. However, measuring these values directly in the Ricoh Theta S is not feasible even if the camera is disassembled, because the accuracy of these measurements will be under the image measurement precision. Furthermore, absolute constraints do not consider the natural uncertainties when estimating PC coordinates. An alternative solution to this problem, which is assessed in experiments C and D, is the use of Relative Orientation Stability Constraints (ROSC) for the calibration of this poly-dioptric system. Since the cameras are in the same structure, it is possible to assume that a stability relation between the cameras' PCs exist during a sequential image acquisition, as applied for perspective dual-cameras by Bazan et al. (2009); Tommaselli et al. (2013) and Lichti et al. (2015). In this regard, the principal motivation for the use of this method, which applies stability constraints in the bundle adjustment based on relative orientation elements, is the possibility of accepting some random variations for these elements, as

described in Tommaselli et al. (2013). In this way, the method proposed did not introduce a new weighted constraint or new parameters as unknowns. Also, implementing constraints of stability of relative orientation can be done by summing matrices to the existing normal equations.

3.4.2.2 Experiment C - Simultaneous camera calibration with PCs' distance and relative rotation matrix as stability constraint.

In experiment C, a simultaneous bundle adjustment with relative orientation stability constraints was used, considering that the relative rotation matrix and distance between PCs are stable, but small variations are admitted as random variables. Therefore, new absolute constraints were not introduced in the bundle adjustment, avoiding the need for direct measurements. This approach is based on the condition that R_R (Equation 20) and D_{PC} (Equation 21) remain stable during the same image acquisition process. Therefore, the values calculated by Equation 20 and Equation 21, for instance (i) and (i+1), can be considered similar, in which v is a residual in the constraint equation (BAZAN et al., 2009 and TOMMASELLI et al., 2013).

$$R_i^R - R_{i+1}^R = 0 + v \quad (20)$$

$$D_{PC(i)}^2 - D_{PC(i+1)}^2 = 0 + v \quad (21)$$

As described in Tommaselli et al. (2013) four constraint equations can be developed in the function of these stability conditions, which were applied in experiment C. The relation between the relative orientation matrix presented in Equations 21 gives rise to three constraints (Equation 22 to Equation 24), considering the rotation matrices of each sensor (R^1 and R^2). The fourth constraint equation is the stability between the PCs' distance (Equation 25), which is derived from the squared differences in Equation 21. The zero difference is treated as a pseudo-observation with a certain standard deviation, which is established empirically, and which is related to the errors in the estimate of EOP. The smaller these standard deviations are, the higher the residuals will be, so there is a compromise value that is acceptable. The constraints standard deviations (σ) assumed for this experiment were 0.005° for each R^R elements ($\Delta\omega$, $\Delta\phi$, $\Delta\kappa$), which correspond to $18''$, and 0.1 mm for D_{PC} , considering as an initial value the results obtained in Experiment B. These σ values were estimated empirically after a set of weight combinations. The

weights to be assumed in the Least Squares Adjustment were estimated by error propagation in Equation 22 to Equation 25 from the values admitted for the standard deviations of the ROP. Note that Equations 22 to 25 functionally connect the EOP of both cameras in two consecutive different stations.

$$G_1 = (R_{21}^1 R_{11}^2 + R_{22}^1 R_{12}^2 + R_{23}^1 R_{13}^2)^i - (R_{21}^1 R_{11}^2 + R_{22}^1 R_{12}^2 + R_{23}^1 R_{13}^2)^{i+1} = 0 + v_1 \quad (22)$$

$$G_2 = (R_{31}^1 R_{11}^2 + R_{32}^1 R_{12}^2 + R_{33}^1 R_{13}^2)^i - (R_{31}^1 R_{11}^2 + R_{32}^1 R_{12}^2 + R_{33}^1 R_{13}^2)^{i+1} = 0 + v_2 \quad (23)$$

$$G_3 = (R_{31}^1 R_{21}^2 + R_{32}^1 R_{22}^2 + R_{33}^1 R_{23}^2)^i - (R_{31}^1 R_{21}^2 + R_{32}^1 R_{22}^2 + R_{33}^1 R_{23}^2)^{i+1} = 0 + v_3 \quad (24)$$

$$G_4 = [(X_0^2 - X_0^1)^2 + (Y_0^2 - Y_0^1)^2 + (Z_0^2 - Z_0^1)^2]^i - [(X_0^2 - X_0^1)^2 + (Y_0^2 - Y_0^1)^2 + (Z_0^2 - Z_0^1)^2]^{i+1} = 0 + v_4 \quad (25)$$

3.4.2.3 Experiment D - Simultaneous camera calibration with base elements and relative rotation matrix as stability constraint.

The constraints between cameras PCs can also be expressed in terms of base components (B_x, B_y, B_z), which can be derived from EOP, as shown in Equation 18. Then, these elements can also be considered stable during the acquisition (Equation 26). Therefore, Experiment D is based on C, but considering that Equation 25 is replaced by three equations, based on Equation 27 to 29 for B_x, B_y, B_z elements respectively.

Similar to Experiments C, the standard deviations for the constraints were computed by covariance propagation assuming standard deviations for $\Delta\omega, \Delta\phi, \Delta\kappa$ with a value of 18". The variation allowed (v) between cameras base components during a sequential image acquisition was 0.001 mm. This value was empirically defined after several trials with bundle adjustment.

$$([B_x \ B_y \ B_z]^T)^i - ([B_x \ B_y \ B_z]^T)^{i+1} = 0 + v \quad (26)$$

in which,

$$B_x^{(i)} = [R_{11}^{1(i)}(X_0^{2(i)} - X_0^{1(i)}) + R_{12}^{1(i)}(Y_0^{2(i)} - Y_0^{1(i)}) + R_{13}^{1(i)}(Z_0^{2(i)} - Z_0^{1(i)})] \quad (27)$$

$$B_y^{(i)} = [R_{21}^{1(i)}(X_0^{2(i)} - X_0^{1(i)}) + R_{22}^{1(i)}(Y_0^{2(i)} - Y_0^{1(i)}) + R_{23}^{1(i)}(Z_0^{2(i)} - Z_0^{1(i)})] \quad (28)$$

$$B_z^{(i)} = [R_{31}^{1(i)}(X_0^{2(i)} - X_0^{1(i)}) + R_{32}^{1(i)}(Y_0^{2(i)} - Y_0^{1(i)}) + R_{33}^{1(i)}(Z_0^{2(i)} - Z_0^{1(i)})] \quad (29)$$

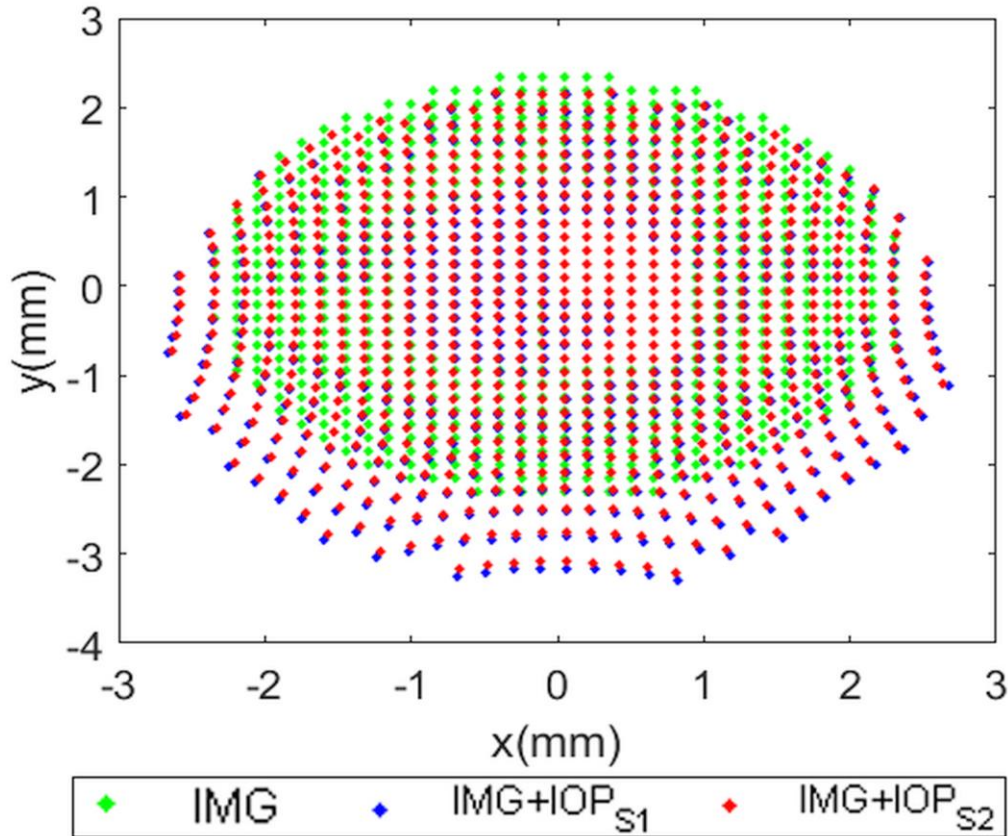
3.5 RESULTS AND DISCUSSION

Results and discussion about the methodology proposed are summarized by a preliminary study that highlight the importance of the estimation of two set of IOP for polydioptric system (Section 3.5.1), the performance of the four calibration approaches tested on IOP estimation (Section 3.5.2), the accuracy of ROP estimation (Section 3.5.3), and finally the accuracy assessment of Ricoh Theta S camera calibration (Section 3.5.4).

3.5.1 Preliminary analysis

A preliminary experiment was performed to assess the similarity between IOP of both sensors. The ZROT method (HABIB et al., 2006) enables the analysis of similarity between IOP obtained for Sensor 1 (S1) and Sensor 2 (S2), reproducing the bundle of rays that should be generated with each of these two sets of IOP. A synthetic regular grid with dimensions of the spherical image (960 x 1080) was generated and presented in Figure 17 in green. Then, the two sets of IOP estimated in Experiment A (Table 8 and Table 9) are applied in simulated grid points to the corrected principal point position and lens distortions. The effect of these corrections is shown in blue for Sensor 1 and red for Sensor 2. The degree of similarity is given by the RMSE of discrepancies between the corrected distortion grid for each sensor (blue x red). The RMSE obtained was 0.0144 mm or 2.88 pixels, which represents six times the standard deviation of the image coordinate measurement. Therefore, IOP for sensor 1 and sensor 2 cannot be considered similar, demonstrating the relevance of estimating both sets of IOP for poly-dioptic systems as the Ricoh Theta S, aiming at photogrammetric applications.

Figure 17. ZROT method results for the two sets of IOPs estimated in Experiment A. Synthetic regular grid of the spherical image's 960 x 1080 pixels shown in green. Principal point and lens distortion corrections shown in blue for Sensor 1 and red for Sensor 2.



Source: Campos et al, (2018b)

3.5.2 Calibration with Bundle Adjustment

Since, the ZROT experiment showed the relevance of estimating both sets of IOP for the Ricoh Theta S, the IOP and EOP results of Experiments A to D will be presented in this Section considering two sets of parameters. Table 8 shows the IOP c , x_0 and y_0 estimated for these experiments and the respective estimated standard deviations (σ). Standard deviations of the estimated principal distance and estimated principal point coordinates, for all approaches, were less than 1 pixel (1 pixel = 0.005 mm) for both sensors, as recommended for close range photogrammetry applications (BRYAN et al., 2009; CAMPOS et al., 2015). These results have an accuracy comparable to the conventional camera calibration with perspective cameras, which shows the importance of a correct mathematical model to describe the camera's inner geometry, mainly for non-conventional sensors. Then it is possible to consider that these four approaches are

viable. However better results in terms of estimated standard deviations (σ) for the IOP components were obtained when a simultaneous bundle adjustment with relative orientation stability constraints was performed. For instance, in Experiment D, the estimated σ for principal distance and for principal point coordinates, x_0 and y_0 had an improvement in comparison to Experiment A of 20% on average for Sensor 1 and 30% on average for Sensor 2.

Table 8. Estimated values of principal distance (c) and principal point coordinates (x_0, y_0) and standard deviations for Sensor 1 and Sensor 2, obtained in experiments A to D.

Exp.	SENSOR 1			SENSOR 2		
	c (mm)	x_0 (mm)	y_0 (mm)	c (mm)	x_0 (mm)	y_0 (mm)
A	1.4310	-0.0016	0.3013	1.4303	0.0042	0.2942
	± 0.0012	± 0.0010	± 0.0011	± 0.0019	± 0.0014	± 0.0015
B	1.4310	-0.0016	0.3011	1.4298	0.0039	0.2946
	± 0.0010	± 0.0009	± 0.0010	± 0.0015	± 0.0012	± 0.0010
C	1.4308	-0.0019	0.3014	1.4299	0.0041	0.2936
	± 0.009	± 0.0008	± 0.0009	± 0.0013	± 0.0009	± 0.0010
D	1.4306	-0.0018	0.3017	1.4295	0.0042	0.2934
	± 0.0009	± 0.0007	± 0.0009	± 0.0012	± 0.0008	± 0.0010

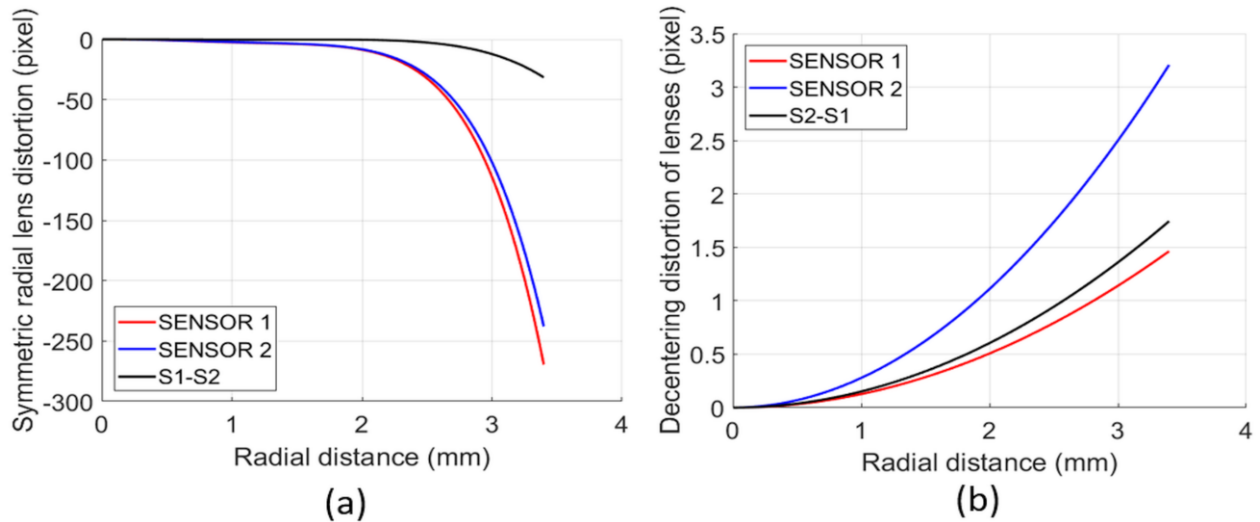
Regarding lens distortion, the differences between values estimated in Experiments B, C and D (simultaneous bundle adjustment) are smaller than the estimated standard deviation (σ) for these coefficients. Comparing Experiment A (single camera calibration) to the simultaneous camera calibration approaches, the variation between lens distortion coefficients are smaller than 2σ . The effect of the variations in coefficients values is less than one pixel in the image border. For this reason, the estimated parameter values, k_1 , k_2 , k_3 , p_1 and p_2 can be considered similar. Considering this similarity, the lens distortion behaviour is presented in Table 9 expressed by the symmetric radial distortion coefficients, k_1 , k_2 and k_3 , and the decentring distortion coefficients, p_1 and p_2 , obtained in Experiment D, which showed the least standard deviation in the parameters estimation. Figure 18 shows the effect of symmetric radial lens distortion (a) and decentring lens distortion (b) in pixels for Ricoh Theta S image according to the radial distance in mm, in Sensor 1 (S1 in red) and Sensor 2 (S2 in blue) The difference between the effect of lens distortions in Sensor 1 and Sensor 2 (S1-S2 in black) was computed and presented in Figure 18. Especially regarding k_1 parameter, it is noticed that this parameter not only model lens distortion, but also assimilate part of the residuals of the adjustment due to the equidistant model used. Ricoh Theta S lenses do not follow equidistant behavior in the image border ($\alpha > 90^\circ$), which is compensated by k_1 parameter. This Ricoh Theta S lenses behavior is better discussed in Section 5.2.2.1 (Figure 32).

Similarly to the IOP (Table 8 and Table 9), improvements can also be noticed in the EOP standard deviations when a simultaneous bundle adjustment approach was applied, which were reduced with the use of stability constraints, mainly for the rotation matrix angles (ω , ϕ , κ). Table 10 presents the average values of the EOP standard deviations estimated for each image obtained from covariance matrix of experiments A to D.

Table 9. Estimated values and standard variations of lenses distortions parameters for both sensors.

Distortion coefficients	Sensor 1		Sensor 2	
	Value	σ	Value	σ
$k_1(mm^{-2})$	-18.4 E-3	7.0 E-4	-16.0 E-3	1.2 E-3
$k_2(mm^{-4})$	7.50 E-3	3.0 E-4	6.26 E-3	5.0 E-4
$k_3(mm^{-6})$	-1.39E-3	4.73E-5	-1.19 E-3	6.63 E-5
$p_1(mm^{-1})$	-2.37E-4	5.07E-5	-7.13 E-4	6.12 E-5
$p_2(mm^{-1})$	6.07E-6	5.98E-5	-1.29 E-4	7.32 E-5

Figure 18. (a) Effects of symmetric radial lens distortion, and (b) decentring lens distortion for the lenses of the Ricoh Theta S camera for Sensor 1 (S1 in red), Sensor 2 (S2 in blue) and the difference between both sensors (S1-S2 in black).



Source: Campos et al, (2018b)

Table 10. Average values of the estimated standard deviations of the EOP: omega (σ_ω), phi (σ_ϕ) and kappa (σ_κ) in degrees and camera perspective center (σ_{x0} , σ_{y0} , σ_{z0}) in metres.

Exp.	σ_ω (°)	σ_ϕ (°)	σ_κ (°)	σ_{x0} (mm)	σ_{y0} (mm)	σ_{z0} (mm)
A	0.4730	0.0526	0.4733	1.4	1.5	1.3
B	0.4167	0.0464	0.4168	1.1	1.2	1.2
C	0.3275	0.0372	0.3271	1.0	1.1	1.1
D	0.2824	0.0321	0.2817	1.0	1.0	1.0

Another concern during camera calibration process is the correlation between IOPs and EOPs. High correlations can be avoided if a proper configuration is used during image acquisition (e.g. convergent and rotated camera positions and depth variation). Table 11 presents the minimum and maximum correlation coefficients (ρ) estimated in Experiment D for Sensor 1 (20 images) and Sensor 2 (16 images) used in the camera calibration process. The maximum correlation is within 80% between IOPs and EOPs. Considering only $\rho \approx 0.70$ (70%), the highest correlations were obtained between y_0 and Y_0 and between y_0 and ϕ in both sensors. Regarding image dataset, the correlation between c and Z_0 was significantly reduced, ranging between 0 and 0.76, due to the large depth variation of the 360° calibration field, minimizing the linear dependence of these parameters.

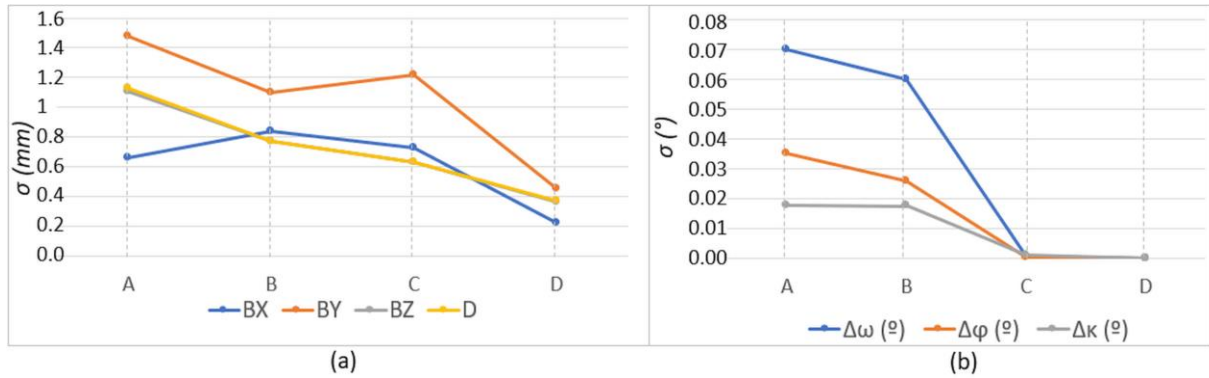
Table 11. Minimum and maximum coefficient of correlation between IOP and EOP for Ricoh Theta S camera calibration

		ω		ϕ		κ		X_0		Y_0		Z_0	
IOP		S1	S2	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
c	min	0.01	0.00	0.01	0.00	0.00	0.02	0.02	0.00	0.01	0.00	0.00	0.00
	max	0.09	0.08	0.20	0.12	0.09	0.08	0.69	0.62	0.10	0.15	0.53	0.54
x_0	min	0.00	0.03	0.00	0.01	0.00	0.01	0.00	0.00	0.01	0.01	0.00	0.00
	max	0.67	0.70	0.62	0.62	0.63	0.63	0.11	0.09	0.71	0.69	0.63	0.68
y_0	min	0.00	0.01	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
	max	0.66	0.49	0.74	0.76	0.67	0.66	0.48	0.45	0.69	0.78	0.69	0.76

3.5.3 Dual camera system relative orientation

Stability constraints presented significant advantages for the relative orientation parameters (R^R) estimation. Figure 19.a presents the standard deviations of the estimated base elements and the PC distance (D_{PC}) in Experiments A to D, while, Figure 19.b shows the standard deviations of R^R components ($\Delta\omega$, $\Delta\phi$, $\Delta\kappa$). Experiments A (blue) and B (orange) presented the biggest standard deviation for the distance D_{PC} . The standard deviations of $\Delta\omega$ and $\Delta\phi$, are close to 0.04° and 0.02° , respectively, which are comparable to the magnitude of these parameters. Although experiments A and B presented similar results to experiment C and D for IOP estimation (Table 8), the values estimated for the EOP and ROP in experiments A and B were not consistent (mainly the R^R components) with Ricoh Theta S physical structure.

Figure 19. Standard deviation of the estimated base components and distance between PCs (a) and standard deviations of rotation elements of the relative rotation matrix (b), computed for Experiments A to D.



Source: Campos et al, (2018b)

Introducing Relative Orientation Stability Constraints improved ROP estimation, as can be seen in experiments C and D. Experiment C showed better results than A and B, achieving standard deviations for base distance smaller than 1 mm. The same behaviour was noticed for the B_z component, which is the element that most influences D_{PC} , since cameras are in a back to back position. A significant improvement was also obtained for the $\Delta\omega$, $\Delta\phi$, $\Delta\kappa$ components when approaches C were applied, in which standard deviations were reduced to arcseconds (Table 12). The base components B_x and B_y vary more than B_z with a behaviour comparable to the two previous experiments, without ROSC. Experiment D presented the best result because stability constraints were imposed on each base component. The standard deviation for B_x , B_y and B_z were less than 0.5 mm, indicating stable results. Standard deviations for $\Delta\omega$, $\Delta\phi$ and $\Delta\kappa$ also presented better results, being smaller than 1". Table 12 presents the values of the estimated ROP and the respective standard deviations obtained in Exp. D. These values can be used in future works as accurate estimates for the ROP to connect the EOP of both cameras.

Table 12. Estimated values and standard deviation for the RR components ($\Delta\omega$, $\Delta\phi$ and $\Delta\kappa$) in degrees, the PCs distance (D_{PC}) in mm and the base elements B_x , B_y and B_z in mm resulted from Exp. D.

Exp.	$\Delta\omega$ (°)	$\Delta\phi$ (°)	$\Delta\kappa$ (°)	D_{PC} (°) mm	B_x (°) mm	B_y (°) mm	B_z (°) mm
D	0.0149° (0.025")	0.0526° (0.024")	179.66° (0.025")	19.08 (0.37)	-0.79 (0.22)	0.41 (0.45)	19.04 (0.36)

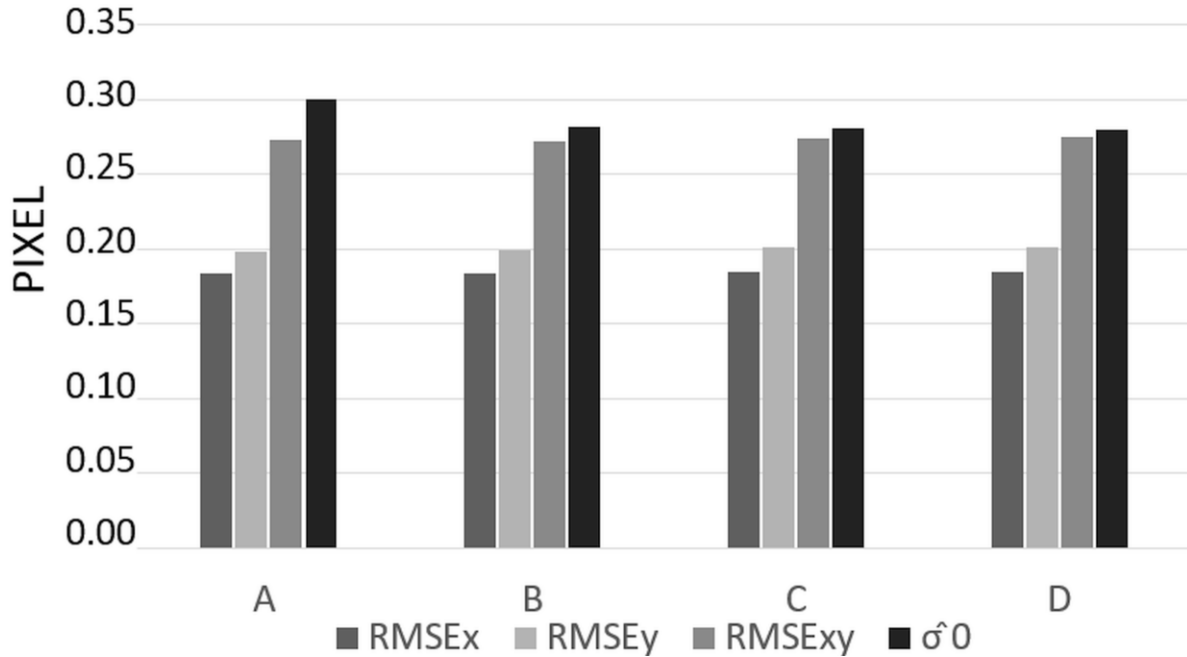
3.5.4 Accuracy assessment

Figure 20 presents the RMSE of image coordinates residuals in x and y directions, in resultant direction ($RMSE_{xy}$) and the a posteriori variance factor (σ_0) obtained in the

camera calibration with experiments A to D. The residuals increased slightly (0.34%) when stability constraints were imposed, which was expected since the PC position variations are reduced due to these constraints. Since these approaches admit random variations, this increased residual still results in an RMSE of image observations smaller than the measurement error (0.5 pixel), thus being a good alternative. The a posteriori variance factor σ_0 indicates a consecutive improvement in the bundle adjustments from experiments A to D. Experiment A consists of two single calibration processes achieving an estimated σ_0 for each sensor ($\sigma_{0S1}=0.348$ pixel and $\sigma_{0S2}=0.375$ pixel), but Figure 20 only shows the mean of these σ_0 values. The other three experiments have a simultaneous solution, in which a σ_0 of 0.318 pixel, 0.315 pixel and 0.314 pixel were obtained for experiments B, C and D, respectively. Therefore, despite the minor increase in image coordinates residuals, general improvements on the results were noticed, especially in the parameters estimation (IOP, EOP, ROP) when an integrated approach for both sensors is adopted, as presented before in Table 8 and Table 9 for IOP, Table 10 for EOP and Figure 19 for ROP.

In line with these results, the consistency of the bundle adjustment approaches was also assessed by computing the discrepancies between distances measured directly in the 360° calibration field and distances estimated in each camera calibration approach from ground coordinates. Six well-distributed horizontal and vertical control distances were measured for both coverage areas of each sensor in the 360° calibration field using a 2 m caliper, with submillimetre precision (0.02 mm). Table 13 presents the RMSE of these discrepancies for each experiment. It can be observed that the RMSE of discrepancies between control and estimated distances was 4.8 mm for Exp. A, 4.6 mm for Exp. B and 4.7 mm for C and D, which is compatible with the average pixel size in object space units (5 mm). Despite a posteriori variance of less than 1 pixel has been obtained in all experiments, it is important to notice that these discrepancies are a combined effect of direct distance measurement error and errors in each ground coordinate point estimated in the bundle adjustment, which corresponds in average to 1 mm.

Figure 20. RMSE of the image coordinates residuals in the x and y direction, in the xy resultant direction and the aposteriori variance factor σ_0 for the camera calibration approaches A, B, C and D, in pixel.



Source: Campos et al, (2018b)

Table 13. Statistics from discrepancies between control and estimated distances in the calibration field.

EXP.	A	B	C	D
RMSE (mm)	4.839	4.614	4.657	4.678

3.6 CONCLUSION

This chapter presented a practical study on the geometric calibration of a low-cost dual 360° image system, considering different bundle adjustment approaches. Equidistant projection with the Conrady-Brown distortion model was considered a suitable mathematical model for the Ricoh Theta S internal geometry, which is in line with the RMSE of image coordinates residuals, check distances discrepancies and the standard deviation of the IOP for all experiments. A limitation in this mathematical model was observed when the incident angle (α) is higher than 90°, where Ricoh Theta S lens behavior does not follow the equidistant projection (Figure 32, Section 5). In future works, an extension to this model will be developed to cope with this problem.

The focus of this study was the assessment of the influence of the ROSC for 360°

dual-systems calibration. The main advantage of this approach is that no direct measurement information about ROP is initially required. The approach that better suits the Ricoh Theta S calibration was presented in Experiment D, with a simultaneous camera calibration with base elements and relative rotation matrix stability constraints. Experiment D achieved better results than other experiments (A to C) for IOP estimation (Table 8), but mainly, for EOP (Table 10) and ROP estimation (Figure 19). The correlation between IOPs and EOPs using approach D and the 360° calibration field was smaller than 80% (Table 11). The parameters (IOP and ROP) estimated in Experiment D were used as input data for the other photogrammetric processes (image matching and bundle adjustment) studied in this work.

The use of low-cost, light-weight poly-dioptric spherical imaging cameras has been growing for non-metric purposes. But the results presented in this chapter showed that these systems when associated with a suitable mathematical model and good practices, can provide acceptable results in the photogrammetric process, such as camera calibration. The Ricoh Theta S is just an example of these systems and the methodology proposed can be extended to other dual-systems, such as the Samsung Gear 360, the LG 360 and the Nikon KeyMission 360.

CHAPTER IV

A RECURSIVE FEATURE SEARCH METHOD FOR FISHEYE IMAGE MATCHING

4. SCOPE

Omnidirectional vision requires Feature based Matching (FBM) in multiple images, which is a challenge, even for state-of-the-art methods, such as SIFT (Scale-Invariant Feature Transform), because of the strong scale change from image to image. This chapter summarizes the studies performed on FBM techniques for fisheye images. An overview of the omnidirectional image matching trends and the motivation of the proposed method are first discussed (Section 4.1). Then, the method based in a simplified technique for epipolar constraint in the sphere domain considering the equidistant projection model and a recursive approach to reduce the search window is presented (Section 4.2). Finally, experimental assessments (Section 4.3) and analysis (Section 4.4) of the proposed method for outdoor Close Range Photogrammetry (CRP) applications were performed considering real data sets collected with the PMTS in forest and urban areas. This chapter is based on the work presented by Castanheiro et al. (2017).

4.1 OMNIDIRECTIONAL IMAGE MATCHING APPROACHES

Many studies have focused on extending the well-known FBM algorithms from perspective to omnidirectional images. Three trends in omnidirectional image-matching issues can be noted (HASSEN et al., 2010). The first possibility (approach A) is conversion of a wide-angle image into a geometrically corrected perspective image (HASSEN et al., 2007; TOMMASELLI and BERVEGLIERI, 2014; CHUANG and PERNG, 2018), and thus standard FBM algorithms can be considered geometrically valid. However, these reprojections can add radiometric and geometric distortions and image information losses.

Alternatively, the original wide-angle images can be used with consideration of performed consistent geometric modifications in the image-matching approach, such as in feature extraction or feature comparison. Therefore, a second option (approach B) adapts existing interest operators, such as SIFT (LOWE, 2004), using a spherical surface as a nondeformable domain for image operations (DANIILIDIS, 2002). This approach requires an adaptive scale-space function defined on the sphere. Various research studies were developed in this direction, with the aim of efficiently solving the problem of feature detection of the sphere surface without introducing image

artifacts (BULOW, 2004; HASEN et al., 2010; CRUZ et al., 2012; ARICAN and FROSSARD, 2012; LOURENÇO et al., 2012; PUIG et al, 2014). It is well known that FBM techniques are limited by the quality of the interest point and the features extracted. The works mentioned proposed geometrically rigorous methods for interest point detection in omnidirectional images, which are of wide interest for 3D reconstruction. However, complex implementation and high computation efforts are required.

Generally, the number of correct matches tracked by standard FBM algorithms on an omnidirectional image is beyond the minimum number of points required for robust sensor orientation estimation. A third possibility (approach C) focuses on the improvement of the step of comparing SIFT descriptors to reduce mismatches, mainly in outdoor mapping, for which additional information such as camera positions determined with GNSS (Global Navigation Satellite System) are commonly available. Currently, imaging and navigation systems are combined in mobile mapping platforms. Therefore, considering that the imaging system model is known and that the initial camera position and attitude are supplied by the navigation systems, the position of the searched feature can be predicted, and the feature extraction process can be applied to a small template window enclosing the feature, instead of to the entire image.

Scaramuzza et al. (2008) used a calibrated omnidirectional camera arranged in a vehicle roof for outdoor mobile mapping. The SIFT operator was combined with a filter based on RANSAC to remove outliers, and to retain the reliable homologous points for sensor position estimation. In this approach, outliers were successfully removed by RANSAC, however, features distribution throughout the keyframes in the data set was still a problem. Valgren and Lilienthal (2010) evaluated the performance of two interest operators (SIFT and Speeded-Up Robust Features-SURF) for outdoor panoramic images acquired in different seasons with a mobile mapping system composed of an omnidirectional camera and a GNSS receiver integrated in a robotic platform. Improvements in the SIFT and SURF performances were obtained by changing the matching metric and the acceptance threshold for descriptor comparison according to the data sets evaluated. Additionally, the operators were tested in different panoramic image resolutions and with the use of epipolar constraints to eliminate false matches.

The computational time and match correctness benefits of reducing the search space, e.g., based on epipolar geometry, are well known for perspective images (KRAUSS et al., 2007). Therefore, certain authors extended the classical mathematical model of the perspective

epipolar concept to omnidirectional sensor geometry (SVOBODA et al., 1998; BUNSCHOTEN and KROSE, 2001; ABRAHAM and FORSTNER, 2005; FANGI and NARDINOCCHI, 2013; LUHMANN et al., 2013). More recently, Valiente et al. (2017) proposed a similar strategy to improve image matching with an adaptation of the epipolar constraint for an omnidirectional system composed of a calibrated omnidirectional camera, internal odometer, and a LASER range finder mounted on a four-wheel robotic platform. A search window was defined based on the epipolar line position, propagation of the current omnidirectional system uncertainty, and Extended Kalman Filter (EKF) prediction.

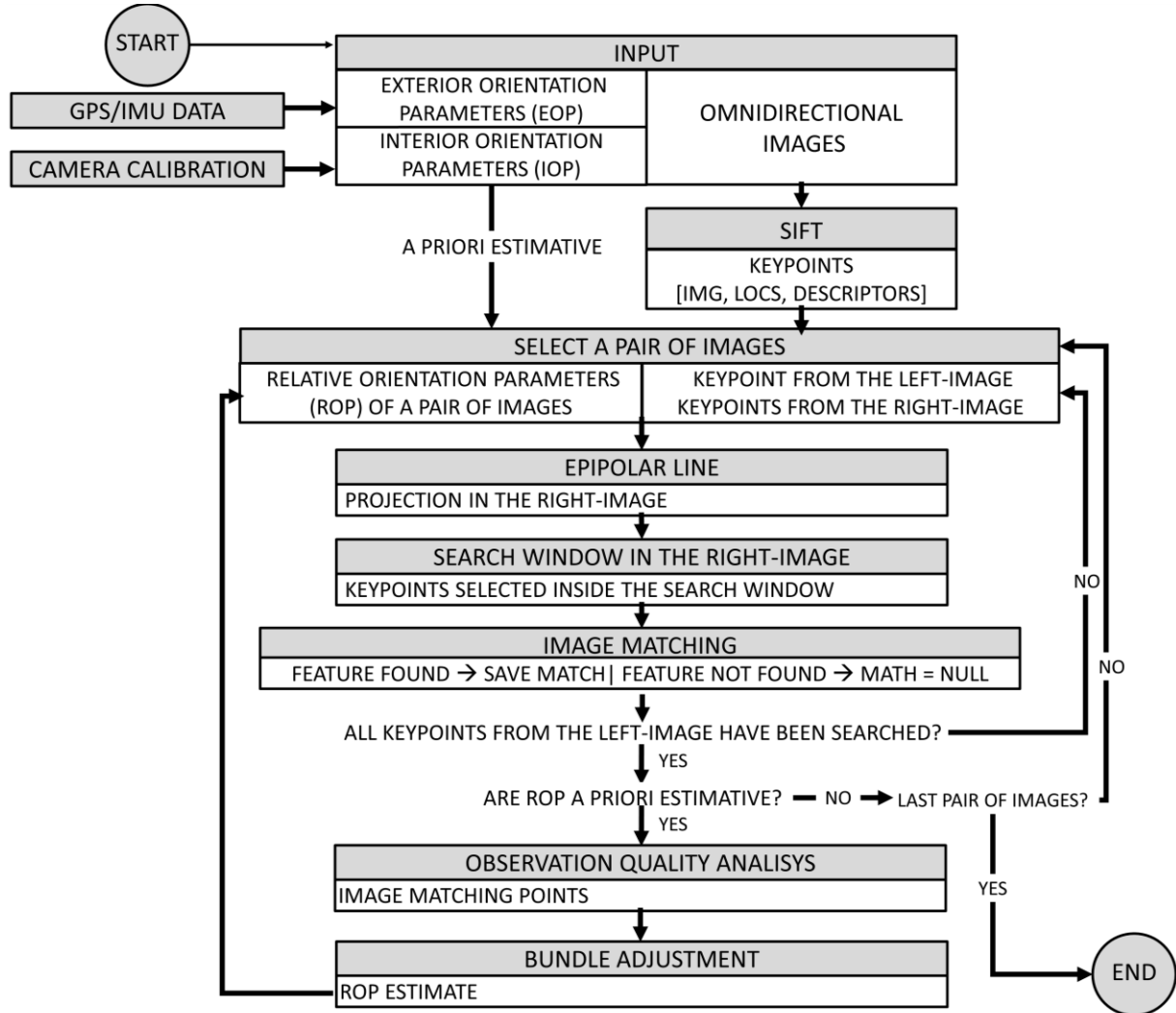
The approaches proposed by Tommaselli and Tozzi, (1996), Scaramuzza et al. (2008), Valgren and Lilienthal (2010) and Valiente et al. (2017) have inspired the method for fisheye image matching proposed in this chapter (approach C). To this end, we propose a technique that reduces the search space window recursively in the fisheye images based on the epipolar geometry constraint in the sphere domain and applies the equidistant mathematical model. Experiments and assessments were performed, by combining the search space technique with the interest operator SIFT (LOWE, 2004), which is one of the most popular interest operators, particularly due to robust matching even in affine distortion cases. A preliminary assessment of the performance of standard interest operator (e.g. SIFT and SURF) on fisheye images was performed, supporting the motivation for SIFT use. Data sets were collected using the low-cost mobile mapping system, described in Chapter 2. In summary, the main contributions of our study presented in this chapter are: (1) a simplified technique for epipolar constraint of the sphere domain considering the equidistant projection model, (2) a recursive approach to reducing the search window, (3) a study of the similarity functions to compare key points in omnidirectional image matching, and (4) analysis of outdoor CRP applications using real data collected with the proposed PMTS (Chapter 2) in forest and urban areas.

4.2 SIFT_{RFS}: A PROPOSAL TO IMPROVE FBM IN FISHEYE IMAGES

Figure 21 shows a flow chart of the proposed method for Recursive Feature Search (RFS) in fisheye images. The input data include the interior orientation parameters (IOPs) from a previous camera calibration process, the initial exterior orientation parameters (EOPs) acquired by a navigation system (Global Positioning System-GPS receiver and IMU) and omnidirectional images. The Relative Orientation Parameters (ROP) encompassing the base

elements (BX, BY, BZ) and relative rotations ($\Delta\omega, \Delta\phi, \Delta\kappa$) between consecutive images acquired at instance (i) and ($i+1, i+2, \dots, i+n$) were estimated based on the initial EOPs. ROP between the cameras that composes Ricoh THETA S system and IOPs values estimated in Chapter 3 are input data for the proposed method presented in this section.

Figure 21. Flow chart of the recursive search space method for fisheye image matching.



Considering a minimum of 2 fisheye images, the SIFT technique was applied to detect interest points (key points) in all images covering the scene. Section 4.3.2 discusses the motivation of SIFT choice. An iterative process is performed from each possible pair of images (left and right images). For each key point extracted in the first image (left image) a search window is defined in the right image based on the epipolar line projection (Section 4.2.1), and the invariant descriptor vector from the key point searched is compared only with the key points belonging to

that search window. The search window is computed as a function of the IOPs and ROPs between a pair of images, and its dimensions depend on the initial EOP uncertainty. Subsequently, recursive bundle adjustment is performed to improve the ROP estimation, thus reducing the search window dimensions (Section 4.2.2). The homologue points used in the first iteration are evaluated and selected, using the minimum standardized Euclidean distance and the maximum correlation coefficient between key point features (Section 4.2.3) as quality metrics. The assumptions made in the following sections refer to a pair of images, but they can be generalized to successive pairs in a sequence of images.

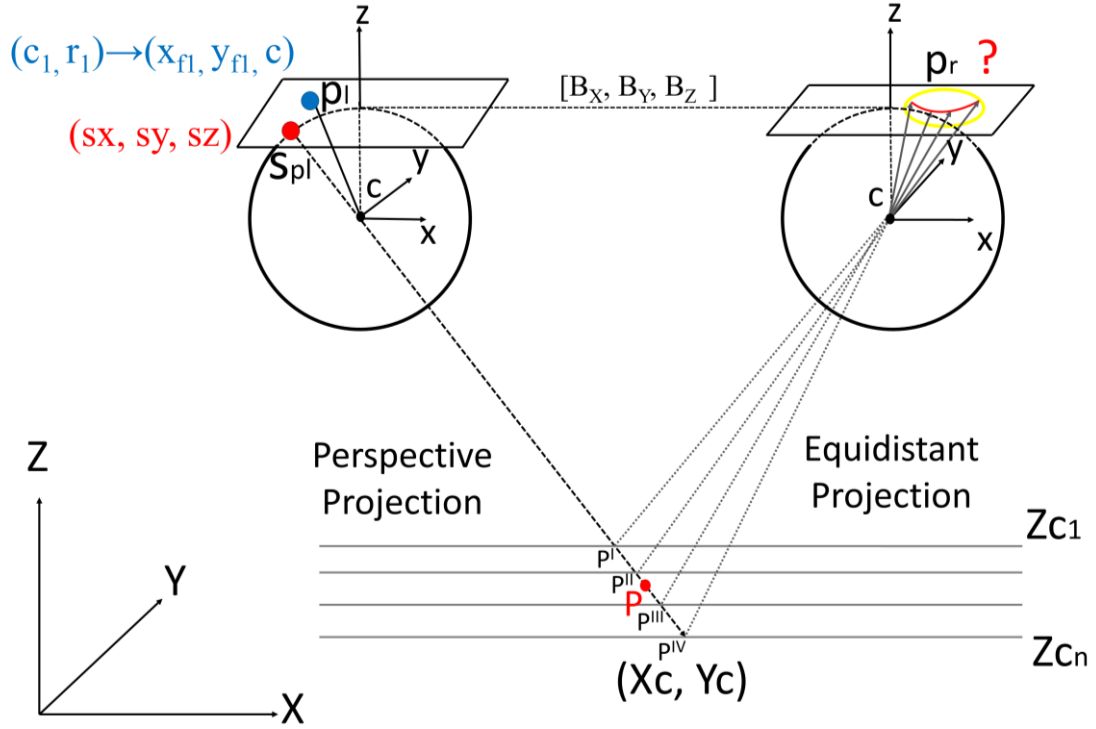
4.2.1 Epipolar geometry on the sphere domain

Considering a pair of images, the image location of a key point from the left image can be predicted in the right image using a geometric search approach, such as those based on epipolar geometry. The principle of epipolar geometry consists of coplanarity between a point P with coordinates (X, Y, Z) in the object space and the perspective centers (PCs) of the cameras forming the image pair with coordinates $(X_{PCL}, Y_{PCL}, Z_{PCL})$ and $(X_{PCR}, Y_{PCR}, Z_{PCR})$ that cover point P . Considering perspective geometry, the intersection between the epipolar plane formed by point P and the PCs and the image planes results in two lines, known as conjugated epipolar lines. The projections of point P in the left (p_l) and right image (p_r) are constrained to fit on the conjugated epipolar lines in the respective images. The epipolar line can be estimated with different methods based on the relative orientation between images (HELLWICH et al., 1994; ZHANG, 1998; GALO and TOZZI, 2004), such as the solution with 8 or more point matches, an analytic method with a rank-2 constraint, a gradient-based technique and least median of squares (LMedS). In summary, consecutive estimation of the ROP is an iterative process that requires a set of correspondent points ($n \geq 7$) and an estimation technique.

Certain modifications in the classical approaches must be considered for estimation of the epipolar line shape and position in the fisheye images. For instance, considering the omnidirectional image geometry, the epipolar lines are curves (ABRAHAM and FORSTNER, 2005). We propose a simplified approach based on projection of an interest point in the left image to the object space using the inverse collinearity condition on the sphere domain, and subsequently, the reprojection to the right image using the equidistant equation. The epipolar line position can be recursively refined with the improvements in the relative orientation estimation (ZHANG, 1998;

GALO and TOZZI, 2004), which will be presented with more details in Section 4.2.2. Figure 22 summarizes the proposed method.

Figure 22. Epipolar curve estimation using the proposed method.



If the IOPs and ROPs between consecutive images are known, the position of the epipolar lines can be estimated. The IOPs were previously determined in the camera calibration process presented in Chapter 3. As mentioned, the ROP base elements (B_x, B_y, B_z) and attitude changes ($\Delta\omega, \Delta\phi, \Delta\kappa$) between consecutive images were estimated based on the initial EOPs measured by the navigation system (GNSS/IMU) of the PMTS proposed (Chapter 2). The left image is considered as the origin of the local system, and therefore, the PC coordinates ($X_{PCL}, Y_{PCL}, Z_{PCL}$) and the rotations ($\omega_L, \phi_L, \kappa_L$) can be assumed equal to zero or to the nominal values acquired by the navigation system. With the former option, the PC coordinates of the right image ($X_{PCR}, Y_{PCR}, Z_{PCR}$) are the base elements (B_x, B_y, B_z), and the rotations ($\omega_R, \phi_R, \kappa_R$) are the relative angles ($\Delta\omega, \Delta\phi, \Delta\kappa$) between consecutive images.

Because the parameters are known, it is possible to project an interest point p_l from the left image to the object space (P) and subsequently reproject to the right image (p_r). The projection from p_l in the image space to the object space can be accomplished by projecting from the sensor plane to the sphere and also to the object space using the inverse collinearity equation.

Because fisheye images do not follow perspective geometry, the collinearity equation is not geometrically valid for the original fisheye coordinates in the image plane, as discussed in Chapter 3. However, the collinearity equation can be considered valid after a projection of the image coordinates from the image plane to a sphere, which is a spatial domain free from deformation. Therefore, the coordinates measured in the left image (c_1, r_1) are computed in the photogrammetric image system (x_{f1}, y_{f1}), as presented in Equation 30, where (s) is the pixel size in mm, (x_0, y_0) are the principal point coordinates, (c_x, c_y) are the coordinates of the geometric image center and ($\Delta x, \Delta y$) are the corrections of lens distortions. Thus, $p_l(x_{f1}, y_{f1})$ is projected to a sphere (s_x, s_y, s_z) with a radius equal to the principal distance (c) using Equation 31, which considers the equidistant property of Ricoh Theta S lenses (HUGHES et al., 2010).

$$x_{f1} = [s * (c_1 - c_x)] - x_0 + \Delta_x \quad (30)$$

$$y_{f1} = [-s * (r_1 - c_y)] - y_0 + \Delta_y$$

$$s_x^{(i)} = c * \cos \theta * \sin \alpha; \quad s_y^{(i)} = c * \sin \theta^{(i)} * \sin \alpha^{(i)}; \quad s_z^{(i)} = -c * \cos \alpha^{(i)} \quad (31)$$

where α is the incident angle, and θ is the angle formed by the radius (r) of an image point (i) and the x-axis in the image (Equation 32).

$$r^{(i)} = \sqrt{x_{f1}^2 + y_{f1}^2}; \quad \alpha^{(i)} = \left(\frac{r^{(i)}}{c} \right) \text{rad}; \quad \theta^{(i)} = \arctan \left(\frac{y_{f1}}{x_{f1}} \right) \quad (32)$$

The coordinates of p_l (from an image point i) on the sphere ($s_x^{(i)}, s_y^{(i)}, s_z^{(i)}$) can be projected to the object space $P(X, Y, Z)$ using the collinearity principle in a mono-restitution process. Because the depth of point P is also unknown, a set of planes (defined by $Z_{(j)}$) is considered. The number of planes used (j) and the ranging of Z values needs to be previously defined (empirically), according to the dataset used. Therefore, for each $Z_{(j)}$ supplied, the $X_{(j)}$ and $Y_{(j)}$ coordinates are estimated as presented in Equation 33 in which $\lambda_{(j)}$ are scale factors (Equation 34), and r_{ij} (i and j from 1 to 3) are rotation matrix elements related to the left-image attitude (origin of the local system), resulting in an identity matrix.

$$X_{(j)} = X_{CPL} + \lambda_{(j)} * [r_{11} \cdot s_x^{(i)} + r_{21} \cdot s_y^{(i)} + r_{31} \cdot s_z^{(i)}]$$

(33)

$$Y_{(j)} = Y_{CPL} + \lambda_{(j)} * [r_{12} \cdot s_x^{(i)} + r_{22} \cdot s_y^{(i)} + r_{32} \cdot s_z^{(i)}]$$

$$\lambda_{(j)} = (Z_{(j)} - Z_{CPL}) / [r_{13} \cdot s_x^{(i)} + r_{23} \cdot s_y^{(i)} + r_{33} \cdot s_z^{(i)}] \quad (34)$$

Finally, the computed $X_{(j)}$ and $Y_{(j)}$ coordinates are projected to the right image ($x_{f2}^{(i)}$, $y_{f2}^{(i)}$) using the ROPs for the right image and the equidistant equation (HUGHES et al., 2010). The equidistant equation was used because this is the fisheye model on which the Ricoh Theta S lenses are based. Another fisheye mathematical model can also be used to model the lens geometry (HUGHES et al., 2010; SCHNEIDER et al., 2009). Image distortions (IOP - Chapter 3) are added to these coordinates in the photogrammetric image system ($x_{f2}^{(i)}$, $y_{f2}^{(i)}$). The image coordinates from the left image were corrected from lens distortion, however, lens distortions still affect the position of the homologous point in the right image. Therefore, the projection of P ($X_{(j)}$, $Y_{(j)}$, $Z_{(j)}$) from the object space to the image space needs to consider lens distortion effect. The photogrammetric image coordinates are, then, converted to a set of coordinates in the column-row system ($c_2^{(i)}$, $r_2^{(i)}$). Finally, the set of computed images coordinates ($c_2^{(i)}$, $r_2^{(i)}$) generated the conjugated epipolar curve in the right image (Figure 22)

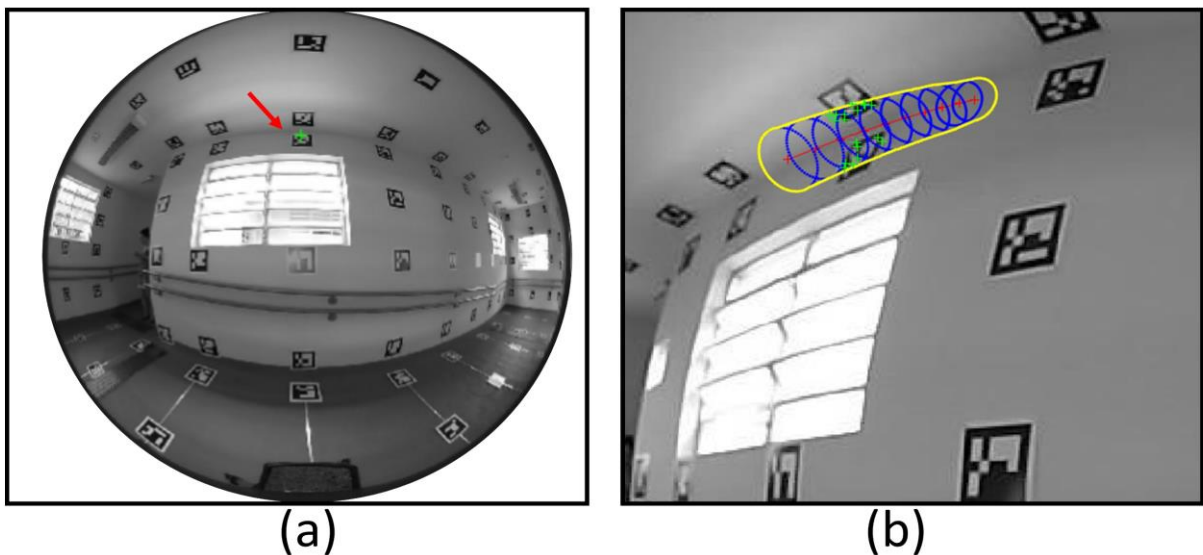
4.2.2 Search space window and bundle adjustment

The initial relative orientation between images was estimated with the data from the navigation system. The navigation system is commonly composed of a GNSS receiver and an IMU. These sensor measurements are affected by errors related to the equipment and its physical components, the synchronization of the components, and the method of data processing. In a first iteration of the proposed method, the accuracy of the epipolar line position was related to the ROP accuracy and consequently was affected by navigation system errors. Therefore, a window surrounding the epipolar curve can be defined as a search space. For each point (j) projected, an error ellipse was computed based on the error propagation of the ROP. An irregular polygon was fitted surrounding these ellipses (convex hull), thus determining the search space. Considering an interest point defined in the left image (Figure 23.a), the corresponding point was searched only inside this search window (Figure 23.b), thus reducing false matches and the computation time. Figure 23 illustrates this process in which (a) presents an interest point in the left image (green) and

(b) shows the epipolar line (red), the error ellipses (blue) and the search window (yellow) defined in the right image.

The errors in the ROP initially estimated from navigation data can be minimized by the bundle adjustment based on the least squares method. This bundle adjustment was implemented using the least squares method according to Mikhail and Ackerman (2011) and Gemaël et al., (2015). The equidistant equation was considered as a mathematical model in the bundle adjustment. The approximated parameter values and observations (tie points) used as input data in the bundle adjustment were selected from the first image-matching iteration in which the search window dimensions were related to the navigation data error. For each pair of images, at least seven well distributed pairs of matches were automatically chosen as observations using the minimum standard Euclidean distance and maximum correlation between key point vector features as quality criteria, as discussed in Section 4.2.3. Furthermore, a statistical test (Pope test) was performed during the bundle adjustment to eventually reduce the weight of potential wrong observations (mismatches) based on the normalized standard deviation and the t-student distribution of 95% sigma (POPE, 1976). Control points were not used in this bundle adjustment process. After ROP refinement, the epipolar line position can be improved in an iterative process, and the search window dimensions can be updated based on the standard deviation of the estimated ROP.

Figure 23. (a) Interest point in the left-image and (b) search window in the right-image.



4.2.3 Image Matching

Homologous points are defined by comparing the SIFT features of a point in the left image $H1(i)$ with the SIFT features of the points inside of the search space window ($H2(j..m)$) in the right image (Figure 23). The correct selection of homologous features requires an appropriate metric for key point comparison. Therefore, the similarity between SIFT features $H1(i..n)$ and $H2(j..m)$ can be computed with different metrics.

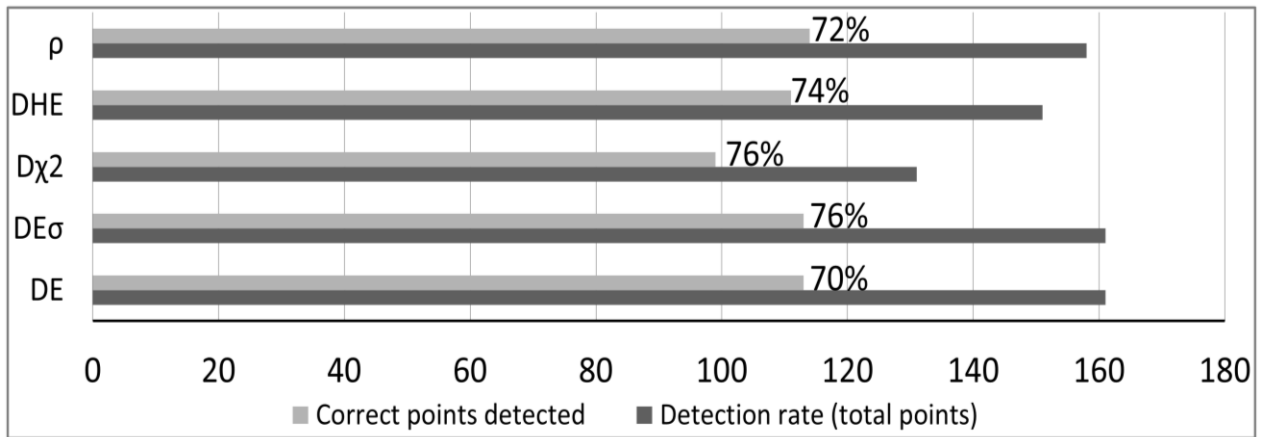
The standard SIFT algorithm (LOWE, 2004) considers the minimum Euclidean distance (D_E , Equation 35.a) as the similarity criterion. However, a different metric for image matching could improve the SIFT performance (ARANDJELOVIC and ZISSERMAN, 2012). Other criteria can also be used to select the nearest neighbor (Equation 35a to 35.e), such as the standardized Euclidean distance ($D_{\sigma E}$, Equation 35.b), chi-square distance (D_{χ^2} , Equation 35.c), Hellinger distance (D_{HE} , Equation 35.d) and histogram correlation (ρ , Equation 35.e). The nine metrics presented in Equation 35 have been combined with SIFT implementation to assess which approach fits best for FBM with fisheye images from outdoor environments.

$$\begin{aligned}
 D_E &= \sqrt{(H1_{(i)} - H2_{(1..m)})^2} & (a) & \quad D_{\sigma E} = \sqrt{\frac{(H1_{(i)} - H2_{(1..m)})^2}{s}} & (b) \\
 D_{\chi^2} &= \sum_i^n \frac{(H1_{(i)} - H2_{(1..m)})^2}{H1_{(i)} + H2_{(1..m)}} & (c) & \quad D_{HE} = (H1_{(i)}^5 - H2_{(1..m)}^5)^2 & (d) \quad (35) \\
 \rho(H1, H2) &= 1 - \frac{(H1_{(i)} - \bar{H1})(H2_{(1..m)} - \bar{H2})}{\sqrt{(H1_{(i)} - \bar{H1})^2 + (H2_{(1..m)} - \bar{H2})^2}} & (e)
 \end{aligned}$$

Based on the computed distances, the nearest neighbor of a feature descriptor was selected and assessed to decide whether the second-closest neighbor distance is smaller than a threshold value. This criterion used to select or reject a match is more effective than a global threshold. Originally, a value of 0.8 was suggested as a threshold for the relationship between the nearest and second nearest neighbor key points (LOWE, 2004). However, this value must be empirically defined according to the data set. In this study, fisheye images from Data set III (Figure 25 and Figure 26.c) were used. This dataset was acquired with the proposed PMTS (Chapter 2) in a test area with sparse vegetation and urban features (e.g. buildings). More details will be presented in Section 4.3.1 (Data set III). To cope with the large variation in surroundings of a feature in fisheye images, thresholds of 0.5 for linear distances and $(0.5)^2$ for squared distances and correlation were empirically defined in our experiments. Figure 24 presents the number of

successful matches, proportion of correct matches and proportion of incorrect or missed matches for each metric approach (distances) used as matching criteria.

Figure 24. Number of successful matches and proportion of correct and incorrect or missed matches using the following metrics: Euclidean (D_E), standardized Euclidean ($D_{\sigma E}$), chi-square (D_{χ^2}), Hellinger (D_{HE}) and correlation (ρ).



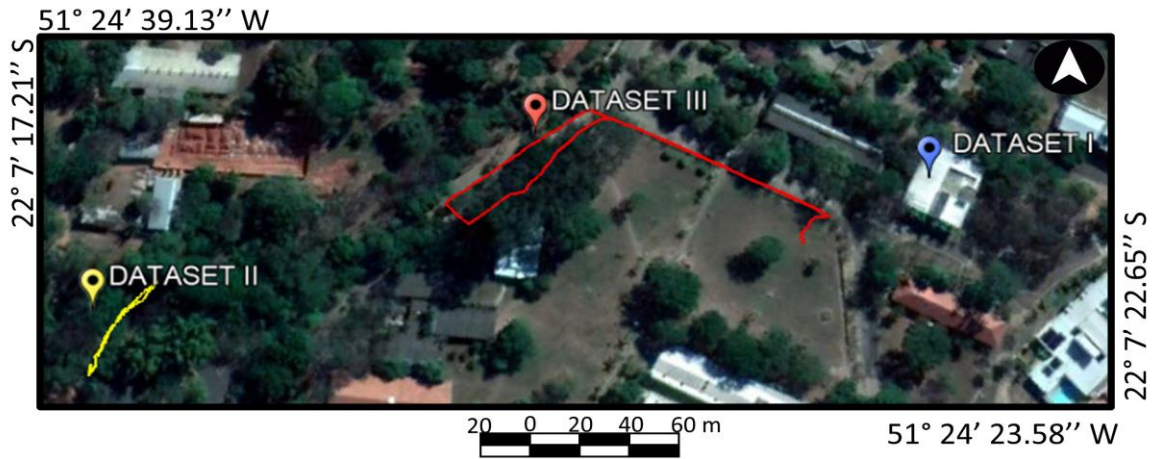
Three approaches can be highlighted in terms of performance compared with that of the classical Euclidean distance (70% of correct detection). The standardized Euclidean (76%), chi-square (76%) and Hellinger (74%) improved the number of corrected match detection along the fisheye image, as also noted in (ARANDJELOVIC and ZISSERMAN, 2012). This result means that in a pair of images of the data set assessed (data set III, Section 4.3.1), 10 matches on average are better selected using one of the metrics other than the Euclidean distance. These results showed that an appropriate metrics for descriptor comparison can improve the results. Considering this preliminary result, the standardized Euclidean distance was used as a metric for feature comparison in the proposed method.

4.3 MATERIALS AND EXPERIMENTS

4.3.1 Data sets

The experimental assessments presented in this section were performed in the test area 1 located on the São Paulo State University (UNESP) campus (Figure 25). Three data sets (I, II and III) were acquired in test area 1, using the poly-dioptric system Ricoh Theta S.

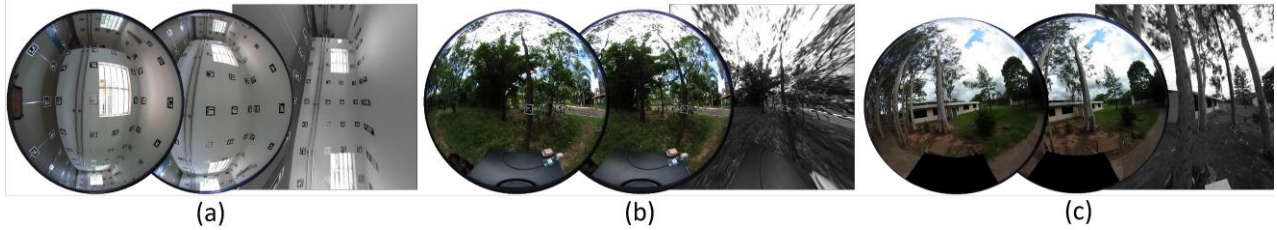
Figure 25. Test area at São Paulo State University: Indoor (data set I), forest (data set II) and urban (data set III) scenes.



- Dataset I was obtained for a controlled scene (indoor) composed of ARUCO coded targets with known coordinates (camera calibration field, Section 3.3), as shown in Figure 26.a. This data set was initially used to assess the proposed method advantages and disadvantages because the image analysis is more controlled in artificial targets. A set of 5 omnidirectional images was acquired with the Ricoh Theta S arranged in a leveled tripod. The EOPs were estimated in a previous bundle adjustment (Chapter 3), and errors (0.08 m and 1°) were added to simulate initial values similar to the PMTS navigation system (Section 2.4 .1). The pixel size in object space units ranges between 4 and 6 mm (the approximate camera-to-object distance is 1.2 m from the central targets), and the image baseline varied between 1 to 4 m.
- Dataset II corresponds to a set of 77 omnidirectional images with EOPs from the navigation system obtained in a small forest area with the proposed PMTS (Chapter 2). The relative accuracies of the ROPs were 1° in attitude and 0.3 m in position because the GPS measurements are more heavily affected by multipath errors and signal obstruction inside of the forest areas. From this dataset, 5 images were selected with an average pixel size in object space units of 1 cm and an image base ranging from 0.73 m to 2.3 m.
- Dataset III describes an area covered with sparse vegetation areas of eucalyptus trees, ground vegetation (grass and bush), small buildings and urban objects (e.g., traffic signals). A set of 5 images were selected in this area with pixel size in object space units ranging between 2 to 5 cm and an image base ranging from 0.74 m to 3.15 m. Initial navigation data were obtained with average accuracy of 1° per angle (ω , ϕ , κ) and 0.3 m for planimetric coordinates.

Figure 26 shows a sample of the omnidirectional and rectified images from data sets I, II and III. Ricoh Theta S omnidirectional images were rectified using the method proposed by Tommaselli and Berveglieri (2014).

Figure 26. Sample of omnidirectional and rectified images from the test areas: (a) data set I-indoor, (b) data set II-forest and (c) data set III-urban scenes.



4.3.2 Preliminary assessment: Why SIFT?

SIFT, FAST (Features from Accelerated Segment Test), SURF and MOPS (Multi-Scale Oriented Patches) had been extensively tested and analyzed for perspective images. However, only few studies discussed the applicability of these operators in omnidirectional images (MURILLO et al., 2007; VALGREN and LILIENTHAL, 2010; CASTANHEIRO et al., 2017). In order to verify the performance of these interest operators for fisheye images in close range applications, a triplet of images from dataset III was selected. These four operators were applied in the triplet of fisheye images using the MATLAB functions for SIFT, SURF and FAST and the free-software ImageJ for MOPS. A threshold of 0.5 for image matching acceptance was employed. The detection rate (number of points detected) and number of successful matches (location correctness) were considered as performance criteria. Table 14 shows the number of total points detected and the percentage of correct matches and mismatches for each standard SIFT, SURF, FAST and MOPS algorithms applied for the triple of fisheye images.

Table 14. SIFT, SURF, FAST and MOPS performance (detection rate and location correctness) in Ricoh Theta S fisheye images

Interest Operator	Detection rate	Location correctness	Mismatches
SIFT	164	70% (114/164)	30% (50/164)
SURF	155	63% (98/155)	37% (57/155)
FAST	59	64% (38/59)	36% (21/59)
MOPS	123	14% (19/123)	86% (104/123)

From this data set, SIFT obtained the most successful detection rate among the four operators with 70% of correct match, while, FAST and SURF presented around 60% of location correctness. The less favorable performance was obtained with MOPS operator, in which the mismatch is higher than the correct rate. In summary, SIFT and SURF presented compatible performances for fisheye images, as also noticed by Murillo et al., (2007). However, considering the evaluated data set, some advantages and disadvantages can be noted. For instance, SURF was computationally faster (at least four times more than SIFT), on the other hand, SIFT provided a higher number of matches and a better location accuracy than SURF for Ricoh Theta S fisheye images. Furthermore, the points distribution was more homogeneous. These results are in line with previous works, such as Scaramuzza et al., (2008) and Valgren and Lilienthal, (2010), which also applied SIFT in omnidirectional images from outdoor scenarios with satisfactory results. This preliminary result motivated the SIFT choice in this study.

4.3.3 Experiment I - A comparison of omnidirectional matching approaches based on SIFT

This experiment was designed to compare and discuss the performance of the three research trends (Section 4.1) on omnidirectional image matching (A, B and C) with consideration of data sets I to III and to analyze the feasibility of these approaches for close-range applications. The proposed method (SIFT_{RFS}) detailed in Section 4.2 (approach C) was compared with the following approaches:

- Approach A (SIFT_{RTF}): The standard SIFT method (LOWE, 2004) was applied in fisheye images geometrically converted to perspective images, as presented in Figure 26. These images will be called as rectified fisheye images in this Section. The rectified fisheye images were generated as presented and implemented by (TOMMASELLI and BERVEGLIERI, 2014).
- Approach B: pSIFT (HASEN et al., 2010) was used to compare the proposed method with an omnidirectional interest operator for close range applications. This method is a SIFT extension that uses the original omnidirectional images and considers a stereographic projection as an approximation for the diffusion on the sphere, which was previously developed and tested for outdoor sensor location (HASEN et al., 2010). These approaches were implemented in MATLAB scripts.

In this experiment, for all three data sets, a threshold of 0.4 was selected for SIFT_{RFS} and pSIFT applied on the original fisheye images because of the large scale variation. Because radiometric distortions are higher in rectified images, a threshold of 0.6 was empirically defined for standard SIFT on the rectified fisheye images (SIFT_{RTF}), with a minimum number of matches per pair of images (> 6). In each data set, image matching was performed with SIFT_{RTF} , SIFT_{RFS} and pSIFT applied to 5 consecutive images, which represents 10 sequential pairs of images (1 to 5, 2 to 5, and 3 to 5 and 4-5) to be analyzed. The range of 1 to 5 sequential images was empirically determined to avoid exhaustive matching by considering the maximum base distance (baseline) between images obtained with the PMTS, which enables the detection of a minimum number of matches (> 6) for a bundle adjustment. Because of the large field of view of the fisheye images ($\approx 210^\circ$), the overlap between images (1 to 5) is almost 100%. However, the main limitation is the drastic change of target appearance when increasing the baseline, especially for the targets close to the sensor. Therefore, the detection rate decreased as a function of the baseline between images. This range can change according to the environment or the speed of platform displacement, which needs to be pre-evaluated.

4.3.4 Experiment II – Sensor Position and attitude estimation

The applicability of the proposed method (SIFT_{RFS}) was tested with a postprocessing approach that considered the set of matches obtained as observations in a bundle adjustment to estimate the EOPs during dataset III acquisition. The bundle adjustment was implemented (C/C++ language, CMC) based on the unified approach to the least squares method (MIKHAIL and ACKERMANN, 1976) using the equidistant fisheye lens equations (Ricoh Theta S model) as a mathematical model (Chapter 3).

A standard deviation of 1.5 pixels was assumed for image observations in both components, i.e., column and row (c, r). The 3D coordinates of five ground control points (GCPs) and five checkpoints were obtained from a high-accuracy terrestrial LASER point cloud measured in the same test area using the Leica ScanStation P40. A standard deviation of 0.05 m was set for GCPs in the bundle adjustment. The initial values for the EOPs were applied with standard deviations of 0.5 m for the camera position (X_0, Y_0, Z_0) and 10° for the attitude parameters (ω, ϕ, κ). The IOPs were considered as absolute constraints, which were estimated in a previous camera calibration process (Chapter 3).

4.3.5 Performance Assessments

The detection rate (number of matches), number of successful matches (location correctness), repeatability (defined in this work as the number of conjugate points detected in sequential pair of images) and the geometric distribution of points over the images were considered as criteria for performance assessment in Experiment I. These criteria were defined because they are closely related to the accuracy of estimated values for the sensor exterior orientation (Experiment II).

In Experiment I, the matches were manually checked by three independent operators, and matching errors greater than 3 pixels were considered erroneous. Repeatability was estimated by computing the number of image pairs in which the same point was successfully matched. The geometric distribution was assessed using the number of matches located in the four quadrants of an image. The standard SIFT performance in omnidirectional and perspective images was considered as a reference to quantify the improvements achieved with approaches A, B and C. To this end, SIFT was applied in the original fisheye images from all data sets. In addition to the fisheye images, a set of perspective images was acquired using a perspective digital camera (Nikon D3100, 4608×3072 pixels) in dataset III with a B/D (base-to-distance ratio) geometry similar to the omnidirectional images.

In Experiment II, the residuals in the image coordinates obtained in the bundle adjustment and the *a posteriori* standard error of unit weight (sigma naught) were used to assess the quality of match locations. The estimated precision (standard deviations) of the EOPs was analyzed based on average values estimated from the covariance matrix. Finally, the accuracies achieved with the bundle adjustment on the ground points were assessed by considering the discrepancies between the estimated and directly measured coordinates of the checkpoints.

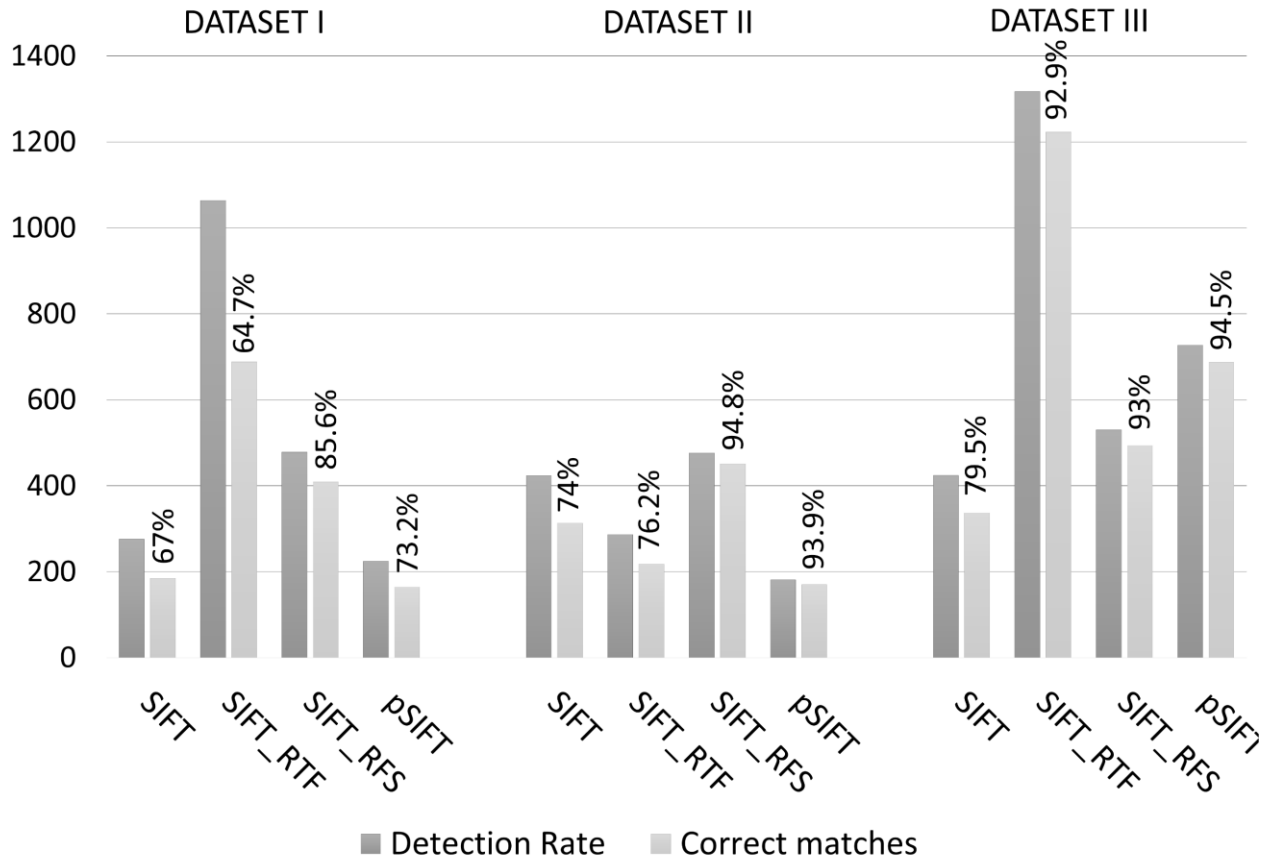
4.4 RESULTS AND DISCUSSION

This section presents the results of Experiment I (Section 4.4.1) and Experiment II (Section 4.4.2). The discussions are based on the performance assessment criteria defined on Section 4.3.5 and the results obtained in related works (Section 4.1).

4.4.1 Comparative analysis

Figure 27 summarizes the detection rate and the number of successful matches (location accuracy in percentage) obtained from 10 pairs of fisheye images with standard SIFT, SIFT_{RTF} (approach A), pSIFT (approach B) and SIFT_{RFS} (approach C) for data sets I to III

Figure 27. Experiment I: Total of detection rate and location correctness of SIFT, SIFT_{RTF}, SIFT_{RFS} and pSIFT for indoor (data set I), forest (data set II) and urban (data set III) scenes.



4.4.1.1 Detection rate and location correctness

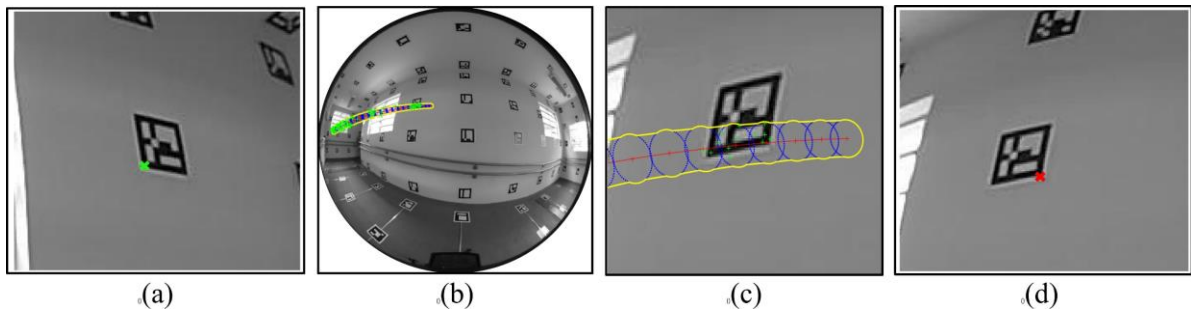
The main challenge in terms of location correctness was multiple matches (nonuniqueness) due to repetitive patterns. Although approach A (SIFT_{RTF}) had a higher detection rate (Figure 27) in datasets I and II, the image projection to a plane introduced radiometric artifacts, mainly caused by illumination variations and nonstatic environments such as forests (data set II). Thus, higher thresholds for matching acceptance are required for SIFT_{RTF}, which enhanced the

similarity between different targets, thus augmenting the ambiguity problem and resulting in the lowest location correctness among the three approaches.

The approaches that worked with the original fisheye images presented more consistent results. pSIFT (approach B) resulted in a detection rate similar to that of standard SIFT but with an improvement of 6.2% in correctness on data set I. Better results were obtained with SIFT_{RFS} (approach C). Improvements of 73% in terms of number of detected points and of 18.6% for correct matches were obtained with SIFT_{RFS} in the same dataset compared with the standard SIFT on fisheye images. For dataset II, SIFT_{RFS} and pSIFT presented compatible results in terms of location correctness (i.e., 94.8% and 93.9%) with improvements of 20.8% and 19.9% respectively, compared with the standard SIFT. However, SIFT_{RFS} achieved a higher detection rate.

In summary, the results from data sets I (indoor) and II (forest) demonstrated that the search space reduction minimized the image-matching ambiguity problem and the impact of omnidirectional image deformation. Therefore, SIFT_{RFS} is a good alternative in areas with repetitive patterns and homogeneous areas, which is often the case in outdoor applications of photogrammetry. The results are consistent with the discussions presented by Valiente et al. (2017) for the omnidirectional case, who also showed that epipolar geometry constraint reduced false matches in wide angle images by establishing a limited feature search space and improved sensor location. It should be noted that, as well as in the perspective case, the epipolar curve is only a geometric restriction, which does not assure a correct match. Most of the false matches detected with SIFT_{RFS} were located in the same epipolar curve or were occluded, which is an intrinsic characteristic of this approach. Figure 28 illustrate this limitation, where (a) presents a key point detected in the left-image, (b) the epipolar curve and the search window defined in the right-image, (c) example of key points select inside the search window and (d) the mismatch in the right image.

Figure 28. Example of mismatch with SIFT_{RFS}: (a) interest point in the left-image, (b) epipolar curve and search window, (c) zoom in close to the interest target and (d) mismatch in the right-image.



Dataset III offers a different awareness. $SIFT_{RFS}$ (93%) presented comparable location correctness results to $SIFT_{RTF}$ (92.9%) and pSIFT (94,5%). However, in terms of detection rate, the geometrically rigorous approaches A and B outperformed approach C, mainly in more stable environments such as data set III. The detection rate is an important criterion because the sensor location and attitude (EOPs) can be better solved with a higher number of matched points, which results in a higher degree of freedom in the bundle adjustment (VALIENTE et al., 2017).

Considering as a reference the standard SIFT performance applied to perspective images ($SIFT_{PERSPECTIVE}$) from the same test area of data set III, a location correctness of 97% was achieved, which illustrates the impact of wide angle image distortions on SIFT efficiency when applied to fisheye images ($SIFT_{FISHEYE}$ - 79%), as summarized in Table 15. SIFT performance was reduced by approximately 20% for fisheye images compared with perspective images. An increased number of false matches in nonperspective images was also noted by Scaramuzza et al. (2008) using a catadioptric system. The three approaches (A, B and C) based on SIFT for fisheye images presented successful matching rates superior to 90% in data set III (Figure 27), thus improving the interest operator performance and partially overcoming the SIFT limitations. The results obtained for analysis of the detection rate and location correctness show the relevancy of introducing modifications compatible with fisheye geometry.

Table 15. Summarized results of SIFT for data set III.

Method	Detection rate (total/ corrected points)	Successful matching rate
$SIFT_{FISHEYE}$	424/ 337	79.5%
$SIFT_{PERSPECTIVE}$	858/ 832	97%
$SIFT_{RFS}$	530/ 493	93%

4.4.1.2 Repeatability

The repeatability performance criterion is the main limitation of the proposed method ($SIFT_{RFS}$, approach C) because the interest points are located with standard SIFT. According to Lowe (2004), an average repeatability of 80% was obtained for key point detection and final matching with SIFT for a perspective pair of images with viewpoint rotation (affine distortion) of less than 30°. Repeatability is significantly reduced when the viewpoint rotation is increased. For instance, viewpoint rotations greater than 50° result in a repeatability of approximately 40% (LOWE, 2004). A similar aspect change occurs from image to image in the

omnidirectional case, thus directly affecting SIFT performance. The repeatability assessment presented by Cruz et al. (2012) showed a mean repeatability of 39% for standard SIFT in omnidirectional images. The experiments performed in Lowe (2004) and Cruz et al. (2012) consider the stability between key point generation in the same pair of images computationally rotated to simulate different points of view. This analysis was extended to real data sets to assess SIFT, SIFT_{RFS}, SIFT_{RTF} and pSIFT for outdoor CRP applications.

In this experiment, the viewpoint rotation between fisheye images changes according to the Ricoh Theta S point of view during PMTS acquisition. Therefore, the viewpoint rotation increases as a function of the base between image pairs. The repeatability was computed for 10 pairs of images in data set III. Repeatability values of 13% were obtained with SIFT_{RFS} (70/528 points) and pSIFT (97/727 points), against an average repeatability of 17% (235/1307 points) for approaches A (SIFT_{RTF}). Methods, such as SIFT_{RTF}, which applied SIFT in the wide-angle images converted to a geometrically corrected perspective image, presented better results in terms of repeatability because the appearance of the target is more preserved along the PMTS acquisition.

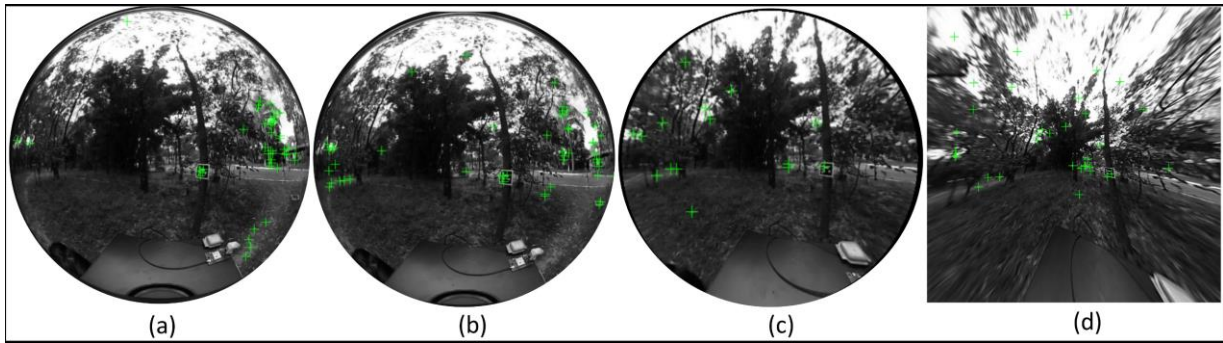
Despite this disadvantage, the number of consecutive repeated points detected in the sequential fisheye images with SIFT_{RFS} can be considered acceptable for EOP estimation in a bundle adjustment (Experiment II). The repeatability of key point detection in standard SIFT can be improved, whereas, it directly increases as a function of the standard deviation (σ) of the Gaussian function (LOWE, 2004; CRUZ et al., 2012) and the acceptance threshold for descriptor comparison (CRUZ et al., 2012). However, a cost exists for use of a large σ in terms of efficiency (LOWE, 2004) and a large threshold in terms of location correctness (CRUZ et al., 2012), thus requiring further assessments to establish optimal values.

4.4.1.3 Geometric distribution

Another consequence of scale and perspective variations in fisheye images for standard SIFT is the weak geometry of point distribution. Figure 29 presents examples of match distributions in a fisheye image from dataset II obtained with standard SIFT (a), SIFT_{RFS} (b), pSIFT (c) and SIFT_{RTF} (d). Because the SIFT key points were selected based on their stability (LOWE, 2004), many matches were selected in the fisheye image border (Figure 29.a). These outcomes occur because targets far from the sensor have less variation in scale, viewpoint rotation and

orientation during image acquisition along the PMTS trajectory, and thus their appearance changes less than that of points near the sensor (usually in image center). Point distribution were improved relative to standard SIFT using the three approaches. Considering all data sets, better point distributions were obtained with approaches A and B (SIFT_{RTF} and pSIFT), as depicted in Figure 29 for dataset II, because the point extraction processes are geometrically compatible.

Figure 29. Experiment I: Example of match distribution for dataset II with (a) standard SIFT, (b) SIFT_{RFS} , (c) pSIFT and (d) SIFT_{RTF} .



The geometric distribution of the matched points in the block of images (tie points in the bundle adjustment) directly affects the EOP estimation, especially the attitude angles due to B/D geometry. An optimal base-to-distance ratio geometry is particularly challenging when using a fisheye lens because the number of point matches decreases as a function of the baseline. This difficulty increases in close-range applications because significant changes in the target aspect and illumination in a sequence of images are more likely to occur, especially if using mobile platforms. Therefore, in certain cases, a pair of images with optimal B/D value (≈ 1) does not supply a suitable number of matches. This problem was also discussed in Hasen et al. (2010) and Lourenço et al. (2012) using interest operators for omnidirectional images (approach A) based on SIFT.

The quality of the image matching procedure depends mainly on the type and content of the image signal and the features imaged. Therefore, as a concluding comment in this section, the results from datasets I, II and III show that image-matching performance was strongly related to the dataset environmental features, mostly in CRP applications, which can influence location correctness, repeatability and the geometric distribution.

4.4.2 Bundle adjustment results

The matches obtained with the proposed method (SIFT_{RFS}) in dataset III were used in Experiment II with the goal of estimating the EOPs and ground coordinates of key points in a post-processing bundle adjustment (with in-the-house software CMC). The mathematical model used in this bundle adjustment process was depicted in Chapter 3. The challenges in determining the EOPs from omnidirectional systems include accuracy of the image point measurements and their geometric distribution. Automatic image observations in nonstatic environments such as forests can select nonstable features such as tree leaves, shadows and clouds, which hinder the convergence of the solution, mainly when small displacements occur. These points determined in the matching step with large displacements can be identified and removed by analyzing the image coordinate residuals in the x and y directions. With respect to the geometric distribution of the observations, the 3D coordinates of the detected image points aligned with the PMTS trajectory cannot be estimated because the intersection of the rays vanish to infinity. Therefore, these points can be removed automatically (MATLAB script), considering the angles formed between the bundles of rays (approaching zero). After observation filtering, a set of 198 match points was used. The final RMSE of the image coordinate residuals obtained in the bundle adjustment process was 0.6 pixels, with a posteriori standard error with a unit weight (sigma naught) of 0.11 pixels. This result is compatible with the SIFT subpixel location accuracy.

The average values of the estimated standard deviations of the exterior orientation angles, omega (σ_ω), phi (σ_ϕ) and kappa (σ_κ), and the camera perspective center (σ_{x0} , σ_{y0} , σ_{z0}) for all images obtained from the covariance matrix are presented in Table 16. Overall, the EOPs were estimated with standard deviations of 0.1° and 2.2 cm for the sensor attitude and position, respectively. Table 16 presents the statistics, including the mean ($\bar{x}$), standard deviation (σ) and RMSE, of the discrepancies between the five check point 3D coordinates measured in a LASER point cloud (reference) and the estimated coordinates after bundle adjustment, resulting in centimeter accuracy compatible with the estimated values for the EOPs. The precision of the 3D terrain coordinates (E, N, h) of the points detected with SIFT_{RFS} (tie points) were estimated considering the standard deviation obtained from covariance matrix. An average planimetric precision of 24 cm and a height precision of 13 cm were achieved, which can be considered compatible with the Ricoh Theta S image resolution.

Table 16. Estimated standard deviation of the exterior orientation parameters: omega, phi and kappa in degrees and camera perspective center in meters.

σ_{ω} (°)	σ_{ϕ} (°)	σ_{κ} (°)	σ_{x0} (m)	σ_{y0} (m)	σ_{z0} (m)
0.14	0.09	0.12	0.019	0.018	0.028

Table 17. Statistics: mean ($\bar{x}$), standard deviation (σ) and RMSE of the checkpoints.

Statistics	E (m)	N (m)	h (m)
$\bar{x}$	-0.009	-0.071	0.054
σ	0.068	0.116	0.199
RMSE	0.061	0.126	0.187

4.5 CONCLUSIONS

A new recursive feature search method for fisheye image matching (SIFT_{RFS}) in close range photogrammetric applications was presented. As proposed, the reduction of the search space can be easily implemented, reproduced and combined in future work with other interest operators such as SURF, pSIFT and ASIFT (Affine-SIFT; MOREL and YU, 2009). The proposed method for epipolar line estimation is an alternative to conventional implementations for omnidirectional images (VALIENTE *et al.*, 2017; FANGI and NARDINOCCHI, 2013). The introduction of a geometric constraint supplied compatible results in real data challenges relative to other geometrically rigorous approaches such as SIFT_{RTF} (TOMMASELLI and BERVEGLIERI, 2014) and pSIFT (HASSEN, *et al.*, 2010), which require higher implementation efforts. For instance, SIFT_{RFS} presented better results in areas with repetitive patterns, whereas SIFT_{RTF} and pSIFT outperformed SIFT_{RFS} in terms of point geometric distribution. The applicability of low-cost GNSS/IMU sensors for outdoor application with integrated sensor orientation methodologies can be improved using approaches based on the relative orientation parameters, such as SIFT_{RFS}, thus enabling more flexible mobile mapping systems, for instance the presented PMTS. Low-cost systems are noise-sensitive, which can result in a low absolute accuracy, as discussed in Chapter 2. However, the relative measurements are more stable because certain sources of error are mitigated, such as inaccuracies in the mounting parameters (BLÁZQUEZ and COLOMINA, 2012) and systematic errors in the GNSS position and the attitude computed from IMU data. The relative positions and angles between GNSS/IMU observations can be used as relative constraints (BLÁZQUEZ and COLOMINA, 2012) to improve sensor location estimation in the bundle adjustment. In this study, the EOPs were estimated in postprocessing bundle adjustment with centimeter positional accuracy. Real-time trajectory estimation, as performed by

Hasen, et al. (2010), Cruz et al. (2012), and Valiente et al. (2017), using SIFT_{RFS} is recommended as a complement to assessment of the proposed method in CRP applications.

CHAPTER V

AN INCREMENTAL BUNDLE ADJUSTMENT BASED ON WIDE ANGLE IMAGE GEOMETRY

5. OVERVIEW

The 3D reconstruction of an unknown environment with the simultaneous determination of the sensors orientation is one of the main problems to be solved in real-time Close Range Photogrammetry (CRP), specially using omnidirectional systems. The 3D reconstruction is obtained with the location and estimation of tie points coordinates, and by this way generating a sparse point cloud. The orientation of the sensor (also referred as pose) can be described by its position and attitude (Exterior Orientation Parameters - EOP). The EOPs and a 3D sparse point cloud have been simultaneously estimated using similar techniques based on multiple perspective views of a scene obtained from image sensors motion developed by the Photogrammetry, Computer Vision and Robotics communities with diverse denominations, such as Simultaneous Localization and Mapping (SLAM) and Structure from Motion (SfM). A comprehensive overview of SLAM and SfM techniques and terminologies are presented by Durant-White and Bailey, (2006) and Granshaw et al. (2018), respectively.

Regarding real-time solutions, an incremental approach needs to be considered since new measurements are made available at each time period during sensor displacement. The most used estimation techniques for achieving real-time capabilities are Kalman-filter (sub-optimal solution) and incremental bundle adjustment (optimal non-linear least-squares solution), which is the main topic of this chapter. Indeed, since the origins of analog Photogrammetry, a solution based on an incremental bundle adjustment are used and analytical solution were proposed by Gruen (1988), also called as on-line triangulation or real-time photogrammetry. Bundle adjustment approaches are well recognized as providing a high accurate estimation of sensor position and orientation. But in those early works, this approach was considered unaffordable for real-time applications, since bundle adjustment is very time-consuming and therefore, a sequential estimation would demand high-computational resources. Recently, we are witnessing the transition of this classic method from academic to industrial applications, specially motivated by the emergent real-word problems of navigation in indoor and outdoor environments related to mobile systems automation. The feasibility of an incremental bundle adjustment for real-time estimations was enabled by the advancement of

computational photogrammetric processing (especially image matching), and the diversity of navigation and imaging sensors, such as omnidirectional ones, available in the market (INDELMAN et al., 2012; MUR-ARTAL et al., 2015; WANG et al., 2019). For instance, bundle adjustment have been used with omnidirectional images, aiming a dynamic data acquisition, due to the large field of view of the images obtained with these systems (SIBLES et al., 2009; MOURAGNON et al., 2009; EUDES et al., 2010; SCHNEIDER et al., 2013; LUKIERSKI et al., 2015; MUR-ARTAL et al., 2015; CHEN et al., 2018; FROST et al., 2018; MATSUKI et al., 2018; WANG et al., 2018). Furthermore, Integrated Sensor Orientation (ISO) approaches have opened many possibilities to optimized real-time photogrammetry considering a set of observations from different sources in a bundle adjustment. Initial information provided by Global Navigation Satellite Systems receiver and Inertial Measurement Unit (GNSS/IMU) mainly improved the processes with high computational requirements, such as image matching and bundle adjustment. The increasing number of low-cost navigation sensors and the diversity of image acquisition systems have inspired new ISO concepts based on the flexibility and interoperability of these sensors.

The motivation of this chapter comes with this background. We present an ISO approach to estimate omnidirectional sensor orientation and a 3D sparse point cloud during a mobile mapping acquisition (PMTS – Chapter 2), using an incremental bundle adjustment strategy and low-cost sensor data. The mobile mapping system used is composed by a poly-dioptric camera (dual-fisheye lenses) and a GPS/IMU navigation system, as presented in Chapter 2. As discussed by Blázquez and Colomina (2012), the feasibility of an ISO approach is strictly related to a rigorous mathematical model and a proper stochastic modelling of the sensor's properties. Several challenges remain in the treatment of image observations acquired by sensors with non-perspective inner geometry, such as omnidirectional cameras, which were already discussed in related works (Section 5.1). Therefore, in Section 5.2, we suggest a mathematical model for the bundle adjustment process consistent with the geometry of the poly-dioptric camera used (Ricoh Theta S – Chapter 3) and a stochastic treatment of the observations based on the fisheye lens properties. The incremental bundle adjustment methodology, starting from data acquisition to final parameters estimation, is described in Section 5.3. Experimental assessments simulating a real-time CRP application was performed to discuss the feasibility of the method (Section 5.4).

5.1 OPEN PROBLEMS ON REAL-TIME CLOSE RANGE PHOTOGRAMMETRY WITH OMNIDIRECTIONAL SENSORS

Incremental bundle adjustment can be understood as an estimation technique based on a sequential adjustment for image system orientation and 3D point coordinate estimation, in order to obtain an optimal least squares solution (GRUN, 1988; MAAS, 1997). An accurate solution on the bundle adjustment requires a strong network configuration of tie points and Ground Control Points (GCP). The network configuration requires high accuracy in image measurements, suitable geometric distribution in the images, an optimal number of tie points and GCP, and a consistent geometry of intersection rays. Therefore, the main challenge relies on the automatic detection of image observations (image matching assessment and geometry) to support a consistence network of matches and the auto-decision of whether to include an observation or not (outliers and weak geometry points disclosure). These open problems were pointed out in recently related works focused on real-time CRP applications (MICHELINI and MAYER, 2014; CADENA et al., 2016; WANG et al., 2019), which can be considered more severe when using omnidirectional systems (HASEN et al., 2010; LOURENÇO et al., 2012; SCHNEIDER et al., 2013; MUR-ARTAL et al., 2015). We can summarize these problems in 2 main topics: (1) omnidirectional match network and bundle adjustment geometry and (2) real-time strategies.

5.1.1 Omnidirectional match network and bundle adjustment geometry

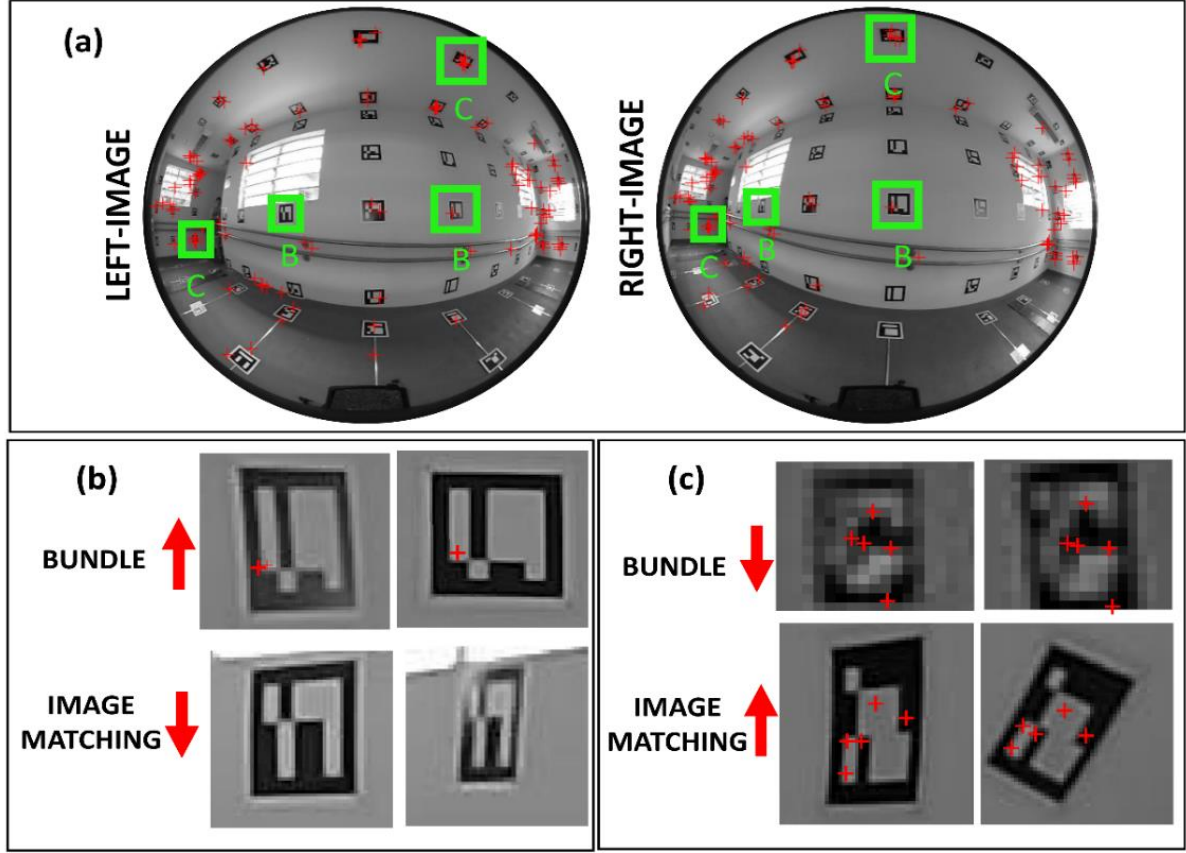
Feature-based Matching (FBM) on omnidirectional images and the geometry of observations for bundle adjustment can be considered a competing task. FBM techniques, such as SIFT, often depend on the space and scale stability of key points, while, bundle adjustment requires a consistent intersection geometry from rays generated by image observations in different images. These requirements lead us to two points of discussion. First, the geometry of the detected points as a function of the huge depth and scale variation in wide angle images. And second, the intersection geometry from rays related to the baseline length and the depth of the scene (Base/Depth ratio).

Concerning the first one, the consistency of space and scale invariances of SIFT-based key points decreases with viewpoint changes higher than 50° (LOWE, 2004; MIKOLAJCZYK and SCHMID, 2005), which is quite usual in sequential wide-angle images. The viewpoint rotation in a pair of images increases as a function of the baseline, especially for targets close to the sensor (Figure 30.b). Consequently, these targets with better ray intersection geometry

for the bundle adjustment, change appearance significantly as a function of the displacement, affecting key points stability and comparison for match. In the same context, key points with invariant scale are located far away from the sensor. The intersection of rays of these points usually provide a weak observation geometry for the bundle adjustment (Figure 30.c). Furthermore, the geometry problem can be considered even more critical for the 3D coordinates of the detected image points aligned with the sensor's trajectory, which cannot be estimated because the intersection of the rays vanish to infinity (Figure 30.c). This geometric problem using wide angle images was also discussed by Schneider et al., (2013).

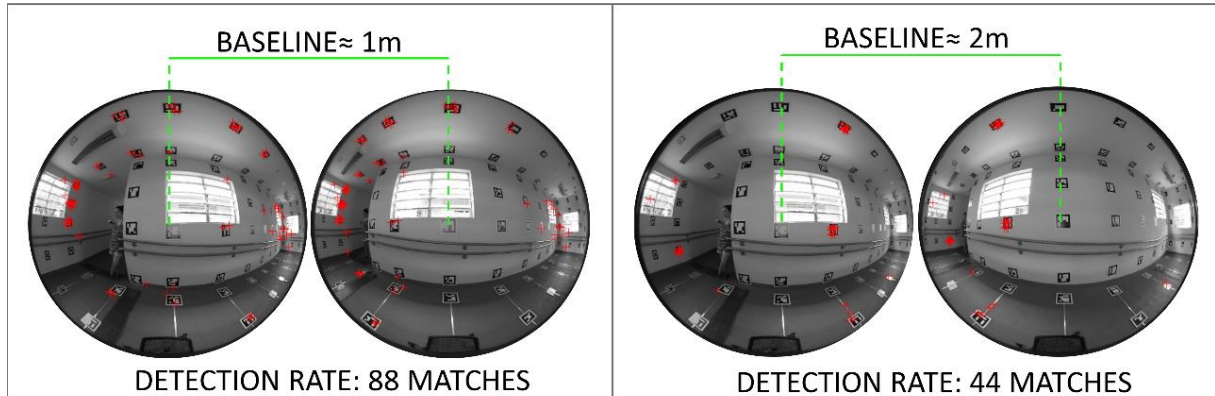
These problems involving image matching detection and bundle adjustment geometry were illustrated in Figure 30 and Figure 31. Figure 30.a shows a pair of fisheye images (Ricoh Theta S camera) with a baseline of 1 m, acquired in a 360° camera calibration field (Chapter 3). The image key points were automatically detected using SIFT_{RFS}. In summary, many optimal key points for FBM techniques are non-optimal for the bundle adjustment reliability (Figure 30.b) and vice-versa (Figure 30.c). Therefore, despite of the great advances on omnidirectional image matching using FBM techniques (HASSEN et al., 2010; CRUZ et al., 2012; ARICAN and FROSSARD, 2012; LOURENÇO et al., 2012; PUIG et al, 2014), the network of matches using wide angle images, mainly in mobile mapping systems, can present a weak geometry intersection of rays, hindering the conversion of the solution and decreasing the quality of the estimation. For instance, the weak geometric distribution of matches using omnidirectional image is pointed out by Scaramuzza et al. (2009) as one of the main reasons to the drift of the incremental solution.

Figure 30. Example of matching and bundle adjustment geometry problems: (a) shows the matching distribution detected with SIFT_{RFS} in a pair of fisheye images, in which (b) exemplify points with good geometry for the bundle adjustment but with low stability for the image matching detection and (c) illustrate the opposite case.



Regarding the second point of discussion, a weak intersection geometry in the bundle adjustment also results from a small ratio (<1) between the baseline length and the depth of the scene. This limitation on mobile mapping omnidirectional system has been discussed by many authors (HASSEN et al., 2010; LOURENÇO et al., 2012; MUR-ARTAL et al., 2015). The increase in baseline length has significant advantages in improving the estimation accuracy of the sensor pose and 3D reconstruction in the bundle adjustment, because wide-baseline observations can reduce depth uncertainty. However, the detecting and matching of the same key points in images taken from wide-baselines are a difficult task, especially in CRP applications with mobile mapping platforms. The key point space-scale consistency limits the baseline in the image matching and consequentially bundle adjustment geometry. Figure 31 shows the decrease of matches (SIFT_{RFS} - Chapter 4) in a sequential pair of fisheye images as a function of the increasing baseline (+ 1 m) and the viewpoint change.

Figure 31. Example of detection rate decreasing due to viewpoint change and image pairs baseline.



5.1.2 Real-time strategies

Real-time estimation requires automatic processes and the ability to include or exclude an observation. Auto-decisions can be defined by constraints and optimization strategies, which generally require a priori knowledge of the problem and the environment to be mapped. The main difficulties include (1) to avoid exhaustive matching, unnecessary observations and outliers, (2) to provide suitable initial information, (3) to detect failure and recovery, and (4) observations and parameters storage. The storage problem is well-known as the capability of learning, remembering and forgetting an observation when required, in which it is necessary to decide when this information is essential to be maintained, need to be updated or becomes outdated and can be discarded (CADENA et al., 2016). The discussion on this section will be focused in the points (1), (2) and (3).

(1) An exhaustive matching approach for tie-point extraction is still the most computationally expensive process in incremental real-time implementations, such as incremental SfM and SLAM approaches. In the conventional exhaustive approach, all images are considered as possible feature pair combinations, resulting in $(N \cdot (N - 1)/2)$ image pairs in total, given N images (WANG et al., 2019). This methodology enables the connections between all parts in a sequential image data. However, many of those image pairs do not have a suitable overlap or B/D ratio, resulting in outliers or unnecessary observations for the estimation improvement. In the omnidirectional case, the inclusion of non-optimal observation can be more recurrent because of the huge depth and scale variation in wide angle images, as discussed in Section 5.1, degrading the quality of the estimated parameters, the capability to detect outliers and the recovering of the

solution (CADENA et al., 2016). However, an exhaustive matching approach can be avoided with some restrictions. First, some constraints can be set when initial information is provided. For instance, the order of images and camera position observed by external sensors (e.g. GNSS receiver), which is often the case with mobile mapping systems, enables the selection of a sequential and consistent pair of images (keyframes). Exhaustive matching can also be avoided without a prior knowledge, considering an unordered sort of images and using graph-based optimization, as discussed by Hartmann et al. (2016) and Wang et al., (2019). In our work we will focused in the first alternative (Section 5.4).

(2) According to Mur-Artal et al. (2015), the feasibility of an incremental bundle adjustment with reasonable computational costs for real-time applications requires initial information of the sensor pose (EOPs) and tie point coordinates for the nonlinear optimization along the sensor trajectory. This scenario can be achieved by combining photogrammetric techniques with GNSS/IMU-derived EOP observations, so called ISO. Other approaches without the use of external sensor information are also possible (LORENÇO et al., 2012; CADENA et al., 2016), but this is not the focus of this work. As mentioned, an ISO approach is strictly related to a rigorous mathematical model and a proper treatment of sensors and observations stochastic properties (BLÁZQUEZ and COLOMINA, 2012). Concerning initial EOPs using low-cost navigation sensors, the accuracy of the measurement is affected by different sources of errors and sensor stability (Chapter 2). As discussed by Clark (2018), microelectromechanical systems (MEMS) present a large gap between their predicted accuracy and actual performances. Therefore, the standard deviation for initial EOPs needs to be carefully considered. Aiming at the 3D reconstruction, the initial values of ground coordinates of tie points can be estimated using the initial EOPs from GNSS/IMU and a suitable mathematical model for wide angle images (SCHNEIDER et al., 2009; HUGHES et al., 2010). These initial values will be computed as presented in Section 3.2. Furthermore, tie points require a stochastic treatment. Especially considering the points with non-consistency geometry, such as those aligned with the system trajectory or far away from the sensor, as described in Section 5.1.

(3) The difficulties mentioned-above determine the system proneness of failure and capability of recovery (MAAS, 1997), which involves initial information, observation filtering and landmarks to close loops in real-time. The ability of detect outliers and omnidirectional image pairs with weak B/D ratio is still an open problem in real-time application with wide-angle images.

As well described by Cadena et al. (2016), the failure in outlier detection can lead to a wrong solution of the methods to detect weak geometry pairs, and consequently also reduces the capability of discerning outliers later on and recovery stability. Outliers can be avoided or/and rejected from the set of matching in omnidirectional images using different methods, such as based on RANSAC algorithm (BOLLES and FISCHLER, 1981), match point geometry or/and epipolar restriction. Methodologies based on RANSAC algorithm are the most used and had been performed by Scaramuzza et al. (2008); Scaramuzza et al. (2009); Jäger et al. (2013) and Song et al. (2018). Another possibility is to analyze match point geometry in function of the reprojection error (LUKIERSKI et al., 2015) or the number of intersection rays, as proposed by Mur-Art et al. (2015). These authors presented a restrictive test approach to remove outliers and wrongly triangulated points, in which an estimated tie point is only retained in the bundle adjustment if it was tracked at least in 25% of the images predicted to be visible and if it is observed from less than three points of view. Moreover, epipolar geometry restrictions are also commonly used to avoid outliers, as proposed by Davison et al. (2003), Valiente et al. (2012), Valiente et al. (2017), Matsuki et al. (2018) and Wang et al. (2018). These methods can be combined to achieve more satisfactory results.

5.2 MATHEMATICAL MODEL AND SENSOR STOCHASTIC PROPERTIES

This section aims to discuss and propose solutions to minimize the problems discussed in Section 5.1.1, considering suitable functional and stochastic models. The presented assumptions are valid to approaches using fisheye lenses.

5.2.1 Functional model

This section discusses the mathematical models for a bundle adjustment based on observations from omnidirectional systems with dual-fisheye lenses, also classified as poly-dioptric solutions (MAAS, 2008). As mentioned in Chapter 3, the image acquisition using polydioptric systems based on dual-fisheye lenses does not follow the collinearity condition, requiring a suitable functional model to reflect the behavior of the fisheye lenses (SCHNEIDER et al., 2009; HUGHES et al., 2010). Furthermore, considering that these systems are composed by multiple cameras, the relative orientations between multiple cameras needs to be considered for a

rigorous approach. The functional models to preserve fisheye geometry usually fall into one of two categories: either original fisheye image observation is preserved or not.

The first category follows the “*Do not touch observation*” principle (BLÁZQUEZ and COLOMINA, 2012), in which it is recommended to fit the model in function of observations geometry, avoiding “*corrections*”. Therefore, original observations from fisheye images are combined to the intrinsic mathematical model of these lens, such as stereographic, equidistant, orthogonal or equi-solid-angle projections. A review of the fisheye projection models can be found in Ray (2002), Abraham and Forstner (2005), Hughes et al. (2010), and in Chapter 3. Section 3.4.2 described the mathematical model used in the bundle adjustment as a function of the equidistant projection (Equation 13), which is the fisheye lens model that best fits Ricoh Theta S lenses. Concerning the offsets between fisheye lens cameras that compose the polydioptric system, the mathematical model can be combined with stability or absolute constraints based on the Relative Orientation Parameters (ROP) between them. In order to maintain the principle mentioned-above, our work combined the equidistant mathematical model (Equation 13), with ROP stability constraints in the bundle adjustment. Therefore, the estimations of sensor 1 and 2 poses were performed almost independently. But, they are connected by the ROP stability constraints. The bundle adjustment mathematical model with relative orientation stability constraints (Equations 20 to 29) is also described in Section 3.4.2. This mathematical model will be used in the incremental bundle adjustment approach presented in this work.

The focus of the second category of models is to fit the observations to a general functional model, such as the perspective one (collinearity equations) combined with adaptative strategies to fisheye lenses (KANNALA and BRANDT, 2006; SCHNEIDER et al., 2012; URBAN et al., 2017). One possibility is to consider that in full spherical images (dual-fisheye) the intermediary image reference surface is not a plane but a sphere. Therefore, the collinearity equation can be considered valid for fisheye lenses after a projection of the original fisheye image coordinates from the image plane to a sphere, which is a spatial domain free from deformation. The original observations projected to the sphere surface (radius equal to the principal distance - c) result in spherical coordinates, as presented in Section 4.2.1 (Equations 30 to 32). Once, the spherical coordinates (s_x , s_y , s_z), corrected from lens distortions have been computed, 3D coordinates in the object space (X , Y , Z) can be obtained considering an inverse collinearity projection, which correspond to the equation of a line in 3-D space. In a 3D spherical space, it is not

possible to consider that s_z is a constant equal to c , as in conventional collinearity equations for the flat image plane. Therefore, s_z coordinate is not a constant value, as expressed by Equation 36, in which (X_0, Y_0, Z_0) are the coordinates of the perspective center in the object space and $m_{(i,j)}$ are the components of the 3 x 3 rotation matrix, obtained in function of the Euler angles $(\omega_i, \varphi_i, \kappa_i)$. However, this mathematical model is singular for points with s_z higher than zero ($s_z > 0$). The development of a bundle adjustment (either simultaneous or sequential) with this mathematical model is left as a suggestion for future research and it will not be approached in this work.

$$sx_i = sz_i * \frac{m_{11} * (X_i - X_{0i}) + m_{12} * (Y_i - Y_{0i}) + m_{13} * (Z_i - Z_{0i})}{m_{31} * (X_i - X_{0i}) + m_{32} * (Y_i - Y_{0i}) + m_{33} * (Z_i - Z_{0i})}$$

$$sy_i = sz_i * \frac{m_{21} * (X_i - X_{0i}) + m_{22} * (Y_i - Y_{0i}) + m_{23} * (Z_i - Z_{0i})}{m_{31} * (X_i - X_{0i}) + m_{32} * (Y_i - Y_{0i}) + m_{33} * (Z_i - Z_{0i})}$$
(36)

5.2.2 Stochastic treatment

Section 5.1.1 discussed the problems involving image observations geometry, which can be handled with a proper stochastic treatment based on the fisheye lens geometry. The simultaneous estimation of EOPs and 3D coordinates in the object space is performed with the unified approach of least squares adjustment and parameters constraints (MIKHAIL and ACKERMAN, 1976), which consists of a combined adjustment of observations and parameters. Indeed, the differentiation between observations and parameters (variables) in the mathematical formulation is done by the covariance, or weight, matrix of observations. For instance, if a variable weight is zero, it is varying freely in the adjustment, otherwise, if the variable weight tends to infinity, it is not allowed to change (residuals = 0), which characterizes a constant (MIKHAIL and ACKERMAN, 1976). Therefore, a stochastic treatment of the parameters and observations are required. According to Mikhail and Ackerman (1976), the variables involved in the mathematical formulation are basically, the image coordinates (x, y) , the 3D coordinates of these image observations in the object space (X, Y, Z) and the positional (X_0, Y_0, Z_0) and attitude $(\omega, \varphi, \kappa)$ parameters of each exposure station. This section suggests a stochastic treatment of these variables based on fisheye lenses geometry.

5.2.2.1 Image coordinates in the image space

Two considerations have been done in our work regarding image coordinates weighting (columns and rows standard deviations). First, fisheye image observations can be classified as a function of the nonuniform image resolution, considering the radial distance of the image point as criterium (Figure 32). Lorenzo et al. (2012) and Puig et al. (2014) considered different weights in function of the radial distance from the image center to describe the omnidirectional image geometry in the image matching process. We extend this assumption of fisheye image geometry to determine the standard deviation of image observations (σ_{OBS}) in the incremental bundle adjustment.

Figure 32 presents an experimental analysis of the quality of the images acquired by Ricoh Theta S fisheye lenses. A 360° circular target ($r = 30$ cm) composed by sequential black and white strips (each one with 4 mm-width) was used (Figure 32.a and Figure 32.b). A fisheye image covering this 360° target was acquired using the Ricoh Theta S system (Figure 32.c). It was observed that the image resolution decreases as a function of the radius (r) with respect to the image center. For instance, a 4 mm feature in the object is mapped as 4 pixels in the center of the image (1mm/1pixel). Then, as the radius increase, resolution gradually decreases to 3 pixels (1.5mm/1pixel) and 2 pixels (2 mm/1pixel) in the image periphery, showing the information compression due to fisheye imaging geometry (Figure 32.d). The loss of resolution can be observed from a gray scale profile (row 476, column 0:1080, Figure 32.c – blue line) of the 360° target (Figure 33). Figure 33.a presents this profile (blue line, Figure 32.c), in which we can notice a high distinction of the black and white strip borders at the center of the image (2) and blurring towards the image edges (1 and 3).

Figure 32. Fisheye lenses behavior: (a) Top view of the 360° target, (b) example of acquisition with Ricoh Theta S system, (c) Fisheye image (sensor 1) acquired with Ricoh Theta S in the 360° target and (d) example of nonuniformity resolution in fisheye images.

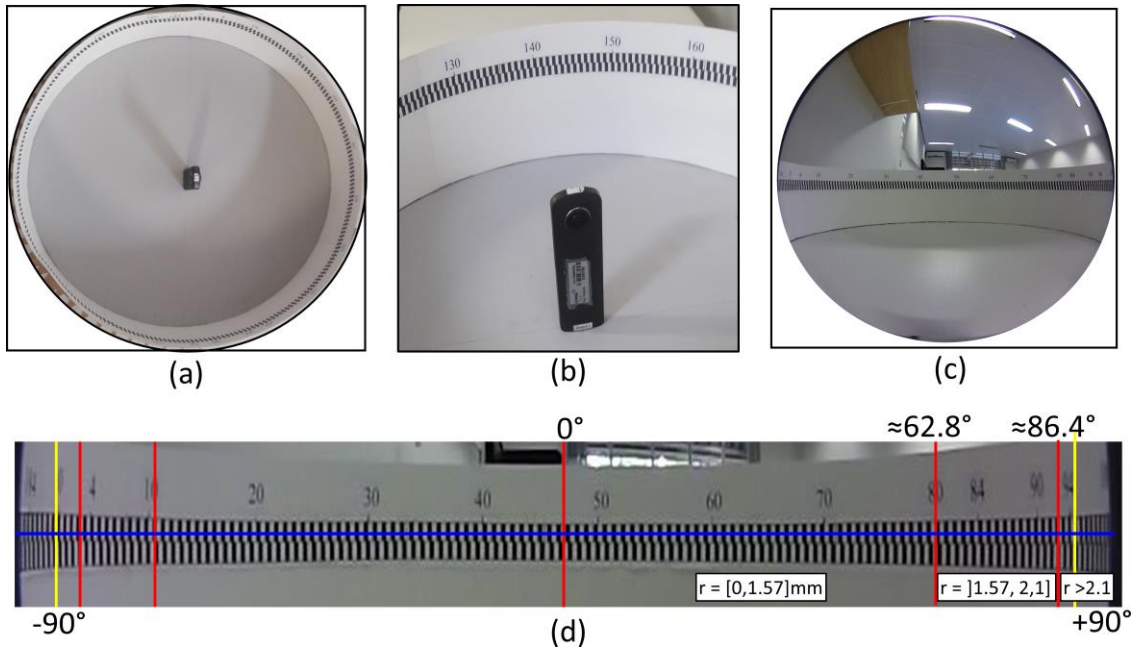
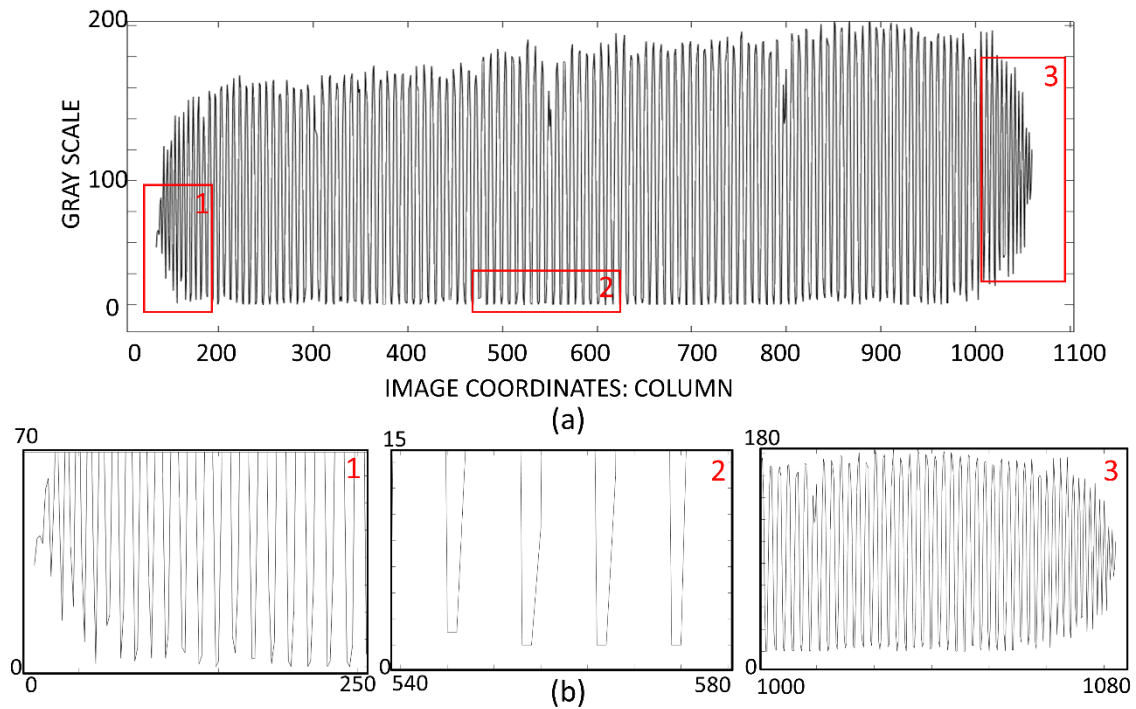


Figure 33. (a) Gray scale profile (row 476, column 0:1080) of the 360° target imaged by Ricoh Theta S system and (b) examples of image resolution considering the distinction of the black and white strip borders from the 360° target, in the center and limits of the fisheye image.



A discrete stochastic model (Equation 37) was suggested to cope with the non-uniform image resolution (Figure 32.d) and its effects on SIFT detection accuracy. SIFT technique usually achieves subpixel detection accuracy (≈ 0.5 pixel) for perspective images (LOWE, 2004). However, SIFT interest point extraction in omnidirectional images is non-optimum, especially due to the non-uniform image resolution, which has been discussed in some related works (HASEN et al., 2010; LORENÇO et al., 2012; CRUZ et al., 2012; PUIG et al., 2014). Therefore, a k value was used to weight image observations as a function of the image resolution, which depends on the radial distance from the image center (Figure 33). A reprojection error of 1 pixel for SIFT detection accuracy (σ_{SIFT}) in perspective images with a standard deviation of ± 0.5 pixels was considered (Equation 37). Then, the k value varies between 1 and 3 according to the radius range. This radius range was defined based on in Figure 32.d. As discussed by Cruz et al. (2012), SIFT detection accuracy is locally uniform, mainly in the center of the image ($r \in \mathbb{R} \mid 0 < r \leq 1.57$ mm), therefore, for a radius close to 1 mm a k value equal to 1 was used. However, as soon as wider regions of the image are considered, SIFT detection accuracy decreases, due both to the deformation that the omnidirectional sensor geometry introduces (Figure 32.d) and the loss of resolution due to blurring. Therefore, a k value of 2.0 can be assumed for the image region between the center and the intermediate radius ($1.57 < r \leq 2.1$ mm), and k equal to 3 for image border ($2.1 < r \leq 3.14$ mm). Figure 32 shows that Ricoh Theta S has a field of view larger than 180° $[-90^\circ, +90^\circ]$. As discussed in Chapter 3, for these points with a radius larger than 3.14 mm, Z has a negative value and a double mapping problem occurs, which cannot be modelled by the mathematical models proposed in Section 5.2.1. Therefore, these points were classified as outliers and eliminated as a function of the radius. Equation 37 describes the proposed discrete stochastic model to estimate the standard deviation of image observations in both components (σ_{OBS}), in which k is a value that varies according to the image resolution, which is function of the radius. The definition of continuous function is recommended for future works.

$$\sigma_{obs} = \sigma_{SIFT} + k * 0.5 \quad (37)$$

in which, $k = 1.0 \leftrightarrow r \in]0, 1.57]$; $k = 2.0 \leftrightarrow r \in]1.57, 2.1]$; $k = 3 \leftrightarrow r \in]2.1, 3.14]$.

A second criteria to weight the fisheye image observations can be defined, considering the angle between the intersection of rays in the object space, aiming the treatment of

points located far away from the sensor or near the sensor trajectory (infinity, Section 5.1.1). Since a point $P(x, y, z)$ will be calculated as an intersection of two rays, its accuracy depends on the angle (θ) between both rays, which is also known as the parallax angle. The smallest the angle θ , the weakest is the point geometry. This stochastic treatment is introduced only when the computed intersection angle θ (Equation 38) is smaller than 1° (points at the infinity, SCHNEIDER et al., 2012). Otherwise, the standard deviation is set according to the first criterium (nonuniform image resolution) previously presented. As discussed by Schneider et al. (2013), points at the infinity ($\theta < 1^\circ$) are usually excluded from the bundle adjustment. However, these points are imaged by the sensor over long periods of time, especially in the omnidirectional case. Thus, despite of their low quality in intersection geometry, these points have been proven to contribute in the indirect estimation of the cameras orientation of, especially their rotations (SCHNEIDER et al., 2012; SCHNEIDER et al., 2013; MUR-ARTAL et al., 2015). For instance, Mur-Artal et al. (2015) discussed that camera rotation accuracy can be improved by points at infinity, suggesting the incorporation of these points in the methodology as future works.

The stochastic treatment of infinity points required proper weights both for observations in the image and initial values in the object space. The angle θ can be computed (Equation 38) considering the projecting (also referred as viewing) vectors of right (u^i, v^i, w^i) and left-images (u^j, v^j, w^j), in x, y , and z directions, respectively and the norm of this both vectors (Equation 38, denominator). The projecting vectors (u^i, v^i, w^i and u^j, v^j, w^j) are estimated as presented in Equation 39, in which (x, y, z) are the image coordinates in the photogrammetric reference system (point projected to the sphere) and $m_{(i,j)}$ (i and j from 1 to 3) are rotation matrix elements for the right and left-images. If θ is smaller than 1° , a standard deviation of 3 pixels ($k=4$) for this point coordinates (column, row) in the image space was applied. However, the admitted variation in the object space should be limited by a weighted constraint, to enable conversion of the solution, which is discussed as follow.

$$\theta = \arccos \left(\frac{[u^i \ v^i \ w^i] \cdot [u^j \ v^j \ w^j]^T}{\|u^i \ v^i \ w^i\| \cdot \|u^j \ v^j \ w^j\|} \right); \quad 0 < \theta < \pi \quad (38)$$

in which,

$$\begin{aligned} u^{i,j} &= m_{11}^{i,j} \cdot x^{i,j} + m_{21}^{i,j} \cdot y^{i,j} + m_{31}^{i,j} \cdot z^{i,j} \\ v^{i,j} &= m_{12}^{i,j} \cdot x^{i,j} + m_{22}^{i,j} \cdot y^{i,j} + m_{32}^{i,j} \cdot z^{i,j} \\ w^{i,j} &= m_{13}^{i,j} \cdot x^{i,j} + m_{23}^{i,j} \cdot y^{i,j} + m_{33}^{i,j} \cdot z^{i,j} \end{aligned} \quad (39)$$

5.2.2.2 3D coordinates in the object space

Real-time solutions, mainly SLAM, are usually based on the principle that environment (ground coordinates) and sensor pose (EOPs) are unknown. In our case, an ISO approach is proposed, in which ground coordinates of some GCP and initial values for the EOPs are available. Therefore, approximate values for the 3D coordinates of the tie points in the object space (X , Y , Z) can be computed by space intersection (SHENK et al., 2005), using the measurements of the PMTS navigation system (GNSS/IMU) as EOPs. Two approaches were proposed to adapt the classic intersection space model considering that the observations are related to reference sphere surface. These approaches will be detailed in Section 5.3, as the focus of this section is to describe the stochastic properties of the variables in the bundle adjustment. We propose here a standard deviation for the initial 3D coordinates of the tie points computed with the space intersection model (Section 5.3) and the initial EOPs from PMTS acquisition (Chapter 2), considering the estimated error for the method and the GSD (Ground Sample Distance) of each observation. The GSD in this work is understood as the pixel size in object space units, as proposed by Newby (2012) for CRP. These values and their standard deviations will be used as constraints in the bundle adjustment.

First, an estimative of the precision of the 3D ground coordinates calculated by intersection using initial EOPs were performed considering the existing ground control points as reference. The discrepancies between the estimated 3D coordinates of 8 well-defined GCPs (4 for each sensor side) were compared with the 3D coordinates directly measured. A mean absolute error of 1.5 m, 1.4 m and 0.9 m were obtained for X , Y and Z components, respectively, with a resultant error of 2.3 m, for this case study. This precision depends on errors in the initial EOP values, which can change according to the equipment used and the environment. To avoid being too optimistic, a general estimated error (ϵ_r) of 3 m was considered for the proposed intersection method for each component (X , Y , Z) using the particular system. However, the observations have different intersection geometries and some of them are weak (small intersection angles). A standard deviation (σ_{XYZ}) value is proposed by considering the scale variation (GSD) in fisheye images. Equation 40 present the function proposed, in which the GSD is computed with Equation 41. The constraints on the XYZ coordinates of a point close to the sensor has a higher standard deviation in the bundle adjustment than the points far away of the sensor with weak intersection geometry. For instance, the observations with small θ ($< 1^\circ$), i.e. infinity points, cannot vary freely in the object

space, which can result in an ill-conditioned system. A maximum pixel size in object space units was considered to avoid infinity weight values ($\sigma_{XYZ} \approx 0$). Therefore, if the estimated GSD is higher than 0.8 m, this threshold (0.8) is used in Equation 40.

$$\sigma_{XYZ}^i = \varepsilon_r (1 - GSD^i); \quad 0 < GSD \leq 0.8m \quad (40)$$

$$\frac{1}{s_d^i} = |c_{red}^i| / D_i; \quad GSD^i = (s * S_d^i) 1000; \quad (41)$$

where: GSD is the pixel size in object space units (in meters) computed for each tie point (i), D is the distance in the 3D space between the camera perspective point (initial values) and the initial 3D coordinate (X, Y, Z), c_{red} is a reduced principal distance calculated for each tie point considering a spherical surface (Equation 42), S_d is the image scale (denominator) and (s) is the pixel size in mm (0.005). For instance, a GSD close to 0.01 m will result in a σ_{XYZ} close to 3 m, while, a maximum GSD of 0.8 m outcomes a minimum σ_{XYZ} of 0.6 m.

5.2.2.3 Initial Exterior Orientation Parameters

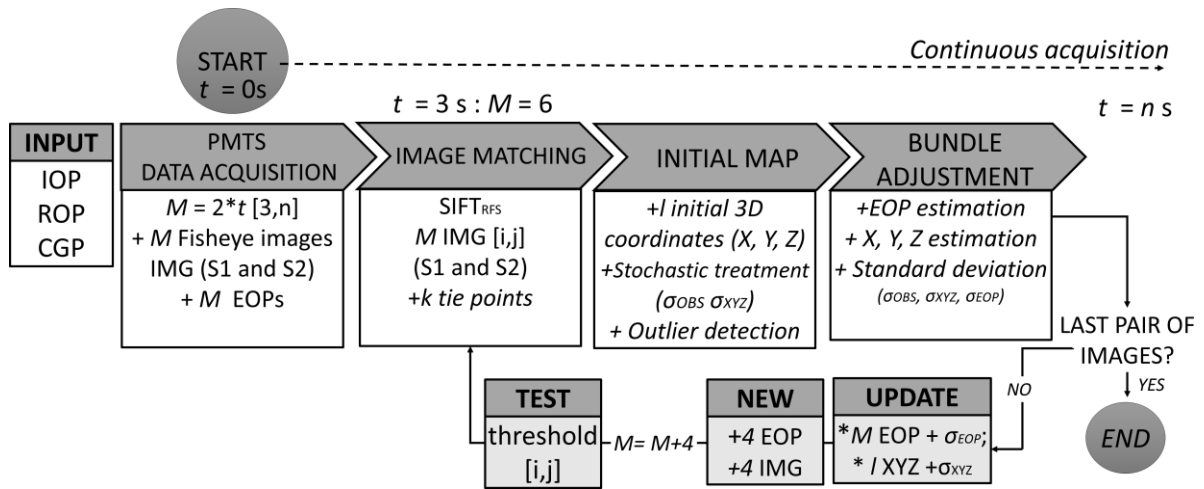
After defining the observations and their stochastic properties, the ISO approach requires the definition of the initial EOPs. The inertial navigation system data can be used as constraints with proper weights in the bundle adjustment. The GPS receiver (Ublox NEO- 6M) and the IMU (MPU6050) used in the PMTS navigation system proposed (Section 2.2.2) are low-cost systems based on MEMS technology. As well explained by Clark (2018), MEMS have a large uncertainty and unknown accuracy, requiring systematic system calibration and external accuracy evaluation. A large gap between their predicted and actual performances still exists. Usually, two MEMS have not been able to perform identically and even a same MEMS do not behave constantly in all data acquisitions (CLARK, 2018). An evaluation of the PMTS navigation system was performed in Chapter 2, comparing the PMTS navigation system performance with more stable devices used as reference. The results obtained in Section 2.4.1 was used as reference for the standard deviation applied in the incremental bundle adjustment. More details are presented in Section 5.4.

5.3 INCREMENTAL BUNDLE ADJUSTMENT METHODOLOGY

The structure of the incremental bundle adjustment proposed relies on: (1) a sequential acquisition of the data set; (2) image matching process; (3) estimation of initial 3D

coordinates of the image observation (map points); (4) outlier and inliers classification and stochastic treatment; (5) local bundle adjustment with incremental observations. This process is performed iteratively, starting with a triplet of images until the end of the trajectory. At the end point a final incremental solution is obtained. The methodology described in this section was developed focusing in real-time strategies presented in Section 5.2. Figure 34 presents a workflow of the incremental bundle adjustment methodology. The steps are detailed as follow, considering the photogrammetric processing described in Chapters 2 to 4.

Figure 34. Workflow: Incremental bundle adjustment.



5.3.1 PMTS sequential data acquisition

The proposed methodology considers that IOPs (Chapter 3), ROPs (Chapter 3) and system boresight angles are known from a previous calibration process. The location of a set of landmarks (or GCP), especially in the begging of the trajectory, or an accurate position of the starting point is recommended to enable a suitable initial solution. This external information is an input data for the incremental process. The sequential data acquisition is performed using a mobile mapping system embedded in a backpack, well known as PMTS. The PMTS used is composed by an omnidirectional camera (Ricoh Theta S) and a low-cost navigation system, which combines a GPS receiver Ublox Neo 6M and a MEMS-IMU (MPU 6050) integrated to an Arduino microprocessor. The setup of the PMTS used was described in Chapter 2. This system enables the acquisition of omnidirectional images (dual-fisheye – Chapter 3) and initial sensor pose (EOP) for each one second in a sequence of data collection, ranging in a period of time (t) between 0 (*start*) and n seconds (*end*). Therefore, a set of M fisheye images and M initial EOPs are obtained in

function of the time ($M = 2*t$). The photogrammetric process starts with $t = 3$, that means, after an image triplet has been acquired.

5.3.2 Image Matching

The image matching process starts considering an initial triplet of omnidirectional images (6 fisheye images with more than 180° , forming 3 pairs of images with 360°) and 6 initial EOPs. The SIFT technique was used to detect interest points (key points) in a subset of images between i and j from a set of M images, in which i starts in 1 and j with 3. The values of i and j change during the incremental process. A test considering the baseline and the number of matches detect is performed to define the new set of keyframes $[i,j]$. Keyframes is defined in this work as suitable image pairs for the image matching processing. This test is performed to avoid exhaustive matching; thus, the image matching process is performed only for fisheye image pairs within the defined range. However, the consistency of the geometry of the keyframes to support a strong match network configuration was not evaluated. The approaches presented by Michelini and Mayer, (2014) and Wang et al. (2009) to select keyframes with consistent geometry in perspective images can be extended to the omnidirectional case, but this step was not implemented in the proposed methodology, being left as a suggestion for future works.

A search space strategy considering the epipolar geometry on the sphere domain was combined with SIFT matching to avoid outliers. This approach was named as SIFT_{RFS}, which had been described in more details in Chapter 4. In summary, for each key point extracted in the left image a search window is defined in the right image, based on the epipolar line projection. Then, the invariant descriptor vector from the key point searched is compared only with the key points belonging to that search window. In a first moment, the search window dimensions depend on the initial EOPs uncertainty. Subsequently, a recursive adjustment is performed to improve the search window position and to reduce its dimensions. The definition of the search window and the parameters estimation used the ROPs values computed from the initial EOP. The use of ROPs parameters bring many benefits for an optimized process, for instance, there is no need of system calibration parameters estimation (i.e. boresight angles and lever arm), systematic errors are eliminated, the adjustment implementation is simplified, and the relative accuracy is higher than

absolute, especially using low-cost sensors. This iterative process was performed from each possible pair of images (keyframes) during image acquisition.

SIFT_{RFS} was applied in the two set of fisheye images from sensor 1 and sensor 2 independently. At the end of an iteration, a set of k match points coordinates in the image space are obtained, which are used as tie points in the incremental bundle adjustment. However, initial 3D coordinates of the tie points in the object space (terrain coordinates) are required.

5.3.3 Initial map (approximated values for the ground coordinates of the tie points)

An initial estimation of the ground coordinates of the tie points detected by SIFT_{RFS} was obtained with a space intersection model adapted with a mathematical model for fisheye geometry and using the data provided by the PMTS navigation system as initial EOPs. Two approaches with similar results are suggested to adapt the classical space intersection model (SHENK et al., 2005).

Considering that fisheye image observations come from an intermediary spherical surface, the c value is not constant and thus varies as a function of the point position in the sphere (in this case c cannot be considered the principal distance, but a z component). Therefore, the c value is computed for each observation in the photogrammetric system (x, y) in function of the ratio (r) and the incident angle (α), as presented in Equation 42, obtaining a reduced c value (c_{red}). This value (c_{red}) is used to compute the projecting vectors u, v, w , for the right (i) and left images (j), as presented in Equation 43 adapted from Equation 39. Then, the approximated values for the scales (λ^i, λ^j) for each match point can be estimated solving a system of 3 linear equations with two unknowns (Equation 44 to 48). If the scales are known, the x and y coordinates in the object space from the right (i) and left images (j) can be projected to the object space ($X_{proj}, Y_{proj}, Z_{proj}$) by Equations 48. The mean of $X_{proj}^i, Y_{proj}^i, Z_{proj}^i$ and $X_{proj}^j, Y_{proj}^j, Z_{proj}^j$, resulting in the 3D coordinates of the match point in the object space (X, Y, Z) computed by the space intersection of rays (Equation 44).

$$r = \sqrt{(x^2 + y^2)}; \quad \alpha = \frac{r}{c}; \quad c_{red} = r * \tan\left(\frac{\pi}{2-\alpha}\right) \quad (42)$$

$$\begin{aligned} u^{i,j} &= m_{11}^{i,j} \cdot x^{i,j} + m_{21}^{i,j} \cdot y^{i,j} + m_{31}^{i,j} \cdot c_{red}^{i,j} \\ v^{i,j} &= m_{12}^{i,j} \cdot x^{i,j} + m_{22}^{i,j} \cdot y^{i,j} + m_{32}^{i,j} \cdot c_{red}^{i,j} \\ w^{i,j} &= m_{13}^{i,j} \cdot x^{i,j} + m_{23}^{i,j} \cdot y^{i,j} + m_{33}^{i,j} \cdot c_{red}^{i,j} \end{aligned} \quad (43)$$

in which, m^{ij} are the 3 x 3 rotation matrix elements (from 1,1 to 3,3) for the right (i) and left (j)-images, calculated with the EOP angles measured with MPU 6050.

Equation 45 is used to estimate the scales (λ^{ij}), which is obtained rearranging Equation 44 according to the base elements (Bx , By , Bz). Equation 46 presents the vector of observation (Lb) composed by the base elements (Bx , By , Bz) computed from the initial EOPs, the Jacobian matrix (A) with the projecting vectors calculated by Equation 43 and the parameter vector (Xa). Equation 47 presents the solution from the least squares method, estimating approximate scale values (Xa), in which the weight matrix P is an identity matrix.

$$X_{proj}^{i,j} = X_0^{i,j} + \lambda^{i,j} \cdot u^{i,j}; \quad Y_{proj}^{i,j} = Y_0^{i,j} + \lambda^{i,j} \cdot v^{i,j}; \quad Z_{proj}^{i,j} = Z_0^{i,j} + \lambda^{i,j} \cdot w^{i,j} \quad (44)$$

$$B_x = \lambda^i \cdot u^i - \lambda^j \cdot u^j; \quad B_y = \lambda^i \cdot v^i - \lambda^j \cdot v^j; \quad B_z = \lambda^i \cdot w^i - \lambda^j \cdot w^j \quad (45)$$

$$Lb = \begin{bmatrix} Bx \\ By \\ Bz \end{bmatrix}; \quad A = \begin{bmatrix} u^i & -u^j \\ v^i & -v^j \\ w^i & -w^j \end{bmatrix}; \quad Xa = \begin{bmatrix} \lambda^i \\ \lambda^j \end{bmatrix} \quad (46)$$

$$Xa = (A^T \cdot P \cdot A)^{-1} (A^T \cdot P \cdot Lb) \quad (47)$$

$$X = \frac{x_{proj}^i + x_{proj}^j}{2}; \quad Y = \frac{y_{proj}^i + y_{proj}^j}{2}; \quad Z = \frac{z_{proj}^i + z_{proj}^j}{2} \quad (48)$$

Another possibility is to project the fisheye image observations from the plane to a sphere (Equations 30 to 32 at Section 4.2.1), computing the spherical coordinates (s_x , s_y , s_z) and then apply the space intersection model presented in Equation 43, replacing x , y and C_{red} for the variables s_x , s_y , s_z . The two approaches lead to the same results since they are mathematically compatible, but with a different syntax to describe the problem. In our study, we employed the first approach described using the C_{red} .

Initial coordinates in the object space are computed for each set of new image observation (l) from the sequential image matching process with SIFT_{RFS}. Concomitantly to this step, the stochastic treatment of the observations is performed to provide initial standard deviations (σ_{OBS} , σ_{XYZ}), as presented in Section 5.2.2. The match points with radius higher than 3.14 mm are classified as outliers and removed. Moreover, in our works, some outliers were removed by

analyzing the image coordinate residuals in the x and y directions from a previously bundle adjustment. This process was performed manually to ensure the convergence of the solution, but it will be automated in a future work, considering the method proposed to Scaramuzza et al. (2009), which uses RANSAC. The dataset cannot be considered optimal, because some outliers still remains, since only the points that caused the divergence of the solution had been removed.

5.3.4 Incremental bundle adjustment estimation

In the incremental bundle adjustment process, parameters are organized in two sets with different weights: one with parameters (EOP and 3D ground coordinates) and respective standard deviations that had been already estimated in the bundle adjustment in an earlier instance and a second one with new parameters observed in the scene in the current instance with initial values. As mentioned, the incremental approach starts with image points in a triplet of images using $SIFT_{RFS}$. Then, approximate values of 3D ground coordinates are computed for these image points using the space intersection method adapted for fisheye lens geometry. These observations, initial EOPs and the calibrated IOPs and ROPs (considered as constants), are input for the first iteration of the incremental bundle adjustment. Observations and parameters are combined in an adjustment performed with the unified approach of least squares method and parameters constraints (MIKHAIL and ACKERMAN, 1976), which enables the simultaneous estimation of EOPs and 3D terrain coordinates for these three images. The equidistant mathematical model with stability constraints in the base and rotation elements between images is used (Section 5.2.1).

The initial values are updated with the estimated values from the bundle adjustment for the EOPs, the 3D ground coordinates (X, Y, Z) and the respective standard deviations obtained from the covariance matrix. Therefore, the adjusted parameters are not set as absolute constraints after a first estimation. In the second iteration, 4 more fisheye images are added to the process with new initial information. A preliminary test is performed to identify the new keyframes. $SIFT_{RFS}$ is applied for the new keyframes, and the initial map and the set of EOPs are now composed by adjusted and initial values. The normal equation matrix and the covariance matrix receive the new measurements, and the bundle adjustment is performed again. Parameters and respectively standard deviation are update and 4 new images are sequentially added to the block. This incremental process is performed until the end of the dataset.

5.4 EXPERIMENTAL ASSESSMENTS

The feasibility of the methodology proposed in Section 5.3 is assessed simulating a real-time photogrammetric process. This section describes the dataset used (Section 5.4.1), the setups applied for SIFT_{RFS}, initial map estimation and bundle adjustment (Section 5.4.2) and the accuracy assessment of the positional trajectory and the estimated 3D ground coordinates (Section 5.4.3).

5.4.1 Dataset

A sequential data acquisition with the proposed PMTS was performed in a test area located inside the São Paulo State University (UNESP) at Presidente Prudente (22° 07' S, 51° 24' W). This area is a narrow path covered with sparse vegetation areas (*Eucalyptus* trees and ground vegetation - grass) and urban features (small building, light poles and traffic signs). A 10 m trajectory was traversed, resulting in a set of 28 images (14 dual-fisheye images) and 28 groups of EOPs from the PMTS navigation system. The PMTS data was synchronized as detailed in Section 2.2.3. Therefore, each fisheye image has its position and attitude obtained by the navigation system. The pixel size in object space units of the fisheye images in this dataset range between 1 and 80 cm and the initial navigation data were obtained with a planimetric accuracy (latitude and longitude) ranging from 7 cm to 1 m, with an average discrepancy of 34 cm. Despite of the consistency relative accuracy between positional measurements, an absolute error varying between 1 and 5 m is usually obtained in height with Ublox NEO-6M, thus, an initial correction ($\Delta_z \approx -5\text{m}$) was applied in the heights of all camera positions (Z_0) estimated with Ublox NEO-6M. The Δ_z value (constant) was computed considering the height of GCPs in the begging of the trajectory aligned with the camera center position. This correction does not affect the image matching process, which uses the relative orientation, but it is enable a better estimation of the initial 3D terrain coordinates and EOPs closer to an effective camera position. The Ublox NEO-6M data were acquired in NMEA format and converted to UTM (E, N, h), which is compatible with the coordinate system that is used for GCPs.

A set of 12 GCPs (6 for each sensor) were obtained from an existing terrestrial LASER point cloud (collected with Leica ScanStation P40). A standard deviation of 5 cm was set for the GCPs in the object space which is compatible with the point cloud accuracy and resolution. The 3D ground coordinates were measured in the LASER point cloud using Cloud Compare and

Fugro Viewer software. All GCPs were identified and manually measured in the images. The coordinates in the image and object spaces were previously informed for the incremental process. The 12 GCPs were located in the beginning of the trajectory. At least 9 GCPs were imaged by the first triplet of images, but along the trajectory the visibility of the reference points decreased until the total absence of GCPs. At the end of the trajectory a global bundle adjustment was performed with 6 check points measured in the last six images to estimate the accuracy of the 3D point cloud generated (Section 5.4.3). The coordinates of the check points used as reference were also measured in the LASER point cloud (Leica ScanStation P40).

Another input data is the interior orientation parameter and the relative orientation parameter, connecting the two Ricoh Theta S sensors. These parameters were estimated in a previous camera calibration process (Chapter 3), which is presented in Table 16 and Table 17, respectively. The fisheye images from sensor 1 (S1) and sensor 2 (S2) were rotated (-90° , $+90^\circ$) in relation to the original dual-fisheye images used in the Ricoh Theta S calibration (Figure 13). Therefore, the perspective center values were changed according to this rotation, which can be noted comparing the values from Table 8 and Table 18. The linear and angular misalignments (mounting parameters) between sensors are also required to fully calibrate the system. The distance between the camera projection centers and GPS antenna phase center (lever arm) is smaller than the measurement error of the GPS receiver used. Consequentially, the standard deviation set in the bundle adjustment is higher than lever arm correction. Therefore, these linear offsets were not considered to be significant. The misalignments between the IMU and cameras (boresight angles) were estimated as presented in Section 2.2.3 (Equation 1 and Equation 2). These boresight angles are constantly re-estimated because they change according to the system assembly. The angular misalignments used for this acquisition are presented in Table 19.

Table 18. Principal distance (c) and principal point coordinates (x_0 , y_0) with their respective standard deviations for Sensor 1 and Sensor 2 as used in the bundle adjustment.

IOPs	SENSOR 1			SENSOR 2		
	c (mm)	x_0 (mm)	y_0 (mm)	c (mm)	x_0 (mm)	y_0 (mm)
Value	1.4307	-0.3013	-0.0019	1.4297	0.2935	-0.0040
+ σ	± 0.001	± 0.0001	± 0.0001	± 0.001	± 0.0001	± 0.0001

Table 19. Angular and positional offsets in the Personal Mobile Terrestrial System.

OFFSETS	ΔX (mm)	ΔY (mm)	ΔZ (mm)	$\Delta\omega$ (°)	$\Delta\varphi$ (°)	$\Delta\kappa$ (°)
S1/Platform	0	0	0	-85.79	-20.61	174.14
ROP	-0.79	0.41	19.04	0.0149	0.0526	179.66

5.4.2 Photogrammetric process setup

A simulation of a real-time photogrammetric process was performed considering the dataset described-above and the workflow defined in Section 5.3. The setups of each step are presented in this section.

In the image matching process, $SIFT_{RFS}$ was set with a threshold of 0.5 for the Euclidean distance between the nearest and second nearest neighbor key points, which is used as criterium to accept a match (LOWE, 2004). Besides that, the depth of each tie point in the object space ($Z_{(i)}$) is unknown, thus a set of planes is required for the epipolar line (curve) estimation, as described in Section 4.2.1 (Equations 33 and 34). The number and range of the planes depends mainly on the dataset features, which need to be tested and defined a priori and it can be considered a limitation of the $SIFT_{RFS}$. A $Z_{(i)}$ ranging between 1.1 m and 10.1 m with an interval of 0.1 m (distance between consecutive planes) were considered suitable for this dataset. As final considerations for this step, the same set of EOPs without boresight correction was used for Sensor 1 and Sensor 2, because only the relative orientation between the keyframes is used in the proposed image matching approach. The search space window (Section 4.2.2) is computed according to the error propagation of the ROPs. A relative accuracy of 1° for the angles and 0.2 m for the position was set.

The initial map is calculated using the matches from $SIFT_{RFS}$, the IOPs from Table 3 (constants) and two set of EOPs corrected from the linear and angular misalignments (Table 17). The initial values for the EOPs of sensor 2 were computed with the EOPs from sensor 1 and the ROP between them. In this step, the EOPs provided by the PMTS navigation system are considered as constant (absolute value). The initial standard deviation (σ_{XYZ}) of the 3D terrain coordinates estimated in this step will range between 3 m and 0.9 m, whilst a pixel size in object space units of the fisheye images range between 1 and 80 cm.

A very similar configuration is used in the bundle adjustment. Nevertheless, the 3D terrain coordinates and the EOPs are parameters constraints to be imposed. A initial standard deviation of 10° for the angles ($\sigma_{\omega\varphi\kappa}$) and 0.3 m for the camera position ($\sigma_{X0,Y0,Z0}$) were applied.

The image observations were weighted as described in Section 5.2.2. The convergence criterion for the sequential bundle adjustment considers a parameters correction smaller than 0.0001 and a maximum number of iterations (<50). The image matching and the initial map estimation were developed in MATLAB scripts, while, the incremental bundle adjustment was performed in the in-the-house software (CMC) implemented in C/C++ language.

5.4.3 Quality Control: EOPs and 3D Map accuracies

The performance assessment of the incremental bundle adjustment using PMTS data was done mainly focusing on the accuracy of the initial information provided, the consistency of the bundle adjustment, the positional accuracy of the trajectory and the accuracy of the 3D coordinates map.

The bundle adjustment consistency was analyzed considering the residuals in the image coordinates (match points), the a posteriori standard error of unit weight (σ_{naught}) and the estimated precision (standard deviations) of the EOPs and 3D ground coordinates (X, Y, Z) obtained from the average values estimated in the covariance matrix. The correlation between the number of observations and 3D map precision is also discussed.

The positional accuracy of the trajectory was estimated comparing the results from the incremental bundle adjustment with a reference (ground truth). During the traversing, the trajectory was simultaneously tracked with a Topcon HIPER SR (L1/L2) GNSS receiver installed in the backpack platform. However, as discussed in Chapter 2, the precision of the GNSS measured under the vegetation is reduced, resulting in large gaps in the data and very noisy solution. Therefore, the ground truth was obtained from a consistent bundle block adjustment performed in PhotoScan with the same set of fisheye images and GCPs well-distributed along the trajectory. The reprojection error in this estimation was smaller than 1 pixel and the RMSE of the GCP is 0.096 m. The cameras positions from the reference and the bundle adjustment estimative were compared to estimate trajectory accuracy. The accuracy of the initial values provided by the PMTS navigation system was also evaluated considering the PhotoScan data as reference.

The accuracy achieved with the incremental bundle adjustment on the 3D ground coordinate were assessed by considering the discrepancies between the values estimated in the incremental bundle adjustment and the reference coordinates of the 6 checkpoints. The 3D ground coordinates of the checkpoints used as reference were also measured in the terrestrial LASER point

cloud with centimetric accuracy. After 10 m of traversing, a global bundle adjustment including the check points was performed. The check points are located at the end of the trajectory and they were measured in the last three images to estimate the accuracy of the 3D point cloud generated and the solution drift, considering the sequential reduction of GCPs from 12 to 0.

5.5 RESULTS

The results of the incremental bundle adjustment process described in Section 5.4 are presented in this section, evaluating the accuracy of the initial information provided by the PMTS navigation system (Section 5.5.1), the consistency of the bundle adjustment (Section 5.5.2) and the positional accuracy of the trajectory and the 3D coordinates map (Section 5.5.3). The reference values were generated in a simultaneous bundle adjustment with PhotoScan.

5.5.1 Performance assessment of PMTS initial values

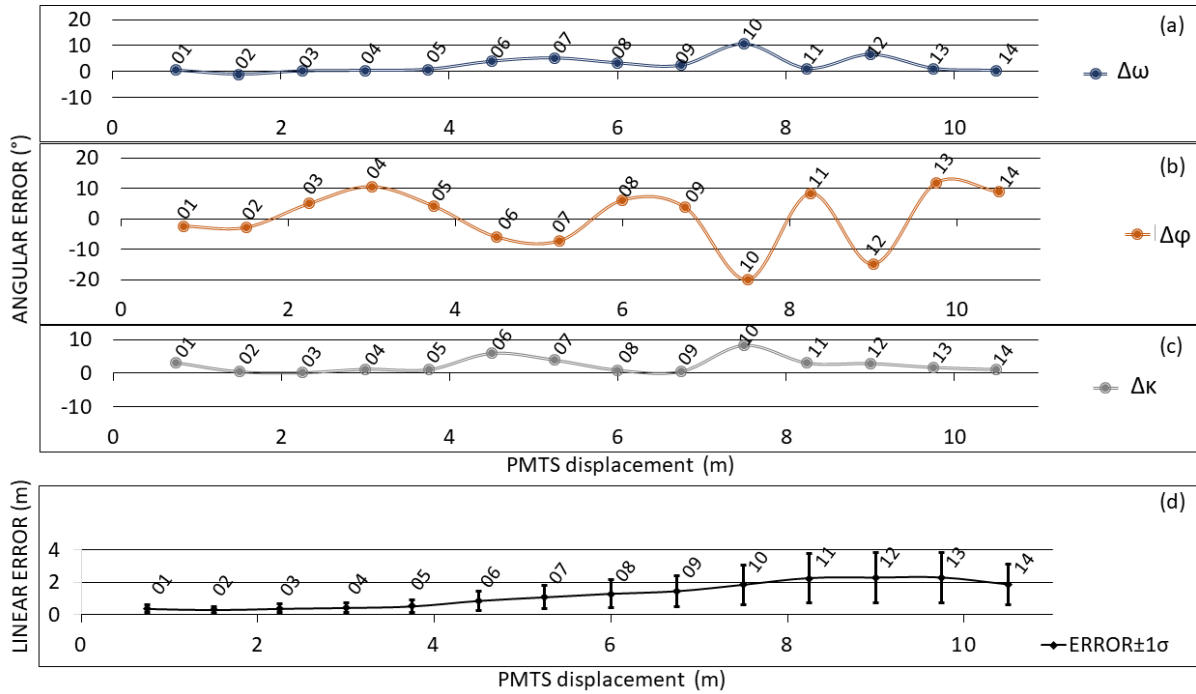
Figure 35 presents an initial assessment on the accuracy of the navigation system of the PMTS for this dataset. The values measured by the GPS receiver and the MEMS-IMU were compared to the ground reference computed with PhotoScan. Despite of the correction of misalignments angles (boresight), no additional treatment in MPU6050 data was performed. The angular discrepancies in omega, phi and kappa ($\Delta\omega$, $\Delta\phi$, $\Delta\kappa$) between MPU6050 and PhotoScan datasets were computed from the rotation matrices, as presented in Equation 1 and 2 (Chapter 2). MEMS-MPU6050 provides high level of errors and noise. Thus, the greatest source of error in the initial estimation of sensors pose and 3D maps comes from the attitude, especially after 6 m of acquisition (sensor pose 10), where the IMU solution start to drift significantly (Figure 35.a to Figure 35.c). After this point, the MPU6050 data presented errors higher than the expected standard deviation ($>10^\circ$) for the ϕ angle estimation (Figure 35.b), which was understood as noise caused by the high sensitivity of this sensor. In Chapter 2 a moving-average filter was applied in a postprocessing step to minimize the outliers (Table 4), but in this experiment the raw data was used to simulate a real-time application.

Ublox GPS receiver presented a suitable performance, with a RMSE smaller than 1 m for each component, for instance, 0.91 m and 0.47 m for X and Y, respectively. A maximum resultant error (X, Y and Z) smaller than 3 m (Figure 35.d) was obtained, which was expected for trajectory in covered areas. As previously assessed in Chapter 2, Ublox GPS receiver provides a

trajectory under vegetated areas with meter level accuracy. The GPS accuracy decreases along the trajectory because of the increasingly density of the vegetation, which causes the loss of satellites observations, multipath errors and poor satellite geometry (NAESSET, 2001).

Despite of the low-accuracy of the GPS/IMU measurements obtained, these initial values, especially the platform position, combined with the GCPs in the initial image triplet provide suitable information for the solution convergence (Table 18), reducing the systematic errors. Therefore, the GPS/IMU errors are minimized by the existing GCPs at the beginning of the trajectory, which improve the next sensor position estimative during the incremental process.

Figure 35. Estimated angular and linear errors of MPU6050-IMU and UBLOX NEO 6M receiver for sensor pose (1 to 14) along the PMTS trajectory, in which (a) to (c) show the angular errors ($\Delta\omega$, $\Delta\phi$, $\Delta\kappa$) independently for ω , ϕ and κ angles, respectively. And (d) presents the resultant of the linear errors (X, Y, Z) following to the PMTS displacement (reference values came from a simultaneous bundle adjustment with PhotoScan).



5.5.2 Incremental bundle adjustment results

Table 20 present some statistics of the incremental bundle adjustment for each step at an instance (t), such as, the number of images and observations used, degrees of freedom (DoF), the a posteriori standard error of the a unit weight ($\hat{\sigma}_0$), the RMSE of the image coordinates residuals from SIFT_{RFS} (x and y directions), the average standard deviation (precision) of the

estimated 3D ground coordinates and the average standard deviation of the estimated EOPs, both from the covariance matrix. The estimated standard deviation of EOPs and 3D ground coordinates, presented in Table 20, were computed considering only the precisions of the parameters of the new images inserted in the incremental bundle adjustment process in the respective instant (t).

Table 20. Incremental bundle adjustment statistics for each step in an instant (t).

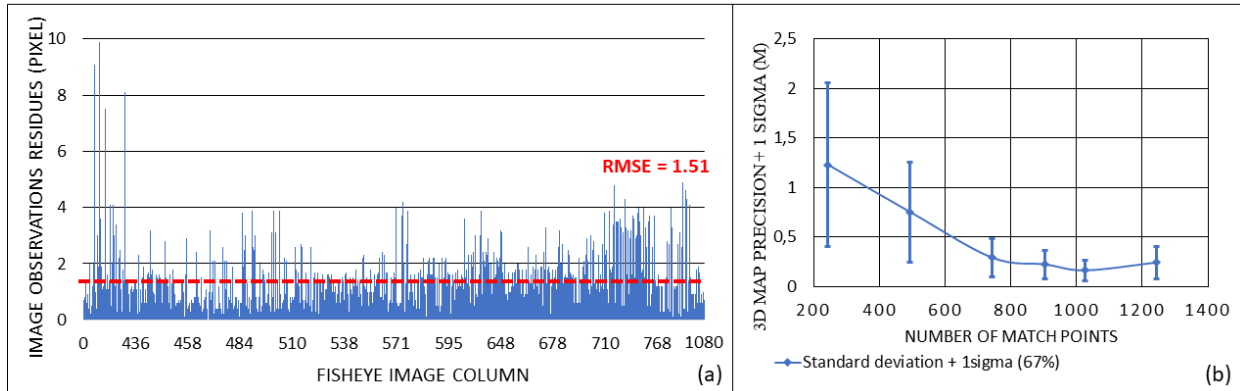
t (s)	N_{IMG}	New IMG	N_{OBS}	DoF	$\hat{\sigma}_0$	RMSE _{obs} (pixel)	$\sigma^{3D \text{ MAP}}$ (m)	σ^{EOP} (°)	σ^{EOP} (m)
3	6	-	488	482	0.673	1.64	1.23	0.429	0.045
5	10	4	988	1000	0.671	1.45	0.91	0.189	0.020
7	14	4	1484	1508	0.604	1.51	0.50	0.146	0.014
9	18	4	1808	1844	0.557	1.42	0.55	0.181	0.013
11	22	4	2054	2102	0.785	1.50	0.46	0.235	0.029
14	28	6	2488	2566	0.589	1.51	0.91	0.090	0.012

A set of 1244 match points (2488 image observations) from SIFT_{RFS} were used in the incremental bundle adjustment. From this dataset, 2% were classified in the stochastic treatment as infinity point (22), 20% had a radial distance higher than 1.57 mm (250 points) and 78% were located in the center of the image with a radial distance ranging between 0 and 1.57 mm. The final RMSE of the image coordinates residuals obtained in the incremental bundle adjustment process was 1.51 pixels (Figure 36.a), with a $\hat{\sigma}_0$ of 0.59 pixels. The final $\hat{\sigma}_0$ value was obtained considering a high DoF and absolute constraints in the parameters at the beginning of the trajectory due to the incremental process and weight updates. Therefore the $\hat{\sigma}_0$ obtained is not compatible with the final RMSE of the image coordinates residuals.

Figure 36.a shows the image coordinates residuals distribution along the x-axis of the fisheye image (Columns). The higher residuals are located in the image borders, where image distortions are larger and the resolution is poorer. Figure 36.b presents the average standard deviation of the estimated 3D ground coordinates as a function of the number of matches used in the incremental bundle adjustment, showing that the precision increases as a function of the number of match points used. This assumption is confirmed when the same sensor pose is estimated with different number of matches, since the estimated error decreases when a greater number of matches are included, as presented and discussed by Valiente et al. (2017). However, the processing time increased significantly with the introduction of more observations, showing that this topic deserves a previous assessment to determine an optimal number of matches from SIFT_{RFS}.

considering precision versus processing time. Furthermore, a higher DoF can provide an overestimated sigma posterior ($\hat{\sigma}_0$).

Figure 36. (a) residuals in the image observation and (b) standard deviation from the covariance matrix of 3D ground coordinates as a function of the number of SIFT_{RFS} match points in the incremental bundle adjustment.



For instance, the standard deviation of the 3D ground coordinates and the EOPs estimated from the covariance matrix (precision) can be considered overestimated (Table 20). The $\hat{\sigma}_0$ was computed considering a high DoF. The DoF increases during the incremental process according to the increase number of images included and consequently observations (match points). Then, the covariance matrices in each instant (t) were computed considering a non-realistic $\hat{\sigma}_0$ value and the precision increases (Figure 36.b). However, the estimated precisions along the trajectory were consistence with the expected error behavior. Regarding the 3D map precision, a lower precision for the estimation of the 3D ground coordinates ($\sigma^{3D\text{ MAP}}$) was obtained in the end of the trajectory (close to 1 m) with the addition of the last 6 images (3 image pair for each sensor) with a reduced number of GCPs. Furthermore, the angular error presented in Figure 35, especially after sensor attitude in the instant $t=10$, was detected and weighted with a higher standard deviation. Therefore, the proposed method was able to predict the error and give a suitable weight in the estimation process, recovering the solution, which was also noticed in the estimation of sensor position (Figure 39, Section 5.3.3).

Figure 37 shows (a) the predicted error of Ricoh Theta S attitude ($^\circ$) and (b) Ricoh Theta S position (m) as a function of the PMTS displacement. The camera position and attitude errors have the same behavior, which can be associated not only with the initial weighted value from GPS/IMU sensor but mainly to the geometric quality of the image observations acquired at each sensor pose, since SIFT_{RFS} performance is directly affected by IMU noises.

Therefore, in a more detailed evaluation, it is possible to observe that the match points (image observations) in the images acquired at instance 7, 10 and 11 were almost aligned and poorly distributed throughout the image (weak geometry), which can be considered the main reason for the drift of the incremental solution. Figure 38 illustrates the weak geometry of the keyframes formed with images 10 and 11 acquired with sensor 2.

Figure 37. Estimated standard deviations (precision) of ω , ϕ and κ angles from the covariance matrixes computed during the incremental bundle adjustment process (instance 1 to 14) accordingly to PMTS displacement.

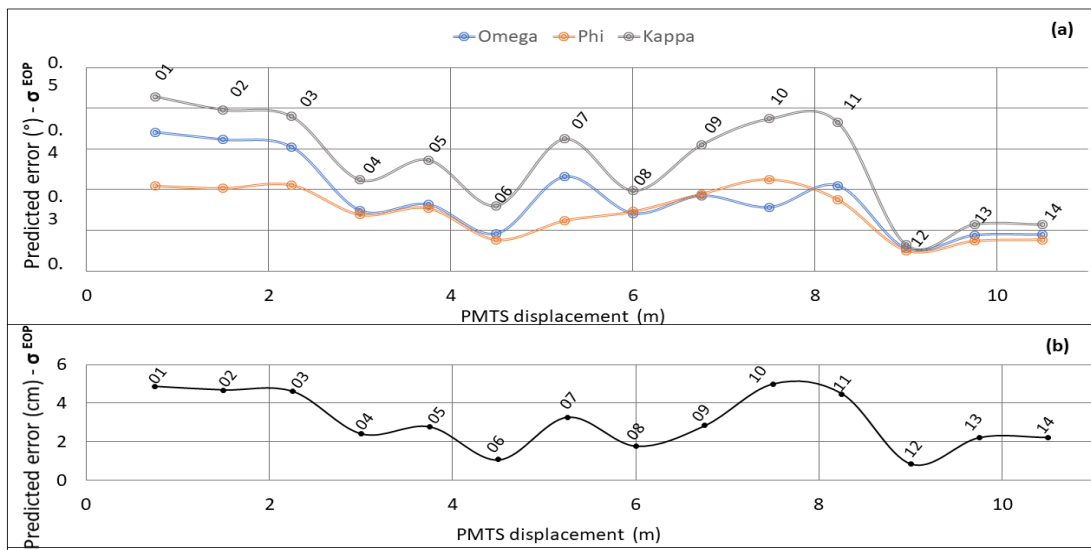
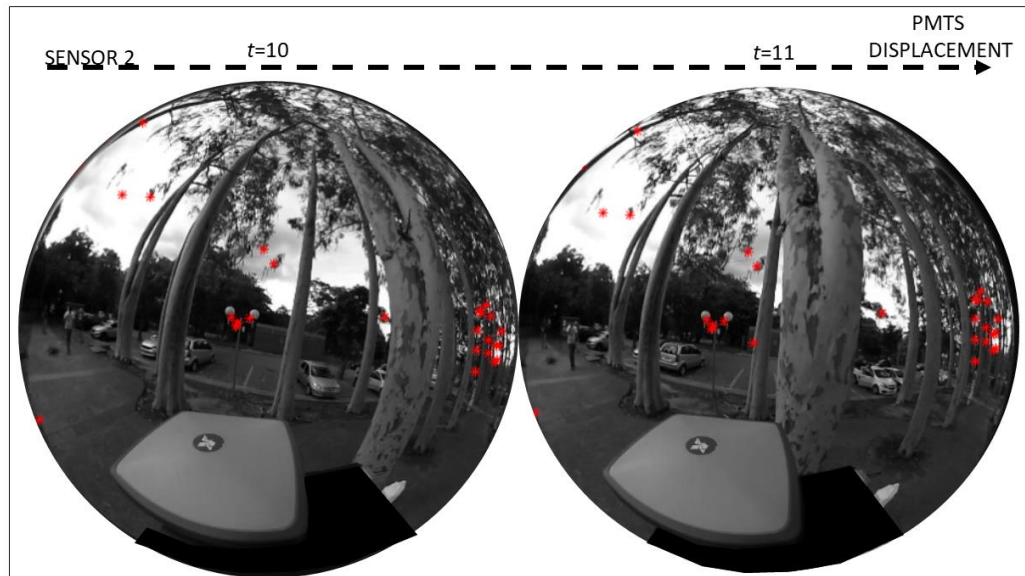


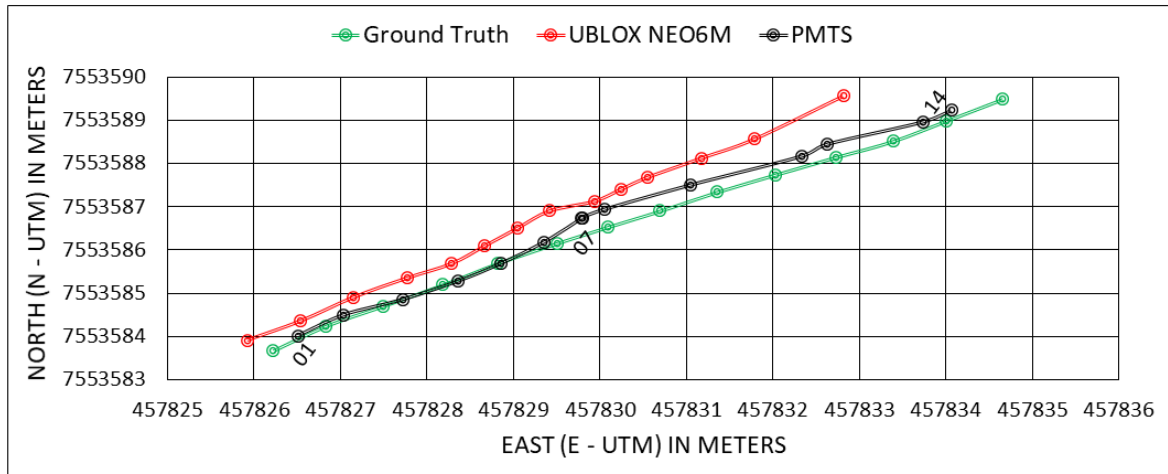
Figure 38. Example of weak geometry distributions of match points in the keyframes composed by the images acquired by sensor 2 in the instances 10 and 11, which presents higher errors in the EOPs estimative with the incremental bundle adjustment process.



5.5.3 PMTS trajectory and the 3D ground coordinates accuracy

The accuracy of the camera positions (trajectory) was estimated by comparing the results from the incremental bundle adjustment with reference values (Bundle adjustment using PhotoScan), as previously mentioned. Figure 39 presents the trajectory estimated with the Ublox GPS receiver (red), the estimated trajectory with the bundle adjustment (black) and the ground reference (green). The trajectory was estimated with a positional accuracy (RMSE) of 0.59 m, 0.20 m and 0.21 m for E, N and h respectively, resulting in a planimetric accuracy of 0.627 m and planialtimetric accuracy of 0.661 m.

Figure 39. Sensor trajectory estimated by PhotoScan (Ground Truth), Ublox GPS receiver and in the incremental bundle adjustment using PMTS data along the trajectory.



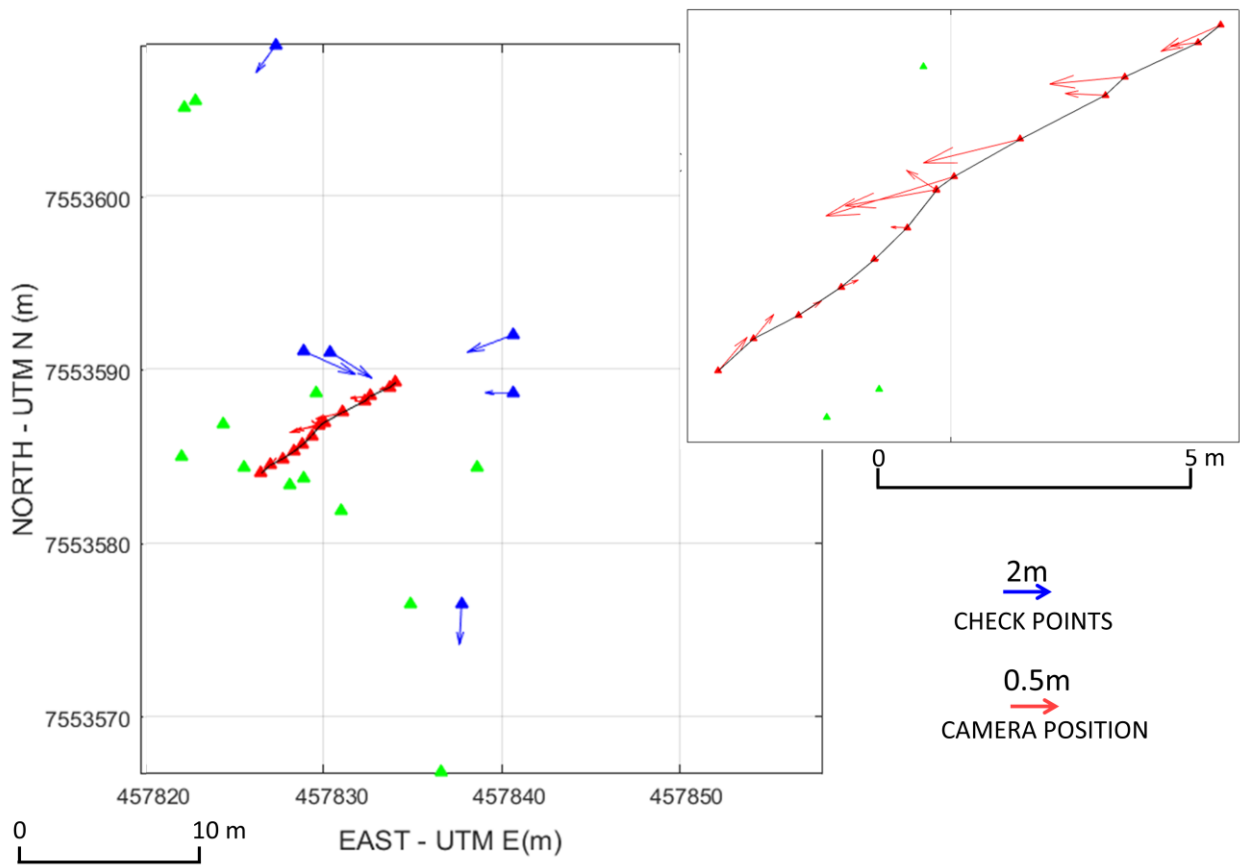
The accuracy of the 3D ground coordinate was assessed with 6 checkpoints, specially positioned at the end of the trajectory (last 6 images). Table 21 presents the statistics, including the mean ($\bar{x}$), standard deviation (σ) and RMSE of the discrepancies between the 3D coordinates of the 6 check points measured in a LASER point cloud (reference) and the estimated coordinates at the end of the bundle adjustment. Therefore, the 3D terrain coordinates (E, N, h) of the points detected with SIFT_{RFS} (tie points) were estimated at the end of the trajectory with an average planimetric accuracy of 1.6 m.

Table 21. Statistics: mean, standard deviation and RMSE of the checkpoints.

Statistics	E (m)	N (m)	h (m)	Planialtimetric (m)
$\bar{x}$	-0.018	-0.865	0.416	-
σ	1.475	0.511	0.316	-
RMSE	1.347	0.982	0.506	1.74

Figure 40 shows the behavior of the residuals obtained in the sensor position estimation (red) and for the six checkpoints, according to the PMTS trajectory. The planimetric error vector (magnitude in scale) and directions are presented as a function of E and N coordinates. The GCPs used in the incremental bundle adjustment are illustrated in green and the trajectory also zoomed (b).

Figure 40. Residues distribution of the camera positions (red) and the check points (blue) along the PMTS trajectory (black) estimated by the incremental bundle adjustment. The GCPs are illustrated in green.



These results presented in Figure 39 and Figure 40 show the drift of the solution along the trajectory caused by different sources of errors. First of all, the MPU5060 data was very noisy after 6 m (after sensor pose n. 7) of PMTS displacement, as presented in Figure 35. The error in the initial EOPs directly affected SIFT_{RFS} performance (due to the reduced search space) and the initial map estimation. Regarding SIFT_{RFS} performance, it can be noted in Figure 36.a that the reprojection errors of image observations are smaller than 2 pixels, which reflects a suitable

location accuracy of the FBM method. However, the error in the initial EOPs can affect the precision of the epipolar line projection, which usually results in a poor geometric distribution of the match points in the keyframes. This problem occurs mainly in the images acquired after sensor pose in the instant $t=7$. Therefore, the keyframes formed by these images hinder the solution. Furthermore, the space intersection model used to compute the initial 3D ground coordinates of the tie points (Section 5.3) depends directly on initial EOPs accuracy, which decreases in the end of the trajectory. At the end points (instant $t=10$ to $t=14$), the number of GCPs is reduced, for instance 2 GCPs for each sensor side (maximum), which hinders a more accurate solution in the incremental process. Despite of the GPS/IMU noise and the weak geometry of some images, the proposed methodology was able to weight the observations (σ_{OBS}) and the initial values for the parameters (σ_{XYZ} , σ_{EOP}) properly in the bundle adjustment and recover the solution, as presented in Figure 39 (b). For instance, the residuals in the sensor position start to increase between instant $t=7$ and $t=10$ (nearly to 1 m in E), since some keyframes with weak geometry were included. However, the solution for instants $t=11$ to $t=14$ are close to the reference values with a planimetric error of 0.5 m, approximately, which is consistent with the average trajectory accuracy (RMSE) estimated.

5.5 DISCUSSION AND CONCLUSIONS

Section 5.1 discussed the open problems in real-time close range photogrammetry with omnidirectional sensors, focusing in omnidirectional image matching and bundle adjustment geometry. Based on this discussion, the proposed incremental bundle adjustment methodology used a mathematical model consistent with fisheye lens geometry combined with a stochastic treatment based on the non-uniform resolution of the fisheye images and the geometry of ray intersection (infinity points). An experimental assessment was performed to evaluate the feasibility of the proposed methodology using the PMTS data. The results showed that the use of low-cost embedded system, such as the proposed PMTS, for outdoor close range applications is viable and promising. However, many improvements are still required, which will be detailed in Chapter 6

The PMTS trajectory was estimated with the incremental bundle adjustment with a planimetric accuracy (RMSE) of 0.63 m and planialtimetric accuracy of 0.67 m. This accuracy obtained with the PMTS data in a real-time simulation is consistent with other recent works that also proposed low-cost real-time embedded systems. For instance, Surh et al. (2017) achieved a

sensor pose accuracy around 0.7 m, using a low-cost GNSS/IMU sensor, a wheel speed sensor and a single front camera. Chen et al. (2018) combined a low-cost GNSS receiver with an omnidirectional camera for real-time outdoor applications, achieving a sensor pose accuracy of 0.6 m when the GNSS positioning accuracy is close to 1 m. However, as discussed by these authors, if the GNSS receiver's error is higher than 3 m, the trajectory error estimated in the real-time bundle adjustment is usually higher than 0.7 m, which is often the case in vegetation areas, such as in the dataset used.

A RMSE of 1.74 m was obtained in the independent check points, which express the 3D map accuracy and the drift of the solutions at the end of the trajectory. Regarding the precision of the low-cost sensors used, the result obtained meets the expectations, especially considering a real-time incremental process, since new observations are applied mainly to estimate the current sensor pose without refining significantly the previous ones. The trajectory drift and the accuracy of the 3D map can be improved in a final global bundle adjustment including GCPs at strategic points, which shows the importance of a consistent self-diagnostic to improve the PMTS solution. An extended discussion and future research directions are presented in Chapter 6.

CHAPTER VI

FINAL REMARKS

In this work a PMTS with omnidirectional system for forest and urban mapping was proposed. A performance assessment of this system focusing in the photogrammetric process using fisheye images was presented. This final chapter discuss the open problems for future works based on the achieved results. The PMTS provided centimetric accuracy for object reconstruction and trajectory estimation in a postprocessing approach (Chapter 2). In a real-time application, a meter level accuracy for object reconstruction and a trajectory accuracy of 0.66 m was obtained. Therefore, passive optical systems can be considered a feasible alternative in covered areas, where GPS signal is not often available. However, system features (Chapter 2) can be improved (remarks 1 to 3).

(1) Some concerns about the imaging system used (Ricoh Theta S) remains. The 8s time gap between successive still images with the Ricoh Theta S prevents the dynamic sequential acquisition of images, leading to the decision of using frame videos with lower resolution in our system, which impaired the imaging range of the PMTS. Other dual systems can be assessed to improve the proposed PMTS and the data acquisition procedure. Furthermore, in the study of the Ricoh Theta S inner geometry (Chapter 3), it was noticed that the field of view of each camera covers more than 180° (the same targets appear in images of both sensors), which was not well-modeled by the proposed mathematical model. A suitable treatment of this geometry is required.

(2) Remarking the navigation system used (Chapter 2), MEMS measurement still provides high level of errors and noise. The MEMS-sensors uncertainty must be considered in the stochastic model. Especially aiming real-time application, these errors can result in a quick drift of the solution. Higher quality navigation systems should be assessed in future improvements of this system.

(3) The proposed PMTS was assessed mainly in outdoor environments. Therefore, the data acquisition methodology can be evaluated for other datasets and the PMTS can be modified, aiming for instance, indoor or agriculture applications. Other sensors, such as TLS, can be incorporated to the PMTS platform, resulting in many possibilities of multi-sensor systems and data processing methodologies.

The classical photogrammetric processes were extended from the perspective to omnidirectional case to improve PMTS results (Chapter 3 to 5). A recursive feature search method for fisheye image matching (SIFT_{RFS}) was presented in Chapter 4. Chapter 5 presented a methodology for incremental bundle adjustment to evaluate the potential of the proposed PMTS in real time CRP applications. Many improvements are still required to automate the PMTS process for robust real-time applications. The main challenge is to reach a strong network of matches to the incremental bundle adjustment, using automatic image points measurements considering fisheye image geometry, which open many research opportunities, as following (remarks 4 to 8).

(4) FBM techniques, such as SIFT, provided a minimum number of conjugate points detected in a sequential pair of fisheye images (defined in this work as repeatability). The low repeatability can be explained by the significant and quick appearance change of the features in function of the PMTS displacement using fisheye images, affecting key point stability and match comparison. The appearance changes can be mitigated with a local rectification, as initially proposed by Daniilidis (2002). For instance, fisheye image templates can be projected from the sphere surface to a vertical plane, considering different distances of projection. These templates can be defined using the proposed epipolar line technique.

(5) The evaluation of the consistency of the geometry of the keyframes to support a robust point network configuration is required. In our work, a range of fisheye images was defined according to the baseline and a minimum number of matches to avoid exhaustive matching. However, the detection and exclusion of poor or degenerated camera configurations was not performed. Usually in the perspective case, a poor camera configuration results from a small ratio between the baseline length and the depth of the scene, as discussed by Micheline and Mayer (2014). The challenge for fisheye images comes with the huge depth and scale variation. For instance, a keyframe with optimal Base/Depth ratio can provide a low number of matches, and thus, a criterion to analyze the geometric consistency of keyframes needs to consider the specific fisheye features. An approach considering graph-based optimization is suggested.

(6) The presence of outliers and the weak geometry of match points are point out in related works as the main reasons for the drift of the bundle adjustment solution (SCARAMMUZZA et al., 2009). Automatic outlier detection is a complex task in omnidirectional systems. Different sources of error can be mentioned from the image matching process. The number of false matches detected using FBM techniques is larger in fisheye images than those

usually achieved on standard perspective images, due to the huge variations both in illumination and appearance. Furthermore, points with weak geometry, poor intersection of rays or in the infinity, i.e. clouds (especially in outdoor applications), need to be identified. In our work, the outliers were manually detected based on the image coordinate residuals in the x and y directions from a previously postprocessing bundle adjustment. Therefore, the dataset used in Chapter 5 can be considered optimal and further experiments considering an automatic selection of inliers and outliers for the bundle adjustment is recommended.

(7) The proposed methodology for the photogrammetric process focused in the fisheye lens geometry. The mathematical model used in the camera calibration process and in the incremental bundle adjustment was based on the equidistant projection. Considering the experiments performed with the most popular fisheye lenses models, the equidistant equation best fits for Ricoh Theta S. However, this mathematical model (equidistant) does not describe the geometry of all observations, since some image compressions were noticed (Chapter 5), requiring further investigations. The methodology proposed can be extended to other omnidirectional sensors considering other fisheye lens equations. Furthermore, mathematical models based in the collinearity condition in the sphere domain (Equation 36) and general mathematical model as proposed by Kannala and Brandt (2006) and Urban et al. (2017) should also be assessed.

(8) The reliability of the 3D reconstruction of the environment depends on the quality of the data acquired, which are affected by different sources of errors. Therefore, real time assessment of data quality and error prediction, especially regarding autonomous systems, are necessary to help users to understand the level of uncertainty during the data acquisition and intervene, whenever necessary. In future works an expected PMTS outcome can inform the user about the quality throughout the whole data collection.

REFERENCES

- ABRAHAM, S.; FORSTNER, W. Fish-eye-stereo calibration and epipolar rectification. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 59, n. 5, p. 278–288, 2005.
- AGHAYARIA, S.; SAADATSERESHTA, M.; OMIDALIZARANDIB, M.; NEUMANN, I. Geometric calibration of full spherical panoramic Ricoh-Theta camera. **International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. IV-1(W1), p. 237-245, 2017.
- AGISOFT LLC. **Agisoft PhotoScan User Manual**: Professional Edition, Version 1.2. 2014. Available online: http://www.agisoft.com/pdf/photoscan-pro_1_2_en.pdf (accessed on 6 December 2017).
- ALAM, F.; ZHAIHE, Z.; JIA, H. A Comparative Analysis of Orientation Estimation Filters using MEMS based IMU. In: International Conference on Research in Science, Engineering and Technology, 2014. **Proceedings [...]**. Dubai: UAE, 2014. p. 21–22.
- ARANDJELOVIC, R.; ZISSERMAN, A. Three things everyone should know to improve object retrieval. In: IEEE Conference on Computer Vision and Pattern Recognition, 2012. **Proceedings [...]**. Providence: IEEE, 2012. p. 2911-2918.
- ARICAN, Z.; FROSSARD, P. Scale-invariant features and polar descriptors in omnidirectional imaging. **IEEE Transactions on Image Processing**, v. 21, n. 5, p. 2412-2423, 2012.
- BAZAN, W. S.; TOMMASELLI, A. M. G.; GALO, M.; RUY, R. D. S. A influência das injunções de orientação relativa na calibração de um sistema dual de câmeras digitais. **Boletim de Ciências Geodésicas**, p. 444-466, 2009.
- BLÁZQUEZ, M.; COLOMINA, I. Relative INS/GNSS aerial control in integrated sensor orientation: Models and performance. **ISPRS journal of photogrammetry and remote sensing**, v. 67, p. 120-133, 2012.
- BOLLES, R. C.; FISCHLER, M. A. A RANSAC-Based Approach to Model Fitting and Its Application to Finding Cylinders in Range Data. In: International Joint Conference on Artificial Intelligence (IJCAI), **Proceeding [...]**. Vancouver: University of British Columbia, 1981. p. 637-643.
- BOUGUET, J. **Camera Calibration Tool-box for MATLAB**, 2002. Available online: http://www.vision.caltech.edu/bouguetj/calib_doc/ (Accessed: 10th August 2017).
- BOURKE, P. Converting Dual Fisheye Images into a Spherical (Equirectangular) Projection, 2016. Available online: <http://paulbourke.net/dome/dualfish2sphere/> (Accessed: 28th September 2017).
- BROWN, D. C. Close-Range Camera Calibration. **Photogrammetric Engineering**, v. 37, n. 8, p. 855-866, 1971
- BRYAN, P. G.; BLAKE, B.; BEDFORD, J. **Metric survey specification for English Heritage**. 2 ed. Inglaterra: Ed. English Heritage, 2009. 111 p.
- BULOW, T. Spherical diffusion for 3D surface smoothing. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 26, n. 12, p. 1650-1654, 2004.
- BUNSCHOTEN, R.; KROSE, B. Range estimation from a pair of omnidirectional images. In: International Conference on Robotics & Automation (ICRA), 2001. **Proceedings [...]**. Seul: IEEE, 2001. p. 1174-1179.
- CADENA, C.; CARLONE, L.; CARRILLO, H.; LATIF, Y.; SCARAMUZZA, D.; NEIRA, J.; REID, I.; LEONARD, J. J. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. **IEEE Transactions on Robotics**, v. 32, n. 6, p. 1309-1332, 2016.
- CAMPOS, M. B.; TOMMASELLI, A. M. G.; HONKAVAARA, E.; PROL, F. D. S.; KAARTINEN, H.; EL ISSAOUI, A.; HAKALA, T. A Backpack-Mounted Omnidirectional Camera with Off-the-Shelf Navigation Sensors for Mobile Terrestrial Mapping: Development and Forest Application. **Sensors**, v. 18, n. 827, p. 1-18, 2018a.
- CAMPOS, M. B.; TOMMASELLI, A. M. G.; MARCATO JUNIOR, J.; HONKAVAARA, E. Geometric model and assessment of a dual-fisheye imaging system. **The Photogrammetric Record**, v. 33, n. 162, p. 243-263. 2018b.
- CAMPOS, M. B.; TOMMASELLI, A. M. G.; HONKAVAARA, E. Feasibility Study of a Backpack Mobile Terrestrial System with a Low-Cost Omnidirectional Camera. In: 10th International conference on Mobile Mapping Technology (MMT), 2017. **Proceedings [...]** Cairo: ISPRS, 2017. p. 308-314.

- CAMPOS, M. B.; TOMMASELLI, A. M.; IVÁNOVÁ, I.; BILLEN, R. Data Product Specification Proposal for Architectural Heritage Documentation with Photogrammetric Techniques: A Case Study in Brazil. **Remote Sensing**, v. 7, n. 10, p. 13337-13363, 2015.
- CASTANHEIRO, L. F.; TOMMASELLI, A. M. G.; CAMPOS, M. B.; BERVEGLIERI, A. Uma Técnica para Determinação de Linhas Epipolares em Imagens Panorâmicas Terrestres. *In: XXVII Congresso Brasileiro de Cartografia e XXVI Exposita*, 2017. **Proceedings [...]**. Rio de Janeiro: SBC, 2017. p. 908-912.
- CASTANHEIRO, L. F.; CAMPOS, M. B.; TOMMASELLI, A. M. G. Análise da Viabilidade de um Sistema de Mapeamento Móvel Terrestre de Baixo Custo para Cadastro Territorial Urbano. *In: XIII Congresso Brasileiro de Cadastro Técnico Multifinalitário e Gestão Territorial (COBRAC)*, 2018. **Proceedings [...]** Florianópolis: UFSC, 2018.
- CHANDLER, J. H.; FRYER, J. G.; JACK, A. Metric capabilities of low-cost digital cameras for close range surface measurement. **Photogrammetric Record**, p.12–26, March, 2005.
- CHEN, X.; HU, W.; ZHANG, L.; SHI, Z.; LI, M. Integration of Low-Cost GNSS and Monocular Cameras for Simultaneous Localization and Mapping. **Sensors**, v. 18, n. 7, p. 2193, 2018.
- CHUANG, T. Y.; PERNG, N. H. Rectified Feature Matching for Spherical Panoramic Images. **Photogrammetric Engineering & Remote Sensing**, v. 84, n. 1, p. 25-32, 2018.
- CLARK, J. Self-Calibration and Performance Control of MEMS with Applications for IoT. **Sensors**, v. 18, n. 12, p. 4411, 2018.
- CONRADY, A. Decentered Lens Systems. **Monthly Notices of the Royal Astronomical Society**, v. 79, p. 384–390, 1919.
- CORSO, N.; ZAKHOR, A. Indoor localization algorithms for an ambulatory human operated 3d mobile mapping system. **Remote Sensing**, v. 5, n. 12, p. 6611-6646, 2013.
- CRAMER, M. Performance of GPS/inertial solutions in photogrammetry. *In: Photogrammetric Week*, 2001. **Proceedings...** Stuttgart, German, pp. 49–62, 2001.
- CRUZ-MOTA, J.; BOGDANOVA, I.; PAQUIER, B.; BIERLAIRE, M.; THIRAN, J. P. Scale invariant feature transform on the sphere: Theory and applications. **International journal of computer vision**, v. 98, n. 2, p. 217-241, 2012.
- DANIILIDIS, K.; MAKADIA, A.; BULOW, T. Image processing in catadioptric planes: Spatiotemporal derivatives and optical flow computation. *In: Third Workshop on Omnidirectional Vision*, 2002. **Proceedings[...]**. Copenhagen: IEEE, 2002. p. 3-10.
- DAVISON, A. J. Real-time simultaneous localisation and mapping with a single camera. *In: Ninth IEEE International Conference on Computer Vision (NIPSV)*, 2003. **Proceedings[...]**. Nice: IEEE, 2003. p. 1403.
- DEWEZ, T.J.B.; LEROUX, J.; MORELLI, S. Cliff collapse hazard from repeated multicopter UAV acquisitions: Return on experience. **International Archives of Photogrammetry and Remote Sensing Spatial Information Science**. 2016, 41, 805–811
- DURRANT-WHYTE, H.; BAILEY, T. Simultaneous localization and mapping: part I. **IEEE robotics & automation magazine**, v. 13, n. 2, p. 99-110, 2006.
- EOS Systems. **PhotoModeler User Guide**. Sixth edition. EOS Systems, Vancouver, Canada. 52 pages, 2014. Available online: http://photomodeler.com/order/PhotoModeler_QuickStartGuide.pdf (Accessed: 10th August 2017).
- EUDES, A.; LHUILLIER, M.; NAUDET-COLLETTE, S.; DHOME, M. Fast odometry integration in local bundle adjustment-based visual slam. *In: 20th International Conference on Pattern Recognition (ICPR)*, 2010. **Proceedings [...]**. Istanbul: IEEE, 2010. p. 290-293.
- ESPINHOSA, D. R. S.; SILVA, J. F. C.; GALO, M. Efeitos da injeção de base nas orientações interior e exterior e nas coordenadas do espaço objeto em imagens obtidas com a estéreo vídeo-câmara de um sistema móvel de mapeamento. **Boletim de Ciências Geodésicas**, v. 14, p. 504-522, 2008.
- FANGI, G.; NARDINOCCHI, C. Photogrammetric processing of spherical panoramas. **The Photogrammetric Record**, v. 28, n. 143, p. 293-311, 2013.

- FISHER, A. Google's road map to global domination. **N. Y. Times**, v.11, p.1–13, 2013. Available online: <http://www.nytimes.com/2013/12/15/magazine/googles-plan-for-global-domination-dont-ask-whyask-where.html> (accessed on 6 December 2017).
- FORSMAN, M.; BÖRLIN, N.; HOLMGREN, J. Estimation of Tree Stem Attributes Using Terrestrial Photogrammetry with a Camera Rig. **Forests**, v. 7, n. 3, p. 61, 2016.
- FRASER, C.S., Automatic camera calibration in close range photogrammetry. **Photogrammetric Engineering and Remote Sensing**, v. 79, n. 4, p. 381-388, 2013
- FREITAS, V.L.S.; REIS, B.M.F.; TOMMASELLI, A.M.G. Automatic shadow detection in aerial and terrestrial images. **Bulletin of Geodetic Sciences**, v. 23, n. 4, p. 578-590, 2017.
- FROST, D.; PRISACARIU, V.; MURRAY, D. Recovering Stable Scale in Monocular SLAM Using Object-Supplemented Bundle Adjustment. **IEEE Transactions on Robotics**, v. 34, n. 3, p. 736-747, 2018.
- GALO, M.; TOZZI, C. L. Feature-point based matching: a sequential approach based on relaxation labeling and relative orientation. **Journal of WSCG**, vol. 12, no. 1-3, p. 113-120, 2004.
- GALO, M.; TOMMASELLI, A. M. G.; HASEGAWA, J. K. The influence of subpixel measurement on digital camera calibration. **Revue Française de Photogrammétrie et de Télédétection**, v. 199, p. 62-70, 2012.
- GARRIDO-JURADO, S.; MUÑOZ-SALINAS, R; MADRID-CUEVAS, F. J; MARÍN-JIMÉNEZ, M. J. Automatic generation and detection of highly reliable fiducial markers under occlusion, **Pattern Recognition**, v.47, n.6, p. 2280-2292, 2014.
- GIRARDEAU-MONTAUT, D. **Cloud Compare**: 3D Point Cloud and Mesh Processing Software: Open-Source Project. Available online: <http://www.danielgm.net/cc> (accessed on 6 December 2017).
- GRANSHAW, S. I. Structure from motion: origins and originality. **The Photogrammetric Record**, v. 33, n. 161, p. 6-10, 2018.
- GRUEN, A. Towards real-time photogrammetry. **Photogrammetria**, v. 42, n. 5-6, p. 209-244, 1988.
- GUARNIERI, A.; PIROTTI, F.; VETTORE, A. Low-cost MEMS sensors and vision system for motion and position estimation of a scooter. **Sensors**, v. 13, n. 2, p. 1510-1522, 2013.
- HABIB, A. F.; MORGAN, M. F. Automatic calibration of low-cost digital cameras. **SPIE Journal of Optical Engineering**, v.42, n.4, p.948-955, 2003.
- HABIB, A., PULLIVELLI, A., MITISHITA, E., GHANMA, M., KIM, E. M. Stability analysis of low-cost digital cameras for aerial mapping using different georeferencing techniques. **The Photogrammetric Record**, v. 21, n. 113, p. 29-43, 2006.
- HANSEN, P.; CORKE, P.; BOLES, W. Wide-angle visual feature matching for outdoor localization. **The International Journal of Robotics Research**, v. 29, n. 2-3, p. 267-297, 2010.
- HANSEN, P. I.; CORKE, P.; BOLES, W.; DANIILIDIS, K. Scale-invariant features on the sphere. *In*: 11th International Conference on Computer Vision 2007. **Proceedings [...]**. Rio de Janeiro: IEEE, 2007. p. 1-8.
- HARTMANN, W.; HAVLENA, M.; SCHINDLER, K. Recent developments in large-scale tie-point matching. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 115, p. 47-62, 2016.
- HE, G.; NOVAK, K.; FENG, W. Stereo camera system calibration with relative orientation constraints. *In*: International Society for Optics and Photonics (Videometrics), 1993. **Proceedings [...]**. Boston: International Society for Optics and Photonics, v. 1820, p. 2-9, 1993.
- HEDGECK, W.; MAROTI, M.; SALLAI, J.; VOLGYESI, P.; LEDECZI, A. High-accuracy differential tracking of low-cost GPS receivers. *In*: 11th Annual International Conference on Mobile Systems, Applications and Services, 2013. **Proceedings [...]**. Taipei: ACM DL, 2013. v. 11, pp. 221–234.
- HELLWICH, O.; HEIPKE, C.; TANG, L.; EBNER, H.; MAYR, W. Experiences with automatic relative orientation. *In*: ISPRS Commission III Symposium on Spatial Information from Digital Photogrammetry and Computer Vision: International Society for Optics and Photonics, 1994. **Proceedings [...]**. Munich: SPIE DL, 1994. v. 2357 p. 370-379.

- HYYPÄ, J.; VIRTANEN, J. P.; JAAKKOLA, A.; YU, X.; HYYPÄ, H.; LIANG, X. Feasibility of Google Tango and Kinect for crowdsourcing forestry information. **Forests**, v. 9, n. 1, p. 6, 2017.
- HUGHES, C.; DENNY, P.; JONES, E.; GLAVIN, M. Accuracy of fish-eye lens models. **Applied Optics**, v. 49, n. 17, p. 3338-3347, 2010.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). **ISO 19157:2011 Geographic Information—Data quality**; ISO: Geneva, Switzerland, 2011; p. 160.
- IP, A.; EL-SHEIMY, N.; MOSTAFA, M. Performance analysis of integrated sensor orientation. **Photogrammetric Engineering & Remote Sensing**, v. 73, n. 1, p. 89-97, 2007.
- INDELMAN, V.; ROBERTS, R.; BEALL, C.; DELLAERT, F. Incremental light bundle adjustment. In: British Machine Vision Conference (BMVC), 2012. **Proceedings [...]**. Guildford: University of Surrey, 2012. DOI: <http://dx.doi.org/10.5244/C.26>.
- JÄGER, B.; MAIR, E.; BRAND, C.; STÜRZL, W.; SUPPA, M. Efficient navigation based on the Landmark-Tree map and the Z_{∞} algorithm using an omnidirectional camera. In: International Conference on Intelligent Robots and Systems (IROS), 2013. **Proceedings[...]**. Tokyo: IEEE, 2013 p. 1930-1937.
- KANNALA, J.; BRANDT, S. S. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. **IEEE transactions on pattern analysis and machine intelligence**, v. 28, n. 8, p. 1335-1340, 2006.
- KENEFICK, J. F.; GYER, M.S.; HARP, W.F. Analytical Self-Calibration. **Photogrammetric Engineering and Remote Sensing**, v.40, 1974
- KING, B. A. **Methods for the Photogrammetric Adjustments of Bundles of Constrained Stereopairs**. PhD thesis, University of Newcastle, New South Wales, Australia, 1994, 221 p.
- KING, B. Bundle adjustment of constrained stereopairs-mathematical models. **Geomatics Research Australasia**, p. 67-91, 1995.
- KLINGENSMITH, M.; SIRINIVASA, S. S.; KAESSE, M. Articulated robot motion for simultaneous localization and mapping (arm-slam). **IEEE Robotics and Automation Letters**, v. 1, n. 2, p. 1156-1163, 2016.
- KRAUS, K. **Photogrammetry: Geometry from Images and LASER Scans**. Berlin, Germany: Walter de Gruyter, 2007, p. 1-459.
- KUKKO, A., KAARTINEN, H., HYYPÄ, J., CHEN, Y. Multiplatform mobile LASER scanning: Usability and performance. **Sensors**, v. 12, n. 9, p. 11712-11733, 2012.
- LAUTERBACH, H. A.; BORRMANN, D.; HEß, R.; ECK, D.; SCHILLING, K.; NÜCHTER, A. Evaluation of a backpack-mounted 3D mobile scanning system. **Remote Sensing**, v. 7, n. 10, p. 13753-13781, 2015.
- LEHTOLA, V. V.; KAARTINEN, H.; NÜCHTER, A.; KAIJALUOTO, R.; KUKKO, A.; LITKEY, P.; KURKELA, M. Comparison of the selected state-of-the-art 3D indoor scanning and point cloud generation methods. **Remote Sensing**, v. 9, n. 8, p. 796, 2017.
- LEICA GEOSYSTEMS. **Leica Pegasus: Backpack Datasheet** (Mobile Reality Capture). 2016. Available online: <https://leica-geosystems.com/products/mobile-sensor-platforms/capture-platforms/leica-pegasus-backpack> (accessed on 6 December 2017).
- LIANG, X.; KUKKO, A.; KAARTINEN, H.; HYYPÄ, J.; YU, X.; JAAKKOLA, A.; WANG, Y. Possibilities of a personal LASER scanning system for forest mapping and ecosystem services. **Sensors**, v. 14, n. 1, p. 1228-1248, 2014.
- LICHTI, D. D.; SHARMA, G. B.; KUNTZE, G.; MUND, B.; BEVERIDGE, J. E.; RONSKEY, J. L. Rigorous geometric self-calibrating bundle adjustment for a dual fluoroscopic imaging system. **IEEE Transactions on Medical Imaging**, v. 34, n. 2, p. 589-598, 2015.
- LOURENÇO, M.; BARRETO, J. P.; VASCONCELOS, F. sRD-SIFT: keypoint detection and matching in images with radial distortion. **IEEE Transactions on Robotics**, v. 4, p. 5, 2012.
- LOWE, D. G. Distinctive image features from scale-invariant keypoints. **International journal of computer vision**, v. 60, n. 2, p. 91-110, 2004.

- LUHMANN, T.; FRASER, C.; MAAS, H. Sensor modelling and camera calibration for close-range photogrammetry. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 115, p. 37-46, 2016.
- LUHMANN, T.; ROBSON, S.; KYLE, S.; BOEHM, J. **Close-range photogrammetry and 3D imaging**. 2. ed. Berlin: Walter de Gruyter, 2013.
- LUKIERSKI, R.; LEUTENEGGER, S.; DAVISON, A. J. Rapid free-space mapping from a single omnidirectional camera. In: European Conference on Mobile Robots (ECMR), 2015. **Proceedings[...]**. Lincoln: IEEE, 2015. p. 1-8.
- MAAS, H. Concepts of real-time photogrammetry. **Human Movement Science**, v. 16, n. 2-3, p. 189-199, 1997.
- MAAS, H. Close-range photogrammetry sensors. In: LI, Z.; CHEN, J.; BALSAVIAS, E. **Advances in Photogrammetry, Remote Sensing and Spatial Information Science: 2008 ISPRS Congress Book**. 2008. New York: CRC Press, 2008. p. 63-72.
- MARCATO JUNIOR, J; MORAES, M.V.A., TOMASELLI, A. M. G. Geometric calibration of a Stereographic fisheye lens camera. **Boletim de Ciências Geodésicas**, Curitiba, 2015.
- MARINO, J.; FREY, M.; GAUTERIN, F.; SHARMA, R. Development of a highly accurate and low-cost measurement device for field operational tests. In: IEEE International Symposium on Inertial Sensors and Systems, 2016. **Proceedings...** Laguna Beach, CA, USA, pp. 74-77, 2016.
- MATSUKI, H.; VON STUMBERG, L.; USENKO, V.; STÜCKLER, J.; CREMERS, D. Omnidirectional DSO: Direct sparse odometry with fisheye cameras. **IEEE Robotics and Automation Letters**, v. 3, n. 4, p. 3693-3700, 2018.
- MICHELINI, M.; MAYER, H. Detection of critical camera configurations for structure from motion. **The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 40, n. 3, p. 73, 2014.
- MIKHAIL, E. M.; ACKERMAN, F., **Observations and Least Squares**, 1st ed; IEP, New York NY, USA, 1976. 497p.
- MIKOLAJCZYK, K.; SCHMID, C. A performance evaluation of local descriptors. **IEEE transactions on pattern analysis and machine intelligence**, v. 27, n. 10, p. 1615-1630, 2005.
- MORAES, M. V. A.; TOMMASELLI, A. M. G.; SANTOS, L. D.; RUBIO, M. F; CARVALHO, G. J.; TOMMASELLI, J. T. G. Monitoring Bank Erosion in Hydroelectric Reservoirs with Mobile LASER Scanning. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 9, n. 12, p. 5524-5532, 2016.
- MOREL, J. M.; YU, G. ASIFT: A new framework for fully affine invariant image comparison. **SIAM journal on imaging sciences**, v. 2, n. 2, p. 438-469, 2009.
- MOSTAFA, M. M. R. Bore sight calibration of integrated inertial/camera systems. In: The International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation (KIS):Baff, 2001. **Proceedings [...]**, Calgary: University of Calgary. 2001. p. 5-8.
- MOURAGNON, E.; LHUILLIER, M.; DHOME, M.; DEKEYSER, F.; SAYD, P. Generic and real-time structure from motion using local bundle adjustment. **Image and Vision Computing**, v. 27, n. 8, p. 1178-1193, 2009.
- MUÑOZ-SALINAS, R., 2012. ARUCO: a minimal library for Augmented Reality applications based on OpenCv, Disponível em: <http://www.uco.es/investiga/grupos/ava/node/26> Acesso em: 08 dez. 2016
- MUR-ARTAL, R.; MONTIEL, J. M. M.; TARDOS, J. D. ORB-SLAM: A versatile and accurate monocular SLAM system. **IEEE Transactions on Robotics**, v. 31, n. 5, p. 1147-1163, 2015.
- MURILLO, A. C.; GUERRERO, J. J.; SAGUES, C. Surf features for efficient robot localization with omnidirectional images. In: IEEE International Conference on Robotics and Automation, 2007. **Proceedings [...]**. Roma: IEEE, 2007. p. 3901-3907.
- NAESSET, E. Effects of differential single-and dual-frequency GPS and GLONASS observations on point accuracy under forest canopies. **Photogrammetric Engineering and Remote Sensing**, v. 67, n. 9, p. 1021-1026, 2001.
- NEWBY, P. RT. Photogrammetric terminology. **The Photogrammetric Record**, v. 27, n. 139, p. 360-386, 2012.
- OVELAND, I.; HAUGLIN, M.; GOBAKKEN, T.; NAESSET, E.; MAALEN-JOHANSEN, I. Automatic estimation of tree position and stem diameter using a moving terrestrial LASER scanner. **Remote Sensing**, v. 9, n. 4, p. 350, 2017.

- POPE, A. J. The Statistics of Residuals and the Detection of Outliers. **NOAA Technical Report.**, Rockville, MD, USA, Tech. Rep. NOS-65-NGS-1, May. 1976.
- PUIG, L.; GUERRERO, J.; DANIILIDIS, K. Scale space for camera invariant features. **IEEE Transactions on Pattern Analysis & Machine Intelligence**, v.36 n. 9, p. 1832-1846, 2014.
- RAY, S. F. **Applied Photographic Optics: Lenses and Optical Systems for Photography, Film, Video and Electronic Imaging**. Oxford: Focal Press, 2002
- REISS, M.L.L ; TOMMASELLI, A. M. G. A low-cost 3D reconstruction system using a single shot projection of a pattern matrix. **Photogrammetric Record**, v. 26, p. 91-110, 2011
- ROCA, D.; LAGÜELA, S.; DÍAZ-VILARIÑO, L.; ARMESTO, J.; ARIAS, P. Low-cost aerial unit for outdoor inspection of building façades. **Automation in Construction**, v. 36, p. 128-135, 2013.
- RODRÍGUEZ-PÉREZ, J.R.; ÁLVAREZ, M.; SANZ-ABLANEDO, E. Assessment of low-cost GPS receiver accuracy and precision in forest environments. **Journal of Surveying Engineering**, v. 133, n. 4, p. 159-167, 2007.
- ROWBERG, J. **I2Cdev Device Library**: code is placed under the MIT License (Open Source Code), 2012. Available online: <https://github.com/jrowberg/i2cdevlib/blob/master/dsPIC30F/I2Cdev/I2Cdev.h> (accessed on 6 December 2017).
- SCARAMUZZA, D.; MARTINELLI, A.; SIEGWART, R. A toolbox for easily calibrating omnidirectional cameras. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, 2006. **Proceedings [...]**. Beijing: IEEE, 2006. p. 5695-5701.
- SCARAMUZZA, D.; SIEGWART, R. Appearance-guided monocular omnidirectional visual odometry for outdoor ground vehicles. **IEEE transactions on robotics**, v. 24, n. 5, p. 1015-1026, 2008.
- SCARAMUZZA, D.; FRAUNDORFER, F.; SIEGWART, R. Real-time monocular visual odometry for on-road vehicles with 1-point ransac. In: IEEE International Conference on Robotics and Automation (ICRA), 2009. **Proceedings[...]**. Kobe: IEEE, 2009. p. 4293-4299.
- SCHENK, T. **Introduction to photogrammetry**. 1.ed. Columbus: The Ohio State University (Department of Civil and Environmental Engineering and Geodetic Science), ,2005. 95 p.
- SCHNEIDER, D.; SCHWALBE, E.; MAAS, H. –G. Validation of geometric models for fisheye lenses. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 64, n.3, p. 259-266, 2009.
- SCHNEIDER, J.; SCHINDLER, F.; LÄBE, T.; FÖRSTNER, W. Bundle adjustment for multi-camera systems with points at infinity. **ISPRS Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 3, p. 75-80, 2012.
- SCHNEIDER, J.; LÄBE, T.; FÖRSTNER, W. Incremental real-time bundle adjustment for multi-camera systems with points at infinity. **ISPRS Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. XL-1/W2, p. 1-7, 2013.
- SCHWARZ, K. P.; CHAPMAN, M. A.; CANNON, M. W.; GONG, P. An integrated INS/GPS approach to the georeferencing of remotely sensed data. **Photogrammetric engineering and remote sensing**, v. 59, n. 11, p. 1667-1674, 1993.
- SCHWARZ, K. P.; EL-SHEIMY, N. Mobile mapping systems–state of the art and future trends. **International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 35, n. Part B, p. 10, 2004.
- SHEN, S.; MICHAEL, N.; KUMAR, V. Autonomous indoor 3D exploration with a micro-aerial vehicle. In: IEEE International Conference on Robotics and Automation (ICRA), 2012. **Proceedings [...]**. Saint Paul: IEEE, 2012. p. 9-15.
- SIBLEY, D.; MEI, C.; REID, I. D.; NEWMAN, P. Adaptive relative bundle adjustment. **Robotics: science and systems**, v. 32, p. 1-33, 2009
- SILVA, S. L. A.; TOMMASELLI, A. M. G.; ARTERO, A. O. Utilização de Alvos Codificados do tipo Aruco na Automação do Processo de Calibração de Câmaras. **Boletim de Ciências Geodésicas**, Curitiba, v. 20, n. 3, p. 636-656, 2014

- SONG, M.; WATANABE, H.; HARA, J. Robust 3D reconstruction with omni-directional camera based on structure from motion. *In: International Workshop on. IEEE Advanced Image Technology (IWAIT)*, 2018. **Proceedings[...]**. Chiang Mai: IEEE, 2018. p. 1-4.
- SUHR, J. K.; JANG, J.; MIN, D.; JUNG, H. G. Sensor fusion-based low-cost vehicle localization system for complex urban environments. **IEEE Transactions on Intelligent Transportation Systems**, v. 18, n. 5, p. 1078-1086, 2017.
- SVOBODA, T.; PAJDLA, T.; HLAVÁČ, V. Epipolar geometry for panoramic cameras. *In: European Conference on Computer Vision*, 1998. **Proceedings [...]**. Berlin: Springer, 1998. p. 218-231.
- TAKASU, T.; YASUDA, A. Development of the low-cost RTK-GPS receiver with an open source program package RTKLIB. *In: international symposium on GPS/GNSS*. Jeju, Korea: International Convention Centre, 2009. p. 4-6.
- THETA DEVELOPERS. **Proprietary Technical Information: Lens Parameter Information**. 2017. Available online: <http://theta360.guide/community-document/community.html> (accessed on 6 December 2017).
- TOMAŠTÍK, J.; SALOŇ, Š.; TUNÁK, D.; CHUDÝ, F.; KARDOŠ, M. Tango in forests—An initial experience of the use of the new Google technology in connection with forest inventory tasks. **Computers and Electronics in Agriculture**, v. 141, p. 109-117, 2017.
- TOMMASELLI, A. M. G.; TOZZI, C. L. A recursive approach to space resection using straight lines. **Photogrammetric Engineering and Remote Sensing**, p. 57-66, 1996.
- TOMMASELLI, A. M. G.; GALO, M.; MARCATO JUNIOR, J.; RUY, R. S.; LOPES, R. F. Registration and fusion of multiple images acquired with medium format cameras. **International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. 38, n. 1, p. 1-6, 2010.
- TOMMASELLI, A. M. G.; GALO, M.; DE MORAES, M. V. A.; MARATO JUNIOR, J.; CALDEIRA, C. R. T.; LOPES, R. F. Generating virtual images from oblique frames. **Remote Sensing**, v. 5 n. 4, p.1875–1893, 2013.
- TOMMASELLI, A. M. G.; BERVEGLIERI, A. Automatic orientation of multi-scale terrestrial images for 3D reconstruction. **Remote Sensing**, v. 6, n. 4, p. 3020-3040, 2014.
- TSENG, Y.H.; LIN, K.Y.; CHEN, Y.C. Bundle Adjustment of Spherical Images Acquired with a Portable Panoramic Image Mapping System (PPIMS). **Photogrammetry Engineering and Remote Sensing**, v. 82, p. 935–943, 2016.
- URBAN, S.; WURSTHORN, S.; LEITLOFF, J.; HINZ, S. MultiCol bundle adjustment: a generic method for pose estimation, simultaneous self-calibration and reconstruction for arbitrary multi-camera systems. **International Journal of Computer Vision**, v. 121, n. 2, p. 234-252, 2017.
- VALIENTE, D. G.; FERNÁNDEZ, L. R.; APARICIO, A. G.; CASTELLÓ, L. P.; REINOSO O. G. Visual odometry through appearance-and feature-based method with omnidirectional images. **Journal of Robotics**, v. 2012, 2012.
- VALIENTE, D.; GIL, A.; FERNÁNDEZ, L.; REINOSO, Ó. A comparison of EKF and SGD applied to a view-based SLAM approach with omnidirectional images. **Robotics and Autonomous Systems**, v. 62, n. 2, p. 108-119, 2014.
- VALIENTE, D.; GIL, A.; REINOSO, Ó.; JULIÁ, M.; HOLLOWAY, M. A. Improved omnidirectional odometry for a view-based mapping approach. **Sensors**, v. 17, n. 2, p. 325, 2017.
- VECTORNAV EMBEDDED NAVIGATION SOLUTIONS. **VN-100 Rugged IMU User Manual**. Available online:[https://www.vectornav.com/docs/default-source/documentation/vn-100-documentation/vn-100-usermanual-\(um001\).pdf?sfvrsn=b49fe6b9_18](https://www.vectornav.com/docs/default-source/documentation/vn-100-documentation/vn-100-usermanual-(um001).pdf?sfvrsn=b49fe6b9_18) (accessed on 6 December 2017).
- VALGREN, C.; LILIENTHAL, A. J. SIFT, SURF & seasons: Appearance-based long-term localization in outdoor environments. **Robotics and Autonomous Systems**, v. 58, n. 2, p. 149-156, 2010.
- WANG, X.; ROTTENSTEINER, F.; HEIPKE, C. Structure from motion for ordered and unordered image sets based on random kd forests and global pose estimation. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 147, p. 19-41, 2019.
- WANG, S.; YUEL, J.; DONG, Y.; SHEN, R.; ZHANG, X. REAL-time Omnidirectional Visual SLAM with Semi-Dense Mapping. *In: IEEE Intelligent Vehicles Symposium (IV)*, 2018. **Proceeding [...]**. Changshu: IEEE, 2018. p. 695-700.

WELLS, D. E.; KRAKIWSKY, E. J. The Method of Least Squares. Lecture Notes 18, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, New Brunswick, Canada, 1971, 180p.

WEN, C.; PAN, S.; WANG, C.; LI, J. An Indoor Backpack System for 2-D and 3-D Mapping of Building Interiors. **IEEE Geoscience and Remote Sensing Letters**, v. 13, n. 7, p. 992-996, 2016.

WEST, P. W.; WEST, P. W. **Tree and forest measurement**. New York: Springer, 2009.

XIU MING, **Teardown of Ricoh Theta 360 Degree Spherical Panorama Camera**. Disponível em: <https://zavax.wordpress.com/2015/11/12/teardown-of-ricoh-theta-360-degree-spherical-panorama-camera/> Acesso em: 30 jan. 2017

ZHANG, Z. Determining the epipolar geometry and its uncertainty: A review. **International journal of computer vision**, v. 27, n. 2, p. 161-195, 1998.

APPENDIX A

PMTS INTEGRATION

```

=====
//                      PMTS INTEGRATION CODE (ARDUINO) - GPS AND IMU DATA
//=====
==
//SD*
#include <SPI.h>
#include <SD.h>
File myFile;
//*****GPS*****
/*
    RX is digital pin 9 (connect to TX of other device)
    TX is digital pin 8 (connect to RX of other device)
*/
#include "SoftwareSerial.h"
#include "TinyGPS.h"
SoftwareSerial mySerial (9,8); // RX, TX
TinyGPS gps;
int i, j, filecontrol=0;
long lat, lon;
float alt;
unsigned long age, tempo, date, speed, course;
//*****IMU*****
#include "I2Cdev.h"
#include "MPU6050_6Axis_MotionApps20.h"
#if I2CDEV_IMPLEMENTATION == I2CDEV_ARDUINO_WIRE
#include "Wire.h"
#endif
MPU6050 mpu;
bool dmpReady = false; // set true if DMP init was successful
uint8_t mpuIntStatus; // holds actual interrupt status byte from MPU
uint8_t devStatus; // return status after each device operation (0 = success, !0 = error)
uint16_t packetSize; // expected DMP packet size (default is 42 bytes)
uint16_t fifoCount; // count of all bytes currently in FIFO
uint8_t fifoBuffer[64]; // FIFO storage buffer
// orientation/motion vars
Quaternion q; // [w, x, y, z] quaternion container
VectorInt16 aa; // [x, y, z] accel sensor measurements
VectorInt16 aaReal; // [x, y, z] gravity-free accel sensor measurements
VectorInt16 aaWorld; // [x, y, z] world-frame accel sensor measurements
VectorFloat gravity; // [x, y, z] gravity vector
float euler[3]; // [psi, theta, phi] Euler angle container
float ypr[3]; // [yaw, pitch, roll] yaw/pitch/roll container and gravity vector
int countReset;

```

```

//*****OUTPUT IMU *****

// uncomment "OUTPUT_READABLE_QUATERNION" if you want to see the actual
// quaternion components in a [w, x, y, z] format (not best for parsing
// on a remote host such as Processing or something though)
#define OUTPUT_READABLE_QUATERNION

// uncomment "OUTPUT_READABLE_EULER" if you want to see Euler angles
// (in degrees) calculated from the quaternions coming from the FIFO.
// Note that Euler angles suffer from gimbal lock (for more info, see
// http://en.wikipedia.org/wiki/Gimbal\_lock)
#define OUTPUT_READABLE_EULER

// uncomment "OUTPUT_READABLE_YAWPITCHROLL" if you want to see the yaw/
// pitch/roll angles (in degrees) calculated from the quaternions coming
// from the FIFO. Note this also requires gravity vector calculations.
// Also note that yaw/pitch/roll angles suffer from gimbal lock (for
// more info, see: http://en.wikipedia.org/wiki/Gimbal\_lock)
#define OUTPUT_READABLE_YAWPITCHROLL

// uncomment "OUTPUT_READABLE_REALACCEL" if you want to see acceleration
// components with gravity removed. This acceleration reference frame is
// not compensated for orientation, so +X is always +X according to the
// sensor, just without the effects of gravity. If you want acceleration
// compensated for orientation, use OUTPUT_READABLE_WORLDACCEL instead.
#define OUTPUT_READABLE_REALACCEL

// uncomment "OUTPUT_READABLE_WORLDACCEL" if you want to see acceleration
// components with gravity removed and adjusted for the world frame of
// reference (yaw is relative to initial orientation, since no magnetometer
// is present in this case). Could be quite handy in some cases.
#define OUTPUT_READABLE_WORLDACCEL

// uncomment "OUTPUT_TEAPOT" if you want output that matches the
// format used for the InvenSense teapot demo
#define OUTPUT_TEAPOT

```

```
//
=====
// =====          FUNCTION – SENSOR INITIALIZATION          =====
// =====
void Init()
{
// Open serial communications and wait for port to open:
myFile.println(" Inicializando GPS - uBlox Neo 6M");
// set the data rate for the SoftwareSerial port
mySerial.begin(9600);
Wire.begin();
mpu.initialize();
myFile.println(F(" Inicializando IMU - Testing device connections..."));
myFile.println(mpu.testConnection() ? F("MPU6050 connection successful") :
F("MPU6050 connection failed"));
myFile.println(F("Initializing DMP..."));
devStatus = mpu.dmpInitialize();

//default – MPU6050 CALIBRATION PARAMETERS
mpu.setXGyroOffset(112);
mpu.setYGyroOffset(121);
mpu.setZGyroOffset(10);
mpu.setXAccelOffset(-3028);
mpu.setYAccelOffset(1481);
mpu.setZAccelOffset(799); // 1688 factory default for my test chip
//Your offsets: -3028 1481 799 112 121 10
myFile.println(F("Enabling DMP..."));
mpu.setDMPEnabled(true);
mpuIntStatus = mpu.getIntStatus();
dmpReady = true;
// get expected DMP packet size for later comparison
packetSize = mpu.dmpGetFIFOPageSize();
countReset=1;
}

// =====
// =====          INITIAL SETUP          =====
// =====

void setup()
{
//Serial.begin(9600);
SD.begin(4);
// open the file. note that only one file can be open at a time,
// so you have to close this one before opening another.
myFile = SD.open("DATA.txt", FILE_WRITE);
Init ();
}
}
```

```
//
=====
===== MAIN PROGRAM LOOP =====
// =====

void loop() // run over and over
{
    delay(100);
    i=0;
    while (mySerial.available()){ //check GPS data
        if (gps.encode(mySerial.read())){ //encode GPS data
            gps.get_position(&lat,&lon); //get latitude and longitude
            gps.get_datetime(&date, &tempo, &age);
            alt=gps.f_altitude();
            // time in hhmmsscc, date in ddmmyy
            myFile.print(tempo); myFile.print("\t");
            myFile.print(lat); myFile.print("\t");
            myFile.print(lon); myFile.print("\t");
            myFile.println(alt); myFile.print("\t");
            mpu.resetFIFO();
            i=1;
            filecontrol=filecontrol+1;
        }
    }

    mpu.resetFIFO();
    mpuIntStatus = mpu.getIntStatus();
    //Serial.print("fifo: "); Serial.println(fifoCount);
    fifoCount = mpu.getFIFOCount();
    if(fifoCount==0){ countReset=countReset+1; }
    if(fifoCount == packetSize){
        mpu.getFIFOBytes(fifoBuffer, packetSize);
        fifoCount -= packetSize;
        countReset=1;
    }

    #ifdef OUTPUT_READABLE_YAWPITCHROLL
        mpu.dmpGetQuaternion(&q, fifoBuffer);
        mpu.dmpGetGravity(&gravity, &q);
        mpu.dmpGetYawPitchRoll(ypr, &q, &gravity);
        if(i==1){
            myFile.print("ypr\t"); myFile.print(ypr[0] * 180/M_PI);
            myFile.print("\t"); myFile.print(ypr[1] * 180/M_PI);
            myFile.print("\t"); myFile.println(ypr[2] * 180/M_PI);
        }
    #endif

    #endif
    /* DATA FILE CONTROL : O_APPEND */
    if(filecontrol==10) {
        myFile.close();
        Serial.println("terminou");
        myFile = SD.open("DATA.txt", O_CREAT | O_APPEND | O_WRITE);
        filecontrol=0;
    }

    if(countReset>100){ mySerial.end(); Wire.end(); Init();}
```

Figure 41 – PMTS - Electronic schema of measurements units and SD card adapter integration in an Arduino microprocessor.

