PhD program in Aerospace Science and Technology

On hybrid positioning using non-semantic image information

Thesis submitted to obtain the title of Doctor by the Universitat Politècnica de Catalunya

Eduard Angelats Company

Advisor:

Dr. Ismael Colomina Fosch, Geonumerics

Tutor:

Prof. Josep Gili Ripoll, Department of Civil and Environmental Engineering, Universitat Politècnica de Catalunya (UPC)

Castelldefels, July 2023

“Sovint s’ha dit que la felicitat mai no
és plena, sinó més aviat relativa,
sobtada i fugissera”

Francesco Piccolo, Moments d’inadvertida felicitat

Abstract

Hybrid or multi-sensor-based positioning has been a research topic actively investigated in the last decade. In this context, the possibility of using information extracted from imaging sensors, for positioning, is very appealing to mitigate the problems that GNSS or INS/GNSS-based trajectories have in terms of robustness and accuracy. On the other hand, different processing workflows, sensor positioning quality or system calibration errors, may also produce even in GNSS-friendly conditions, that multiple geospatial datasets are not properly co-registered. This thesis proposes the use of non-semantic information, this is, the use of a set of geometric entities or features, to improve the trajectory estimation in a multi-sensor-based approach. This thesis covers the mathematical modelling of non-semantic information, implements several hybrid-based trajectory estimation approaches that use this kind of information with the appropriate modelling, and also explores the use of non-semantic features to model the trajectory error modelling.

The implementation of combined models allowing to use of observations from camera or LiDAR sensors is the first contribution of this thesis. The proposed models have enabled improved trajectory determination in both urban post-processing and airborne environments with good accuracy (cm level). The implemented INS/GNSS trajectory error models are relatively simple but proved to be efficient. The combined models have been tested, in post-processing, using a bundle adjustment approach, with real data from metric cameras and aerial laser mapping systems as well as in Terrestrial Mobile Mapping systems (TMM).

The second contribution of this thesis is the characterization of trajectory errors that TMM may have in GNSS urban scenarios. The non-semantic information extracted from the images has allowed, using an integrated sensor orientation approach, to model these errors in an urban environment. This modelling opens the door to the development of new, more advanced trajectory error models that go beyond the deterministic models currently used.

The determination of trajectories in real time, in GNSS unfriendly environments, is also explored in this thesis using non-semantic features. An approach has been implemented based on a tightly coupling sequential nonlinear least squares using GNSS positions, image coordinates and raw inertial measurements. The proposed approach exploits a sliding window bundle adjustment technique to use the image coordinates of tie points and the positions and attitudes derived from the last epochs to determine the position and attitude parameters of the most recent epoch. The approach has been evaluated using both real and simulated data from a mobile mapping campaign over an urban area with long GNSS outage periods, with promising results.

This thesis also presents an approach to improve the determination of Remotely Piloted Aircraft Systems (RPAS) trajectories using open aerial data obtained in the framework of a national mapping project (PNOA). The development of this methodology is another contribution aiming to ensure the geospatial coherence between the orthophotos, and digital elevation models obtained with an RPAS and the orthophotos and digital models of the PNOA. The results, applied in the context of a multi-temporal and multi-sensor high-resolution archaeological documentation, show that photogrammetric products can be generated with a similar accuracy (cm level accuracy) to the ones generated with more complex approaches.

Last but not least, this thesis presents a seamless indoor-outdoor positioning approach with encouraging results (meter-level accuracy) in several scenarios. This contribution opens the door for enhanced tracking of members of civil protection and emergency teams. This is an open field of research with not widely accepted /adopted solution yet.

Resum

El posicionament híbrid o multisensorial ha estat un tema de recerca molt estudiat durant la darrera dècada. En aquest context, la possibilitat d'utilitzar informació extreta de sensors d'imatge per al posicionament és molt prometedora per mitigar els problemes que presenten les trajectòries basades en GNSS o INS/GNSS en termes de robustesa i precisió. D'altra banda, els diferents fluxos de treball durant el processament, la qualitat dels sensors utilitzats per al posicionament o els errors de calibratge del sistema, també poden produir, fins i tot en condicions favorables de visibilitat de satèl·lits GNSS, que múltiples conjunts de dades geoespacials no estiguin correctament co-registrats. Aquesta tesi proposa l'ús d'informació no semàntica, és a dir, l'ús d'un conjunt d'entitats o característiques geomètriques, per millorar l'estimació de la trajectòria en un enfocament multisensorial. Aquesta tesi aborda el modelatge matemàtic de la informació no semàntica, implementa diverses aproximacions híbrides per a l'estimació de trajectòria basada en aquest tipus d'informació amb el modelatge apropiat i també explora l'ús de característiques no semàntiques per modelar l'error d'una trajectòria.

La implementació de models combinats que permeten utilitzar observacions de càmeres o sensors LiDAR és la primera contribució d'aquesta tesi. Els models proposats han permès millorar la determinació de trajectòries tant en entorns urbans, en post processat, com en aerotransportats amb una bona precisió (nivell centimètric). La modelització de l'error de les trajectòries INS/GNSS implementada és relativament senzilla, però ha demostrat ser eficient. Els models combinats s'han provat, en post procés, utilitzant un ajust de xarxes, amb dades reals de càmeres mètriques i de sistemes LiDAR tant aerotransportats com embarcats en sistemes de cartografia mòbil terrestre (TMM).

La segona contribució d'aquesta tesi és la caracterització dels errors de les trajectòries TMM en escenaris urbans poc propicis per al GNSS. La informació no semàntica extreta de les imatges ha permès, utilitzant una orientació integrada, modelitzar aquests errors en aquest entorn. Aquesta modelització obre la porta al desenvolupament de nous models d'error més avançats que van més enllà dels models deterministes utilitzats actualment.

La determinació de trajectòries en temps real fent servir característiques no semàntiques, en entorns poc propicis per a GNSS, també s'explora en aquesta tesi. Per a això s'ha implementat una aproximació basada en mínims quadrats seqüencials no lineals acoblats profundament, i que utilitza posicions GNSS, coordenades d'imatge i mesures inercials. L'enfocament proposat es basa en una tècnica d'ajust de xarxes utilitzant les coordenades imatge dels punts d'enllaç i les posicions i actituds derivades de les últimes èpoques, per determinar els paràmetres de posició i actitud de l'època més recent. L'aproximació s'ha avaluat utilitzant dades reals i

simulades d'una campanya de cartografia mòbil sobre una zona urbana amb llargs períodes d'interrupció del senyal GNSS, amb resultats prometedors.

Aquesta tesi també presenta una aproximació per millorar la determinació de les trajectòries dels vehicles aeris no tripulats (RPAS) utilitzant dades aèries d'accés obert, obtingudes en el marc d'un projecte cartogràfic nacional (PNOA). El desenvolupament d'aquesta metodologia és una altra contribució destinada a garantir la coherència geoespacial entre les ortofotos i els models digitals d'elevació obtinguts amb un RPAS i les ortofotos i els models digitals del PNOA. Els resultats, utilitzats per la documentació arqueològica multitemporal i multisensorial d'alta resolució, mostren que es poden generar productes fotogramètrics amb una precisió similar (precisió a nivell de cm) als generats amb aproximacions més complexes.

Finalment, però no menys important, aquesta tesi presenta un mètode de posicionament continu tant en interior com exterior amb resultats encoratjadors (precisió a nivell de metres) en diversos escenaris. Aquesta contribució obre la porta a la millora del seguiment dels membres dels equips de protecció civil i emergències. Es tracta d'un camp de recerca obert on encara no hi ha una solució àmpliament acceptada o adoptada

Resumen

El posicionamiento híbrido o multisensorial ha sido un tema de investigación muy estudiado en la última década. En este contexto, la posibilidad de utilizar información extraída de sensores de imagen para el posicionamiento es muy prometedora para mitigar los problemas que presentan las trayectorias basadas en GNSS o INS/GNSS en términos de robustez y precisión. Por otra parte, los distintos flujos de trabajo durante el procesamiento, la calidad de los sensores utilizados para el posicionamiento o los errores de calibración del sistema, también pueden producir, incluso en condiciones favorables de visibilidad de satélites GNSS, que múltiples conjuntos de datos geoespaciales no estén correctamente co-registrados. Esta tesis propone el uso de información no semántica, es decir, el uso de un conjunto de entidades o características geométricas, para mejorar la estimación de la trayectoria en un enfoque multisensorial. Esta tesis aborda el modelado matemático de la información no semántica, implementa diversas aproximaciones híbridas para la estimación de trayectoria basada en este tipo de información con el modelado apropiado y también explora el uso de características no semánticas para modelar el error de una trayectoria.

La implementación de modelos combinados que permiten utilizar observaciones de cámaras o sensores LiDAR es la primera contribución de esta tesis. Los modelos propuestos han permitido mejorar la determinación de trayectorias tanto en entornos urbanos, en post procesado, como en aerotransportados con buena precisión (nivel centimétrico). La modelización del error de las trayectorias INS/GNSS implementada es relativamente sencilla, pero ha demostrado ser eficiente. Los modelos combinados se han probado, en postproceso, utilizando un ajuste de redes, con datos reales de cámaras métricas y de sistemas LiDAR tanto aerotransportados como embarcados en sistemas de cartografía móvil terrestre (TMM).

La segunda contribución de esta tesis es la caracterización de los errores de las trayectorias TMM en escenarios urbanos poco propicios para el GNSS. La información no semántica extraída de las imágenes ha permitido, utilizando una orientación integrada, modelizar estos errores en ese entorno. Esta modelización abre la puerta al desarrollo de nuevos modelos de error más avanzados que van más allá de los modelos deterministas utilizados en la actualidad.

La determinación de trayectorias en tiempo real utilizando características no semánticas, en entornos poco propicios para GNSS, también se explora en esta tesis. Para ello se ha implementado una aproximación basada en mínimos cuadrados secuenciales no lineales acoplados profundamente, utilizando posiciones GNSS, coordenadas de imagen y medidas inerciales. El enfoque propuesto se basa en una técnica de ajuste de redes utilizando las coordenadas imagen de los puntos de enlace y las posiciones y actitudes derivadas de las últimas épocas, para determinar los

parámetros de posición y actitud de la época más reciente. La aproximación se ha evaluado utilizando datos reales y simulados de una campaña de cartografía móvil sobre una zona urbana con largos períodos de interrupción de la señal GNSS, con resultados prometedores.

Esta tesis también presenta una aproximación para mejorar la determinación de las trayectorias de los vehículos aéreos no tripulados (RPAS) utilizando datos aéreos de acceso abierto, obtenidos en el marco de un proyecto cartográfico nacional (PNOA). El desarrollo de esta metodología es otra contribución destinada a garantizar la coherencia geoespacial entre las ortofotos y los modelos digitales de elevación obtenidos con un RPAS y las ortofotos y modelos digitales del PNOA. Los resultados, utilizados por la documentación arqueológica multitemporal y multisensorial de alta resolución, muestran que pueden generarse productos fotogramétricos con una precisión similar (precisión a nivel de cm) a los generados con aproximaciones más complejas.

Por último, pero no menos importante, esta tesis presenta un método de posicionamiento continuo tanto en interior como exterior con resultados alentadores (precisión a nivel de metros) en varios escenarios. Esta contribución abre la puerta a la mejora del seguimiento de los miembros de los equipos de protección civil y emergencias. Se trata de un campo de investigación abierto donde todavía no existe una solución ampliamente aceptada o adoptada.

Contents

List of Figures

List of Tables

Definitions

Attitude is the orientation of a system in relation to a reference frame.

A **geodetic coordinate system** is a coordinate system used to describe a position on or near the Earth's surface.

A **geodetic reference frame** is a realization of a reference system through a set of points whose coordinates are monumented or otherwise observable in or near the Earth's surface. In other words, it is a ``list'' of points with their position and velocity coordinates (e.g. ITRF96 of ITRS).

A **geodetic reference system** is a reference system used to define at any time a triad of axes on or near the Earth's surface (e.g. ITRS).

Global Navigation Satellite System (GNSS) is a system of satellites that are continuously transmitting (electromagnetic) signals that allow the users to determine their position over the Earth's surface.

Hybridization / Fusion. Sensor fusion is the process of merging data from multiple sensors to reduce the amount of uncertainty that may be involved in trajectory estimation. Several sensor fusion approaches have been proposed such as uncoupled, loosely, or tightly coupled approaches. An **uncoupled** algorithm considers VIO position, velocity and attitude plus GNSS computed position and determines position and velocity. A **loosely coupled** system also uses VIO position and velocity plus GNSS sensor computed position and velocity. It computes not only position and velocity but also VIO systematic errors (biases). A **tightly coupled** system uses VIO raw data (tie points, linear accelerations and angular velocities) and GNSS positions. It computes position and velocity and VIO systematic errors.

An **Inertial Measurement Unit** (IMU) is a device composed of one clock, several angular rate sensors (usually three), several accelerometers (usually three) and a mechanical holding structure plus the needed electronics and communication systems to provide data. The data provided by an IMU are: time, angular rates (one for each angular rate sensor) and linear accelerations (one for each accelerometer).

An **Inertial Navigation System** (INS) or Inertial Navigation Unit (INU) is a device composed of one IMU plus a navigation processor. The data provided by an INS is: time, position, velocity and attitude.

Inertial positioning/navigation is positioning/navigation based on inertial sensors.

An **inertial sensor** is a device that reacts on the basis of Newton's laws of motion. There are two types of motion, translational and rotational motion, thus, there are two types of inertial sensors: accelerometers and angular rate sensors (such as gyroscopes.) The accelerometers measure the linear accelerations (forces) or the instantaneous rate of change of velocity suffered by the sensor (translational movement.) The angular rate sensors measure the instantaneous angular velocities suffered by the sensor (rotational movement).

Measurement is the assignment of a number to a characteristic of an object.

Model is a mathematical description of a system or a process.

Navigation is real-time positioning. Sometimes navigation means the process of reading and controlling the movement of a craft or vehicle from one place to another. (Wikipedia)

Observation is a numerical property of a physical system that can be determined by a sequence of physical or mathematical operations,

Position is the place where something or someone is, often in relation to other things.

Positioning is to know a system's position, velocity, and attitude.

The **orientation parameters**, orientation elements, transformation or transfer parameters are a set of values which define the relationship between two reference systems (or frames.)

Real-Time Kinematics (**RTK**) is a differential GNSS technique which provides highly accurate and precise positioning by using carrier measurements and the transmission of corrections in real time from a base station (Adapted from Navipedia).

A **reference frame** is a realization of a reference system through a set of points whose coordinates are observable.

A **reference system** is a definition, a set of prescriptions and conventions together with the modelling required to define at any time a set of axes.

Velocity is a vector that denotes the change of position within a time rate with a directional component.

Visual-Inertial odometry (VIO) is the process of estimating the trajectory (position, attitude, and velocity) of a platform (e.g., car, UAV) by using only the input of one or more cameras plus one or more Inertial Measurement Units (IMUs) attached to it. (Adapted from Scaramuzza and Zhang, 2019).

Chapter 1

INTRODUCTION

1.1. BACKGROUND AND RESEARCH JUSTIFICATION

3D georeferenced data are widely used for many applications such as 3D city modelling, cadastral mapping, cultural heritage, facility management, and autonomous driving to mention a few examples. Technology progress, society's needs and also a limited availability of funds to pay for them, have changed the way these data are collected, from the sensor point of view but also from the platform point of view. Nowadays, these data are collected using a wide range of imaging and positioning sensors, sometimes statically but mainly kinematically. Examples of these manned or unmanned platforms are planes, cars, bikes or more recently rovers, trolleys, backpacks, or even mobile phones (Figure 1). Moreover, robust, quick, and ideally automated data processing algorithms must be developed to derive useful information. In this context, getting precise and accurate trajectories, that is, a time series of positions, velocities and attitudes, is a key step to generate 3D georeferenced data.

Figure 1. Examples of platforms for outdoor and indoor mapping. Commercial state-of-the-art terrestrial mobile mapping system Leica Geosystems AG); (b) Google street view platforms (Google Inc.); (c) Commercial rotary wing RPAS (Parrot Drones SAS) (d) Vertical Take Off and Landing RPAS (CATUAV S.A).

Terrestrial mobile mapping (TMM) is a relatively young technology, complementary to aerial and satellite mapping, that allows 3D georeferenced data generation from terrestrial moving platforms. TMM was born in the 1990s' within a research context, and the first commercial integrated system appeared in 2007 (Fernández, 2012). TMM has gained popularity allowing easy access to geoinformation, although with

low accuracy, thanks to the Google Street view family systems (Figure 1 a)) and it has been boosted recently with experiences such as the Google Street view backpack (Figure 1 a) or the Leica backpack system for indoor mapping.

Unmanned Aerial Vehicles (UAV) (Figure 1 c and d), also known as Remotely Piloted Aircraft Systems (RPAS), or drones, are a type of airborne platform that has emerged over the last 15 years (Colomina and Molina, 2014), (Nex et.al, 2022). Nowadays, UAVs represent one of the most prominent technologies for geospatial and remote sensing data generation thanks to their capability to carry out multiple miniaturized remote sensing (RGB, multispectral, LiDAR) and positioning (GNSS, IMU) sensors, keeping the platform and sensors' costs moderate enough. Thanks to much lower cost than aerial (planes, helicopters) and satellite platforms, and their capability to fly close to ground, very high spatial but also high temporal (daily) resolution mapping products can be generated (Nex et.al, 2022).

Common outputs from any TMM system are a set of oriented images, usually in the visible spectrum, or a point cloud. An oriented image is an image whose position and attitude are known. A point cloud is an unordered set of three-dimensional points expressed in a coordinate reference frame. In a mapping context, a point cloud is a 3D representation of a cartographic service, that is, a digital elevation model (DEM). Point clouds may also include some additional features such as laser return intensity or number of returns. They may also include RGB information if point clouds are generated from RGB camera images, also known as photogrammetric point clouds. This information can be also transferred from an RGB camera, in case of LiDAR-based point clouds, if the geometric information between LiDAR and RGB camera systems is known. Point clouds derived from TMM platforms, are denser than aerial or satellite point clouds due to object proximity, allowing improved or new applications, especially in urban environments (Figure 2). Figure 2 shows a set of applications where TMM is a key technology as well as it suggests the required point cloud accuracy and point density for each of the applications (NCHRP, 2013).

Primary outputs from RPAS or UAV mapping systems vary depending on the kind of imaging sensor included in the payload. UAVs that carry out RGB, multispectral or thermal cameras provide, after the processing of the raw imagery, very high-resolution orthophotos, points clouds, Digital Surface Models (DSM) and DEMs. (Colomina and Molina, 2014), (Nex et al., 2022). Point clouds derived from a RPAS camera, besides the 3D coordinates of the points, including for each point additional features such as RGB values, multiple band values or thermal information depending on the primary camera sensor. Orthophotos may also include multiple bands or channels depending as well of the primary sensors. For instance, RGB imagery includes 3 band or channels while thermal imagery has only band. The applications detailed in Figure 2, except for autonomous driving, are also valid for RPAS mapping and have been used in many studies in the literature.

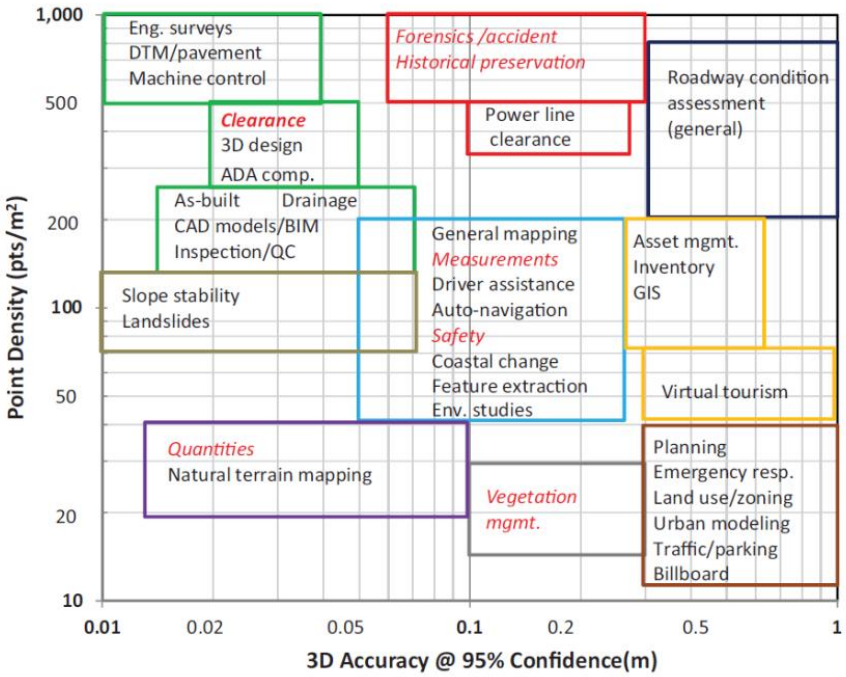

Figure 2. Terrestrial mobile mapping applications and suggested point cloud accuracy(NCHRP, 2013). The colour of the frames in the figure conveys no meaning at all; they are used just to ease the identification of several groups of applications.

These mobile mapping applications mentioned above can be grouped into three levels according to their point cloud accuracy requirements: high, medium and low (Fernández, 2012). Professional applications require high (σ<5cm, 1-sigma) and medium (σ<50cm-1m 1-sigma) (Fernández, 2012) accuracies. (NCHRP, 2013) suggests a similar point accuracy requirement expressed in 2-sigma (Figure 2). In this figure, the term 3D accuracy is used to remark that expected accuracy corresponds to the norm of the error of the set of three coordinates, not to that of each individual component. Trajectory in TMM systems is mainly estimated by combining GNSS and inertial data (INS/GNSS). Point accuracies below 0.5 m can be translated to a few centimetres platform's positions and to mili degrees for attitude.

Trajectory estimation workflows in UAV/RPAS mapping vary significantly depending on the on-board imaging and positioning sensors (GNSS, INS/GNSS); on the capability to perform RTK or differential GNSS processing; on the ground control support and also it on budget constraints (Nex et.al, 2022). If the primary mapping sensor is a LiDAR, the trajectory is usually estimated using INS/GNSS and later the point cloud is generated using a direct georefencing approach.

Meanwhile, for UAS/RPAS camera-based systems, the approaches are similar to the ones used for trajectory estimation in aerial metric cameras, that is Indirect Sensor Orientation (ISO), Direct Sensor Orientation (DSO) (Blázquez and Colomina, 2012), (Colomina and Molina, 2014), (Nex et.al, 2022)

Robust and precise positioning with INS/GNSS-based multi-sensor systems in urban and/or non-friendly environments, faces additional challenges than those in an open sky scenario, such as partial or total GNSS signal obstruction, GNSS signal multipath and low dynamics motion (Elhashash et. al, 2022). High-end TMM systems, with a cost in the range of several hundred dollars, cannot achieve the high and medium accuracy point cloud requirements in urban environments, mainly due to platform positioning errors. In post-processing, several SW solutions have been proposed allowing to mitigate, to a certain degree, the positioning errors. However, achieving a high-performance positioning in urban scenarios, especially in multi-temporal and long-term mapping ones is still challenging, motivating this PhD thesis.

A key challenge for the RPAS mapping is the integration (i.e. the co-registration) of newly acquired RPAS imagery with aerial datasets, other geospatial datasets, or already generated RPAS orthophoto and Digital Elevation Maps. Due to platform positioning, system and sensor calibration errors, multi-temporal and multi-platform co-registration between the different geospatial datasets is complicated to achieve, also motivating this PhD thesis.

In mobile mapping applications, modern systems can include one or several laser scanners and cameras. One of the aims of this thesis is to explore the benefits in terms of robustness, precision and accuracy that can be achieved thanks to using INS/GNSS-based positioning aided with non-semantic image information from these sensors, especially in GNSS-unfriendly environments. Moreover, this thesis aims to explore the benefits in terms of RPAS positioning that may be obtained thanks to using external positioning data to improve RPAS positioning and mapping, in a combined, network adjustment using common non-semantic geometric entities. In particular, non-semantic information refers to the process of extracting relevant information or salient features from imaging sensors which can be useful for positioning, but useless for understanding what is depicted in the image.

This thesis takes all steps necessary to derive useful observations from imaging sensors, to analyse them to detect and isolate outliers, and to model them appropriately to fuse them with inertial and GNSS observations or to fuse them with a derived trajectory from an auxiliary system.

It is important to highlight and remark that this thesis is not focused on the development of algorithms for non-semantic information extraction. The topic is already being investigated by the Computer Vision community and it deserves by

itself a PhD thesis. In contrast, this thesis relies on and makes use of already available well-known algorithms and tools for this step.

1.2. NON-SEMANTIC IMAGE INFORMATION

Semantic information refers to the process of extracting relevant information or salient features from imaging sensors which can be useful for understanding what is in an image/s or a point cloud. Figure 3 a) and b) show two examples of semantic segmentation from an image and from a point cloud. Thanks to the semantic information, the reader can identify that the images or point clouds comprise a set of buildings, cars, a road and other elements. On the contrary, non-semantic image information refers to information that is present in an image but does not convey meaning or context. This can include things like geometric entities such as points, lines, planes, etc. Non-semantic information is not directly related to the content of the image and does not contribute to our understanding of what is depicted in the image but it can be very useful for positioning purposes.

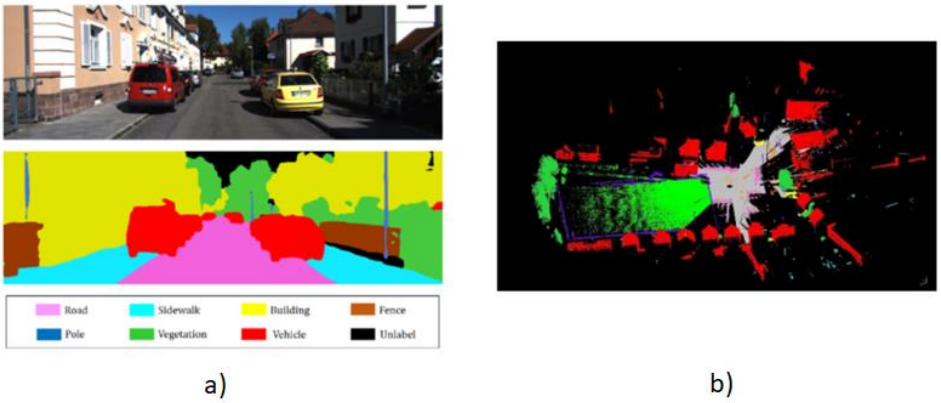

Figure 3. Examples of semantic segmentation. a) (Park et.al, 2018). b) (Hackel et.al, 2017).

The key point of the approximation lies in the fact that a geometric entity can be modelled in such a way that its defining parameters can be related to other spatial parameters, thus creating a relation that allows us to position ourselves. An example of this is the classic collinearity equations.

For the sake of clarity, in this thesis, the meaning of the image feature term is restricted to a geometric entity, while, generally, an image feature is a characteristic or pattern that is present in an image. Image features can include things like lines, shapes, colours, textures, and other visual elements.

1.3. AIMS AND OBJECTIVES

The general objective of this thesis is to explore the potential of hybrid positioning using non-semantic information. The achievement of this goal is subjected to the attainment of several specific objectives, which focus on different topics related to scientific and technical challenges, on multi-sensor-based trajectory estimation, and potential applications.

O1. Develop and implement combined models allowing the use of observations from several imaging sensors from single and multiple platforms for hybrid positioning.

O2. Post-processing trajectory improvement for TMM and RPAS mapping

This includes, firstly, the modelling of the error of sensors or vehicle trajectories, using non-semantic image information in a combined, single network adjustment, using deterministic and stochastic errors. Secondly, the use of external positioning data to improve RPAS positioning and mapping, in a combined, network adjustment using common features/entities is also evaluated.

O3. Trajectory estimation for GNSS-unfriendly environments

This objective led to the following points: the exploration of the potential of hybrid positioning using non-semantic information for GNSS-denied or unfriendly environments; for efficient outlier detection and isolation capability using inertial data or the derived trajectory; and for modelling extracted information from visible camera images to be fused with inertial/GNSS measurements in a state-space approach (SSA) (real-time) and a tightly- coupled scheme.

The proposed research seeks to contribute to the trajectory determination techniques in aerial and terrestrial mapping, by:

1. Improving the image modelling (and so the estimated trajectory) by developing combined models allowing the use of observations from several imaging sensors mounted on one or more platforms (that is, images may originate in different sensors at the same time).
2. Improving, in post-processing, the estimated trajectory using image observations with a deterministic and stochastic model.

3. Improving, in real-time, INS/GNSS trajectory estimation by means of aiding image observations through appropriate modelling and a tightly-coupled approach.
4. Providing tools for detecting indoor/outdoor transitions combining GNSS and Visual Inertial Odometry data.
5. Reducing, in post-processing, the required ground control to achieve target performance.
6. Improving the accuracy of the RPAS positioning (provided by low-cost GNSS sensors) by taking benefit from already existing aerial and reference datasets and high-accurate positioning sensors.

1.4. DOCUMENT OVERVIEW

This document is divided into eight chapters, starting with the current introductory part. The introductory chapter presented the background and motivation of the research, including an overview of the state of art and the scientific challenges tackled in the following chapters (from 2 to 7). These are composed of edited versions of international scientific publications and peer-reviewed conference papers. The research performed in the frame of this thesis covered different aspects of how and where the non-semantic image information can be used for multi-sensor positioning. This translates into four topics (Figure 4), namely mathematical modelling of non-semantic information to be used later in a hybrid-based trajectory estimation, trajectory error modelling, trajectory estimation in different environments and potential applications that can benefit from a hybrid-based positioning using non-semantic information. Each chapter starts with a preface that links the overall goal of the thesis with the topics covered in the specific chapter and its relationship with other chapters.

Briefly, the research started in chapter 2 with the development of mathematical models to enable the use of non-semantic information extracted from LiDAR scenes and camera images and their application on hybrid multi-sensor trajectory estimation of aerial platforms (Objective O1). Chapter 3 explores the use of non-semantic image information for modelling trajectory errors in urban environments (Objective O2). Then, Chapter 4 combines the mathematical modelling already introduced in Chapter 2 and the knowledge about urban trajectory errors gained from Chapter 3 to propose an approach to improve, in post-processing, the estimated trajectory of land vehicles in urban areas (Objective O2). This is done by adapting the mathematical models developed for an aerial case to use it for the urban case and by modelling the trajectory error of land vehicles as a linear positional and attitude shift. Chapter 5 is also devoted to the trajectory estimation of land vehicles in urban environments but explores additional modelling of non-semantic information and

analyses the feasibility of real-time hybrid trajectory using sequential non-linear least squares (Objective O3). The next two chapters (6,7) of the thesis present two examples of the potential of hybrid positioning using non-semantic image information, for applications that respond to technological progress and/or are driven by societal needs, which emerged during this PhD thesis. The first application, presented in Chapter 6, is the co-registration of multi-sensor and multi-temporal RPAS and historical georeferenced datasets, and its application for the documentation and management of archaeological sites (Objective O2). The second application, covered in Chapter 7, is an affordable and robust seamless indoor and outdoor tracking of emergency staff (Objective O3). Finally, Chapter 8, presents the conclusions and the future lines of work.

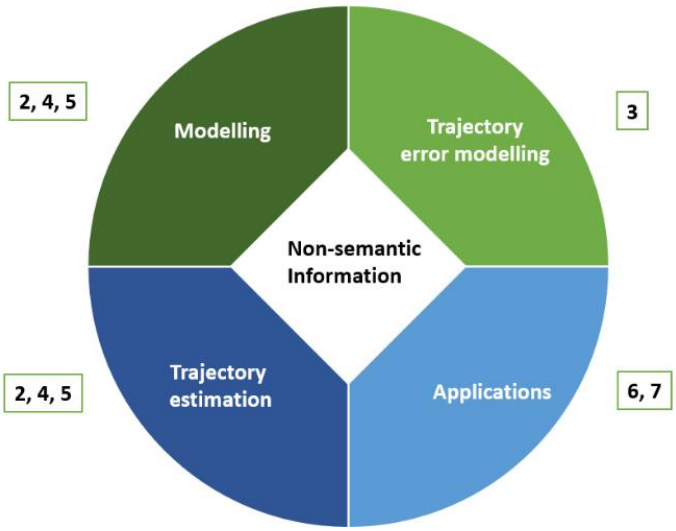

Figure 4. Diagram of the research topics included in this thesis and the relation between topics and chapters (framed numbers)

Chapter 2

POST PROCESSING TRAJECTORY IMPROVEMENT FOR AERIAL MAPPING

A RIGOROUS MODELLING OF NON-SEMANTIC IMAGE INFORMATION FROM CAMERA AND LIDAR

An edited version of:

Angelats, E., Blázquez, M. and Colomina, I., 2012. Simultaneous orientation and calibration of images and laser point clouds with straight segments. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 39(B1), Melbourne, Australia, pp. 91–96.

This chapter sets out the starting point of the research done in the framework of this thesis. It presents a rigorous modelling to introduce the non-semantic information and its relation to the trajectory estimation in post-processing using a set of geometric entities or features (points, planes and lines) (Figure 5). The presented method is based on the use of straight lines as tie features for simultaneous orientation and calibration of Frame Camera (FC) and Airborne Laser Scanner (ALS) images. For the orientation and calibration of FC images, various methods have been investigated in the past. From those, FC Integrated Sensor Orientation (ISO) has been established as the most accurate and robust one. More recently, ISO block adjustment techniques have been successfully applied to ALS orientation and calibration. As of today, for the combined ALS and FC image data sets there are no well-defined procedures as compared to the other two scenarios. The proposed method essentially reduces to the use of straight-line segments as tie features between FC and ALS images. The concept is similar and compatible with the use of tie points in FC ISO and of planar surfaces in ALS ISO thus allowing for the simultaneous orientation and calibration of ALS and FC images that involve, point, lines, and planar surface tie features. In this chapter, the two observation equations that relate (1) FC image point measurements to straight lines and (2) distance-parallelism measures between straight lines and planes, are described and derived. The concept and models are validated with a small FC and ALS block with real data.

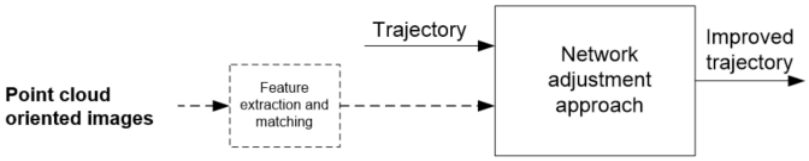

Figure 5. High-level workflow for post-processing trajectory improvement using non-semantic features extracted in a single, combined network adjustment.

2.1. INTRODUCTION

Orientation and calibration systems for photogrammetric and remote sensing instruments are fundamental components of modern mapping data processing systems. Orientation and calibration of frame and line cameras has been thoroughly investigated in the last decades. As a result, precise, accurate and reliable orientation and calibration procedures have been fielded and integrated into the photogrammetric and remote sensing production lines in companies and governmental agencies. With current INS, GNSS and image matching technologies,

the best results are obtained with the well-known Integrated Sensor Orientation (ISO) method. Typically, in addition to the ground and aerial control observations, in ISO, image measurements of ground points — the so-called tie points— are used. ISO, and in particular, sensor orientation and calibration with the bundle adjustment method using "tie straight lines" has also been proven to be feasible and efficient (Tommaselli, 1996, Habib et al., 2002). Analogously, though more recently, ISO has been proven to be the best — in terms of precision, accuracy, and reliability— method for the orientation and calibration of ALS scenes (Friess, 2006, Kager, 2004, Skaloud and Lichti, 2006). These three methods use planar surfaces as tie and control features.

Today, many mapping projects combine photographic and laser scanner data. As an example, the Spanish "Plan Nacional de Ortofotograf´ıa Aerea" (PNOA) seeks to generate and/or update ´ orthophotos and digital elevation models for Spain with an update/revisit cycle of two years with airborne digital cameras and laser scanners. Usually, there are discrepancies between the point clouds derived from the photographs and those derived from the laser scanner data. The data may originate from the same or different missions, but even in the case of simultaneous photographic and laser data acquisition, discrepancies remain. Thus, it is usually the case that, in the same airplane, two different GNSS antenna-GNSS receiver-IMU setups, one for the camera and one for the laser scanner, coexist. The navigation data may later be processed with different software systems. This, together with the system and sensor calibration errors, ends up with inconsistent data sets that require manual editing and are suboptimal from the accuracy and productivity points of view. It is therefore necessary that methods be found, and procedures be fielded to guarantee the correctness and mutual consistency of data derived from the camera and laser scanner images.

The feasibility of combined orientation and calibration of camera and laser scanner images by identifying features in the object space that can be "camera-to-laser scanner" tie features, are investigated. An ISO method where points, planar surfaces and straight-line segments link all types of images, from cameras and laser scanners, is envisioned. Co-registration between photographs and laser scanner point clouds using lines as common tie feature have already been investigated (Ronholm, 2011, Habib et al., 2005). (Habib et al., 2005) proposed two approaches for coregistration. The first approach uses the laser scanner straight lines as control lines in the photogrammetric ISO. The second approach derives a point cloud from photographs and then both point clouds are co-registered using straight lines. This chapter focuses on the orientation and calibration aspects rather than on the co-registration ones and, consequently, on the use of raw laser scanner range-angular data rather than on the point clouds. In the rest of the chapter, the terms "FC images" and "ALS images" are used when referring to photographic images captured with frame cameras and to range-angle data sets captured with airborne laser scanners, respectively.

The chapter is organized as follows. Firstly, the main ideas behind the simultaneous network adjustment are introduced. Then, the proposed models as well as some mathematical concepts behind the uses of tie lines are explained in detail. The concept validation section presents the preliminary results of an adjustment obtained with real data. The last section summarizes the main conclusions of the proposed approach and discusses future improvements.

2.2. PROPOSED APPROACH

A method is proposed where a simultaneous FC, ALS and FC-ALS ISO is performed. In it, the imaging observations are image coordinates measured on FC data (the digital FC images) and selected measurements of the ALS images (ranges and scan-angles). The other observations are surveyed ground control points and aerial INS/GNSS-derived time-Position-Attitude (tPA) control. Ground control points can be used to derive indirect control observations like ground control lines and planes. The tie features are points (FC ISO), planar surfaces (ALS ISO) and straight-line segments (FC-ALS ISO). Points, planes and lines are defined by 3, 3 and 4 independent parameters, respectively.

The classical FC ISO approach is well known. From the observations mentioned above, estimates for the tie points (TP), for the exterior orientation parameters of the FC images (EO), for the self-calibration parameters of the camera, for the INS/GNSS shifts and possibly for other system calibration parameters are computed. In ALS ISO, the unknown parameters of the tie planar bounded surfaces —tie planes (TPL)— are estimated together with INS/GNSS shifts, ALS self-calibration parameters and possibly other system calibration parameters.

In the proposed FC-ALS ISO method, as mentioned, straight line segments —Tie Lines (TL)— are used as tie features between FC and ALS images. TLs tie FC and ALS images as follows. Assume that s is a TL that can be recognized in the FC image c and in the point cloud of an ALS image a, and also assume that s is the intersection of two planar surfaces p_1, p_2 determined by points of the ALS image a. Given s, c, a, p_1 and p_2, then three types of observation equations will be used. The first type of observation equation (a coplanarity equation) relates image measurements of the s TL on the c FC image with s TL parameters, with the EO and self-calibration parameters of c. At least two of these observation equations are needed. The second type of observation equation (a line-in-plane relationship) compels s to lie on planes. In general, there will be two observation equations of this type, for p_1 and p_2, although it is also possible to use just one TPL. The third type of observation equation relates TPLs — in this case p_1 and p_2— to the measured ALS points, i.e., relate the ALS range and scan-angle measurements, to the ALS self-calibration parameters, the INS/GNSS-derived tPA aerial control, the

INS/GNSS shift parameters and possibly other system calibration ones. There will be, for each p_1 and p_2, as many of these observation equations as ALS points that define the plane. Details of the models are given in the next section.

In the combined FC-ALS ISO concept, ALS raw data instead of processed point clouds are used. 3D coordinates of point-cloud points are not explicitly computed: the ALS points that define the TPLs are parameterized with raw range and scan-angle measurements. This is a key aspect of the concept as it allows for self-calibration of ALS measurements. Self-calibration of FC and ALS image observations is the key to correct and consistent results.

It is out of the scope of this research to discuss the measuring (TLs and TPLs extraction) and matching techniques that precede the ISO network adjustment stage. They are summarized for the sake of completeness. TLs are extracted from FC images by combining edge detectors and the Hough transform (Vosselman et al., 2004). Once they are extracted, a matching procedure is performed to identify homologous TLs in different images. In the ALS ISO case, in this research, only planar surfaces belonging to building roofs are extracted. For this purpose, a region-growing algorithm has been used.

2.3. MATHEMATICAL MODELS

In this section observation equations that are more relevant to the proposed approach are discussed and/or introduced. Ground control and INS/GNSS-derived aerial control observations will not be discussed as they do not change from what is already known and published (Blazquez and Colomina, 2012). Firstly, some notation conventions are presented and then the observation equations are developed.

2.3.1. Some naming and notation conventions

The Coordinate Reference Frames (CRFs) involved in the mathematical functional models of the chapter are detailed in Table 1. The CRF l of a variable X is denoted by a superscript symbol like X^l. For a rotation matrix R_b^l , the subscript indicates the origin CRF while the superscript indicates the final CRF so it can be written $X^l = R_b^l X^b$. For the sake of simplicity, $\overrightarrow{X^e} = (x, y, z)^e$ is used instead of the formal $\overrightarrow{X^e} = [(x, y, z)^e]^T$.

l	Local-level terrestrial frame (east-north-up)
e	Earth centered Earth fixed frame
l'	Local-level terrestrial frame (north-east-down)
i	IMU instrumental frame (forward-right-down)
i'	IMU instrumental frame (forward-left-up)
a	ALS instrumental frame
c	FC instrumental frame

Table 1. Reference frames and coordinate systems.

2.3.2. FC point collineartiy observation equations

In FC ISO, the known collinearity observation equations are used to relate the FC image coordinate observations with the EO, tie point and self-calibration parameters. No modifications are required to use them in the current method.

2.3.3. ALS plane observation equations

Again, no modifications are required to use the also known ALS plane observation equation (Skaloud and Lichti, 2006) in the current method. In this equation, the observations are the two ALS image measurements (range and scan-angle) and the six tPA aerial control (position and attitude) ones. The ALS plane observation equation is

$$\overrightarrow{n^l} \cdot \left(\overrightarrow{P^l} - \overrightarrow{P_0^l}\right) = d \tag{1}$$

where

$$\overrightarrow{n^l} = R_x(\delta l) \cdot R_y(\delta p) \cdot \overrightarrow{n_0^l} \tag{2}$$

$$\overrightarrow{P^l} = \overrightarrow{X^l} + R_a^l \cdot \left[R_x(S_\theta \cdot \theta + \Delta\theta) \cdot \left(\vec{r} + \overrightarrow{\Delta r}\right) + \vec{a}\right]^a \tag{3}$$

and where

$$R_a^l = R_e^l \cdot R_{l'}^e \cdot R_i^{l'}(he, pi, ro) \cdot R_{i'}^i \cdot R_a^{i'} \tag{4}$$

For an ALS point, eq. 1 enforces it to belong to a planar surface through a direct georeferencing implicit step (eq. 3). The model extends the one proposed by (Skaloud and Lichti, 2006) with an additional scale factor self-calibration parameter for the scan-angle observation because our experience reveals its significance. The Hessian form of a plane is used to parametrize the planar surface. The plane is then

characterized by a unit normal vector $\overrightarrow{n^l} = (n_x, n_y, n_z)^l$ (eq. 2) and d, the orthogonal distance between the plane and the CRF origin. Table 2 summarizes the mathematical symbols of eqs. 1 to 4.

Observations	
$\vec{r} = (0, 0, r)$	ALS range
θ	ALS scan-angle
$\vec{X}^l = (X, Y, Z)^l$	INS/GNSS position
$R_i^{l'}\ (he, pi, ro)$	INS/GNSS attitude
Parameters	
$\vec{P}^l = (P_x, P_y, P_z)^l$	point object
$\vec{\Delta r} = (0, 0, \Delta r)$	ALS range shift
$\Delta\theta$	ALS scan-angle shift
S_θ	ALS scan-angle scale factor
$(\delta l, \delta p, d)$	tie plane
$\vec{a}^a = (a_x, a_y, a_z)^a$	IMU-to-ALS lever arm
$R_a^{i'}(E_x, E_y, E_z)$	IMU-to-ALS boresight matrix
Constants	
$R_e^l, R_{l'}^e$	rotation matrices from e to l and l' to e CRFs, respectively
$\vec{n}_0^l = (n_{0x}, n_{0y}, n_{0z})^l$	constant approximation of the unit normal vector of tie plane
$\vec{P}_0^l = (P_{0x}, P_{0y}, P_{0z})^l$	local origin coordinates for tie plane.

Table 2. Symbols in the ALS plane observation equation.

2.3.4. 3D straight line parameterization

Straight lines are elementary mathematical objects with many possible parameterizations. Yet another method is proposed which is convenient for numerical computations and whose four universal parameters (p, q, α, β)can be used regardless of the line location and direction. This is achieved with the help of, for each line, some auxiliary values: an auxiliary CRF origin $\overrightarrow{P_{L_0}^l}$ close to the straight line and rough approximations α_0 and β_0 of the straight-line director vector. With this simple strategy, numerical singularities are avoided, and a unique analytical formulation is used.

The point $\overrightarrow{P_{L_0}^l}$ is taken, for instance as one of the points in the interest area, or the mean value of several points in this area. α_0 and β_0 are auxiliary constant angles so that the rotation.

$$R(\alpha_0, \beta_0) = R_z(\alpha_0) \cdot R_y(\beta_0)$$

brings the vector $\vec{u} = (1,0,0)$ to an approximation of the director vector of L (Figure 6). The intersection point $\vec{\gamma} = (0, p, q)$ between the plane $\{X = 0\}$ and the line is obtained by applying the previous translation and the inverse of the rotation $R(\alpha_0, \beta_0)$ to L. A point $\overrightarrow{P_L^l}$ of L can thus be obtained, for some μ as

$$\overrightarrow{P_L^l} - \overrightarrow{P_{L_0}^l} = R(\Lambda) \cdot \vec{\gamma} + \mu \cdot R(\Lambda) \cdot \vec{u} \tag{5}$$

with

$$R(\Lambda) = R(\alpha_0, \beta_0) \cdot R_z(\alpha) \cdot R_y(\beta),$$

where (p, q, α, β) are the characteristic parameters of the straight line. The angles α, β use to be quite small by construction.

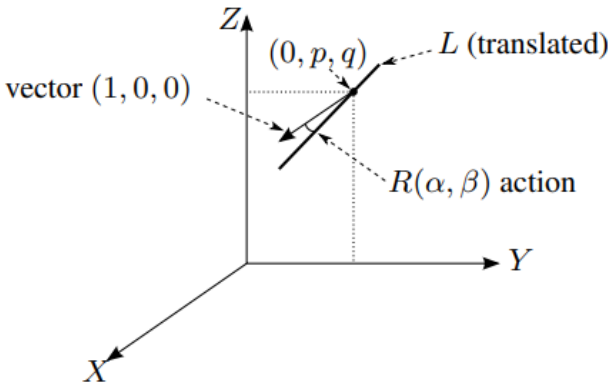

Figure 6. Straight line characterization in object space.

2.3.5. FC line coplanarity observation equation

This model relates the parameters (p, q, α, β) of a 3D line, the EO parameters of an image, and the image observations (x, y) of a point of the line. It has been derived from the collinearity condition and the 3D straight line parameterization presented in the previous section 3.4. Eq. 6, and its equivalent eq. 7, are the coplanarity equations where the straight lines are used as tie features. Table 3 details the mathematical symbols of these equations that are not introduced in this subsection.

Observations	
$\vec{x}^c = (x, y, 0)^c$	image coordinates
Parameters	
$\vec{X}_0^l = (X_0, Y_0, Z_0)^l$	camera position
$\Gamma = (\omega, \varphi, \kappa)$	camera attitude
Constants	
$\vec{f}^c = (0, 0, -f)^c$	camera constant

Table 3. Symbols in the FC line coplanarity equations.

For some λ, μ

$$\overrightarrow{X_0^l} + \lambda R_c^l(\Gamma)\left(\vec{x} + \vec{f}\right)^c = \overrightarrow{P_{L_0}^l} + R(\Lambda)(\mu \cdot \vec{u} + \vec{\gamma}) \tag{6}$$

λ and μ exists, if and only if

$$\det\left(\overrightarrow{X_0^l} - \overrightarrow{P_{L_0}^l} - R(\Lambda)\vec{\gamma},\ R_c^l(\Gamma)\left(\vec{x} + \vec{f}\right)^c,\ R(\Lambda)\vec{u}\right) = 0 \tag{7}$$

Eq 7 is the FC line coplanarity observation equation. In practice, for this model to be of any value, the FC self-calibration function s, $s = (s_x, s_y)$, has to be included to replace $\overrightarrow{x^c}$ by its calibrated value $\left(s_x(x, y), s_y(x, y), 0\right)^c$.

2.3.6. Line-in-plane observation equations

Given the parameterization of section 3.4, the observation equations that express that a straight line belongs to a plane can be easily deduced and result in the two following equations.

$$\left(\overrightarrow{n^l}\right)^T \cdot (R(\Lambda) \cdot \vec{u}) = 0 + v_o \quad (8)$$

$$\left(\overrightarrow{n^l}\right)^T \cdot \left(\overrightarrow{P_{L_0}^l} + R(\Lambda) \cdot \vec{\gamma} - \overrightarrow{P_0^l}\right) = 0 + v_d \quad (9)$$

where the covariance matrix of the "observation" vector $(0,0)$ and the residual vector (v_o, v_d) are a mechanism to respectively tune and assess the actual degree of coincidence of the actual line and plane.

2.4. CONCEPT VALIDATION

In order to validate the proposed concept, real data were collected, and software was developed. The data originate from two flight campaigns over an area near Castellbisbal (Catalonia, Spain). In these campaigns high-end airborne laser scanners and digital metric camera systems were flown. As for the software, an experimental "model toolbox" with the models discussed before was developed at the former Institute of Geomatics (currently CTTC) and run on the generic network adjustment platform GENA (Colomina et al., 2012).

Since the aim of this chapter is not an exhaustive study, a small subset of the original data has been used to validate the feasibility of simultaneous block adjustment for calibration and orientation of FC and ALS images. The main characteristics of the camera and laser scanner systems as well as the test configurations, are detailed in Table 4.

The precision of the observations can also be found in Table 4. Figure 7 shows the block layout. Notice that from the FC-derived data, 4 tie lines have been extracted. These tie lines correspond to building ridges. On the other hand, from the ALS-derived data, 6 planes have been extracted: 4 of them are our tie planes and the other 2 are ground control planes. The tie planes correspond to 2 building roofs and the ground control planes to a football field that was surveyed with standard differential GNSS techniques. The straight lines that result from the intersection of the ALS tie planes of the same building roof are FC tie lines.

FC sub-block	
Equipment	Z/I DMC Applanix POS AV 510
Image size	9 x 17 cm
Pixel size	12 μm
Camera constant (f)	120 mm
Flying height above ground ($\approx$)	1000 m
Scale ($\approx$)	1:6800
Ground sampling distance (GSD) ($\approx$)	8.16 cm
No. of strips	2
No. of images	6
No. of images per strip	3
No. of ground control points (GCPs)	4
No. of ground check points (ChPs)	8
No. of tie points (TPs)	24
No. of tie lines (TLs)	4
Overlap $\approx$	70% $\times$ 60%

ALS sub-block	
Equipment	Leica ALS50 Leica IPAS
Flying height above ground ($\approx$)	1000 m
Density (points/m^2) ($\approx$)	3.8 - 4.3
No. of strips	1
No. of control planes (GCPLs)	2
No. of tie planes (TPLs)	4
Average no. of points per plane	300
No. of tie lines (TLs)	2

Observable	Precision (σ)	Units
FC ic of TPs	$\sigma_{x,y} = 1.5$	μm
FC ic of TLs	$\sigma_{x,y} = 20$	μm
GCPs	$\sigma_{e,n} = 3$, $\sigma_u = 5$	cm
FC INS/GNSS position	$\sigma_{e,n} = 3.5$, $\sigma_u = 5.5$	cm
FC INS/GNSS attitude	$\sigma_{pi,ro} = 5$, $\sigma_{he} = 8$	mdeg
ALS INS/GNSS position	$\sigma_{e,n} = 5$, $\sigma_u = 8$	cm
ALS INS/GNSS attitude	$\sigma_{pi,ro} = 5$, $\sigma_{he} = 8$	mdeg
ALS range	$\sigma_r = 5$	cm
ALS scan angle	$\sigma_\theta = 25$	mdeg

ic: image coordinates

Table 4. Castellbisbal block geometric configuration and precision of the observations.

With the given data, the configurations described in Table 5 have been processed. Block configuration FC TP performs a classical FC ISO using all available tie points and ground control points (GCPs). Test FC TL adds 4 straight lines as tie lines. This configuration allows us to validate the coplanarity equation model with tie lines. As for the ALS tests, an ALS orientation and calibration adjustment without FC data is performed using planes as tie features with two ground control planes (GCPLs). The ALL configuration combines FC data with ALS data using tie points, ground control points for the FC ISO, tie planes and ground control planes for the ALS ISO and tie lines for the FC-ALS ISO. This test allows us to test simultaneously all developed models. Test ALL nc seeks to explore the simultaneous adjustment without using ground control planes in the ALS ISO adjustment. It also helps in understanding how laser scanning can benefit from photogrammetry in order to estimate the ALS self-calibration parameters.

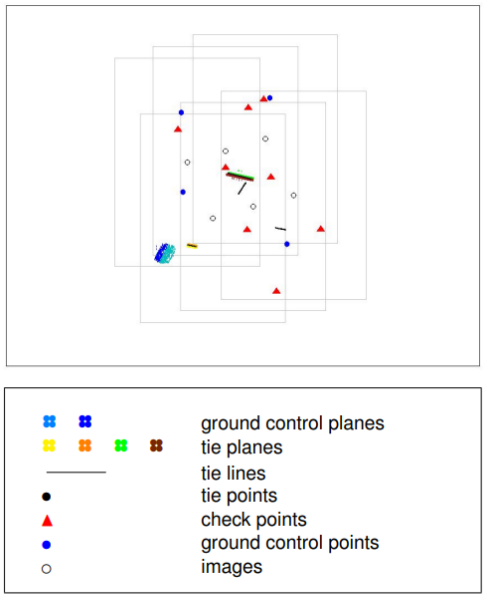

Figure 7. Castellbisbal block layout.

Besides the EO parameters, one INS/GNSS shift for the FC subblock is also estimated as well as corrections in the boresight matrix between the IMU and the camera (Table 6). In the FC ISO, self-calibration parameters are not estimated due to the small number of photographs. As for the ALS ISO calibration, note that the geometry of the sub-block is not strong enough with only one scan line. For the proposed test, self-calibration of ALS refers to the estimation of the self-calibration parameters Δr, $\Delta\theta$ and S. In this case, an INS/GNSS shift cannot be estimated because it is strongly correlated with Δr as well as with the offset between the ALS scan mirror and the IMU. The boresight matrix cannot be estimated because there is only one scan line and the boresight matrix is strongly correlated with ALS self-calibration parameters.

Test	FC ic	GCPs	TPs	FC TLs	ALS rs	TPLs	GCPLs	ALS TLs
FC_TP	YES	4	YES	NO	NO	NO	NO	NO
FC_TL	YES	4	YES	YES	NO	NO	NO	NO
ALS	NO	0	NO	NO	YES	YES	YES	NO
ALL	YES	4	YES	YES	YES	YES	YES	YES
ALL_nc	YES	4	YES	YES	YES	YES	NO	YES

ic: FC image coordinates.
rs: ALS range and scan angle.

Table 5. Observations of the block configurations.

	FC					ALS	
Test	EOs	TPs	BC	Shift	TLs	$(\Delta r, \Delta\theta, S_\theta)$	TPLs
FC_TP	YES	YES	YES	YES	NO	NO	NO
FC_TL	YES	YES	YES	YES	YES	NO	NO
ALS	NO	NO	NO	NO	NO	YES	YES
ALL	YES	YES	YES	YES	YES	YES	YES
ALL_nc	YES	YES	YES	YES	YES	YES	YES

BC: IMU-to-camera boresight matrix calibration.
Shift: INS/GNSS linear shift.

Table 6. Estimated parameters of bloc configurations.

The test results are shown in Table 7 and Table 8. As it was expected, due to the small data set and GCP configuration, FC TP and FC TL results show a similar performance in terms of accuracy (Figure 8) and precision (Table 7) for the EO parameters and tie points. The results of tests cases ALS and ALL show that the simultaneous block adjustment works and, apparently, allows for a slightly better determination of ALS self-calibration parameters in terms of precision although this result is inconclusive given the small differences and the small size of the experiment (Table 8). In general, the tests show the potential of using tie lines to link the camera and laser scanner images. In test ALL nc, the simultaneous adjustment is carried out without using ground control planes. The results indicate that it is possible to estimate Δθ, S and Δr, with the same performance as the ALS case.

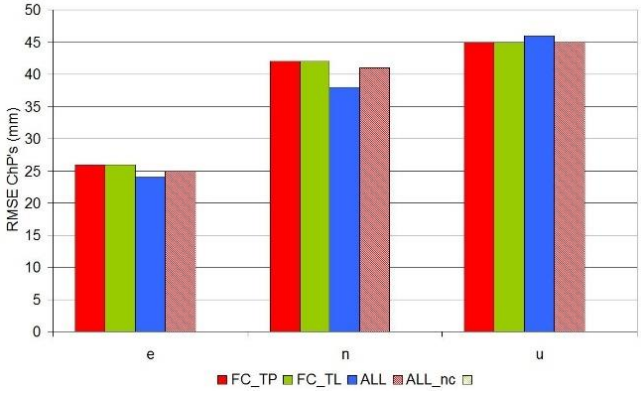

Figure 8. RMSE of ChP's.

	RMSE ChP (mm)			TP (mm)			EO (mm, mdeg)					
Test	e	n	u	e	n	u	e	n	u	ω	ϕ	κ
FC_TP	26	42	45	35	41	77	54	51	57	2.74	3.02	2.74
FC_TL	26	42	45	35	41	76	53	50	56	2.70	2.97	2.70
ALS	-	-	-	-	-	-	-	-	-	-	-	-
ALL	24	38	46	35	41	75	53	50	54	2.71	2.99	2.71
ALL_nc	25	41	45	35	41	76	53	50	56	2.69	2.96	2.68

Table 7. Test results: FC.

Test	Δr (m)	$\sigma_{\Delta r}$ (mm)	$\Delta\theta$ (deg)	$\sigma_{\Delta\theta}$ (mdeg)	S_θ	σ_{S_θ}
FC_TP	-	-	-	-	-	
FC_TL	-	-	-	-	-	
ALS	3.37	46	-2.32	39.7	0.976	0.00339
ALL	3.10	12	-2.07	4.03	1.002	0.00032
ALL_nc	3.10	135	-2.07	5.46	1.002	0.00039

Table 8. Test results: ALS.

2.5. CONCLUSIONS AND FURTHER RESEARCH

The results presented in this chapter show that with this combined FC, ALS and FC-ALS ISO concept, the performance in terms of precision and accuracy is maintained. The results also suggest that the ALS self-calibration can benefit from a combination with photogrammetric data. Nevertheless, the results of this chapter are only a proof of concept because of the size of the data sets. Further work must be done to validate the presented concept and models with larger data sets and to understand the behaviour of FC and ALS self-calibration parameters in combined adjustments as well as the requirements in terms of ground control points and planes.

Chapter 3

TRAJECTORY ERROR MODELLING IN URBAN ENVIRONMENTS

UNDERSTANDING THE BEHAVIOUR OF URBAN TERRESTRIAL MOBILE MAPPING TRAJECTORIES

An edited version of:

Angelats, E., Parés, M.E, and Colomina, I., 2016. Hybrid models for trajectory error modelling in urban environments In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLI-B3, 2016 XXIII ISPRS Congress, 12–19 July 2016, Prague, Czech Republic.

This chapter tackles a initial step of any strategy aiming to improve the trajectory of terrestrial mobile mapping systems in urban environments. An approach to model the error of terrestrial mobile mapping trajectories, combining deterministic and stochastic models is presented. Due to urban specific environment, the deterministic component is modelled with non-continuous functions composed of linear shifts, drifts or polynomial functions. In addition, a stochastic error component for modelling residual noise of the trajectory error function is introduced. The first step for error modelling requires knowing the actual trajectory error values for several representative environments. In order to determine as accurately as possible, the trajectories error, (almost) error less trajectories are estimated using extracted non-semantic features from a sequence of images collected with the terrestrial mobile mapping system and from a full set of ground control points. Once the references are estimated, they have been used to determine the actual errors in terrestrial mobile mapping trajectory. The rigorous analysis of these data sets has allowed us to characterize the errors of a terrestrial mobile mapping system for a wide range of environments. The proposed approach has been evaluated using real data from a mobile mapping campaign over an urban and controlled area, with harmful GNSS conditions.

3.1. INTRODUCTION

Technology progress, society needs and also a limited availability of funds, have changed the way 3D data are collected, not only from the acquisition sensors' point of view but also from the platforms' point of view. Examples of these manned or unmanned platforms are satellites, planes, cars, bikes or more recently rovers, trolleys or even mobile phones. Terrestrial mobile mapping (TMM) is a technology, complementary to aerial and satellite mapping, that allows 3D georeferenced data generation from terrestrial moving platforms. TMM has gained popularity allowing easy access to geoinformation, although with low accuracy, thanks to the Google street view family systems and it might be boosted with experiences such as the Google tango project for indoor mapping.

Nowadays, many applications such as 3D city modelling, cadastral mapping, cultural heritage, facility management, and autonomous driving take benefit of 3D georeferenced data, or point cloud (Kutterer, 2010). The applications mentioned above can be grouped into three levels according to their point cloud accuracy requirements: high (<5 cm, 1-sigma) and medium (<15 cm-50 cm, 1-sigma) professional applications and mass-market applications (<50 cm – 1 m, 1-sigma) (Fernández, 2012).

Getting precise and accurate trajectories, that is, a time series of positions, velocity and attitudes, is a key step to generate precise and accurate 3D georeferenced data. A point cloud is generated by combining the platform trajectory together with laser scanner measurements with direct georeferencing techniques or from a set of images with known position and attitude. Centimetric level discrepancies during trajectory determination can lead to differences between 10 and 50 cm between point clouds of overlapping areas in an urban scenario (Angelats and Colomina, 2014).

Currently, the trajectory of high-end TMM systems is mainly estimated by combining GNSS, inertial and odometer data. Robust and precise positioning in an urban scenario, faces additional challenges than an open sky scenario as partial or total GNSS occlusions, or multipath, may occur. This causes an error in the platform or vehicle trajectory determination. Other sources of error can be Inertial Measurement Unit (IMU) modelling errors or system calibration errors. The system calibration error is an error in the determination of lever arm and boresight between IMU and camera or between IMU and laser scanner. The Integrated Sensor Orientation (ISO) method has been proven to be feasible and efficient for IMU-camera boresight calibration of mobile mapping systems (Kersting et al., 2012), for the IMU-laser boresight calibration with single and multiple laser scanners (Skaloud and Lichti, 2006), (Chan et al., 2013).

The need for high performance in terms of accuracy has made apparent the potential of using measurements derived from imaging sensors for improved trajectory estimation. In other words, imaging sensor measurements can be instrumental in estimating the parameters of error models that extend the trajectory models. In particular, cameras and laser scanners can help improve the trajectory at two different levels: sensor-level error modelling or trajectory-level error modelling. In the context of aiding at the sensor level, visual measurements can be used to correct the drifts of the primary navigation sensors such as an IMU. Depending on a number of factors, one error modelling strategy may be more appropriate than the other. In this research, aiding at trajectory level refers precisely to the modelling of trajectory errors (like error in position or attitude) for a given trajectory, previously estimated by other system, for instance, but not necessarily, an INS/GNSS system (Angelats and Colomina, 2014). The first approach is suitable for applications that require real-time navigation. The second approach is intended for the aforementioned applications to refine the trajectory in post-processing.

The first approach, usually known as visual aiding, refers to the computation of orientation parameters through consecutive, overlapping images by means of measurement of tie points, i.e., photogrammetric observations of the same object point in two or more images. Alternatively, planes or cylinders extracted from laser scanner images, can also be used as tie features, especially in urban or indoor areas, where they are very common. Feature extraction and matching algorithms are

commonly combined, with RANSAC procedures that aim to perform outlier detection and removal using solely image observations by means of position and attitude estimation (Nister, 2013), (Scaramuzza and Fraundorfer, 2011). More in detail, these algorithms estimate the relative orientation between two images, also known as relative pose, using n pairs of matched features, selected randomly among the full matched pairs. (Scaramuzza, 2011) proposes a method where the camera pose can be estimated with only one-point correspondence by exploiting nonholonomic constraints of wheeled vehicles and a histogram-based voting strategy. Other approaches combine derived trajectory or inertial data to predict where a point feature should appear in the second image (Veth, 2011), (Leutenegger et al., 2013).

Once all outliers have been detected and isolated, the camera or vehicle trajectory can be recovered by concatenating the estimated relative positions and orientations using k inliers from overlapping images. Several methods to estimate the navigation states using only images or using image and object observations are reviewed in (Scaramuzza and Fraundorfer, 2011). These approaches are usually referred to visual odometry (VO) or Structure from Motion (SfM) in the robotics and computer vision community. Alternatively, (Taylor et al., 2011) presented two strategies to use an IMU as a primary positioning sensor and to control inertial drift with visual information during the filtering step. The two approaches are implemented using an Unscented Kalman Filter estimation method. The first approach imposes a geometric constraint using image coordinates, while the second one takes the benefit of jointly estimating a set of object coordinates together with navigation states. (Angelats et al., 2014) presents a method to robustly detect and isolate outliers in camera images using inertial-based trajectory in a first step to navigate in GNSS-unfriendly environments, using corresponding tie points measurements together with inertial and GNSS measurements, when available, in a second step. (Schaer and Vallet, 2016) proposes to improve point cloud registration by recomputing the initial platform trajectory with position updates derived from ground control points identified in the point cloud.

Alternatively, the trajectory can be refined, in a single network adjustment. In this approach, the complete trajectory is refined simultaneously for all epochs, using observations extracted from images or point clouds from multiple imaging sensors and also ground control points. The trajectory error can be modelled using simple error models such as linear or polynomial segments (Angelats and Colomina, 2014). (Gressin et al., 2012) proposes a method where the Iterative Closest Point algorithm (ICP) is an initial step for trajectory improvement. The overall registration is performed improving the original platform trajectory. (Elseberg et al., 2013) deals also with laser-to-laser registration by improving the platform trajectory. The trajectory is improved using semi-rigid Simultaneous Localization and Mapping.

The previous techniques rely heavily on the capability to extract and match visual features or on having a dense network of ground control points. In addition, they make the assumption that visual measurements can be extracted and matched properly and that they can be extracted and matched continuously along the image sequence or point cloud. Apart from that, the trajectory error model using in the network-based approach, may model locally well the error but not globally. Regarding the sequential-based approach, a not proper stochastic characterization of observations derived from imaging sensors, might produce a filter divergence, and, so, an estimated trajectory worse than the original one.

For that reason, understanding how the behaviour of TMM trajectory in an urban environment is needed. This chapter is focused on the modelling aspects rather than overall trajectory refinement, that is, to study how is the trajectory error of a TMM in an urban environment. The final goal of this research is to provide knowledge to improve a trajectory of TMM system by rigorous modelling its error in two weak scenarios: when the available ground control points are limited and in scenarios with a weak geometry or non-continuous capability to extract and match non-semantic features.

The chapter is organized as follows. Firstly, the main ideas of the proposed approach are presented. Then, the definition of non-semantic features and their role in trajectory error modelling are introduced. The next subsection describes each of the error components of the trajectory. The experimental results section presents the results using real data from a terrestrial mobile mapping campaign with long GNSS outages. The last section summarizes the conclusions of the proposed approach and discusses future improvements.

3.2. PROPOSED APPROACH

In this chapter, an approach to model trajectory errors of TMM in an urban environment is presetedn. A real vehicle trajectory is a continuous function but the estimated one will probably be a discontinuous one. The error might be caused by GNSS multipath, IMU mismodelling or GNSS satellites occlusions. An approach to model the trajectory error with a hybrid model, that includes a non-continuous deterministic component and a stochastic one, is proposed.

To face this approach, the proposed workflow is presented in Figure 9. Reference trajectories are estimated using the initial platform trajectories estimation together with all raw observations from imaging sensors and all available ground control information. These references will be estimated using block adjustments including observations from multiple imaging sensors. In order to have useful observations

from imaging sensors, non-semantic features from a sequence of images or from point clouds will be used. These non-semantic features describe certain attributes or properties of geometric objects or entities. These features allow us to identify and match common objects between images, between point clouds, or for instance between images and point clouds. These common or tie entities can be related through image, object coordinates or both, with the trajectory components using well-known models such as the collinearity equations.

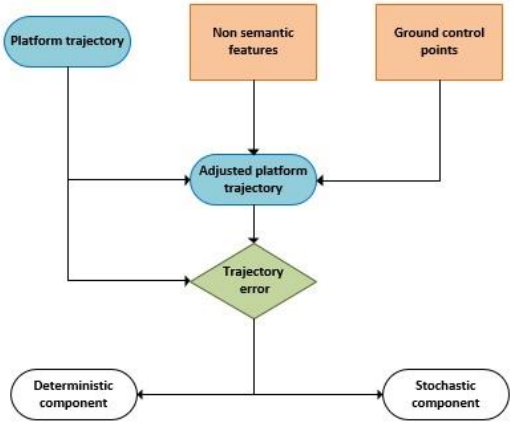

Figure 9. Proposed workflow.

Given the reference trajectories for several environments, the trajectories errors are obtained by subtracting TMM estimated trajectories from reference once.

Finally, for each environment, the trajectory error will be modelled as a non-continuous function. The analysis of the residuals of this identification will provide the stochastic characterization of the error.

3.3. TRAJECTORY ERROR MODELLING

3.3.1. Non-semantic image features

In this research, the capability of identifying common or tie features between images and/or between point clouds is exploited. Instead of exploiting the semantic content, the use of non-semantic features is proposed. By non-semantic features are understood as those that provide useful data to solve a certain task, such as

positioning, but without understanding the image or scene content. Examples of common tie features, used in our approach, are basic geometric primitives such as straight-line segments, points, planes and ellipses.

In the last years, extensive research for extracting and matching primitives such as points and lines from a sequence of camera images has been done. Usually, these features are referred to as visual features (Weinmann, 2012). However, different features can be extracted and matched also from point clouds such as planes, cylinders, toroids, to mention a few. Moreover, new models combining different and joint camera-laser features can be exploited (Angelats and Colomina, 2014). For that reason, the use of the term non-semantic features is preferred instead of visual features to refer to features extracted directly or indirectly from imaging sensors.

These primitives are usually described with a set of attributes such as intensity, colour, the spatial relation with their local neighbours, but such attributes can also describe the aforementioned primitives' response to a certain frequency band or temporal stability. These sets of attributes are used to identify, and match homologous features in a sequence of images, or between overlapping point clouds. Nevertheless, these features themselves cannot provide any relevant information to describe or understand the scene.

3.3.2. Reference trajectory generation

A reference trajectory is generated using the Integrated Sensor Orientation approach. In it, all observations from non-semantic features, coming from single or multiple sensors, are processed together with trajectory (tPA) and ground control observations, in a network adjustment. It is referred to as camera-ISO when the imaging observations are tie points extracted from camera images. Alternatively, the reference trajectory can be estimated using laser only (ranges and scan angles) or combined camera-laser observations (ranges, scan angles and image coordinates).

These observations belong to the following tie features, planar surfaces for the laser ISO, and planar surfaces and straight-line segments for the camera-laser ISO. In laser-ISO and camera-laser ISO, ground control points can be used to derive indirect control observations like ground control lines and planes. The ISO might also provide corrections in the boresight and lever arm between the IMU and the camera or between the IMU and the laser scanner and also correction on camera self-calibration parameters. Not proper geometric calibration of imaging sensors and also erroneous system calibration values might also introduce a systematic error component.

3.3.3. Deterministic model

Temporal geometric variation of the GPS and other GNSS constellations, produces a shift of the estimated positions from overlapping strips. Besides the temporal geometric variation of GNSS, in an urban scenario, GNSS signals may be affected in several ways. For example, some of the signals can be completely blocked in several epochs, causing a GNSS receiver cannot compute a solution and so introducing a drift in the platform trajectory. On the other hand, the GNSS receiver can receive direct or non-line-of sight multipath introducing an additional error into the trajectory.

In order to mitigate the impact of these factors, as part of the trajectory error model, a non-continuous function that minimizes the residuals when fitted with the error data is introduced. Typically, this function will be built upon linear shifts, drifts or polynomial functions (Figure 9, blue and red). The discontinuities of the function will be mainly due to significant changes in constellation or problems with the odometer sensor.

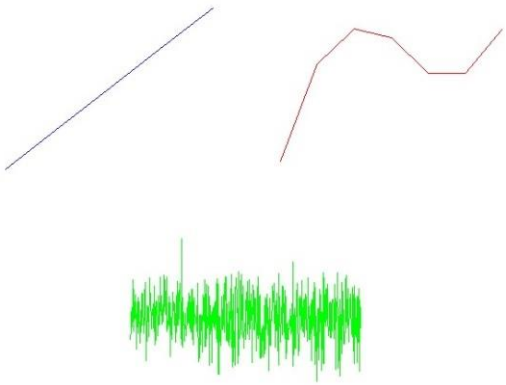

Figure 10. Examples of deterministic and stochastic model. Linear shift (blue), three order polynomio (red), white-noise (green).

3.3.4. Stochastic model

The stochastic model will be determined after the analysis of the residuals generated when fitting the previous non-continuous function to error measurements. It is

expected that these residuals (Figure 9, green) behave as white noise or Gauss-Markov stochastic process.

3.4. CONCEPT VALIDATION

3.4.1. Data set description

The proposed concept was tested and validated using real data from a professional mobile mapping system campaign over a controlled area of Dortmund (Germany). The surveyed area was an urban and GNSS-challenging scenario, with variable and degraded GNSS conditions. However, it was an excellent dataset from the imaging sensors' point of view with many structured buildings and a large variety of features.

The mobile mapping system was an Optech Lynx system, from TopScan GmbH, that included two laser scanners and two cameras. The cameras were mounted looking to each side of the street. The system was mounted on a van, and it was driven over a controlled area for three hours, resulting in 11 overlapping strips. The areas of interest were additionally surveyed to provide a dense network of ground control points. After the survey, the system trajectory was computed using a tightly coupled approach combining differential GNSS, IMU and odometer measurements.

The areas used for testing the approach are shown in Figure 11. Two different sections or blocks, marked with blue and orange ellipses, were identified. Each block configuration was selected to represent different situations/configurations that can occur in urban environments with TMM platforms. In an urban environment, a TMM vehicle can survey the area of interest several times with the same vehicle driving direction (orange case) or it can survey an area in opposite driving directions, and thus, exploiting and taking benefit from geometric diversity (blue).

Table 9 provides the main characteristics for each of the sections, in terms of used equipment, number of strips, ground control support and number of tie points. Images from four different strips have been used to generate the reference trajectory for the blue section. The vehicle surveyed the area in the same direction three times and the last one in the opposite direction. The surveyed street in the orange block is a one-way street; thus, camera images were acquired two times in the same driving direction.

The test areas had a dense network of ground control points and their distribution also varied between sections. For the blue section, 13 points provided ground control, 7 of them placed on the left side of the street. Regarding the orange block,

8 ground control points were used, 5 placed on the right side of the street and 3 of them on the left side. The number of GNSS satellites changed considerably within and between overlapping strips. The number of satellites ranged from one to seven for the blue section, with 70% of the epochs with equal or less than 3 satellites (Figure 12). The GNSS geometry for the orange section was a bit worse than the blue one with all epochs with three or less satellites. Note that in Figure 4, the tracked satellites for a blue strip with elevation angles higher than 15 degrees are shown in blue, cyan and green. The remaining colours show visible satellites below 15 degrees and are not recommended to be used for position estimation. In addition, several GNSS satellites came in or out within a few epochs in the same strip.

Figure 11. Dortmund test areas.

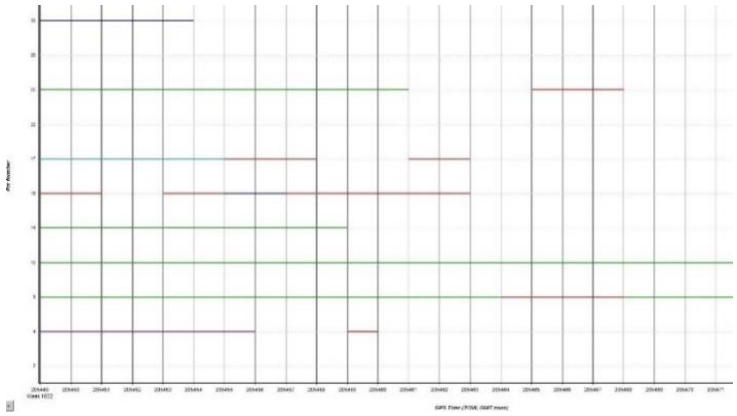

Figure 12. Example of tracked GPS L1 satellites for one strip of the blue section.

TMM equipment	
Positioning system	Applanix POS LV420
Imaging sensors	Optech Lynx cameras
Image size	5.684 x 4.326 mm
Pixel size	3.5 μm
Camera constant 1 (f)	3.864 mm
Camera constant 2 (f)	3.872 mm
TMM blue area	
No. of strips	4
No. of images	271
No. of ground control points (GCPs)	13
No. of tie points (TPs)	55419
Horizontal Overlap ≈	60
TMM orange area	
No. of strips	2
No. of images	121
No. of ground control points (GCPs)	8
No. of tie points (TPs)	22485
Horizontal Overlap ≈	60

Table 9. Dortmund block geometric configuration.

The reference trajectory was estimated by processing the collected information with the Agisoft Photoscan software (Agisoft, 2022). From both camera images, a dense set of points was extracted to be used as a non-semantic feature in the adjustment. The use of both cameras allowed us to work with a stronger geometry. In order to improve the adjustment, the system trajectory was also used as observation.

The geometric calibration values, as well as the lever arm and boresight between the cameras, and the IMU, were previously estimated with Agisoft Photoscan using a good GNSS conditions data set. The estimated geometric calibration values are the focal length, principal point, radial and tangential distortion coefficients. That subset included several strips in different directions to help to decorrelate internal camera parameters from the system calibration and from the exterior orientation parameters. In-house software was developed for estimating trajectory error components, this is, to compare system and reference trajectories.

3.4.2. Experimental results

Using the reference trajectory and the system trajectory, a set of trajectory error values were computed. Figure 13 shows the error of the cross-track axis corresponding to four overlapping strips (blue, red, magenta and green). A position mean bias between strips can be clearly identified due to temporal changes in GNSS constellation. It is important to note here that this bias is smaller between the third and the fourth street because of their time difference (less than 5 minutes). Between the first and the second strip there is a difference of 38 minutes, and between the second and the third it is 12 minutes. Besides the bias, short drift periods broken by several peaks in the trajectory error can also be identified. These peaks are mainly caused by discontinuities in the number of visible GNSS satellites. Moreover, the entrance of a GNSS satellite can produce several epochs of instability during platform trajectory estimation. This effect can be clearly identified in the magenta strip where two big peaks are present.

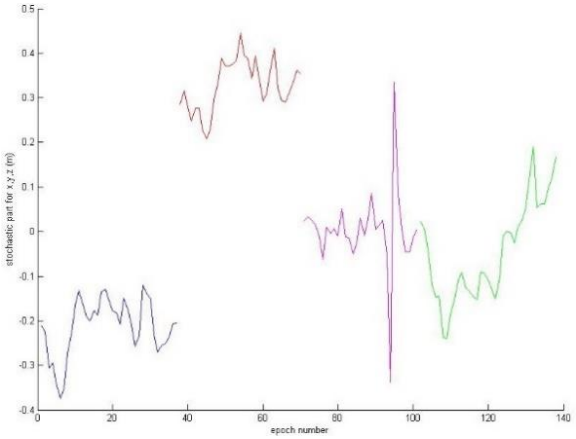

Figure 13. Error component for the heading from the four overlapping strips of blue section.

The error of the heading component of the blue section is shown in Figure 14, following the strip colour scheme of Figure 13. In contrast to the position error, a significant bias between strips cannot be observed. This is because the platform attitude estimation step relies mainly on IMU observations. However, a new GNSS position after several epochs of inertial-based trajectory might introduce also changes in the attitude. This might explain the relevant peak that can be observed in the magenta part of Figure 14.

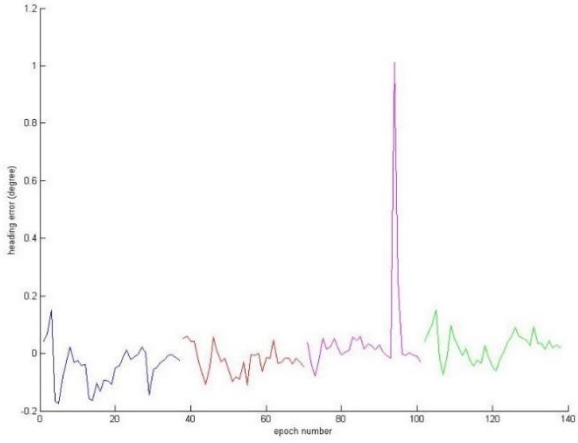

Figure 14. Error component for the heading from the four overlapping strips of blue section.

The analysis of the error leads us to model it as a discontinuous third-order polynomial. The discontinuities within this function were defined regarding changes in strips and also changes in the number of available satellites in each section within a strip. On the other hand, a single stochastic component for each component per strip is computed. Figures 15 and 16 show errors in the track axis marked with blue dots for one strip of the blue and one of the orange block.

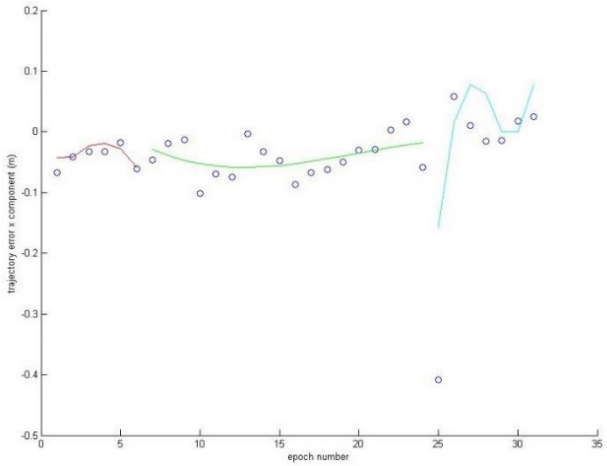

Figure 15. Trajectory error and deterministic part decomposition for the track-axis component of a strip of the blue section.

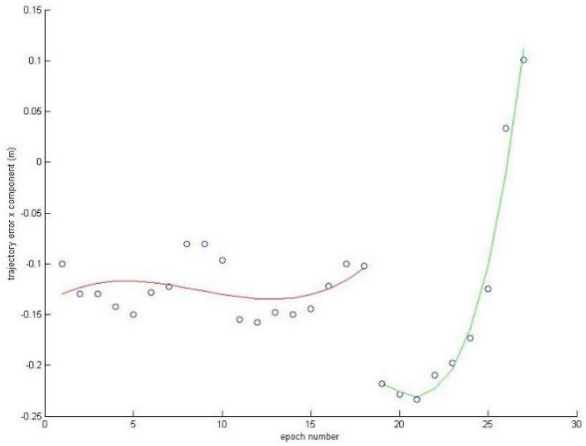

Figure 16. Trajectory error and deterministic part decomposition for the track-axis component of a strip of the orange section.

The figures also show the different segments of the error function. In the examples, the deterministic component was split, within a single strip, into three different segments (red, green, cyan) for the blue section and two (red and green) for the orange section. The comparison of both figures reveals that, at least for the selected sections, the error behaviour is similar, with short drift periods and several discontinuities. The plots for the remaining components are not shown because of their similarity with the presented one.

After the analysis of the residuals generated when fitting the previous deterministic segments to error measurements, the stochastic model was determined. Figures 17 and 18 show the residuals for each of the error components, corresponding to a single strip for both blue and orange sections. The stochastic component of the along-track axis component is shown in red, the cross-track in blue while height is shown in green.

In Table 10 the number of selected continuous segments for each test is presented together with the standard deviation values of the stochastic part, for each strip and each component. The figures reveal the presence of a stochastic component that can be modelled, in a first approximation, as a white noise process. It can be observed that the stochastic component is lower for the height than for the planimetric components. This can be explained by the use of non-holonomic constraints to reduce the height variation during the initial platform trajectory estimation. In addition, it can be observed that stochastic components are similar in terms of standard deviation for the two sections.

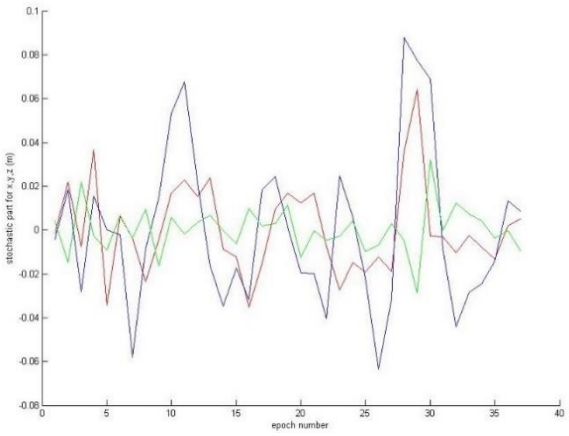

Figure 17. Stochastic components of a single strip of blue section.

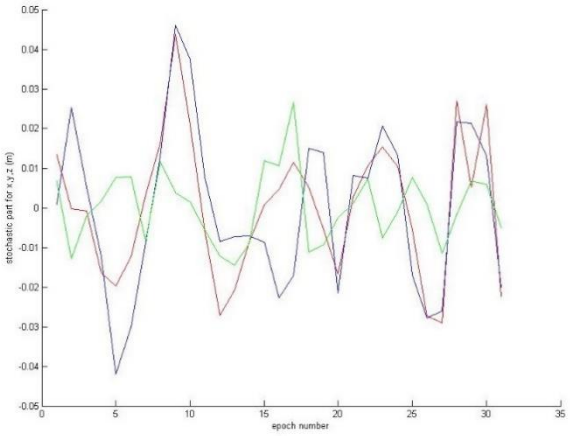

Figure 18. Stochastic components of a single strip of orange section.

Test	Strip	Number of deterministic segments	Standard deviation position (cm) x	y	z	Standard deviation attitude (mdeg) he	pitch	roll
Blue	1	2	2.07	3.63	1.07	43.2	27.2	43.6
	2	3	1.17	2.52	0.94	23.9	22.6	37.8
	3	3	1.76	3.10	1.02	28.6	23.6	37.8
	4	3	5.42	4.45	1.87	59.0	29.0	30.7
Orange	1	2	1.77	2.10	0.93	15.7	21.3	45.0
	2	2	2.21	1.98	0.83	17.9	23.2	49.6

Table 10. Stocastic component characterization.

As it was expected, the stochastic component is independent of the street or area orientation. Besides the comparison between streets, the magnitude of stochastic components, both for position and attitude, remains similar within strips of the same section. The results presented in Table 9 indicate the driving direction has not a significant impact.

3.5. CONCLUSION AND FURTHER RESEARCH

The rigorous analysis of a set of reference trajectories against the TMM estimated trajectories allowed us to characterize the errors of a TMM system for a wide range of environments, and allows exploring in the near term, new and innovative methods for improving TMM trajectory.

The reference trajectory is obtained using extracted non-semantic features from a sequence of images and from a full set of ground control points. The trajectory error is modelled using hybrid, that is, combining deterministic and stochastic components. The deterministic component can be modelled as a non-continuous function composed by linear shifts, drifts or polynomial functions. In addition, a stochastic error component for modelling residual noise of the trajectory error function is introduced.

The proposed approach has been evaluated using real data. The data came from a mobile mapping campaign over an urban and controlled area of Dortmund (Germany), with harmful GNSS conditions. The results show the suitability to decompose trajectory error with non-continuous deterministic and stochastic components. The work presented in this chapter sets the basis for exploring new methods to improve the trajectory of TMM in urban environments.

Further research includes the exploration of methods aiming to improve trajectory in urban scenarios independently of the density and type of non-semantic features (chapter 4) and/or to reduce the number of required ground control points (chapter 5). In addition, the use of vehicle dynamic constraints for trajectory estimation could be explored. Last but not least, the feasibility of applying the same ideas to other mobile mapping platforms that might incorporate consumer-grade sensors, could also be studied.

Chapter 4

POST-PROCESSING TRAJECTORY IMPROVEMENT FOR TERRESTRIAL MOBILE MAPPING IN URBAN ENVIRONMENTS

TOWARDS THE USE OF NON-SEMANTIC IMAGE INFORMATION IN TMM TRAJECTORY DETERMINATION

An edited version of:

Angelats, E. and Colomina, I., 2014. One step mobile mapping laser and camera orientation and calibration. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XL-3/W1, 15–20, https://doi.org/10.5194/isprsarchives-XL-3-W1-15-2014, 2014.

This chapter presents a first strategy to improve, in post-processing, the trajectory of terrestrial mobile mapping systems in urban environments by adapting the error modelling learned in Chapter 3 and by extending and validating the mathematical modelling presented in Chapter 2 for the urban case. The trajectory error modelling proposed in Chapter 2 is simplified by selecting a linear shift function instead of a more complex polynomial one. The models initially developed in Chapter 2 have been modified to be used also for urban environments. These models allowed us to estimate multiple linear shifts in a bundle adjustment approach, and thus improving the estimated trajectory, thanks to identifying common tie features between camera images and LiDAR point clouds and their use in a combined adjustment. These common tie features are straight line segments. The proposed approach has been evaluated using the same dataset used in Chapter 3 coming from a mobile mapping campaign over an urban and controlled area, with harmful GNSS conditions.

4.1. INTRODUCTION

Nowadays, 3D georeferenced data are widely used as the primary data, for many applications such as 3D city modelling, cadastral mapping, cultural heritage, facility management, traffic accident investigation, to mention a few examples. Dense 3D point clouds are the direct output of a mobile mapping system (MMS). An MMS is a particular case of a terrestrial laser scanner, where the sensor is mounted on a moving platform. The point clouds are generated kinematically, with a variable scanner position and orientation during the scanning time (Kutterer, 2010).

High-end mobile mapping systems integrate several laser scanners and several individual cameras or 360°cameras. A mobile mapping campaign may include several overlapping areas, leading to overlapping point clouds. Partial or total GNSS occlusions, or multipath, may occur, especially in urban areas. This causes an error in the platform or vehicle trajectory determination. Other sources of error can be Inertial Measurement Unit (IMU) modelling errors and eventually system calibration errors.

All these errors produced that overlapping point clouds are not properly registered (Figure 19). To solve this, a laser-laser registration must be performed. Also, some applications might require having coloured point clouds. Then, the coloured point cloud is usually derived from another registration process between camera and laser scanner.

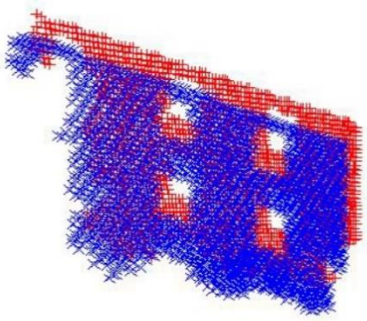

Figure 19. Point cloud of two different strips from overlapping area.

The standard procedure for laser-to-laser and camera-to-laser co-registration includes several steps. The first one is the system calibration where the lever arm and boresight between laser and IMU, and between camera and IMU must be determined. With current INS, GNSS and image matching technologies, the best results are obtained with the well-known Integrated Sensor Orientation (ISO) method. In ISO, tie points measurements are used in combination with ground and platform control observation. The ISO method has been proven to be feasible and efficient for IMU-camera boresight calibration of mobile mapping systems (Kersting et al., 2012). For aerial laser and mobile mapping data, ISO has also been proven to be effective for the IMU-laser boresight calibration with single and multiple laser scanners (Skaloud and Lichti, 2006), (Chan et al., 2013).

After the calibration steps, a camera and LiDAR point cloud can be derived. Then, a co-registration between LIDAR points clouds and between camera point cloud and LiDAR point cloud is computed. Alternatively, the LiDAR point cloud can be projected onto 2D image space, and then several image-to-image registration algorithms can be applied. The Iterative Closest Point algorithm (ICP) (Chen and Medioni, 1992), is widely used for registration of point clouds, enabling many improved variants (Novák and Schindler, 2013).

The ICP has been used for the registration of pairs of point clouds generated from laser, from pairs of overlapping images, or a combination of them. The ICP has been proven to be effective with point clouds that were initially aligned (Novák and Schindler, 2013). (Gressin et al., 2012) propose a method that uses ICP as the initial step for trajectory improvement. The overall registration is performed, improving the original platform trajectory. (Elseberg et al., 2013) deals also with laser-to-laser registration by improving the platform trajectory. The trajectory is improved using semi-rigid Simultaneous Localization and Mapping.

In contrast to the standard approach, in this chapter, we propose to solve the orientation and calibration of laser and camera data in a single, combined adjustment. Solving the orientation and calibration allows us to implicitly deal with the co-registration problem, because in our approach, we do not model the error symptoms but the error sources. The proposed method is based on the identification of common tie features between images and point clouds and their use in a combined adjustment. These common tie features are straight-line segments.

This approach presents several advantages. The use of laser raw and photogrammetric measurements, related through common tie features, allows us to avoid point cloud-based camera to LiDAR registration. In addition, the redundancy of the adjustment is increased, and the geometry is improved by means of including photogrammetric data into a LiDAR adjustment. The overall accuracy can also be improved by adding photogrammetric and LiDAR ground control information.

The chapter is organized as follows. Firstly, the main ideas behind the simultaneous network adjustment are introduced. Then, the proposed models as well as some mathematical concepts behind the use of tie lines are explained in detail. The concept validation section presents the preliminary results of an adjustment obtained with real data. The last section summarizes the conclusions of the proposed approach and discusses future improvements.

4.2. PROPOSED APPROACH

We propose a method where camera, laser and camera-laser ISO are performed in a single, combined adjustment. In it, the imaging observations are image coordinates measured on camera data (the digital camera images) and selected measurements of the laser images (ranges and scan-angles). The other observations are surveyed ground control points and platform time-Position-Attitude (tPA). In state-of-the-art mobile mapping systems, the tPA is previously estimated using INS, GNSS and odometer observations using a Kalman Filter approach. Ground control points can be used to derive indirect control observations like ground control lines and planes. The tie features are points (camera ISO), planar surfaces (laser ISO) and straight-line segments (camera-laser ISO). Points, planes and lines are defined by 3, 3 and 4 independent parameters, respectively.

The classical camera ISO approach is well known. From the observations mentioned above, estimates for the tie points (TP), for the exterior orientation parameters of the camera images (EO), for the self-calibration parameters of the camera, for the tPA shifts and possibly for other system calibration parameters are computed. In laser ISO, the unknown parameters of the tie planar bounded surfaces —tie planes

(TPL)— are estimated together with tPA shifts, laser self-calibration parameters and possibly other system calibration parameters.

In the proposed camera-laser ISO method, as mentioned, straight line segments --- Tie Lines (TL)--- are used as tie features between camera and laser images. TLs tie camera and laser images as follows. Assume that s is a TL that can be recognized in the camera image c and in the point cloud of a laser image a, and also assume that s is the intersection of two planar surfaces p_1, p_2 determined by points of the laser image a. Given s, c, a, p_1 and p_2 we will then use three types of observation equations. The first type of observation equation (a coplanarity equation) relates image measurements of the s TL on the c camera image with s TL parameters, with the EO and self-calibration parameters of c. At least two of these observation equations are needed. The second type of observation equation (a line-in-plane relationship) compels c to lie on planes. In general, there will be two observation equations of this type, for p_1 and p_2, although it is also possible to use just one TPL. The third type of observation equation relates TPLs ---in this case p_1 and p_2 --- to the measured laser points, i.e., relate the laser range and scan-angle measurements, to the laser self-calibration parameters, the derived tPA, the tPA shift parameters and possibly other system calibration ones. There will be, for each p_1 and p_2, as many of these observation equations as laser points that define the plane. Details of the models are given in the next section.

4.3. MATHEMATICAL MODELS

The mathematical models needed to perform a single, combined camera-laser ISO were introduced for the first time in (Angelats et al., 2012). In this work, airborne laser scanner and photogrammetric camera data were simultaneously adjusted using straight lines as tie segments. In order to expand this concept to mobile mapping systems, the presented models are still valid without modifications except for the MMS plane observations equations. In this section, we present in detail this model and the remaining models are briefly introduced. The interested reader may refer to (Angelats et al., 2012) for extended details of mathematical modelling.

We begin by presenting some notation conventions and then develop the observation equations.

4.3.1. Some naming and convention

l	Local-level terrestrial frame (east-north-up)
e	Earth centered Earth fixed frame
l'	Local-level terrestrial frame (north-east-down)
i	IMU instrumental frame (forward-right-down)
s	Laser instrumental frame
c	Camera instrumental frame

Table 11. Reference frames and coordinate systems.

4.3.2. Camera point collinearity observation equations

In camera ISO, the known collinearity observation equations are used to relate the camera image coordinate observations with the EO, tie point and self-calibration parameters. No modifications are required to use them in our method.

4.3.3. MMS plane observation equation

The model extends the one proposed by (Angelats et al., 2012) to be used in mobile mapping systems. The observations are the two MMS laser measurements (range and scan-angle) and the tPA (position and attitude) ones. The MMS plane observation equation is

$$\overrightarrow{n^l} \cdot \left(\overrightarrow{P^l} - \overrightarrow{P_0^l}\right) = d \tag{1}$$

where

$$\overrightarrow{n^l} = R(\lambda_0, \phi_0) \cdot R_y(\delta p) \cdot R_x(\delta l) \cdot \overrightarrow{n_0^l}\, d \tag{2}$$

$$\overrightarrow{P^l} = \overrightarrow{X^l} + \vec{A}X^l + R_i^l \cdot P^i \tag{3}$$

and where

$$P^i = R_s^i \cdot R_y(S_\theta \cdot \theta + \Delta\theta) \cdot \left(\vec{r} + \overrightarrow{\Delta r}\right) + \overrightarrow{a^i} \tag{4}$$

$$R_i^l = R_e^l \cdot R_{l'}^e \cdot R_i^{l'}\left(he + \Delta_{he}, pi + \Delta_{pi}, ro + \Delta_{ro}\right) \tag{5}$$

The model includes an additional scale factor self-calibration parameter for the scan-angle observation because our experience reveals its significance. The Hessian form of a plane is used to parametrize the planar surface. The plane is then characterized by a unit normal vector $\overrightarrow{n^l} = \left(n_x, n_y, n_z\right)^l$ (equation 5) and d, the orthogonal distance between the plane and the CRF origin. λ_0 and ϕ_0 are auxiliary vectors that bring an orthogonal vector $\overrightarrow{n_0^l}$ close to the normal vector. Then, the parameters of the plane are the orthogonal distance d and two rotation angles δl, δp. These rotation angles use to be quite small and are used to avoid numerical singularities.

The model has additional parameters to model tPA errors. These parameters are tPA position and orientation linear shifts. Table 12 summarizes the mathematical symbols of eqs. 1 to 5.

Observations	
$\vec{r} = (0, r, 0)$	MMS range
θ	MMS scan-angle
$\vec{X}^l = (X, Y, Z)^l$	platform position
$R_i^{l'}\ (he, pi, ro)$	platform attitude
Parameters	
$\vec{P}^l = (P_x, P_y, P_z)^l$	point object
$\vec{\Delta} X^l = (\Delta_x, \Delta_y, \Delta_z)$	platform position shift
$\Delta_{he}, \Delta_{pi}, \Delta_{ro}$	platform attitude shift
$\vec{\Delta r} = (0, \Delta r, 0)$	MMS range shift
$\Delta\theta$	MMS scan-angle shift
S_θ	MMS scan-angle scale factor
$\delta l, \delta p, d$	tie plane
$\vec{a}^i = (a_x, a_y, a_z)^i$	IMU-to-MMS lever arm
$R_s^i(E_x, E_y, E_z)$	IMU-to-MMS boresight matrix
Constants	
$R_e^l, R_{l'}^e$	rotation matrices from e to l and l' to e CRFs, respectively
$R(\lambda_0, \phi_0)$	auxiliary rotation matrix
$\vec{n}_0^l = (0, 0, 1)^l$	unit normal vector of tie plane
$\vec{P}_0^l = (P_{0x}, P_{0y}, P_{0z})^l$	local origin coordinates for tie plane.

Table 12. Symbols in the MMS plane observation equation.

4.3.4. 3D straight line parameterization

Straight lines are elementary mathematical objects with many possible parameterizations. We propose yet another one which is convenient for numerical computations and whose four universal parameters (p, q, α, β) can be used regardless of the line location and direction.

4.3.5. FC line coplanarity observation equation

This model relates the parameters (p, q, α, β) of a 3D line, the EO parameters of an image, and the image observations (x, y) of a point of the line. It has been derived from the collinearity condition and the 3D straight line parameterization presented in the previous subsection.

4.3.6. Line-in-plane observation equations

This model relates lines and planes in a 3D space. From a geometrical point of view, a line belongs to a plane if two conditions are satisfied. The first condition is related to the orthogonality, that is, plane's normal vector and line's director vector must be orthogonal. The second condition is related to the distance between line and plane. This distance must be 0 and it is equivalent that the line's point belongs also to the plane.

4.4. CONCEPT VALIDATION

The proposed concept was tested and validated using real data from a high-end mobile mapping system. The data originate from a mobile mapping campaign over a controlled area of Dortmund (Germany). The mobile mapping system was an Optech Lynx system, from TopScan GmbH, that included two laser scanners and two cameras. As for the software, an experimental "model toolbox" with the models discussed before was developed at CTTC that runs on the generic network adjustment platform GENA (Colomina et al., 2012).

The mobile mapping system was mounted on a van, and it was driven over a controlled area for around three hours, resulting in 11 overlapping strips. The controlled area was additionally surveyed to provide ground control points. The points clouds were generated, combining the laser raw measurements, the system calibration and the platform trajectory. The platform trajectory was computed, using a tightly coupled approach, combining differential GNSS, IMU and odometer measurements. The points clouds, for both laser scanners within the same strip, and

also between strips, were not co-registered. Photographs for each of the cameras, were also acquired. The platform trajectory was also used to provide an initial orientation of the photographs. The camera was previously calibrated, and the lens distortion was removed from the photographs.

Since the aim of this work is not an exhaustive study, a small subset of the original data has been used to validate the feasibility of the one step orientation and calibration of laser and camera data from mobile mapping systems. For that reason, only data from one laser scanner and one camera have been used. The main characteristics of the camera and laser scanner systems as well as the test configurations, are detailed in Table 13. The precision of the observations can be found in Table 14. Notice that from the camera-derived data, 4 tie lines have been extracted. These tie lines correspond to some straight elements located on building walls such as cables, window edges, i.e. On the other hand, from the MMS-derived data, 12 planes have been used. 11 of them are present in 3 strips while the remaining plane is only in 1. 4 planes of the first strip are used as ground control planes while the remaining planes are used as tie planes. Tie planes are vertical planes, corresponding to building walls, or horizontal planes, corresponding to street pavement. In this work, the straight lines used as a common feature, belong to the tie planes, but are not breaklines, that is, the ones resulting from the intersection of two tie planes.

Camera sub-block	
Equipment	Optech Lynx camera Applanix POS LV420
Image size	5.684 x 4.326 mm
Pixel size	3.5 μm
Camera constant (f)	3.866 mm
No. of strips	1
No. of images	4
No. of ground control points (GCPs)	1
No. of tie points (TPs)	37
No. of tie lines (TLs)	4
Horizontal Overlap $\approx$	60%
MMS sub-block	
Equipment	Optech Lynx laser Applanix POS LV420
Density (points/m^2) ($\approx$)	NA
No. of strips	3
No. of control planes (GCPLs)	4
No. of tie planes (TPLs)	12
Average no. of points per plane	1700
No. of tie lines (TLs)	2

Table 13. Dortmund block geometric configuration.

Observable	Precision (σ)	Units
Camera ic of TPs	$\sigma_{x,y} = 7$	μm
Camera ic of TLs	$\sigma_{x,y} = 10$	μm
GCPs	$\sigma_{e,n} = 3$, $\sigma_u = 5$	cm
Camera EO position	$\sigma_{e,n,h} = 20$	cm
Camera EO attitude	$\sigma_{\omega,\kappa} = 100$, $\sigma_{\varphi} = 200$	mdeg
MSS tPA position	$\sigma_{e,n} = 15$, $\sigma_u = 15$	cm
MMS tPA attitude	$\sigma_{pi,ro} = 5$, $\sigma_{he} = 8$	mdeg
MMS range	$\sigma_r = 0.05$	cm
MMS scan angle	$\sigma_\theta = 250$	μdeg

ic: image coordinates

Table 14. Precision of the observations.

With the given data, the configurations described in Table 15 have been processed. Block configuration cam_TP performs a classical camera ISO using all available tie points and a ground control point (GCPs) and EOs. Test cam_TL adds 4 straight lines as tie lines. This configuration allows us to validate the coplanarity equation model with tie lines. As for the MMS tests, an MMS orientation and calibration adjustment without camera data is performed using planes as tie features with four ground control planes (GCPLs). The ALL configuration combines camera data with MMS data using tie points, ground control points for the camera ISO, tie planes and ground control planes for the MMS ISO and tie lines for the camera-ALS ISO. This test allows us to test simultaneously all developed models.

Test	camera ic	GCPs	TPs	camera TLs	MMS rs	TPLs	GCPLs	MMS TLs
cam_TP	YES	1	YES	NO	NO	NO	NO	NO
cam_TL	YES	1	YES	YES	NO	NO	NO	NO
MMS	NO	0	NO	NO	YES	YES	YES	NO
ALL	YES	1	YES	YES	YES	YES	YES	YES

ic: camera image coordinates.
rs: MMS range and scan angle.

Table 15. Observations and block configurations.

The EO parameters are estimated for all proposed tests. The proposed approach includes additional parameters, such as a tPA shift, corrections in the boresight matrix between the IMU and the camera (Table 16) and self-calibration parameters. However, in this preliminary study, they are not estimated due to the small number of photographs and weak camera geometry, with only photographs from one strip and parallel to the vehicle movement direction.

For the MMS subblock, a tPA shift for each strip, can be estimated. Vertical tie planes allow us to estimate the e,n component while horizontal planes allow us to estimate the h component. Self-calibration parameters of the MMS, that is Δr, $\Delta\theta$ and S, cannot be estimated, neither the boresight matrix between IMU and laser scanner. These parameters cannot be estimated because they are strongly correlated between them, and the geometry of the sub-block does not allow decorrelating them.

	camera					MMS	
Test	EOs	TPs	BC	Shift	TLs	$(\Delta r, \Delta\theta, S_\theta)$	TPLs
cam_TP	YES	YES	NO	NO	NO	NO	NO
cam_TL	YES	YES	NO	NO	YES	NO	NO
MMS	NO	NO	NO	YES	NO	NO	YES
ALL	YES	YES	NO	YES	YES	NO	YES

BC: IMU-to-camera boresight matrix calibration.
Shift: tPA linear shift.

Table 16. Estimated parameters of block configurations.

The test results are shown in Table 17 and Table 18. As expected, due to the small data set and GCP configuration, FC_TP and FC_TL results show similar performance in terms of accuracy and precision (Table 17) for the EO parameters and tie points.

	RMSE TP (mm)			EO (mm, mdeg)					
Test	e	n	u	e	n	u	ω	ϕ	κ
cam_TP	48.8	220.0	58.7	45.8	66.8	57.4	44.4	89.3	44.2
cam_TL	47.5	211.7	57.8	45.2	66.4	57.3	43.3	87.1	43.4
MMS	-	-	-	-	-	-	-	-	-
ALL	47.6	211.9	57.8	45.0	66.4	57.3	43.5	87.0	43.2

Table 17. Camera test results.

The result of tests case ALL shows that the simultaneous block adjustment works, although this result in terms of precision is inconclusive given the small differences and the small size of the experiment. The results indicate that it is possible to estimate a tPA shift for each strip (Table 18), for MMS and ALL tests. From the results, it can be observed that the h component has a higher C_{xx} than e, n component. This result was expected because more vertical planes than horizontal ones were used as tie planes.

Test	Strip	Estimated Shift (mm) e	n	u	C_{xx} (mm) e	n	u
MMS	1	4.1	-3.0	92.8	1.9	1.1	164.6
	2	-280.2	559.4	410.3	2.4	1.3	164.6
	3	-56.7	159.6	277.0	2.0	1.2	164.6
ALL	1	4.8	-2.8	311.9	1.9	1.2	162.6
	2	-281.7	559.5	629.3	2.4	1.3	162.6
	3	-56.4	159.6	495.3	2.0	1.25	162.6

Table 18. MMS test results.

The co-registration of point clouds is implicitly done by improving the orientation and calibration. In order to evaluate the quality of the co-registration, a new point cloud for each strip has been generated, adding the tPA shift to the original tPA. The point cloud differences, between pairs of strips, before and after the registration has been computed using the point cloud distance tool of Cloud Compare SW (Giradeau-Montaut et al., 2014). The differences, before and after the registration, are shown in Figure 20.

The differences are also computed for each component. We use this SW, only to show the potential of the approach and give a coarse point cloud difference. Further research must be done to compute accurately the distance between point clouds, eventually leading to better results.

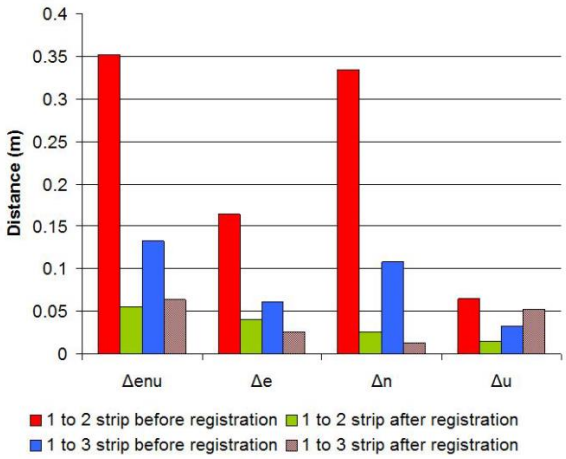

Figure 20. Point cloud differences from overlapping strips.

4.5. CONCLUSIONS AND FURTHER RESEARCH

The mathematical models: 3D straight line parameterization, a coplanarity observation equation and a line-belongs-to-plane, presented in (Angelats et al., 2012) have been proven to be effective, with real data from a high-end mobile mapping system. A new MMS model has been successfully developed and tested. In the MMS model, laser raw measurements are used to allow for system and sensor calibration. In addition, also includes a tPA shift to model platform tPA errors due to GNSS occlusions, multipath.

The results of this chapter are only preliminary because of the small size of the data sets. Photogrammetry results show that with this combined camera, MMS, and camera-MMS, concept, the performance in terms of precision, is at least maintained. The results also suggest that the platform trajectory can be improved with a linear shift. With this trajectory improvement, the co-registration between overlapping point clouds, is improved. The MMS self-calibration has not been performed due to weak geometry and a small dataset. Further research must be done to explore how the combination of photogrammetric data can benefit the MMS the system calibration, and so, the co-registration.

With this approach, a single or several camera and laser data from different systems or from the same system with different acquisition, can be integrated and so, co-registered. The next step is to validate the presented concept and models with larger datasets and to understand the behaviour of camera and MMS self-calibration parameters in combined adjustments. For larger datasets, alternative models than a linear shift to tPA correction, must be explored. Additionally, not only the precision, but also the accuracy of the one step approach must be measured. Further research must be done to explore the number of required ground control points and planes, since previous experience suggests that the combined adjustment could reduce the required ground control information.

Chapter 5

MULTISENSOR NAVIGATION IN GNSS-DENIED ENVIRONMENTS

TOWARDS A REAL-TIME POSITIONING USING NON-SEMANTIC FEATURES

An edited version of:

Angelats, E., Molina, P., Parés, M. E. and Colomina, I., 2014. A parallax based robust image matching for improving multisensory navigation in GNSS-denied environments. In: Proceedings of the 27th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2014), Tampa, USA, pp. 2132–2138.

This chapter explores the potential of using non-semantic information (points) for trajectory estimation in real time. To do so, a two steps methodology has been developed. The first step includes a method to remove outliers in photogrammetric measurements extracted from overlapping images, using relative orientation derived from inertial-based trajectory. By means of relative orientation, the inter-distance between rays traced from corresponding tie points, named parallax, is used for outlier detection and isolation. Once outliers are removed, tie point measurements are used in a second step to improve the trajectory estimation. Image and object space coordinates of the corresponding tie points contribute, through appropriate modelling, as position and attitude updates to the filtering step of a non-linear sequential least square multi-sensor positioning software. The proposed approach has been evaluated using both real and simulated data from a mobile mapping campaign over an urban area with long GNSS outage periods.

5.1. INTRODUCTION

Navigation or real-time positioning is the process of estimating a time series of position, velocity and attitude states for a particular sensor or platform. The quality of navigation in general and, in particular, of hybrid or multisensor navigation depends on the quality of measurements and models among other factors. For convenience, sensors producing measurements for navigation are classified as primary such as an Inertial Measurement Unit (IMU), and aiding sensors such as GNSS receivers, RGB cameras and/or laser scanners. In recent times, the need and complexity of ubiquitous and seamless outdoor/indoor navigation and navigation in GNSS unfriendly environments, have made apparent the potential of measurements derived from aiding sensors. When aiding sensors are cameras and laser scanners, we will refer to these as visual measurements (or observations).

In the context described above, visual measurements can be used to correct the drifts of the primary navigation sensors such as an IMU. However, aiding can be easily turned into "unaiding" if sensors' random, systematic and/or gross errors are too large or too many. When designing a navigation system and/or mission, the time-dependent behaviour of the mentioned errors has to be understood: the propagation of random and systematic errors and the impact of outliers have to be quantified, as well as the means to mitigate or eliminate them. Potential systematic error sources are mainly due to coarse camera calibration or a lever arm and boresight between camera and IMU misalignments.

As previously mentioned, the focus of this work is robust inertial visual-aided navigation. Visual aiding refers to the computation of orientation parameters through consecutive, overlapping images by means of measurement of tie points –

i.e., photogrammetric observations of the same object point in two or more images-. Robustness refers to the insensitivity of the resulting orientation parameters to outliers in the photogrammetric measurements. Alternatively, planes or cylinders extracted from laser scanner images, can also be used as tie features, especially in urban or indoor areas, where they are very common. Yet, the work presented in this chapter does not entail the use of features extracted from laser scanner data.

Extensive research and many algorithms have been proposed to robustly extract, describe, and match common points, ideally invariant to orientation, scale and illumination changes. However, urban or indoor areas, especially buildings, contain repetitive or similar structures, causing a high number of matching outliers even if one of state-of-the-art point extraction, description and matching algorithms are used. Examples of these are the well-known SIFT, SURF algorithms or the more recently BRISK (Lowe, 1999), (Bay et. al, 2008), (Leutenegger et. al, 2013).

Feature extraction and matching algorithms are commonly combined, with RANSAC procedures that aim to perform outlier detection and removal using solely image observations by means of position and attitude estimation. More in detail, these algorithms estimate the relative orientation between two images, also known as pose, using n pairs of matched features, selected randomly among the full N matched pairs ($N > n$). By doing so, outliers are detected and isolated by checking inconsistencies in the process of relative orientation estimation. The 5-point algorithm proposed by (Nister, 2003) is widely used but other approaches use n = 7 or 8 to deal with uncalibrated cameras (Scaramuzza and Fraundorfer, 2011). (Scaramuzza, 2011) proposes a method where the camera pose can be estimated with only one point correspondence by exploiting nonholonomic constraints of wheeled vehicles and a histogram-based voting strategy. Other approaches combine derived trajectory or inertial data to predict where a point feature should appear in the second image (Veth, 2011), (Leutenegger, 2013).

Once all outliers have been detected and isolated, the camera or vehicle trajectory can be recovered by concatenating the estimated relative positions and orientations using k inliers from overlapping images. Several methods to estimate the navigation states using only image or using image and object observations are reviewed in (Scaramuzza and Fraundorfer, 2011). These approaches are usually referred to as Structure from Motion (SfM) or visual odometry among the robotics and computer vision community. Alternatively, (Taylor et.al, 2011) presents two strategies to use an IMU as a primary positioning sensor and to control inertial drift with visual information during the filtering step. The two approaches are implemented using an Unscented Kalman Filter estimation method. The first approach imposes a geometric constraint using image coordinates while the second one takes the benefit of jointly estimating a set of object coordinates together with navigation states.

In contrast to the previous approaches, we propose a strategy to robustly detect and isolate outliers in camera images using inertial-based trajectory in a first step and to navigate in GNSS-unfriendly environments, using corresponding tie points measurements together with inertial and GNSS measurements, when available, in a second step. The filtering step is based initially on the computation of an inter-distance between rays traced from corresponding tie points, named parallax, (Figure 21). Then, a robust statistical analysis based on median and median variance estimators is applied to the inlier set to clean few 'transparent' outliers.

This work is organized as follows. Firstly, the main ideas and some mathematical concepts behind the proposed outlier detection and isolation method are introduced. Then, the proposed model for visual-aided navigation is explained in detail. The experimental results section presents the results for outlier filtering as well as multisensor navigation using real data from a terrestrial mobile mapping campaign with long GNSS outages. The last section summarizes the conclusions of the proposed approach and discusses future improvements.

5.2. NAMING AND CONVENTION

We begin by presenting some notation conventions used in this section and also for the visual aiding modelling. The Coordinate Reference Frames (CRFs) involved in outlier filtering and in visual aiding are detailed in Table 19. The CRF ! of a variable X is denoted by a superscript symbol like X^l. For a rotation matrix R_b^l, the subscript indicates the origin CRF while the superscript indicates the final CRF so it can be written $X^l = R_b^l \cdot X^b$. For the sake of simplicity, $\vec{X}^e = (x, y, z)^e$ is used instead of the formally correct $\vec{X}^e = [(x, y, z)^e]^t$.

!	Local-level terrestrial frame
e	Earth centered Earth fixed frame
b	IMU instrumental frame
c	Camera instrumental frame

Table 19. Notation convention in visual aiding model.

5.3. OUTLIER DETECTION AND ISOLATION

Given a pair of images with a certain overlap over the same scene, N pairs of corresponding points, called tie points, can be found and matched using one point extractor, description and matching algorithm. These are written as:

$$(x, y, -f)^{c_1}, (x, y, -f)^{c_2}, i = 1...N.$$

Note that taking f or $-f$ depends on the selected camera reference frame.

The proposed algorithm uses the image coordinates and the exterior orientation (EO) of both images that is, its position and attitude, to detect and isolate outliers. The EO of the first image is known, corresponding to a processing epoch in the past, while the second EO is predicted using inertial data, corresponding to the current processing epoch. This approach uses directly absolute inertial-based EO in contrast to other approaches based on the estimation of fundamental or essential matrix to retrieve the relative orientation of an image pair (Hartley and Zisserman, 2000).

For each matched pair, using the EOs $(X_{c_1}, Y_{c_1}, Z_{c_1})^l$, $(X_{c_2}, Y_{c_2}, Z_{c_2})^l$ and $R^l_{c_1}, R^l_{c_2}$ expressed in a local frame and the image coordinates, two projecting lines (Figure 20) are constructed using the normalized director vectors $\vec{v}^l_1, \vec{v}^l_2$.

$$l_1 = (X_{c_1}, Y_{c_1}, Z_{c_1})^l + \lambda_1 \cdot \vec{v}^l_1 \quad (1)$$

$$l_2 = (X_{c_2}, Y_{c_2}, Z_{c_2})^l + \lambda_2 \cdot \vec{v}^l_2 \quad (2)$$

Where:

$$\vec{v}^l_1 = (v^{c_1}_x, v^{c_1}_y, v^{c_1}_z)^l = \frac{R^l_{c_1} \cdot (x, y, f)^{c_1}}{\sqrt{x^2+y^2+f^2}} \quad (3)$$

$$\vec{v}^l_2 = (v^{c_2}_x, v^{c_2}_y, v^{c_2}_z)^l = \frac{R^l_{c_2} \cdot (x, y, f)^{c_2}}{\sqrt{x^2+y^2+f^2}} \quad (4)$$

The following equations express the relation between the IMU instrumental frame and the camera instrumental frame:

$$(X_c, Y_c, Z_c)^l = (X_o, Y_o, Z_o)^l + (la_x, la_y, la_z)^l \quad (5)$$

$$R^l_c = R^l_b \cdot R^b_c \quad (6)$$

Where $(X_o, Y_o, Z_o)^l$ is the vehicle position, $(la_x, la_y, la_z)^l$ is the lever arm between IMU and camera, R_c^b is the boresight matrix between the camera frame and the IMU frame while R_b^l is a rotation matrix from body frame to local frame.

Vector magnitude λ_2 is estimated by searching for a point in l_2 that minimizes the distance to line l_1 using the 'point-to-line' distance formula:

$$d(\vec{x} + \lambda_1 \cdot \vec{v}_1, \vec{p}(\lambda_2)) = \left|(\vec{x} - \vec{p}(\lambda_2)) - ((\vec{x} - \vec{p}(\lambda_2)) \cdot \vec{v}_1^t)\vec{v}_1\right| \tag{7}$$

Analogously, knowing λ_2, λ_1 can be found using the same formula. Then, we analyze the function:

$$\begin{aligned} &P : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R} \\ &((x,y)_1),(x,y)_2) \mapsto \left|\vec{x}_1 + \lambda_1 \vec{v}_1^l - \vec{x}_2^l - \lambda_2 \vec{v}_2^l\right| \end{aligned} \tag{8}$$

Which corresponds to the minimum distance between the lines (1) (2). This distance is referred to as parallax. Pair of points with a parallax below a certain threshold are considered to be a correct matching. Alternatively, (Hartley and Schaffalitzky, 2004) propose a method to minimize this distance involving the solution of a sixth-degree polynomial for the case of two views and a L_∞ minimization for the case of n views.

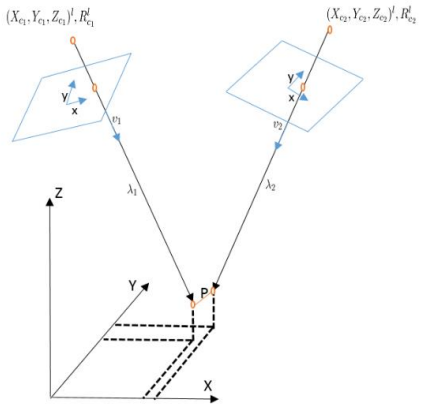

Figure 21. Geometric view of pair of images and the parallax of a pair of given points.

Some 'transparent' outliers may still be present in the inlier set after the parallax-based filtering. Thus, an additional filtering step is applied in order to detect and isolate them from the inlier set. To do that, a ratio value between λ_2 and λ_1 is calculated for all pair of points and a median and median variance are computed. All

points with a distance from median value below a certain threshold are kept as inliers.

5.4. VISUAL AIDING MODELLING

The proposed model relates, object coordinates of a tie point with its image coordinates in an overlapping image pair. To do so, the platform position and attitude of the two images are used, as well as the boresight and lever arm between and the IMU and camera. This model is intended to be used during the filtering step, in a generic non-linear sequential least squares solver together with inertial and GNSS data or with inertial data during GNSS outages. The model contributes with position and attitude updates to the vehicle trajectory estimation but relies on the use of static points only. Note that an object moving between epoch 1 and 2, does not fit within the model assumptions, and thus can significantly degrade the quality of the estimated trajectory.

The observation equations are the well-known photogrammetric collinearity equations (Luhmann et.al, 2006) extended with an additional correction each tie point object coordinates (eq 9 and 10).

$$\vec{P}^e + \vec{\Delta P}^e = \vec{X}^e_{c1} + \mu_1 \cdot R^e_{b1} \cdot R^b_c \cdot (x1, y1, -f) \tag{9}$$

$$\vec{P}^e + \vec{\Delta P}^e = \vec{X}^e_{c2} + \mu_2 \cdot R^e_{b2} \cdot R^b_c \cdot (x2, y2, -f) \tag{10}$$

Where:

$$\vec{X}^e_{c1} = \vec{X}^e_{o1} + R^e_{b1} \cdot \vec{la}^b \tag{11}$$

$$\vec{X}^e_{c2} = \vec{X}^e_{o2} + R^e_{b2} \cdot \vec{la}^b \tag{12}$$

In the above observation equations, the platform position and attitude are estimated, in the second image epoch. In addition, for each point, corrections for object coordinates will also be estimated using observations presented in Table 20. Note that, as we use corrections for the estimation of object tie points coordinates, we introduce a constant value for the tie point object coordinates. These constant values are obtained, as an intermediate step of the parallax-based filtering. The units of tie points image and object coordinates as well as the focal length are in meters.

Observations	
x_1, y_1	Tie point image coordinates of 1st image
x_2, y_2	Tie point image coordinates of 2nd image
$\vec{X}^e_{o1} = (X_{o1}, Y_{o1}, Z_{o1})^e$	Platform position for 1st image
R^e_{b1}	Platform attitude for 1st image
R^b_c	Camera to IMU boresight matrix
$\vec{la}^b = (la_x, la_y, la_z)^b$	Camera to IMU lever arm
Parameters / states	
$\vec{X}^e_{o2} = (X_{o2}, Y_{o2}, Z_{o2})^e$	Platform position for 2nd image
R^e_{b2}	Platform attitude for 2nd image
$\Delta\vec{P}^e = (\Delta P_x, \Delta P_y, \Delta P_z)^e$	Correction of tie point object coordinates
Constants	
f	Focal length
$\vec{P}^e = (P_x, P_y, P_z)^e$	Tie point object coordinates

Table 20. Symbols in visual aiding model.

The use of image coordinates during the filtering step presents some advantages in front of alternative models that provide position and attitude updates derived from solely camera images. Firstly, the image coordinates residuals are available and therefore, one has the capability to detect any remaining outliers with statistical methods. Secondly, there is no need to estimate an uncertainty for the position and attitude derived from camera images which may be unrealistic. Yet, there is a need to estimate an uncertainty for image coordinates that might not be straightforward.

It is important to remark that in order to estimate the EO of a given image, image measurements can be used from any other past other epoch as long as it has overlap with the current image. Therefore, the images may not necessarily be consecutive. Note also that in Table 2, the scale factor μ_1 is not included with the parameters because it can be eliminated dividing 1st and 2nd equation by the 3rd (Luhmann et.al, 2006). The same procedure can be applied for μ_2, dividing 4th and 5th equations by the 6th. In fact, equations (9) and (10) are a system of six equations.

This model does not include camera calibration parameters and assumes that camera has been previously calibrated and lens distortions from images have been removed. The proposed model also assumes that geometry between the camera and the IMU is known.

5.5. EXPERIMENTAL RESULTS

The proposed concept was tested and validated using real and simulated data from a high-end mobile mapping system. The data originate from a mobile mapping campaign over a controlled urban area of Dortmund (Germany). The mobile mapping system was an Optech Lynx system, from TopScan GmbH, that includes two laser scanners and two cameras. The system also incorporates a navigation component, an Applanix POS-LV with two GNSS receivers, odometer and a tactical grade IMU, used in this work as reference trajectory.

The mobile mapping system was mounted on a van and it was driven over a controlled area experiencing long GNSS outages around three hours, resulting in 11 overlapping strips, that is, repeating twice the same street driving in the same direction. Photographs of the two cameras were also acquired. The cameras were previously calibrated and the lens distortion was removed from the photographs. The geometry between the IMU and the cameras was also provided.

A small subset of the original data has been used to validate the feasibility of the proposed point matching filtering method and the hybrid trajectory estimation using GNSS, inertial and photogrammetric data over an urban canyon area. An ideal GNSS trajectory was extracted at 1Hz from the reference INS/differential GNSS trajectory and some outages of 32 and 12 seconds were manually introduced to test the potential of the approach in front of long/moderate GNSS outages (Figure 22). In addition, to evaluate the performance of the approach for various IMU grades, two configurations were studied. Firstly, real data from the Lynx's system IMU was used, consisting of tactical grade IMU. Secondly, simulated data from a tactical grade and from a MEMS grade IMU have been obtained with CTTC's IMU simulator (Parés et.al, 2010). Although the odometer data were available for this mobile mapping campaign, they have not been used for trajectory estimation.

As for the software, for all pairs of consecutive images, tie points were extracted and matched using the OpenCV library implementation (Bradski, 2000) of SURF operator (Bay et.al, 2008). Then, image point matches were filtered using the proposed algorithm (Figure 23).

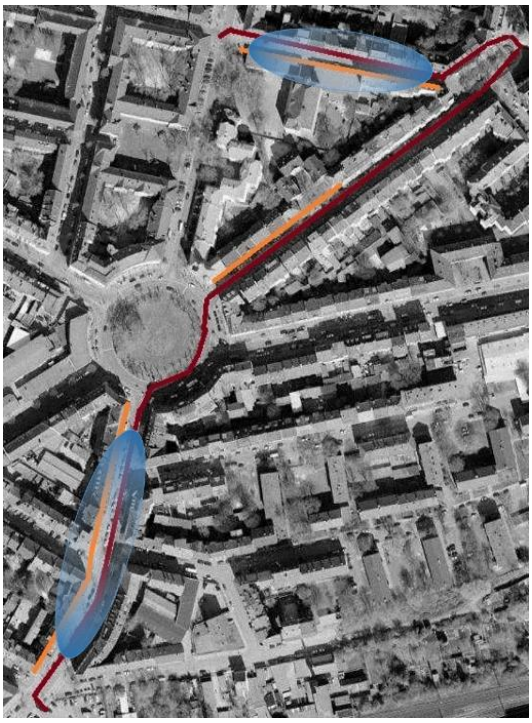

Figure 22. Controlled area used for testing the approach. The platform trajectory is shown in red while GNSS outage areas are marked with blue ellipses. Orange lines correspond to areas where tie points have been extracted, matched and filtered from overlapping, consecutive images.

The proposed visual aiding model has been implemented, validated and tested, using an in-house trajectory estimation SW, named NAVEGA. The platform trajectory has been estimated in post-processing, only in forward mode, emulating a real-time scenario, to evaluate the potential of the approach. NAVEGA provides the estimated trajectory to the matching component, in a close loop (Figure 24). Recall that trajectory is a necessary input for the outlier detection and isolation algorithm.

The estimated precision of a trajectory determined with simulated IMU can be seen in Figure 25. The blue line corresponds to the standard deviation of a trajectory using GNSS/INS while the green line corresponds to camera/GNSS/INS.

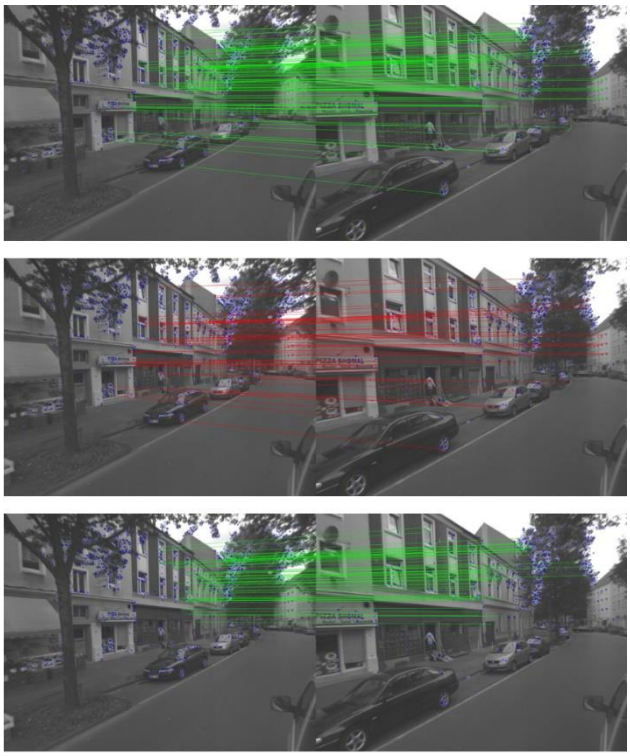

Figure 23. Example of the outlier detection and isolation method for a pair of overlapping images. Initial point matches (top image). Detected and isolated outliers (middle). Final matches or inliers set using for trajectory estimation (bottom).

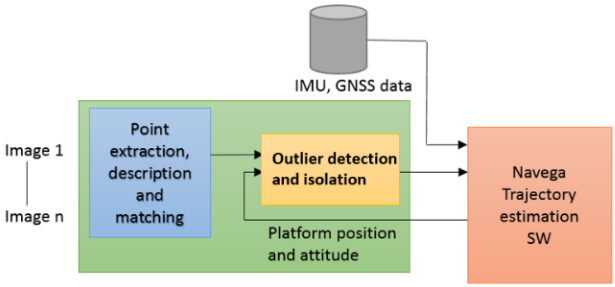

Figure 24. Multisensor navigation architecture.

From the results, the GNSS outages periods can be identified by looking the longitude and latitude standard deviation plot (top and middle plot, blue line). GNSS outages imply an increase in standard deviation from the previous epoch.

The camera was mounted looking to a lateral view. Due to this, we can expect an improvement in the trajectory estimation in the across-track component of the position. The results using image observations (green line) confirm this behaviour. The standard deviation of the across-track component is kept under control even if GNSS outages are present (Figure 25, green line, top and middle). In contrast, it can be seen that image observations have no relevant impact in the attitude standard deviation (Figure 25, bottom), mainly dominated by the IMU. In order to keep the along-track component under control during GNSS outages, odometer data might be also used.

Note that an EKF or non-linear sequential least squares tend to provide optimistic standard deviation values. For that reason, one must take the estimated values only as a way to evaluate the potential of using camera observations during GNSS outages and not as way to evaluate the absolute performance.

The results of the position performance using real data from a tactical grade IMU are shown in Figure 26. The figure shows the difference, for each component, between the estimated position and the reference one. The red line shows the results using inertial and GNSS, when available, observations. The green line also incorporates the image observations to the trajectory estimation. From the results, two different things can be highlighted. Firstly, the visual-aided model significantly improves the height component estimation allowing cm level differences with the reference trajectory during GNSS outages (Figure 26, bottom plot). Secondly, the use of image observations, together with the IMU real data, does not improve yet as expected the across-track component of trajectory.

This could be due to a weak camera geometry configuration. The camera images come from the predefined camera configuration of a mobile mapping system. Their main purpose was to add color to laser point clouds, not to be used for trajectory estimation. For that reason, the camera was not mounted completely perpendicular to the vehicle forward direction, that is looking directly to the building walls. This mounting configuration, together with the fact that the vehicle was driving in narrow streets, led to a non-uniform distribution of tie points. Moreover, the vehicle was driving slowly, producing a short baseline between consecutive, overlapped images.

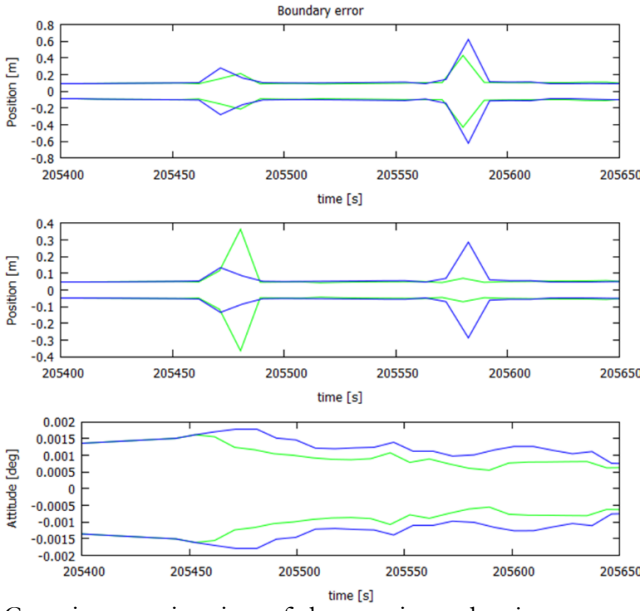

Figure 25. Covariance estimation of three estimated trajectory components using the simulated IMU data. Longitude (top), latitude (middle) and heading (bottom). The figure shows the results with (green line) and without camera observations (blue line).

Figure 26. Position performance using real IMU data. The figure shows the difference between the estimated position and the reference position, for each

component using (green line) and without using image observations (red line). Longitude (top), latitude (middle) and height (bottom).

5.6. CONCLUSIONS AND FURTHER RESEARCH

A strategy to robustly detect and isolate outliers in camera images using inertial-based trajectory has been developed and tested. Imaging measurements of tie points are used to improve the trajectory estimation in a second step. The proposed outlier detection algorithm is also a mechanism to provide initial values of the tie point object coordinates, used later for trajectory estimation.

The performance results with simulated data show the potential of visual-aiding navigation during GNSS outages over urban areas. The performance results with real data are still preliminary and must be refined. However, these preliminary results suggest us several improvements that can be carried out in the near future.

Among them, we are planning to test additional camera mounting configurations and to explore additional models for using measurements from more than two overlapping images. These two ideas would eventually improve the camera geometry. In addition, we would explore the potential of using a second camera and the benefits for trajectory estimation purposes. The use of measurements from a wheel odometer will be incorporated into the trajectory estimation step to improve the along-track component of the position. Last but not least, we will analyze the potential of the approach in multipath environments, not only with full GNSS outages.

Chapter 6

REMOTE SENSING AND AERIAL DATA FUSION USING NON-SEMANTIC FEATURES

THE ROLE OF NON-SEMANTIC FEATURES ON ENHANCED RPAS POSITIONING

An edited version of:

Angelats, E., Mas-Florit, C., Cau-Ontiveros, M.A. “High resolution orthophotos and a digital surface model of the Roman city of Pollentia (Mallorca, Spain) using RPAS imagery, aerial images, and open data archives”, in Proceedings of the 2018 IEEE International Conference on Metrology for Archaeology and Cultural Heritage, 22–24 October 2018, Cassino (Italy)

This chapter introduces a first application of hybrid positioning. This is application is the co-registration of multi-sensor and multi-temporal RPAS and historical georeferenced datasets, and its use in the archaeological domain. This georeferencing consistency is achieved by an enhanced RPAS positioning linking high-accurate aerial reference trajectories, aerial and RPAS mapping using non-semantic image ground features. Two characteristics of this approach can be highlighted. The first is the use of ground control support derived from aerial and publicly accessible aerial data instead of organizing a specific surveying campaign. The second is the capability to implement the proposed approach without modifying the state-of-the-art RPAS photogrammetric workflows and by using RPAS photogrammetric software tools. The approach has been experimentally tested and validated in the Roman city of Pollentia (Alcúdia, Mallorca, Spain).

6.1. INTRODUCTION

Latest advances in Remotely Piloted Aircraft System (RPAS) technology, Commercial-of-the-self miniaturized cameras, and photogrammetric software have allowed the archaeological community the access to this technology beyond research and experimental projects (Campana, 2017), (Nex and Remondino, 2013), (Opitz and Herrmann, 2018). Recent works have shown the potential and utility of the outputs of these technologies (orthophoto, point clouds and Digital Surface Model (DSM) for the management of archaeological sites, both for documenting and a 3D modelling perspective (Campana, 2017), (Cowley et.al, 2018), (Lo Brutto, 2014). Thanks to the capability of flying at low altitudes, dense point clouds and orthophotos with a high-level ground sampling distance (GSD) can be generated. With that, the GSD is increased from meter (satellite imagery) or decimeter (aerial imagery) to few centimeters (Nikolakopoulos, 2016). This improved resolution may allow digitalizing an archaeological site at a level where stones may be clearly distinguished [1,6] or provide a detailed scale of the structures (Figure 27).

These outputs provide a suitable alternative to total stations by relaxing requirements in terms of metric accuracy. Moreover, these technologies may also help to detect potential buried archaeological remains when multispectral orthophotos are generated. The advantage in this case is the capability to acquire data in the optimal time window for detecting soil or cropmarks anomalies (Masini et.al, 2018), (Stek et.al, 2016).

Figure 27. Comparison of aerial orthophoto (25 cm GSD) and RPAS orthophoto (1 cm GSD) resolutions showing some structures in the Roman city of Pollentia.

From an archaeological perspective it is important to have historical data, acquired over many years, properly georeferenced and co-registered. It is essential to have high-resolution orthophoto and DSM, but they might be meaningless if such data, obtained in different time frames, are not properly registered or aligned with aerial orthophotos. An example of these archives are those provided by the Spanish Plan Nacional Ortofotografía Aérea (PNOA) program. The PNOA is a national program led by the Spanish National Geographic Institute. The goal of this program is to generate and make publicly available high-resolution orthophotos and Digital Terrain and Surface Models. These products are generated for the entire country and updated every 2-3 years. The program also provides access to raw imagery and auxiliary data used to generate the aforementioned products. The auxiliary data includes the position and attitude of the raw imagery as well as the camera calibration.

The standard workflow for orthophoto and DSM generation includes the following steps: aerial triangulation, point cloud and digital surface model generation and finally orthophoto generation. The aerial triangulation step involves the estimation of exterior orientation (position and orientation) parameters as well as camera calibration parameters. These parameters are usually estimated using homologous, exterior orientation observations provided by onboard Global Navigation Satellite System (GNSS) receivers and Ground Control Points (GCP). These points are usually surveyed with differential or Real Time Kinematics (RTK) GNSS techniques (Lo Brutto et.al, 2017). However, it may happen that the resulting orthophoto and DSM are non-proper registered with the aerial cartographic archives. Thus, an additional step may be required to register multi-temporal and multi-resolution

orthophotos using an affine or projective model. Registering the DSM may be more difficult or even impossible due to the difficulty of identifying common points.

Alternatively, several solutions have been proposed to deal with the co-registration of multi-temporal datasets during the orientation step, prior to orthophotos and DSM generation (Aicardi et.al, 2016), (Zhuo et.al, 2017). Both solutions focus on the automatic detection of common points between RPAS datasets (Aicardi, et.al, 2016) or between RPAS and aerial datasets (Zhuo et.al, 2017), but differ in the way in which the ground control information is generated. The first solution generates the ground control information by performing initially the image orientation step for one of the dataset (Aicardi et.al, 2016). Then some images in such dataset are used as anchor images to constrain the orientation step of the remaining dataset (or datasets) without using ground coordinates of common points. (Zhuo et.al, 2017) uses planimetric coordinates of common points extracted from available orthophotos and elevations rom Digital Surface or Digital Terrain Models (DTM).

The main limitation of the aforementioned solutions is that it relies on the capability to automatically identify common points between datasets. This cannot be straightforward for datasets with different GSD, long temporal gaps where few common points can be detected. In this context, the reliability of the previous approaches depends of the number of detected points and their distribution.

In this chapter, an approach to generate high-resolution orthophotos from RPAS imagery and DSM is presented avoiding the use of GCPs measured with GNSS techniques. Nevertheless, it is still being possible to co-register orthophotos from aerial archives and RPAS imagery. The innovation of the approach is to avoid the use of GNSS-based ground control points. They are replaced by GCP derived from aerial imagery and exterior orientation data. Note also that these data were used to generate open and broadly accessible aerial orthophoto archives (PNOA). The key point of the approach is the capability to use reliable ground control points triangulated from aerial images that were already used to generate aerial orthophotos. Thanks to this point, the RPAS orthophotos, as well as the DSM, will be co-registered with the PNOA orthophotos avoiding extra co-registration steps. Additionally, the proposed approach may help to potentially reduce the surveying cost by avoiding the use of coordinates of GCP surveyed with GNSS RTK techniques.

6.2. PROPOSED APPROACH

The proposed workflow includes several steps (Figure 28). The first one is the generation of ground points, that is GCP and Checkpoints (CHP). The GCP are used in the triangulation step while the CHP are used only for validation purposes. In this approach, the ground points are observed both in aerial and RPAS imagery.

The coordinates of these points are obtained using image triangulation techniques instead of deriving the ground coordinates of these points from the planimetric coordinates of the orthophoto and the height from the DSM. This can be done in this way because the exterior orientation of the aerial images in the archives as well as camera calibration data are already open and available to anyone upon request.

The planimetric coordinates of the points derived directly from the orthophotos are very good (25 cm resolution GSD in the PNOA program). However, this approach is used because it provides better estimates of the height component than the ones provided in the DTM. Open and available DTM (also from PNOA) provides height in a grid of 5x5 m although it can be slightly improved by interpolating height from nearby cells.

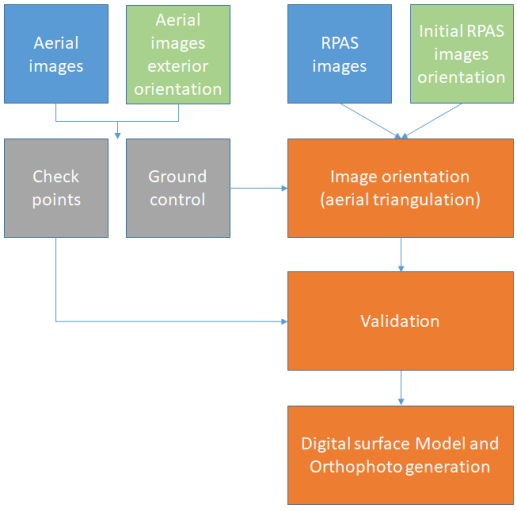

Figure 28. Proposed workflow for processing of RPAS imagery using GCP extracted from aerial and open archives.

The selection of the number and type of GCP is made taking into account two considerations: spatial distribution and temporal stability of points. Regarding the first point, GCP should be distributed covering all the area surveyed by the RPAS imagery. The selected points should be easily recognizable in both aerial and RPAS imagery and invariable in time. Examples of potential candidates are points belonging to fixed structures, and permanent road markings. Examples to avoid are points belonging to vegetation or temporal road marking.

The next step is the aerial triangulation of RPAS imagery involving the following observations: image coordinates of homologous points between images, initial orientation provided by a GNSS receiver and image and ground coordinates from GCP. The exterior orientation of every image as well as the camera calibration parameters are estimated using a bundle adjustment approach. This step is validated using root mean square error (RMSE) between the estimated coordinates of set of checkpoints from aerial triangulation process and the ground coordinates of these points. The final steps, common in many photogrammetric software tools, involve the dense point cloud, DSM and orthophoto generation.

The proposed approach has several advantages: first, the capability to fuse multi-temporal and multi-resolution datasets to obtain a historical perspective of the archaeological excavations. This is done thanks to integrating all data available in a common georeferencing frame. The second advantage is that the structure-from-motion software pipelines may be used with no modification by this approach. The third advantage is the capability to have centimetric precision (less than 10 cm), allowing the clear identification of the different elements composing an archaeological structure, instead of the general form that can be observed from aerial orthophotos. Another advantage is that with the proposed approach, the georeferencing accuracy of the PNOA orthophotos is kept while the resolution might be increased to few centimetres depending on the camera used and the flight altitude. The main limitation of the approach is the lack of automation in the detection of GCP. It relies on the manual identification of GCP and their image coordinates.

The methodology has been tested and validated in the Roman city of Pollentia with the aim of being extended to other archaeological sites.

6.3. CASE STUDY: POLLENTIA

Pollentia is located at the city of Alcudia, in the island of Mallorca (Spain). The ancient city was identified in the 19th century and continuous excavations have been carried out since 1923. Excavations have uncovered a residential area (Sa Portella), a theatre, part of the forum, several necropolises, and other remains of the city (Figure 29). Nevertheless, several areas are still covered, and potential buried remains might be located also in the nearby fields beyond the current limits of the Roman city (Cau-Ontiveros et.al, 2017), (Ranieri et.al, 2016).

6.3.1. Datasets

The datasets include PNOA aerial archives and RPAS imagery. Regarding the aerial archives, three different overlapped aerial images, used to generate the 2015 PNOA orthophoto of the area together with the exterior orientation and camera calibration values, were available. The 4-band images acquired with a high-performance metric aerial camera (Vexcel UltracamXP) were taken providing a GSD of 21 cm. The RPAS imagery (5-band images) were acquired with a Micasense RedEdge multispectral camera that was flown in a fixed-wing RPAS over the Roman city of Pollentia and nearby fields. 150 images were collected corresponding to a GSD of 10 cm. No laboratory camera calibration was available beyond the information provided in the EXIF file.

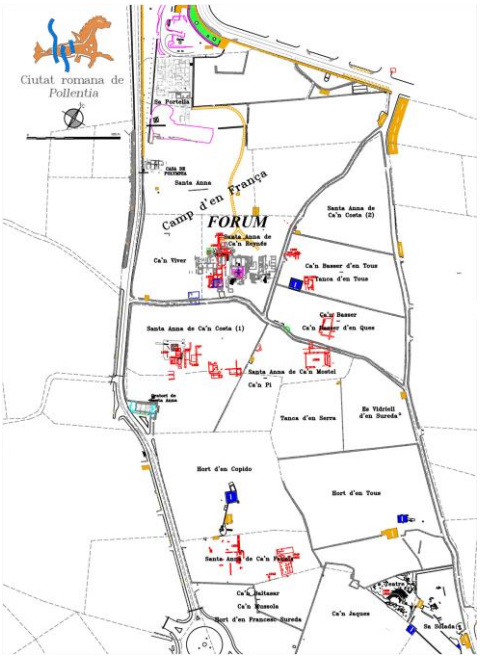

Figure 29. Map of the Roman city of Pollentia (Source: (Ranieri et.al, 2016)).

6.3.2. Experimental results

From the aerial archives, 33 points covering *Pollentia* and its surroundings were selected and their coordinates were triangulated. The triangulation was performed using the image coordinates of these points, the position and attitude of the aerial

images and the aerial camera calibration (Figure 30). 11 of them were selected as GCP while the remaining have been selected as checkpoints (CHP) to evaluate the quality of the RPAS imagery aerial triangulation step.

The typology of the selected ground points was diverse. Most of them belonged to edges or clear visible elements from structures of the Roman city or structures from nearby fields. In addition, some points belonging to structures such as edges of swimming pools, road markings were also selected. Points susceptible to vary such as vegetation were avoided.

The Agisoft Photoscan (Agisoft, 2022) software was used for performing the aerial triangulation step using the ground and image coordinates of the GCP, image coordinates from homologous points together with initial orientation of the images provided by the on-board GNSS receiver of the camera. The image coordinates of the GCP and CHP were manually identified in the RPAS imagery. The ground coordinates of the GCP were input with a very low standard deviation (0.001 m) to constrain the bundle adjustment. Camera calibration parameters (focal length, principal point, radial and tangential distortions) were also estimated by the adjustment.

Figure 30. Distribution of ground control (blue dots) and check points (red triangles) around the Roman city of Pollentia (green area) and nearby fields.

The final step involves the generation of georeferenced maps. Orthophoto and DSM were also generated with the same photogrammetric software. The generated products have a GSD of 10 cm.

In order to evaluate the quality of the results, three different strategies have been used: relative precision evaluation, accuracy evaluation and visual analysis.

The aim of the first strategy is to evaluate the quality of the aerial triangulation, and so, assure the geometric consistency between the aerial and RPAS datasets. To do so, the analysis of the CHP residuals was the tool to evaluate the relative precision quality. The residuals of these points show a planimetric error better than 1 (aerial) GSD (0.1 m and 0.15 m) and slightly higher than 1 GSD for the height component (0.3 m). The residual for the height is worse than the planimetric components and (slightly) higher than 1 GSD. A possible explanation is that the ground points were estimated using only three aerial images belonging to the same flight strip. The use of images from different flight strips would improve the precision of the ground points height component. And thus, it would help to decrease the height component residual below 1 GSD.

These results suggest that aerial triangulation was not only able to estimate the orientation of the images and camera calibration parameters but to provide reliable estimates. In addition, it is also a way to assure that both aerial and RPAS datasets are geometrically consistent.

The proposed methodology relies on the accuracy provided by the Aerial (PNOA) imagery. According to specifications (PNOA, 2022), the planimetric accuracy of the PNOA orthophoto should be better than 0.5 m (RMSE) and 1 m for the height component. However, for the sake of completeness, an assessment of the global accuracy was done. To do so, 5 different points were surveyed with differential GNSS techniques in the area of Forum (Figure 31). The residuals between the triangulated coordinates and the differential GNSS coordinates were computed. The results show residuals of 0.07 and 0.13 m for the planimetric coordinates. The RMSE for X, Y coordinates 0.45, within the expected accuracy of the PNOA program.

The last procedure for the quality checking of results was visual testing. This visual checking seeks to assess the co-registration between PNOA and generated RPAS orthophoto. After the RPAS orthophoto generation, some structures belonging to Sa Portella area were manually digitalized into a vectorial layer. The overlapping of this vectorial layer with the aerial ortophoto was the tool to check visually the co-registration between both orthophotos. Figures 32 and 33 show the geometric consistency of both datasets.

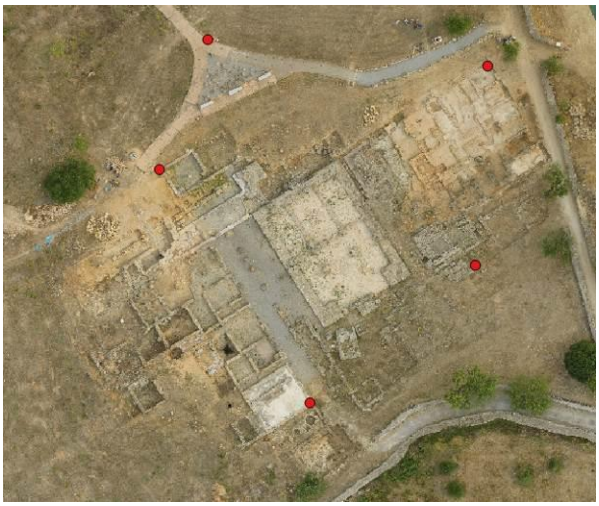

Figure 31. Detail of Forum area from with GNSS surveyed points (red).

Figure 32. Detail of Sa Portella area from most recent PNOA orthophoto and digitilized structures (red) from RPAS orthophoto.

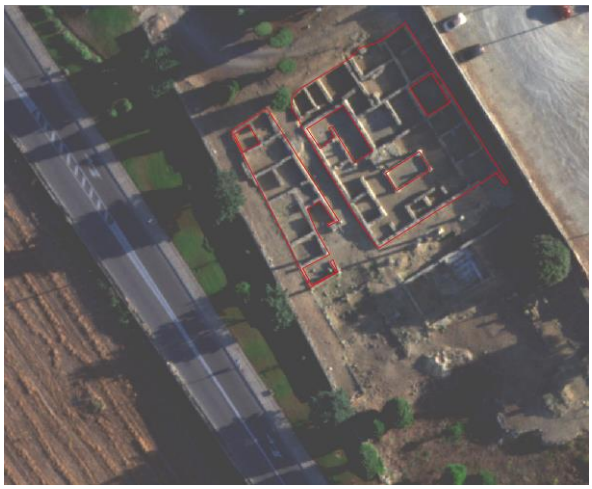

Figure 33. Detail of Sa Portella area from RPAS orthophoto and digitalized structures (red).

6.4. CONCLUSIONS AND FURTHER RESEARCH

An approach to assess not only the proper geoferencing of high-resolution cartography generated with RPAS imagery but also the multi-temporal co-registration using available aerial georeferenced imagery has been presented. The approach has been tested and validated with RPAS and aerial imagery of the Roman city of Pollentia. The aim of this work was not to assess the potential of RPAS imagery, for monitoring and documenting the site, already known by archaeological excavations or remains, but to assess the proper integration/fusion with the available historical georeferenced data.

The approach relies on the use of ground control support derived from aerial and publicly accessible data instead of organizing a specific surveying campaign. The preliminary results show that a relative precision, between aerial and RPAS orthophotos, better than 1 GSD is achieved for the Roman city of Pollentia and their surrounding area. The results also show an absolute accuracy better than 0.5 m (RMSE) for the planimetric coordinates.

Further research will take care of a complete validation to prove the reliability of the approach. This validation will include a comparison of the results obtained with the proposed methodology and the ones obtained using GCP coordinates measured from a specific surveying campaign. Besides this, the same methodology will be applied to additional RPAS imagery from Pollentia as well as other archaeological sites.

Chapter 7

THE ROLE OF NON-SEMANTIC FEATURES IN A RELIABLE AND AFFORDABLE SEAMLESS INDOOR-OUTDOOR POSITIONING

THE IOPES PROJECT APPROACH

An edited version of:

Angelats, E., Espín-López, P. F., Navarro, J. A., and Parés, M. E., 2021. Performance Analysis of the IOPS seamless indoor-outdoor positioning approach. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B4-2021, 229–235, https://doi.org/10.5194/isprs-archives-XLIII-B4-2021-229-2021, 2021.

This chapter explores a second application of hybrid positioning: the tracking of members of civil protection and emergency teams. This is an open field of research with not widely accepted /adopted solution. Although outdoors tracking is routinely performed using well-seasoned techniques such as GNSS, this same problem must still be solved for indoors situations. There exist several approaches for indoor positioning, but these are not appropriate for tracking emergency staff in real time: some of these approaches rely on existing infrastructures; others have not been tested in light devices in real-time; none offers a combined solution. Moreover, the transition from outdoor to indoor areas and vice versa is not usually tackled. This chapter briefly presents a portable, unobtrusive, lightweight device combining GNSS for outdoor positioning and visual-inertial odometry / SLAM for the indoors case, developed in the frame of the IOPES project. This chapter highlights the main steps of the implemented strategy to fuse GNSS and visual-inertial odometry or SLAM-based trajectories. The operational aspects of the prototype, the real-life scenarios where the tests took place and the actual results thus obtained are also described.

7.1. INTRODUCTION

Deciding is a very important component of the management of emergencies: events happen unexpectedly and must be faced promptly; resources, either material or human, must be distributed – hopefully in an efficient manner – throughout the emergency scenario. Deciding means reasoning when, how and what to do, and this reasoning must be backed by information describing the situation in the field. Since situations change constantly, information must follow this very same path, being updated in such a way that the global picture seen by the emergency managers is the right one, at least for as long as possible. Moreover, such information must be as accurate as possible too, to avoid wrong decisions and thus inappropriate responses. In short, information must be both timely and reliable.

Information must be therefore somehow collected, transmitted, and appropriately presented to the emergency managers in order to make their decision process not only possible but also effective. Several systems must be thus distributed throughout the field to perform the aforesaid tasks, which translate to activities such as the deployment of sensors and communication systems, the mapping of the emergency scenario, the processing of data using various types of algorithms to make them easily understandable by humans when they are presented on the appropriate visualization equipment. These systems comprehend Geomatics technologies such as positioning and mapping, either aerial or satellite-based, and advanced information and communication technologies.

As stated above, the collection, transmission and transformation of the relevant information must happen as fast as the situation changes; it should happen, in fact, in real time or in near real time to improve the monitoring, which is crucial (Giordan, 2018) to perform a better risk assessment of the emergency.

Continuous tracking of civil protection and emergency teams working in disaster and post-disaster emergency scenarios is still an open issue. While GNSS is a established solution for outdoors tracking, there is still no suitable solution for the indoors equivalent. Team managers know the building where their teams are working but not in which part of the building they are. The authors are working on developing a full system able to improve this situation. This chapter presents the first results of a low-cost, lightweight positioning system developed in the frame of a project, co-financed by the European commission - Directorate-General Humanitarian Aid and Civil Protection.

Outdoors positioning is routinely performed nowadays by means of Global Navigation Satellite Systems (GNSS) receivers optionally hybridized with Inertial Measuring Unit (IMU) sensors. Despite its reliability, GNSS has some drawbacks such as the need for good environmental conditions. When these are not met (as, for instance, in deep canyons) GNSS is no longer the best technology for precise positioning. Indoor navigation solutions also exist, usually relying on ad-hoc, pre-deployed infrastructures such as Wi-Fi, ultra-wideband or even visual beacons (Mautz and Tilch, 2011; Dardari, 2015), which will not be available (at least, everywhere) in post-disaster scenarios. Alternatively, techniques that combine visual and inertial measurements also exist (Scaramuzza, 2019). (Ramezani et al, 2017) suggest and proposes a visual-inertial odometry approach to improve conventional approaches by using visual measurements derived from omnidirectional cameras and multi-state constraint Kalman filter-based methods.

Solutions already exist for both outdoor and indoor positioning, but these are not appropriate to fulfil the requirements set for the IOPES (Indoor-Outdoor Positioning for Emergency Staff) project. Firstly, some of the indoor technologies rely on complementary infrastructures deployed in advance in the area where the positioning must take place. In the context of IOPES, indoor-outdoor seamless positioning must be available no matter what the situation is, that is, it must be self-sufficient, not depending on pre-existing resources. Secondly, other indoor techniques, based on visual-inertial odometry / SLAM (Ramezani et al, 2017), have not been tested in portable devices and real-time, like the one that this project intends to build. Last, but not least, no combined solution, providing seamless indoor-outdoor positioning, exists for such kind of light devices.

The IOPES concept seeks to fill this gap, building a lightweight, low-cost device for the emergency and civil protection teams working in areas affected by a disaster,

during or after the emergency. Said device, already presented in (Angelats and Navarro, 2017) and (Angelats et al, 2020), combines the well-seasoned GNSS for outdoor positioning and the advances already available in visual-inertial odometry / SLAM materialized in devices such as the T265 tracking device to provide positions indoors. The result is a solution providing seamless indoors-outdoor positioning, suitable for civil protection and emergency teams.

At this point it is worth remarking that the requirements that the IOPES device must match are not as strict as in other situations requiring positioning. This is due to the type of application it is targeted at. Members of civil protection and emergency teams stated, during the phase of requirements collection, that being aware of the room and floor where the members of the team were enough for their purposes. This translates to an accuracy of about 1-2 meters and precision in the range 30-50 cm.

The concept was first introduced in (Angelats and Navarro, 2017). In (Angelats et al, 2020) the proposed methodology to provide seamless outdoor and indoor positioning, detailing hardware, software, and operational aspects, was presented together with the preliminary experimental results. This work presents the results of the performance analysis as well as the conclusions for the first operating prototype performance in real scenarios.

To do so the work has been organized as follows. Firstly, a short introduction to the IOPES project and technologies, together with the hardware, software, and operational aspects of the system, are presented. Then, in section 3, a description of the different real scenarios used to collect the data to estimate the performance of the system together with the experimental results, are presented. Finally, section 4 summarizes the conclusions of the proposed approach and discusses future improvements.

7.2. THE IOPES INDOOR/OUTDOOR POSITIONING SYSTEM

7.2.1. IOPES project

IOPES is a two-year project co-funded by the European Commission involving 7 partners from 5 different European countries. IOPES targets at strengthening the preparedness of emergency personnel by making them more responsive to disasters.

IOPES seeks to improve an already operational Emergency Management System (EMS) – software tool targeted at the handling of emergencies – by providing real-time updates on the position of the teams in the field. Nonetheless, IOPES is not targeted exclusively at improving this specific system but has been designed to interface to any other one by means of a standardized Application Programming Interface (API).

The ability to collect time-tagged positioning information – that may later be related to specific, significant events – facilitates the post-even analysis of the disaster, opening a door to derive new strategies or procedures or the enhancement of these.

The project is funded by the Union Civil Protection Mechanism (UCPM) whose goal is to "improve the quality of EU response capacities" as stated in its Annual Work Programme (2019). Besides that, the IOPES is also fully aligned with priority 4 of the Sendai framework for disaster risk reduction 2015-2030, "Enhancing disaster preparedness for effective response". The project involves the combination of several technologies (Figure 34) including RPAS-based fast mapping, emergency management, portable communications, and positioning technologies. IOPES aims to provide continuous, time-tagged information about the location of Civil Protection Emergency Teams (CPET), either indoors or outdoors.

This chapter does not cover all the technologies involved in IOPES but only those related to the reliable indoor/outdoor positioning of emergency staff.

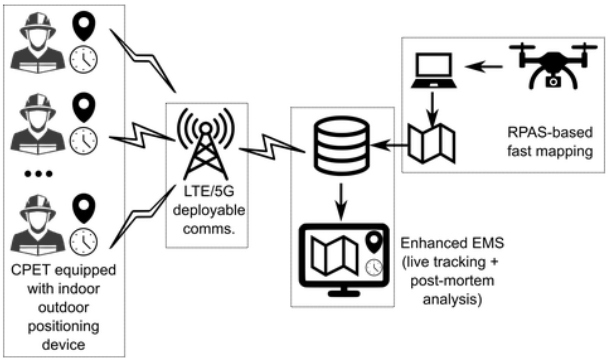

Figure 34. IOPES project technologies.

7.2.2. Hardware

The already developed prototype is a portable, low weight positioning device made of Commercial Off-The-Shelf (COTS) hardware components, mounted on a helmet but also able to be boarded as a drone payload. The positioning sensors used to build the IOPES solution are a GNSS receiver (Drotek DP0803 GNSS module (Drotek, 2021) that includes the U-BLOX NEO-M9N chip (ublox, 2021)), a stereo camera (Intel RealSense T265), (Intel, 2021), which also includes an IMU and a magnetometer (STMicroelectronics LIS3MDL) (ST, 2021). The GNSS receiver has been selected because it is able to provide meter-level accuracy and receive

signals from up to four different single-band GNSS constellations. With this GNSS receiver the system has reliable GNSS position coordinates even in weak GNSS conditions such as urban canyons, exploiting its capability to select the best signals.

The T265 includes a built-in tracking proprietary algorithm running on board. It combines the measurements coming from the IMU and the images produced by the stereo fisheye cameras to provide, by means of visual-inertial odometry/SLAM, a seamless indoor/outdoor solution (Tsykunov, 2020); it provides the current position and orientation with an output rate of 200Hz. The magnetometer delivers raw magnetometer measurements with an output rate of 40 Hz, used to derive absolute heading.

The output frequency of the fused data (GNSS+camera) is 1 Hz. All these components (both hardware and software) are running on a lightweight computer (Raspberry Pi 4 Model B) (Raspberry, 2021), mounted on a helmet designed for work at height and rescue (Petzl Vertex), and powered with a 10 Ah power bank (Anker PowerCore Slim).

The Raspberry Pi is a system on a chip (SOC) with low power requirements that complete and integrate the set of components making the system. Its light weight ensures that it is not a nuisance for its wearers. Its task is to provide the necessary computing resources and storage capacity. Obviously, the low consumption requirements lead to longer operational times, thus reducing the need to replace batteries so often.

From the computing power standpoint, a powerful Graphical Processing Unit (GPU) is not needed, since the computations involved in the visual-inertial odometry solution are performed by the camera device itself. The system also includes a communications module, and a 4G USB dongle with a SIM card (Huawei E3372). Additional features such as a headlamp can easily be attached to the prototype. All the components mounted on the helmet have been installed without modifying its structure to keep the helmet's safety standards (Figure 35).

Figure 35. IOPES positioning system mounted on a Helmet.

7.2.3. Sensor fusion SW approach

The cornerstone of the approach presented is a data fusion algorithm that relates GNSS and camera-based positions providing a single trajectory, regardless of whether it originated indoors, outdoors, or both. The flowchart of the algorithm is detailed in Figure 36. The flowchart is an updated version of the one presented in (Angelats et al, 2020) including additional details on how to derive a combined trajectory.

In areas with low or denied GNSS availability the camera-based tracking system is the main source to provide the positioning solution; conversely, the GNSS positions are used when it is available. A common temporal reference frame is necessary to deal with data coming from these two sources - the internal clock of the SOC is enough for the purposes of the project. The GNSS solution, in friendly GNSS areas, is also used to convert the positions provided by the visual-inertial odometry (VIO) from local to global coordinates.

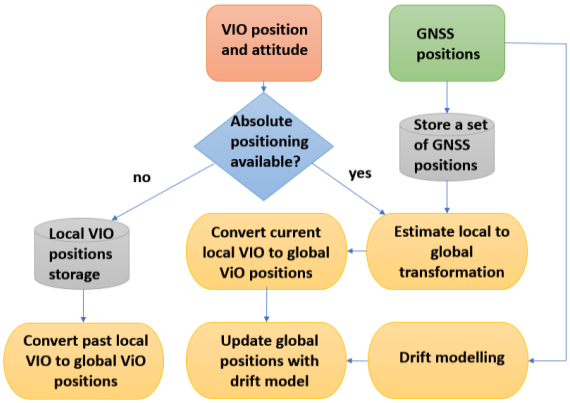

Figure 36. Flowchart for the sensor fusion approach.

If the portable positioning device starts to acquire data in an indoor area, the system is designed to store the VIO positions (in local coordinates) and the three attitude angles in the internal disk till absolute positioning is available. When the positioning device moves to an outdoor area and the number of GNSS satellites allow us to provide a set of reliable GNSS-based positions, a local to global transformation can be estimated, and, consequently, VIO positions in global coordinates can be provided from this moment. The criterion to consider a GNSS position as valid is that the horizontal and vertical errors provided by the GNSS receiver are below a pre-defined threshold (such as 10 m), thus indicating a good GNSS satellite geometry and a good position estimation. The horizontal and vertical errors are the receiver's estimation of one-sigma horizontal and vertical errors.

To estimate the local to global transformation, at least two valid consecutive GNSS positions in different locations are needed to have not only the transformed positions themselves but also to estimate the heading angle derived from GNSS. Alternatively, the absolute heading can be derived from the raw measurements provided by the magnetometer. The criterion to select either heading source is defined by the user. In our approach pitch and roll angles are assumed to be close to zero although this may introduce some positioning errors. Then, a rotation matrix can be computed; afterwards, a global VIO position, the lever arm offset between the GNSS and VIO devices and the local VIO coordinates can be estimated using this rotation matrix.

Although not currently implemented, the GNSS-based positions will be used to model the temporal drift of the VIO data. This action mitigates the drift of the VIO-based solution in indoor environments if the rotation matrix between GNSS and

VIO devices is known and consequently global VIO positions are available. This is done by comparing the coordinates of the global GNSS position and the global VIO estimates for a temporal window of n seconds. After these n seconds, a linear drift is estimated for each positioning component. The positional drift is then applied to the newer global VIO positions till a new positional drift is computed. If no valid GNSS positions are available for the new window (transition from outdoor to indoor), the older positional drift is maintained until new GNSS positions become available (transition from indoor to outdoor). The n number is a parameter that must be set up prior to the use of the system.

Finally, the stored local VIO positions are also converted to global coordinates to keep the historical track of global positions. This track can be used during the management of the emergency or once it is over to perform a post-mortem analysis.

7.2.4. Operational aspects

Ergonomics rules the operational aspects of the IOPES portable device. This means there must be no noticeable difference from the user's standpoint concerning how the system is operated in either indoors or outdoors environment.

Positions and time tags are computed according to the procedure described in section 2.3 and sent to the Emergency Management System so the managers may track the team working in the field – note that the components in the IOPES device as well as the remaining infrastructure required to make communications possible are not described in this chapter, although these are an integral part of the project.

Under some environmental conditions, however, it will not be possible to compute any positions at all – examples of such adverse situations are dust or no lightning; these are limitations of the technology used for indoor positioning and thus affect the performance of the IOPES portable device.

7.3. FIRST RESULTS AND SYSTEM PERFORMANCE

7.3.1. Dataset description

A series of campaigns were carried out to validate the performance of the positioning sensors (visual-inertial odometry and GNSS sensors) and that of the overall system. The campaigns were done at the premises of CTTC in the Parc Mediterrani de la Tecnologia (Castelldefels, Spain) and their surroundings (Figures 37 and 38) and at Garraf town and surroundings (Figure 39). Both scenarios are characterized by

including indoor spaces, clear-sky spaces, and areas with strong multipath conditions.

The performance analysis of the system has been carried out considering several environmental conditions: distance to closest targets (outdoor-clear sky/ outdoor-low GNSS availability/indoor), system dynamics (kinematic/almost static/ static), environmental texture (no texture/texture) and lightning conditions (bright / dark). Four different routes (two per location) were defined and carried out combining the different parameters stated above. The first route, "Sa Falconera", took place in the surroundings of Garraf including a walk near a cliff, and a tunnel. The second route, "Garraf Town", is a walk inside the Garraf town, with narrow streets and ending inside a house with four different rooms on the same floor. Third route, "PMT-LAB", is a walk through the campus including an area with strong multipath and a walk inside the positioning lab of CTTC. The last route "PMT-buildings" is an extension of the previous one but covering all the corridors and main spaces of three floors of the CTTC-B4 building and also two different floors in CTTC-B6 building.

7.3.2. Performance analysis methodology

In order to assess the performance of the system, the estimated positions in the most relevant areas were analysed. Performance indicators are computed using predicted positions from equivalent environments like outdoor clear-sky, indoor or urban corridors. Then, for each of those subsets the mean accuracy and precision were computed; the references were the coordinates provided by PNOA (Plan Nacional Ortofotografía Aérea) maps (outdoors) and plans of the building (indoors). According to specifications (PNOA, 2021), the planimetric accuracy of the PNOA orthophoto should be better than 0.5 m (Root Mean Square Error) and 1 m for the height component.

For the specific case of the "CTTC positioning lab" (indoors), the coordinates of an easily identifiable set of points on the floor were computed by means of other reliable positioning techniques. Such coordinates constituted the reference for this use case. Later, a walk using the IOPES portable device was recorded, and its output coordinates were compared with the aforesaid reference data set.

Figure 37. Parc Mediterrani de la Tecnología where CTTC buildings are located. In red, area with strong multipath conditions. In orange, indoor/outdoor transition areas.

Figure 38. CTTC premises. Positioning lab (left) and corridors (right).

Figure 39. Garraf town and surroundings.

This research does not consider the error estimation provided by the GNSS as a performance indicator since it has been observed that these values are too pessimistic in all trajectories.

7.3.3. Results and discussion

The system output for the "Sa Falconera" route is shown in Figure 40. The system estimates a set of global positions (depicted as dots), and just for clarification, each global position has been coloured red or green depending on the sensor used for its generation (green dots for GNSS estimated positions, red dots for camera-based ones). The summarized performance results for "Sa Falconera" test site are shown in Table 20. The results confirm the good performance of the system under "outdoor clear sky" conditions. Performance gets slightly worse when the system is outdoors surrounded by thick vegetation. When the results are produced by the camera-based sensor the precision is good while the accuracy goes to 3 m.

The system output for the "Garraf Town" route is shown in Figure 41. The summarized performance results for the "Garraf Town" test site are shown in Table 21. In this case the use of the camera-based sensor is higher due to the bad quality of the GNSS signal inside the narrow streets of the village. As in "Sa Falconera" the results under "outdoor clear sky" conditions are good. For most of the outdoor urban section of the route and for the indoor section the camera-based sensor has performed with good precision and 3 meters of accuracy.

The system output for the "CTTC-LAB" route is shown in Figure 42. The summarized performance results for the "CTTC-LAB" test site are shown in Table 22. The most critical section of this route is the ~50 meters walk in a zone where the GNSS receiver suffers from multipath effects. In this zone the performance of the system gets worse, with a mean precision of 1 meter and a mean accuracy of 4 meters. The rest of the route is estimated with the same precision and accuracy performance as "Sa Falconera" and "Garraf Town".

The system output for the "PMT-Buildings" route is shown in Figure 43. The summarized performance results for the "PMT-Buildings" test site are shown in Table 23. In this route the system is tested extensively indoors, under different lighting and space conditions. The outdoor results are similar to the previous study cases. The indoor results show no differences in terms of performance when the lightning conditions are changed. However, it can be observed an important worsening of the performance when the system is tested in big diaphanous spaces (due to the lack of features to be tracked by the camera). In these conditions, the mean precision goes to 2 meters and the mean accuracy to 5 meters.

Figure 40. Estimated "Sa Falconera" trajectory. Green dots indicate positions provided by the GNSS while red dots are the ones provided by camera-based tracking system.

	Main sensor	Mean precision	Mean Accuracy
Outdoor clear sky	GNSS	<0.5m	1m
Outdoor vegetation	GNSS	1m	2m
Tunnel	Camera	<0.5m	3m

Table 21. IOPES system horizontal performance at "Sa Falconera" test site.

Figure 41. Estimated "Garraf -Town" trajectory. Green dots indicate positions provided by the GNSS while red dots are the ones provided by camera-based tracking system.

	Main sensor	Mean precision	Mean Accuracy
Outdoor clear sky	GNSS	<0.5m	1m
Outdoor urban corridor	Camera	~0.5m	3m
House	Camera	<0.5m	3m

Table 22. IOPES system horizontal performance at "Garraf-town" test site.

Figure 42. Estimated "CTTC-LAB" trajectory. Green dots indicate positions provided by the GNSS while red dots are the ones provided by camera-based tracking system.

	Main sensor	Mean precision	Mean Accuracy
Outdoor clear sky	GNSS	<0.5m	1m
Outdoor multipath	GNSS	1m	4m
CTTC lab	Camera	<0.5m	3m

Table 23. IOPES system horizontal performance at "CTTC-lab" test site.

Figure 43. Estimated "PMT-Buildings" trajectory. Green dots indicate positions provided by the GNSS while red dots are the ones provided by camera-based tracking system.

	Main sensor	Mean precision	Mean Accuracy
Outdoor clear sky	GNSS	<0.5m	1m
Outdoor multipath	GNSS	~0.5m	3m
Indoor small spaces	Camera	<0.5m	3m
Indoor big spaces	Camera	2m	5m
Indoor dark rooms	Camera	<0.5m	3m

Table 24. IOPES system horizontal performance at "PMT-buildings" test site.

The results obtained in the different scenarios demonstrate the operability of the idea of the IOPES project. The precision and accuracy obtained for the different scenarios partially cover the user requirements of the IOPES project. However, the authors strongly believe that the user requirements for the outdoors can be easily achieved by improving the quality of the GNSS receiver (for example using a multipath-rejecting GNSS antenna and receiver). The indoor performance can be trickier to improve; the bad quality of the results in indoor big spaces seems to be due to the technology used and hardly improvable; however, the authors have found that the accuracy of the indoor solution in the rest of the cases can be improved by selecting properly the point where the solution changes from outdoor (and thus from the GNSS sensor) to indoor. Taking this point into account, more complex

algorithms for detecting the outdoor/indoor transition could lead to an improvement in the indoor system performance.

7.4. CONCLUSIONS AND FURTHER RESEARCH

This chapter presents the first results for a new, portable, lightweight, unobtrusive positioning device which, combining GNSS and VIO, is offering seamless indoor and outdoor positioning in the context of emergency and disaster management applications and the monitoring of the teams involved in these events.

Such a device is expected to operate in many different scenarios – such as open sky, tunnels, rooms inside building – under different environmental conditions – such as lack of lightning or texture – and subject to changes concerning the elements making the scenarios themselves – as, for instance, due to the presence of moving objects such as people or vehicles.

The expectations concerning the positioning parameters defining the performance of the IOPES device were, literally, "to be able to tell apart both the room and floor where the individual being tracked was". This was a requirement specific to indoor positioning, which the authors translated to more measurable magnitudes, that is, expecting an accuracy of about 1-2 meters and precision in the range 30-50 cm (see section 7.1).

The results commented in section 7.3 show that the requirement concerning accuracy has not been achieved, going up to 4 meters in one of the use cases. The reason explaining this result is directly related to the accuracy obtained by the GNSS receiver itself in the aforesaid scenarios; the overall accuracy of the IOPES device directly depends on this magnitude. It is reasonable to expect, therefore, that this accuracy will be improved when working in more favourable situations. Precision, on its side, meets the expectations stated in section 7.1.

Nonetheless, accuracy and precision are either suitable or not depending on the target application and, especially, on the users involved in the exploitation of such application. It is worth noting that at least one of the users involved in the IOPES project finds these results very interesting, and, consequently, the IOPES device is useful for their purposes. The opinion of the rest of the users is still unknown to us, but they will have the opportunity to express it soon, during the first demonstration of the project.

At any rate, the work on the IOPES device is not yet over. The results presented in this work correspond to the very first assembled prototype. The work to come will concentrate on making the prototype more robust and performant. The foreseen lines of work are the implementation of the correction of the drift when working in

VIO mode and a better mechanism to detect indoors-outdoors (and vice-versa) transitions to reduce the time to switch between two positioning technologies – and thus, reduce errors. Both lines of work will serve to improve the accuracy of the system.

Chapter 8

CONCLUSIONS AND OUTLOOK

Hybrid or multi-sensor-based positioning has been a research topic actively investigated and is still being actively studied by the Geomatics community but also by the Robotics and Computer vision community in the last decade. In this context the possibility of using information, extracted from single or multiple image sensors, for positioning is very appealing to mitigate the problems that GNSS or INS/GNSS-based trajectories have in GNSS-unfriendly environments in terms of robustness and accuracy. Moreover, different processing workflows, sensor positioning quality or system calibration errors, may also produce even in GNSS-friendly conditions, multiple geospatial datasets that are not properly co-registered.

This thesis has proposed the use of non-semantic information, that is the use of a set of geometric entities or features, to improve the trajectory estimation in a multi-sensor-based approach. This thesis has covered the mathematical modelling of non-semantic information, several hybrid-based trajectory estimation approaches that use this kind of information with the appropriate modelling, as well as exploring the use of non-semantic features to model the trajectory error modelling.

The proposed models have enabled improved trajectory determination in both urban post-processing and airborne environments with good accuracy. The implemented INS/GNSS trajectory error models have been relatively simple but still proved to be efficient. The combined models have been tested, in post-processing using a bundle adjustment approach, with real data from metric cameras and aerial laser mapping systems as well as in TMM systems. The simultaneous orientation of LiDAR and optical imagery is a research topic that is not closed, and it has been investigated recently (Glira et.al, 2019, Brun et.al, 2021, Mouzakidou et.al, 2022). These works share some similarities with the aforementioned models but focused on datasets acquired with a different type of platform (RPAS) and the latter two also incorporated the use of Dynamic networks (Colomina and Blázquez, 2004, Cucci et.al, 2017) to study the benefit of using GNSS and raw inertial observations within a single network adjustment instead of introducing INS/GNS derived trajectories plus a trajectory error model. The use of dynamic networks is fully compatible with the proposed combined models, and it should also be part of any forthcoming work. Future works should also evaluate the feasibility of using the previous models for simultaneous orientation of aerial and TMM imagery, and compare them with existing approaches (Jende et.al, 2018), (Molina et.al, 2017).

The non-semantic information extracted from the images has allowed, using an integrated sensor orientation approach, to model the current errors in an urban environment. This modelling is another contribution of the thesis that opens the door to the development of new, more advanced trajectory error models that go beyond the deterministic models currently used. The characterization of errors in urban environments has opened a new line of research to be considered in future work: understanding which context conditions produce higher trajectory error in

urban environments. The determination of the context could be done through semantic segmentation and classification of the acquired imagery and/or considering a temporal window of GNSS raw measurements.

This thesis has also presented an approach to improve the determination of RPAS trajectories using open aerial data obtained in the framework of a national mapping project (PNOA). The development of this methodology is a contribution aiming to ensure the geospatial coherence between the orthophotos, and digital elevation models obtained with an RPAS and the orthophotos and digital models of the PNOA. The proposed approach takes advantage of a technology (in terms of positioning systems) with a much higher cost than the ones used to derive photogrammetric products with RPAS platforms. The results show that photogrammetric products can be generated with similar accuracy to the ones generated with more complex approaches. In this context, future work should investigate the possibility to automatically detecting homologous points between aerial images and RPAS images, not only between pairs of images but also between all existing PNOA images. This would allow the detection of points that persist over time, as the motivation is to generate orthophotos from RPAS images co-registered also with the most recent one but also with historical orthophotos of the PNOA.

The approach that has been proposed in this thesis for the determination of trajectories in real time based on a tightly coupling sequential nonlinear least squares using GNSS positions, image coordinates and raw inertial measurements, does not allow us to derive definitive conclusions beyond demonstrating the potential of the approach in non-friendly GNSS environment. The proposed approach exploited a sliding window bundle adjustment technique to use the image coordinates of tie points and the positions and attitudes derived from the last epochs to determine the position and attitude parameters of the most recent epoch. This concept has evolved a lot in recent years, with many and more robust and reliable approaches proposed by the Robotics and Computer Vision communities. In the last two years, several publications have appeared improving state-of-the-art VIO/SLAM approaches (Campos et.al, 2021) or fusing VIO and GNSS data (Cioffi and Scaramuzza, 2020, Lee et.al, 2020, Boche et.al, 2022, Cao et.al, 2022). These approaches aim to provide (1) trajectory estimation in a global reference frame and (3) to reduce long-term drift of VIO-based trajectories, exploiting loop closing techniques, that is, the detection of already visited areas to reduce accumulated positional drift. These techniques share some similarities also with the already mentioned dynamic network approach. Any further work on that topic should consider including the advances proposed in the previous works.

This thesis has tried to take benefit also of the latest technological advancements, especially in terms of COTS devices such as miniaturized VIO-based trajectory systems. To do so, a seamless indoor-outdoor positioning approach has been

presented with encouraging results in several scenarios. Nevertheless, the results pointed out at the same time the importance of detecting transition areas between indoor and outdoor areas. As for the urban mapping case, the context information should have to be considered to improve the fusion strategy in forthcoming works as well as characterize the features of these transition areas. Moreover, hybrid methods merging the current approach with a trajectory forecasting approach based on Artificial Intelligence (Gupta et.al, 2018) should be considered.

BIBLIOGRAPHY

Agisoft, 2022. Metashape professional edition, version 1.7.3. http://www.agisoft.com. (Accessed: 20th December 2022).

Aicardi I, Nex F, Gerke M, Lingua AM, 2016. An Image-Based Approach for the Co-Registration of Multi-Temporal UAV Image Datasets. Remote Sensing, 8(9):779. https://doi.org/10.3390/rs8090779.

Angelats, E. and Navarro, J. A., 2017. Towards a fast, low-cost Indoor Mapping and Positioning System for Civil Protection and Emergency Teams. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-2/W8, 9–15, https://doi.org/10.5194/isprs-archives-XLII-2-W8-9-2017,2017.

Angelats, E., Navarro, J. A., and Parés, M.E. 2020: Towards seamless indoor-outdoor positioning: the iopes project approach, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B4-2020, 313–319, https://doi.org/10.5194/isprs-archives-XLIII-B4-2020-313-2020, 2020.

Angelats, E., Molina, P., Parés, M. E. and Colomina, I., 2014. A parallax based robust image matching for improving multisensory navigation in gnss-denied environments. In: Procceding of the ION GNSS+ 2014, Tampa, USA, pp. 2132 – 2138.

Angelats, E. and Colomina, I., 2014. One step mobile mapping laser and camera data orientation and calibration. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XL-3/W1, Castelldefels, Spain, pp. 15–20.

Angelats, E., Blázquez, M. and Colomina, I., 2012. Simultaneous orientation and calibration of images and laser point clouds with straight segments. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 39(B1), Melbourne, Australia, pp. 91–96

Bay, H., Ess, A., Tuytelaars, T, Van Gool, L., 2008. SURF: Speeded Up Robust Features, Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346—359. https://doi.org/10.1016/j.cviu.2007.09.014

Blázquez, M. and Colomina, I., 2012. Relative INS/GNSS´aerial control in integrated sensor orientation: Models and performance. ISPRS Journal of Photogrammetry & Remote Sensing 67, pp. 120–133. https://doi.org/10.1016/j.isprsjprs.2011.11.003

Boche, S., Zuo, X., Schaefer, S. and Leutenegger, S., 2022. Visual-Inertial SLAM with Tightly-Coupled Dropout-Tolerant GPS Fusion, In International Conference on Intelligent Robots and Systems (IROS), 2022. https://doi.org/10.48550/arXiv.2208.00709

Bradski, G., 2000. The OpenCV Library, Dr.Dobb's Journal of Software Tools, 2000.

Brun, A, Cucci, D.A, and Skaloud., J., 2022. Lidar point–to–point correspondences for rigorous registration of kinematic scanning in dynamic networks. ISPRS Journal of Photogrammetry and Remote Sensing, Volume 189, July 2022, Pages 185-20. https://doi.org/10.1016/j.isprsjprs.2022.04.027.

Campana, S., 2018. Drones in archaeology. State-of-the-art and future perspectives. Journal of Archaeological Prospection, 24, 275–296. https://doi.org/10.1002/arp.1569

Campos, C., Elvira, R., Gómez Rodríguez, Juan J., Montiel, J.M.M, and Tardós, J. D., 2021. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial and Multi-Map SLAM. IEEE Transactions on Robotics 37(6): pp 1874-1890. https://doi.org/10.1109/TRO.2021.3075644

Cao, X. Lu, and S. Shen, 2022. GVINS: Tightly coupled gnss-visual inertial fusion for smooth and consistent state estimation. IEEE Trans. Robotics 38(4): 2004-2021. https://doi.org/10.1109/TRO.2021.3133730

Cau Ontiveros, M.A, Van Strydonck, M., Boudin, M., Mas Florit, C., Mestres, J.S., Cardona, F., Chávez-Álvarez, E. and Orfila, M., 2017. Christians in a Muslim world? Radiocarbon dating of the cemetery overlaying the Forum of Pollentia (Mallorca, Balearic Islands). Journal of Archaelogical and Anthrolopogical Sciences, 9, 7, 1529-1538. https://doi.org/10.1007/s12520-016-0325-0

Chan, T. O., Lichti, D. and Glennie, C. L., 2013. Multi-feature based boresight self-calibration of a terrestrial mobile mapping system. ISPRS Journal of Photogrammetry & Remote Sensing 82, pp. 112–124. https://doi.org/10.1016/j.isprsjprs.2013.04.005

Chen, Y. and Medioni, G., 1992. Object modelling by registration of multiple range images. Image and Vision Computing 10(3), pp. 145–155. https://doi.org/10.1016/0262-8856(92)90066-C

Cioffi, G., Scaramuzza, D., 2020. Tightly-coupled Fusion of Global Positional Measurements in Optimization-based Visual-Inertial Odometry. In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, 2020. https://doi.org/10.48550/arXiv.2003.04159

Colomina, I., Molina, P., 2014. Unmanned aerial systems for photogrammetry and remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing 92, 79–97. https://doi.org/10.1016/j.isprsjprs.2014.02.013.

Colomina, I., Blazquez, M., Navarro, J. A. and Sastre, J., 2012. The need and keys for a new generation network adjustment software. In Int. Arch. Photogramm. Remote Sens.

Spatial Inf. Sci., XXXIX-B1, 303–308. https://doi.org/10.5194/isprsarchives-XXXIX-B1-303-2012, 2012

Colomina, I, Blázquez, M., 2004. A unified approach to static and dynamic modelling in photogrammetry and remote sensing. In ISPRS International Archives at Photogrammetry, Remote Sensing and Spatial Information Sciences, 35:B1. https://www.isprs.org/proceedings/XXXV/congress/comm1/papers/33.pdf

Cowley, D.C, Moriarty, C., Geddes, G., Brown, G.L., Wade, T. and Nichol, C.J., 2018. UAVs in context: archaeological airborne recording in a national body of survey and record. Drones, 2(1):2. https://doi.org/10.3390/drones2010002

Cucci, D.A, M. Rehak, M., and Skaloud, J., 2017. Bundle adjustment with raw inertial observations in UAV applications. ISPRS Journal of Photogrammetry and Remote Sensing, 130:1–12. https://doi.org/10.1016/j.isprsjprs.2017.05.008

Dardari, D., Closas, P., Djuric, P, 2015. Indoor Tracking: Theory, Methods, and Technologies. IEEE Transactions on Vehicular Technology, 64(4), 2015, 1263-1278. https://doi.org/10.1109/TVT.2015.2403868

Drotek, 2021. DP0803 GNSS module. https://store-drotek.com/919-dp0803-gnss-tiny-m9n.html. (Accessed: 20th April 2021).

Elhashash, M., Albanwan, H., Qin, R., 2022. A Review of Mobile Mapping Systems: From Sensors to Applications. Sensors 2022, 22, 4262. https://doi.org/10.3390/s22114262.

Elseberg, J., Borrmann, D. and Nüchter, A., 2013. Algorithmic solutions for computing precise maximum likelihood 3D point clouds from mobile laser scanning platforms. Remote Sensing 5(11), pp. 5871–5906. https://doi.org/10.3390/rs5115871

Fernández, A., 2012. Final report - ATENEA (Advanced techniques for navigation receivers and applications. https://cordis.europa.eu/project/id/247975/reporting/fr (Accessed: 20th December 2022).

Friess, P., 2006. Toward a rigorous methodology for airborne laser mapping. In: Proceedings of EuroCOW 2006, Castelldefels, Spain, 25-27 January, 7p. https://www.isprs.org/proceedings/2006/eurocow06/euroCOW06.htm

Giordan D, Hayakawa Y, Nex F, Remondino F, Tarolli P, 2018. Review article: the use of remotely piloted aircraft systems (RPASs) for natural hazards monitoring and management. Nat. Hazards Earth. Syst. Sci. 18:1079–1096. https://doi.org/10.5194/nhess-18-1079-2018, 2018.

Giradeau-Montaut, D., Bougacha, S., Bey, A. and Marc, R.,2014. Cloudcompare SW, version 2.5.2, EDF R&D, Telecom ParisTech.

Gressin, A., Cannelle, B., Mallet, C. and Papelard, J., 2012. Trajectory-based registration of 3D lidar point clouds acquired with a mobile mapping system. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 1-3, Melbourne, Australia, pp. 117–122. https://doi.org/10.5194/isprsannals-I-3-117-2012, 2012.

Gupta, A., Johnson, J., Fei-Fei L., Savarese, S., and Alahi, A., 2018. Social gan: Socially acceptable trajectories with generative adversarial networks. In IEEE Conference on Computer Vision and Pattern Recognition, pp. 2255–2264. https://doi.org/10.48550/arXiv.1803.10892

Habib, A., Ghanma, M., Morgan, M. and Al-Ruzouq, R., 2005. Photogrammetric and Lidar Data Registration Using Linear Features. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 71. No.6, pp. 699–707. https://www.asprs.org/wp-content/uploads/pers/2005journal/jun/2005_jun_699-707.pdf

Habib, A., Morgan, M. and Lee, Y., 2002. Bundle adjustment with self-calibration using linear features. The Photogrammetric Record 17(100), pp. 635–650. https://doi.org/10.1111/0031-868X.00211

Hackel, T. Savinov, N., Ladicky, L, Wegner, J. D., Schindler, K. and Pollefeys, M. , 2017. A new large-scale point cloud classification benchmark, In ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, volume IV-1-W1, pages 91—98. https://doi.org/10.5194/isprs-annals-IV-1-W1-91-2017, 2017.

Hartley, R. and Schaffalitzky, F., 2004. L_∞ minimization in geometric reconstruction problems. In Proceedings of Computer Vision and Pattern Recognition conference, 2004.

Hartley, R. and Zisserman, A., 2000. Multiple view geometry in computer vision. Cambridge University Press, 2nd edition.

Huawei, 2021. https://consumer.huawei.com/en/routers/e3372/ (Accessed: 20th April 2021)

Intel, 2021. Intel® RealSense™ Tracking Camera T265, 2019. https://www.intelrealsense.com/tracking-camera-t265/ (Accessed: 20th April 2021).

IOPES project, 2021. https://iopes-project.eu/ (Accessed: 20th April 2021)

Jende, P., Nex, F., Gerke, M., Vosselman, G., 2018. A fully automatic approach to register mobile mapping and airborne imagery to support the correction of platform trajectories in GNSS-denied urban areas. ISPRS Journal of Photogrammetry and Remote Sensing, Volume 141, 2018, Pages 86-99. https://doi.org/10.1016/j.isprsjprs.2018.04.017.

Kager, H., 2004. Discrepancies between overlapping laser scanner strips - simultaneous fitting of aerial laser scanner strips. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 35(Part B1), pp. 555–560. https://www.isprs.org/proceedings/xxxv/congress/comm1/papers/99.pdf

Kersting, A., Habib, A. and Rau, J., 2012. New method for the calibration of multi-camera mobile mapping systems. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Melbourne, Australia. https://doi.org/10.5194/isprsarchives-XXXIX-B1-121-2012, 2012.

Kutterer, H., 2010. Mobile mapping. In: Vosselman and G. and Maas, (Eds.), Airborne and Terrestrial Laser Scanning, first ed. Whittles Publishing, Dunbeath, Scotland, UK, pp. 293–311.

Lee, W., Geneva, P., Yang, Y. and Huang, G. 2022. Tightly-coupled gnss-aided visual-inertial localization. In 2022 International Conference on Robotics and Automation (ICRA), 2022, pp. 9484–9491. https://doi.org/10.1109/ICRA46639.2022.9811362

Leutenegger, S., Furgale, P., Rabaud, V., Chli, M., Konolige, K. and Siegwart, R., 2013. Keyframe-based visual-inertial SLAM using nonlinear optimization. In: Proceedings of Robotics: Science and Systems.

Leutenegger, S., Chli, M., Siegwart, R., 2011. BRISK: Binary Robust Invariant Scalable Keypoints. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) 2011. https://doi.org/10.1109/ICCV.2011.6126542

Lo Brutto, M., Garraffa, A. and Meli., P., 2014. UAV platforms for cultural heritage survey: first results. In ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., II-5, 2014, pp. 227-234. doi: 10.5194/isprsannals-II-5-227-2014.

Lo Brutto, M., Sciortino, R. and Garraffa, A, 2017. RPAS and TLS techniques for archeological survey: the case study of the archeological site of Eraclea Minoa (Italy). In Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-2/W3, 2017, 433-438, https://doi.org/10.5194/isprs-archives-XLII-2-W3-433-2017.

Lowe, D., 1999. Object recognition from local scale-invariant features. In proceedings ICCV. https://doi.org/10.1109/ICCV.1999.790410

Luhmann, T., Robson, S., Kyle, S., Harley, I., 2006. Close Range Photogrammetry: Principles, Methods and Applications. 1st edition, Whittles Publishing, UK.

Masini, N., Marzo, C., Manzari, P., Belmonte, A., Sabia, C., and Lasaponara, R., 2018. On the characterization of temporal and spatial patterns of archaeological crop-marks. Journal of Cultural Heritage. Volume 32, July–August 2018, Pages 124-132. https://doi.org/10.1016/j.culher.2017.12.009

Mautz, R., Tilch, S., 2011. Survey of Optical Indoor Positioning Systems. In: Proceedings of the 2011 International Conference on Indoor Positioning and Indoor Navigation

(IPIN), Guimarães, Portugal, 21-23 September 2011. https://doi.org/10.1109/IPIN.2011.6071925

Molina P, Blázquez M, Cucci DA, Colomina I., 2017. First Results of a Tandem Terrestrial-Unmanned Aerial mapKITE System with Kinematic Ground Control Points for Corridor Mapping. Remote Sensing, 9(1):60. https://doi.org/10.3390/rs9010060

Mouzakidou, K, Cucci, D.A, Skaloud, J., 2022. On the benefit of concurrent adjustment of active and passive optical sensors with GNSS & raw inertial data. In ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., V-1-2022, 161–168, https://doi.org/10.5194/isprs-annals-V-1-2022-161-2022.

Nex, F., Armenakis, C., Cramer, M., Cucci, D.A., Gerke, M., Honkavaara, E., Kukko, A., Persello, C., Skaloud, J., 2022. UAV in the advent of the twenties: Where we stand and what is next. ISPRS J. Photogramm. Remote Sens. 184, 215–242. https://doi.org/10.1016/j.isprsjprs.2021.12.006

Nex, F. and Remondino, F., 2013. UAV for 3d mapping applications: a review. Applied Geomatics, 6, 2013, 1-15. https://doi.org/10.1007/s12518-013-0120-x

Nister, D., 2013. An efficient solution to the five-point relative pose problem. In Proceedings of CVPR03. https://ieeexplore.ieee.org/document/1211470.

Nikolakopoulos, K.G., Soura, K., Koukouvelas, I.K. and Argyropoulos, N.G., 2016. UAV vs classical aerial photogrammetry for archaeological studies, Journal of Archaeological Science: Reports, 2016, doi:10.1016/j.jasrep.2016.09.004.

Novák, D. and Schindler, K., 2013. Approximate registration of point clouds with large scale differences. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. II-5/W2, Antalya, Turkey, pp. 211–216. doi:10.5194/isprsannals-II-5-W2-211-2013.

Opitz, R. and Herrmann, J., 2018. Recent trends and long-standing problems in archaeological remote sensing. Journal of Computer Applications in Archaeology, 1(1), 2018, pp. 19-41. http://doi.org/10.5334/jcaa.11

Parés, M.E. Parés, Rosales, J., Colomina, I., 2008. Yet another IMU Simulator: Validation and Applications. In Proceedings of Eurocow 2008, Castelldefels, Spain. https://www.isprs.org/proceedings/2008/EuroCOW08/euroCOW08_files/papers/20.pdf

Jeong, J.; Yoon, T.S.; Park, J.B. "Towards a Meaningful 3D Map Using a 3D Lidar and a Camera". Sensors 2018, 18, 2571. https://doi.org/10.3390/s18082571

pnoa.ign.es. 'PNOA imagen, especificaciones técnicas' [Online]. Available: http://pnoa.ign.es/caracteristicas-tecnicas. (Accessed: 07- Oct- 2022).

Ramezani, M., Acharya, D., Gu, F., and Khoshelham, K, 2017. Indoor Positioning by Visual-Inertial Odometry. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., IV-2/W4, 371–376, https://doi.org/10.5194/isprs-annals-IV-2-W4-371-2017.

Ranieri, G., Godio, A., Loddo, F, Stocco, S., Capizzi, P., Messina, P., Orfila, M., Cau, M.A., Chávez, M.E, 2016. Geophysical prospection of the Roman city of Pollentia, Alcúdia (Mallorca, Balearic islands, Spain). J. Appl. Geophys. 2016, 134, 125–135. https://doi.org/10.1016/j.jappgeo.2016.08.009

Raspberry, 2021. Raspberry Pi 4. https://www.raspberrypi.org/products/raspberry-pi-4-model-b/. (Accessed: 20th April 2021).

Ronholm, P., 2011. Registration quality - towards integration of laser scanning and photogrammetry. EuroSDR publication. http://www.eurosdr.net/sites/default/files/uploaded_files/eurosdr_publication_ndeg_59.pdf.

Scaramuzza, D. and Zhang, Z, 2019. Visual-inertial odometry of aerial robots. Encyclopedia of Robotics, 2019. https://doi.org/10.48550/arXiv.1906.03289

Scaramuzza, D., 2011. 1-point-ransac structure from motion for vehicle-mounted cameras by exploiting non-holonomic constraints. International Journal of Computer Vision 95(1), pp. 74– 85. https://doi.org/10.1007/s11263-011-0441-3

Scaramuzza, D. and Fraundorfer, F., 2011. visual odometry [tutorial]. IEEE Robotics and Automation Magazine 18(4), pp. 80–92. https://doi.org/10.1109/MRA.2011.943233

Schaer, P. and Vallet, J., 2016. Trajectory adjustment of mobile laser scan data in gps denied environments. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XL-3/W4, pp. 61–64. https://doi.org/10.5194/isprs-archives-XL-3-W4-61-2016, 2016.

Skaloud, J. and Lichti, D., 2006. Rigorous approach to boresight self-calibration in airborne laser scanning. ISPRS Journal of Photogrammetry & Remote Sensing 61(1), pp. 47–59. https://doi.org/10.1016/j.isprsjprs.2006.07.003

ST, 2021. LIS3DL magnetometer. https://www.st.com/en/mems-and-sensors/lis3mdl.html. (Accessed: 20th April 2021).

Stek, T.D, 2016. Drones over mediterranean landscapes. The potential of small UAV's (Drones) for site detection and heritage management in archaeological survey projects: a case study from le Pianelle in the Tappino valley, Molise (Italy). Journal of Cultural Heritage, 2016, 22, 1066–1071. https://doi.org/10.1016/j.culher.2016.06.006

Taylor, C., Veth, M. J., Raquet, J. F. and Miller, M. M., 2011. Comparison of two image and inertial sensor fusion techniques for navigation in unmapped environments. IEEE

Transactions on Aerospace and Electronic Systems 47(2), pp. 946–958. https://doi.org/10.1109/TAES.2011.5751236

Tommaselli, A., 1996. A recursive approach to space resection using straight lines. Photogrammetric Engineering and Remote Sensing 62(1), pp. 57–66. https://www.asprs.org/wp-content/uploads/pers/1996journal/jan/

Tsykunov, E., Ilin, V., Perminov, S, Fedoseev, A., Zainulina, E, 2020. Coupling of localization and depth data for mapping using Intel RealSense T265 and D435i cameras. Pre-print version: https://arxiv.org/abs/2004.00269.

ublox, 2021. NEO-M9N chip. https://www.u-blox.com/en/product/neo-m9n-module. (Accessed: 20th April 2021).

Veth, M. J., 2011. Navigation using images, a survey of techniques. Journal of Navigation 58(2), pp. 127–140.

Vosselman, G., Sester, M. and Mayer, H., 2004. Basic computer vision techniques. In: McGlone, J.C., Mikhail, E.M, Bethel, J., Mullen, R. (Eds.), Manual of Photogrammetry, fifth ed. American Society of Photogrammetry and Remote Sensing, Bethesda, MA, pp. 455–504.

Weinmann, M., 2012. Visual features - from early concepts to modern computer vision. In: G. Farinella, S. Battiato and R. Cipolla (eds), Advanced Topics in Computer Vision. Advances in Computer Vision and Pattern Recognition, Springer, chapter 1, pp. 1–34.

Zhuo, X., Koch, T., Kurz, F., Fraundorfer, F. and Reinartz, P., 2017. Automatic UAV image geo-registration by matching UAV images to georeferenced image data. Remote Sensing, 2017, 9, 376. https://doi.org/10.3390/rs9040376

RESEARCH ACTIVITY

The following references have been produced in the frame of this PhD thesis:

PEER REVIEWED JOURNAL PUBLICATIONS

1. Parés, M. E., **Angelats, E**., Espín, P., Navarro, J.A., Munaretto, D., Salvador, J., Delgado, C., Olbrich, F., Martínez, A., Grudbrandsson, G.O., Martin, D., Bachmann, J., .2022. IOPES, a new approach for seamless indoor/outdoor positioning of emergency teams. Submitted to International Journal of Disaster Risk Reduction.

2. **Angelats, E**.; Conesa, F. C., Mas-Florit, C., Cau-Ontiveros, M. Á., 2022**.** Automated co-registration and digitalization of archaeological features using RPAS imagery, aerial data archives and open-source libraries. In preparation (final editing stage. To be submitted to International Journal of Applied Earth Observation and Geoinformation).

INTERNATIONAL CONFERENCE PUBLICATIONS

3. **Angelats, E**., Espín, P., Navarro, J. A., Parés, M. E., 2021. Performance analysis of the IOPES seamless indoor-outdoor positioning approach. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B4-2021, 229–235, https://doi.org/10.5194/isprs-archives-XLIII-B4-2021-229-2021, 2021.

4. **Angelats, E**., Navarro, J. A., and Parés, M.E. 2020. Towards seamless indoor-outdoor positioning: the iopes project approach. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B4-2020, 313–319, https://doi.org/10.5194/isprs-archives-XLIII-B4-2020-313-2020, 2020.

5. **Angelats, E**.; Parés, M. E., Mas-Florit, C., Cau-Ontiveros, M. Á., 2019. A methodology for semi-automatic documentation of archaeological elements using RPAS imagery. In Proceedings of the 2019 IMEKO TC-4 International Conference on Metrology for Archaeology and Cultural Heritage, 4–6 December 2019, Florence (Italy).

6. **Angelats, E**., Cau-Ontiveros, M. A., Mas-Florit, C., 2018. High resolution orthophotos and a digital surface model of the Roman city of Pollentia (Mallorca, Spain) using RPAS imagery, aerial images, and open data archives. In Proceedings of the 2018 IEEE International Conference on Metrology for Archaeology and Cultural Heritage, 22–24 October 2018, Cassino (Italy) (Best paper award).

7. **Angelats, E**., Parés, M. E., Kumar, P., 2018. Feasibility of smartphone based photogrammetric point clouds for the generation of accesibility maps. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial

Information Sciences, Volume XLII-2, 2018. ISPRS TC II Mid-term Symposium "Towards Photogrammetry 2020", 4–7 June 2018, Riva del Garda (Italy). https://doi.org/10.5194/isprs-archives-XLII-2-35-2018, 2018

8. **Angelats, E**. and Navarro, J. A., 2017. Towards a fast, low-cost Indoor Mapping and Positioning System for Civil Protection and Emergency Teams. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-2/W8, 9–15, https://doi.org/10.5194/isprs-archives-XLII-2-W8-9-2017,2017.

9. **Angelats, E**. Parés, M.E., Colomina, I. Hybrid models for trajectory error modelling in urban environments. In: Archives of the International Society of Photogrammmetry and Remote Sensing, Prague, Czech Republic, 2016.

10. **Angelats, E**., Molina, P., Parés, M. E. and Colomina, I., 2014. A parallax based robust image matching for improving multisensory navigation in gnss-denied environments. In: Procceding of the ION GNSS+ 2014, Tampa, USA, pp. 2132 – 2138.

11. Montaño, J., Molina, P., **Angelats, E**., Colomina, I., Latorre, A., 2014. The PERIGEO: Image processing techniques development & testing for multimodal absolute & relative navigation. Proceedings of the 9th International ESA Conference on Guidance and Navigation Control Systems, 2-6 June 2014, Porto, Portugal

12. **Angelats, E**. and Colomina, I., 2014. One step mobile mapping laser and camera data orientation and calibration. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XL-3/W1, Castelldefels, Spain, pp.

13. **Angelats, E**., Blázquez, M. and Colomina, I., 2012. Simultaneous orientation and calibration of images and laser point clouds with straight segments. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 39(B1), Melbourne, Australia, pp. 91–96

INDUSTRIAL AND RESEARCH PROJECTS

The knowledge gained during this PhD theses and some of their outcomes have been used in several industrial and research projects. Hereafter a list of the main relevant ones is presented.

- Master_MED. Suport a la docència del master de drones de la EETAC. 2017-2022.

- CATERPILLAR-2. Desenvolupament i implementació d'algoritmes per detecció i tracking de nius d'orugues. Projecte industrial per a FitoStinger S.L. 2020-2021.

- ARCHREMOTELANDS. Arqueología, teledetección y arqueometría: una aproximación multidisciplinar al paisaje y las cerámicas desde el periodo romano al medieval en Mallorca (Islas Baleares). Projecte finançat pel Programa estatal de fomento de la investigación científica y técnica de excelencia, subprograma estatal de generación de conocimiento, en el marco del Plan estatal de investigación científica y técnica y de innovación 2013-2016. Convocatoria 2017. (Referència: HAR2017-83335-P). 2018-2021.

- Archaeors. Projecte per la Fundació Bosch i Gimpera. Tasques de Teledetecció que se enmarquen en el projecte "El estudio de la transformación del mundo rural de Mallorca de época romana a la Edad Media mediante la aplicación de técnicas no invasivas" dirigit por M. A. Cau i C. Mas Florit. Arqueologia. 2021.

- IOPES. "Indoor-Outdoor positioning for Emergency Staff"- Coordinador: CTTC. Program UCPM-2019-PP-AG, Funded by DG-ECHO, European Comission. Topic: Preparedness projects. 2020-2021.

- IOPESA. The IOPESA project aims to generate reference trajectories in adverse environments for UAV applications based on combining RTK+INS/stereo camera data. Funded by European Space Agency. 2022-2023.

ACKNOWLEDGEMENTS

Aquesta tesi és part d'un viatge en el qual he tingut sensacions com quan faig excursions a cims dels Pirineus, en molts moments crec que ja arribo al cim però encara queda camí per recórrer. La tesi va tenir les seves arrels a l'antic Institut de Geomàtica i ha acabat al CTTC. Han estat moltes les persones que m'han ajudat, escoltat i donat suport durant aquest procés.

En primer lloc, gràcies Ismael, pels teus consells i supervisió durant aquests anys, i gràcies Josep per ajudar-me a solucionar, sempre amb molta rapidesa i de manera molt fàcil, els diferents tràmits relacionats amb la Universitat. Una persona clau és l'Eulàlia. Aquesta tesi no es pot entendre sense el teu paper actiu. M'has ajudat i aportat idees que apareixen al llarg de la tesi, però sobretot m'has escoltat, has posat llum en la foscor, i m'has ajudat a trobar el camí quan no hi veia clar.

No em voldria oblidar dels companys que tinc i he tingut dins del grup de Geomàtica, he après molt tant en l'àmbit personal com tècnic/recerca d'aquest gran grup de persones. Moltes gràcies Michele per donar-me la llibertat i el suport a l'hora d'explorar línies de recerca molt engrescadores, més enllà d'aquesta tesi. Un suport clau de la tesi ha estat la Marta, no només en el personal, també per tot el suport en la part del modelatge matemàtic, pilar d'aquesta recerca. En paral·lel, part de les idees metodològiques van sortir de diferents discussions tècniques amb Pere. El meu agraïment també a vosaltres. En una etapa final d'aquesta tesi, he rebut també el suport constant i consells del Pedro i el Pep, que han permès desenvolupar les idees i resultats del projecte IOPES. Moltes gràcies per ser-hi, per fer "d'advocat del diable" en les ocasions que calia, per posar la casa com a zones ideal de proves, i per fer les coses molt fàcils. No em voldria oblidar tampoc de la Danielly i el Jesús, dos grans professionals i millors persones, de vosaltres aprenc cada dia.

En paral·lel a la tesi, i al grup de Geomàtica, dins del CTTC he tingut l'oportunitat de treballar colze a colze amb persones de les quals he pogut aprendre molt (Paolo, Marco) i malgrat no participar directament en aquesta, m'han ajudat a complementar els meus coneixements i formació. Fora del CTTC, moltíssimes gràcies al Cuco i la Cati per deixar-me entrar en aquest món de l'arqueologia i obrir-me les portes de la ciutat romana de *Pollentia*. Gràcies també al Francesc per totes les discussions i temps compartit. Baixant cap al sud, moltes gràcies Carles, he après molt, he gaudit i espero continuar gaudint de la recerca que tenim activa amb tu i el Jesús.

Finalment, no em voldria oblidar de la meva família, repartida entre dos continents i l'estrella màgica. Moltes gràcies als meus pares per intentar sempre que no em falti de res, per donar-me totes les eines i mitjans perquè pogués fer el que pensés que era el millor en cada moment. Moltes gràcies Laura per ser-hi, per insistir, ara et toca a tú! Gràcies també als que em vau rebre amb els braços oberts i no heu perdut mai el somriure. I ara sí, per acabar, aquest viatge no es pot entendre sense el suport incondicional de la M.Eulàlia i d'en Biel, que heu aguantat la muntanya russa d'emocions, les preocupacions i tenir el cap en un altre lloc en molts moments. Moltes gràcies, sou la meva lot ... i Biel, ja he acabat el llibre!