Indoor Localization using Google Tango and a 3D Building Model

Graduation plan

Master of Science in Geomatics for the Built Environment
Delft University of Technology by

Francina Johanna Bot

January 11, 2017

Contents

1 Introduction 3
- 1.1 Problem Field 3
- 1.2 Scientific Relevance 3
- 1.3 The Tango Device 4

2 Related Work 5
- 2.1 Research using a Tango Device 5
- 2.2 Image Based SLAM 5

3 Problem Statement 6
- 3.1 Research Question 6

4 Approach 7
- 4.1 Phase 1: Benchmarking the scanning capabilities of the Tango device. 7
  - 4.1.1 Modelling quality 7
  - 4.1.2 Applicability to indoor applications 8
- 4.2 Phase 2: Benchmarking the localization capabilities of the Tango device. 8
- 4.3 Phase 3: Matching the output of the Tango device with a 3D model. 9

5 Schedule 10

6 Preliminary Results 11
- 6.1 Modelling 11
- 6.2 Building / Following the ADF 11

Acronyms

ADF Area Description File 3
AR Augmented Reality 5
GNSS Global Navigation Satellite System 3
GPS Global Positioning System 4
IMU Inertial Measurement Unit 5
SLAM Simultaneous Localization and Mapping 5
VIO Visual-Inertial Odometry 5

Glossary

Positioning: to find a point or area occupied by a physical object. This term is used as to find an xyz-position in space, relative to what is known.

Localization: to find a position or site occupied in space. This term is used as to find a location with a specific meaning, as an enrichment upon the position itself.

Navigation: the act or practice of navigating, as to get around or move. This term is used as being directed from one position or location to another.

These definitions are loosely based on Merriam-Webster's Dictionary. See figure 1.

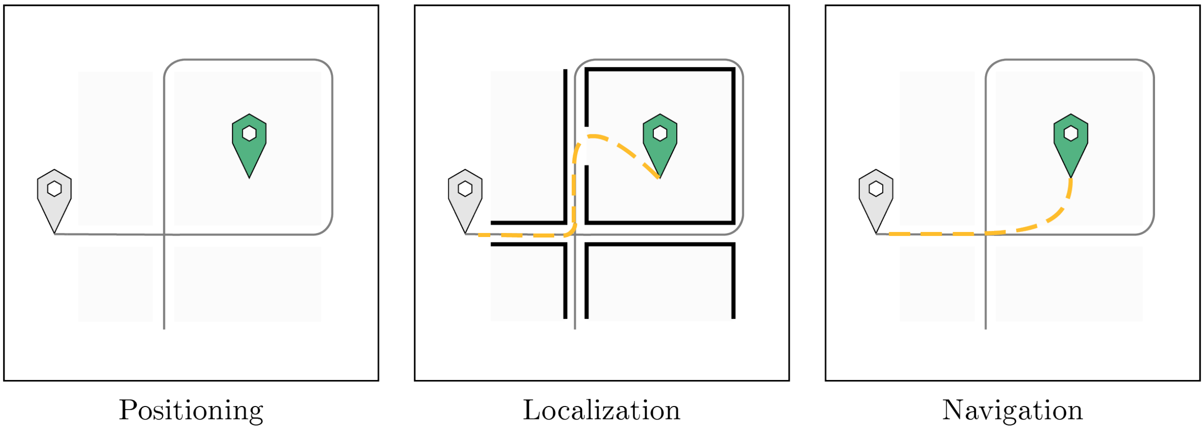

Figure 1: Differences between Positioning, Localization and Navigation for indoor situations.

1 | Introduction

The proposed research will be conducted as part of the graduation process for attaining a Master of Science degree in the field of Geomatics for the Built Environment.

Within the program, I found indoor localization and data gathering techniques particularly interesting, two subjects which are embedded into Project Tango.

1.1 | Problem Field

Outdoor localization represents a well established field, describing methods using Global Navigation Satellite System (GNSS) services and enrichments upon those, meant to increase both accuracy and precision. However, as GNSS signals do not contain the signal strength to penetrate walls, different solutions have to be found for localization indoors.

In order to solve the indoor localization problem, many different systems have been developed, tested and evaluated (Mautz, 2012) and these remain actively under development (Khoshelham and Zlatanova, 2016; Yassin et al., 2016). However, each system that has been implemented thus far is only fit to serve the specific locations these applications have been developed for. Often times these systems even depend on an infrastructure installed for the purpose of indoor localization. These techniques would only become commercially interesting, if they can provide the consumer with an affordable and easily attainable solution, as currently owning a smartphone is the only requirement to utilize outdoor localization techniques (The Economist, 2012).

Project Tango itself is aimed at using "computer vision to give devices the ability to understand their position relative to the world around them" (Google, 2017b). If this is actually the case, then by knowing its position relative to the world around, a Tango enabled device should be able to perform indoor positioning. Connecting this relative output to a specific location, would generate indoor localization. As both Lenovo and ASUS are now producing Tango enabled smartphones as well (Google, 2017a), indoor localization using Tango technologies could become available to the customer as easily as outdoor localization, provided the system performs as desired.

1.2 | Scientific Relevance

Project Tango proposes new possible solutions to the indoor localization problem, through a new combination of sensors into a hand held device. In that way, any application utilizing e.g. stereo imaging or point clouds, can now be made available to the public in a commercially interesting way. Furthermore, the output of these sensors is written into a new type of file, the Area Description File (ADF). This file type stores location specific markers, which should enable the device to know its position relative to the recorded area. However, this newly available output has to adhere to certain requirements in order to be useful for the desired and foreseen applications.

This research will mainly aim at benchmarking the Tango device's capabilities in regard to indoor localization. Benchmarks should be set for positioning itself into, and tracking its movement through, an area stored in an ADF, as well as modelling some characteristics of the indoor space. Furthermore, possibilities should be found to enrich the generated data, to see if the indoor positioning methods proposed by Project Tango can be transcended towards indoor localization methods.

1.3 | The Tango Device

Along with the distribution of Project Tango, a specific tablet was introduced, containing a new combination of sensors which should enable a plethora of new possible applications interesting to the consumer. This Tango device was brought out in 2014 along with an open call for developers to create applications utilizing the device's possibilities. The Tango Tablet Development kit Google (2017b) has a 7.02" 1920x1200 HD IPS display (323 ppi) and runs on Android™ 4.4 KitKat®. The sensors embedded into the device and their use for indoor localization and modelling can be found in table 1.1. Additionally, a Global Positioning System (GPS) receiver is embedded, aided by a Compass and Barometer for a faster and more precise outdoor location fix.

Sensor	**Functions**
4MP 2 μm RGB-IR Camera	Generate Stereo Images
	Generate Point Cloud (Passive)
Fisheye Camera	Generate Stereo Images
	Motion Tracking
IR projector	Generate Point Cloud (Active)
IMU: Accellerometer	Accelleration
IMU: Gyroscope	Pitch, Roll, Yaw

Table 1.1: Sensors embedded into a the Tango Tablet Development kit and their functionality for indoor localization and modelling.

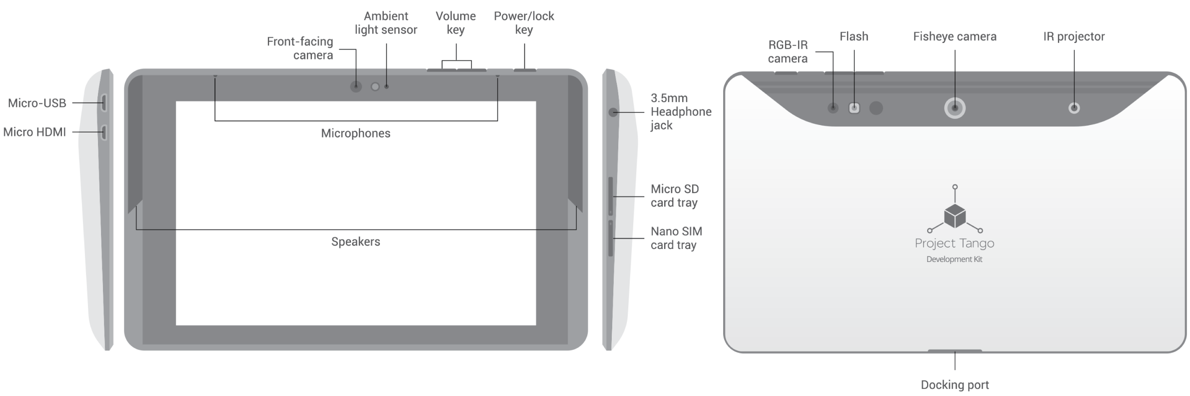

Figure 1.1: Hardware Diagram for the Tango Tablet Development kit.
Source: https://developers.google.com/tango/hardware/tablet

2 | Related Work

Since the release of Project Tango, the device has been used in several research projects for its specific sensor combination in a hand-held device. However, except for the first two mentioned sources, none have specifically aimed at testing the device's capabilities yet.

2.1 | Research using a Tango Device

How well the Tango device is capable of creating 3D models usable for indoor localization was tested by Diakité and Zlatanova (2016), using the device's native applications. Its 3D modelling capabilities in large-scale outdoor scenes have been tested by Schöps et al. (2015), using monocular motion stereo to reconstruct scenes on the fly. This is based on the device's Visual-Inertial Odometry (VIO) and images generated from the fish eye camera by first computing frame by frame depth maps, and then fusing them using volumetric depth map fusion. To improve the accuracy of the model several filtering steps are performed, which are improved in a later set-up (Schöps et al., 2016). Furthermore, the applicability to other devices is discussed, rendering the Tango device to merely a useful tool for the development of this method.

In another outdoor application, Agarwal et al. (2015), have combined the device's VIO with geotagged images from Google Streetview, in order to obtain accurate metric localization. Then, the Tango device is used to implement and evaluate the developed system for the use of personal localization in an urban scenario. However, when performing image matching from such large databases, geometric bursts will occur, a problem to which Sattler et al. (2012) have attempted to find a solution using the Tango device. Both research projects handle localization as an image matching problem.

Ramirez (2015) has used the Tango device to generate data and implement an application for the evaluation of text-based Simultaneous Localization and Mapping (SLAM) algorithms. Sweeney and Turk (2016) developed an Augmented Reality (AR) application for the device to localize itself into a reference 3D point cloud using its absolute position. This is determined through gravity calculations performed using the devices Inertial Measurement Unit (IMU), a technique used to improve the accuracy of imaged based SLAM.

2.2 | Image Based SLAM

In the field of robotics, imaged based SLAM, the technique upon which the ADF is built, is a much researched problem. Therefore, much in-depth information is available. Literature dealing with this and specifically useful for this research will be referenced in the research itself. Some interesting recent developments have been published by Dos Santos et al. (2016); Gálvez-López et al. (2016); Lopez-Antequera et al. (2016); Moteki (2016); Sattler et al. (2016).

3 Problem Statement

In 2014, Google introduced Project Tango, accompanied by a call to developers to "explore physical space around the user, including precise navigation without GPS, windows into virtual 3D worlds, measurement and scanning of spaces, and games that know where they are in the room what's around them" (Google, 2017b). At the moment of writing, no indoor navigation application has been made available yet. However, Google promises that the Tango device contains the right basis to build such applications on. The customer website for project Tango has claimed to make an exploration feature available by the end of 2016 (Figure 3.1). In the way this is promoted, this indoor navigation is based on the device knowing its location.

Therefore, the objective of this research is to define and explore the indoor localization possibilities the Tango device presents. Furthermore, the quality and applicability to different scenes will be tested. An additional goal is to explore how the output of the Tango device can be integrated with a 3D model which will provide in additional and probably more accurate information about the scanned location.

3.1 Research Question

The main research question is defined as follows:

TO WHAT EXTENT CAN THE NATIVE OUTPUT OF THE TANGO DEVICE BE USED FOR INDOOR LOCALIZATION AND NAVIGATION, AND BE CONNECTED TO A 3D MODEL?

This question will be answered by taking the research through three phases and composing sub questions accordingly. The phases are:

- Benchmarking the scanning capabilities of the Tango device.
- Benchmarking the localization capabilities of the Tango device.
- Matching the output of the Tango device with a 3D model in order to perform indoor localization and navigation.

Figure 3.1: Promised additional functionality for the Tango device. (Google, 2017a)

4 | Approach

As the research will consist of three phases, a methodology should be written for all three of these. This due to the fact that each phase demands its own approach. Therefore, the general process can be seen as a method to go through the phases. For each phase, the general Geomatics process as described by Lemmens (1991) should be kept in mind, describing the capture, storage, analysis and visualization of the handled data (See figure 4.1) . Each phase will contain the definition of a phase specific sub question and the design of experiments and theoretical research to answer said question.

4.1 | Phase 1: Benchmarking the scanning capabilities of the Tango device.

The scanning capabilities of a device may be inquired in two ways: testing the quality of the model built with the Tango native applications, and defining the usability of said model for indoor navigation.

4.1.1 | Modelling quality

As mentioned in section 2.1, the modelling capabilities of the Tango device have been tested by Diakité and Zlatanova (2016) and Schöps et al. (2015), though the latter used newly created algorithms. As the objective of this research is set to find the possibilities the native output of the Tango presents, the scanning capabilities will be tested on included algorithms.

In the field of robotics, and therefore in image based SLAM, designed benchmarks are generally aimed at evaluating algorithms, rather than devices. However, Henry et al. (2010) proposed a manner of benchmarking the scanning capabilities of **rgbd!** cameras used for image based SLAM. The experiment performed by Henry et al. consists of three key components:

- The spatial alignment of consecutive data frames.
- Detection of loop closures.
- The globally consistent alignment of the complete data sequence.

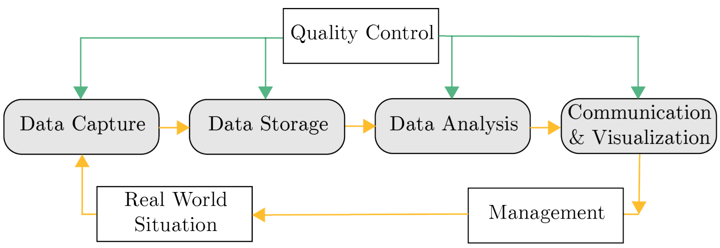

Figure 4.1: Methodology for the Geomatics process, as proposed by Lemmens (1991).

In a similar manner, the output of the Tango device can be benchmarked for its 3D modelling capabilities. This approach can be summarized into the following sub question:

[1.1] WHAT IS THE QUALITY OF 3D MODELS CREATED WITH THE TANGO DEVICE?

4.1.2 Applicability to indoor applications

Another requirement for performing both indoor navigation and localization, is the applicability of the output of the Tango device for indoor navigation. Diakité and Zlatanova (2016) have used the native output to test on the detection of empty space and openings and the extraction of planar patches in order to test this. However, Isikdag et al. (2013) and Zlatanova et al. (2014) have listed several more requirements to a system for indoor applications. The output of the device may be mirrored to these requirements, according to the following sub question:

[1.2] TO WHAT EXTENT IS THE NATIVE OUTPUT OF THE TANGO DEVICE APPLICABLE FOR INDOOR LOCALIZATION AND NAVIGATION?

4.2 Phase 2: Benchmarking the localization capabilities of the Tango device.

Where the second part of the first phase mirrors the output of the Tango device to the literary requirements, the second phase aims to answer this question in a more practical manner. Therefore, the characteristics of the device should be explored, as well as the workings and capabilities of the device and its native applications. This is done on the basis of the three main principles underlying the system: motion tracking, area learning and depth perception (see figure 4.2).

These concepts are presented as separate systems performing different actions, though in reality these concepts are all intertwined. In what way they aid in indoor localization and navigation may be answered through the following sub question:

[2] WHICH CHARACTERISTICS OF THE TANGO DEVICE SUPPORT INDOOR LOCALIZATION AND IN WHAT WAY?

Furthermore, the quality of these operations may be tested, in addition to the quality of the 3D models as described in **??**. The approach for these tests will be defined further into the process.

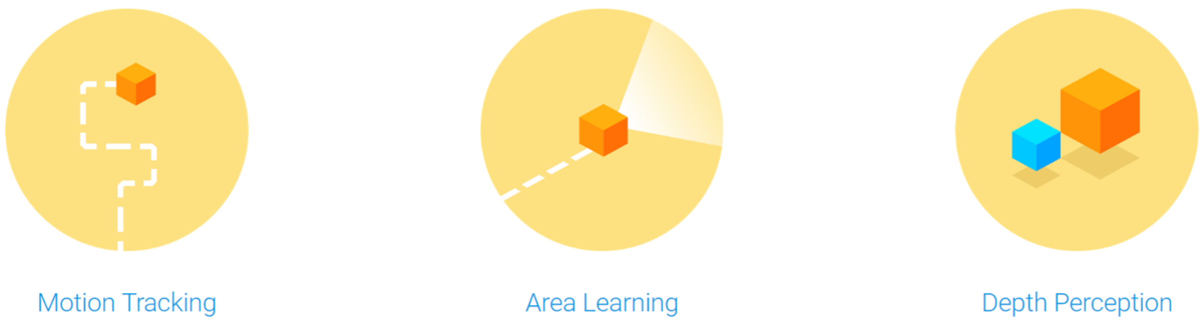

Figure 4.2: Basic Concepts for the Tango Device. (Google, 2017b)

4.3 Phase 3: Matching the output of the Tango device with a 3D model.

The third phase is mainly aimed at performing an overall wrap up of the indoor localization and navigation techniques the Tango device should be capable of. It should become apparent, that in order to transcend the output from positioning to localization, a 3D model containing semantic information should be connected to the data collected by the Tango device. To what extent this may be possible should become incrementally clear throughout the research. In order to reach this goal, the following sub question could function as a guide:

[3] IN WHAT WAYS CAN THE NATIVE OUTPUT OF THE TANGO DEVICE BE CONNECTED TO A 3D MODEL IN ORDER TO PERFORM INDOOR LOCALIZATION AND NAVIGATION?

This phase can be used to show a real world application for which a Tango enabled device could aid consumers in localizing themselves inside, and navigating through an environment new to them.

5 | Schedule

Table 5.1 depicts a general schedule for completion of the research project. This is based on the set P-presentation schedule and the phases proposed in chapter 4.

Period	Date	Activity
Week 1.3		**P0: Register for Graduation**
Week 1.3-1.5		Explore Topics
Week 1.5-1.10		Gather Information about Project Tango
Week 2.1	14-11-2016	**P1: Progress Review**
Week 2.1-2.6		Research Project Tango & SLAM
Week 2.1-2.3		Perform First Experiments
Week 2.4-2.6		Outline Research Scope
Week 2.7	16-01-2017	**P2: Formal Assessment Graduation Plan**
Week 2.7-2.10		Phase 1: Benchmark Scanning
Week 3.1-3.3		Phase 2: Benchmark Indoor Localization
Week 3.4	t.b.a.	**P3: Colloqium Midterm**
Week 3.4-3.5		Phase 3: Match with 3D Model
Week 3.5-3.7		Revisit Phase 1
Week 3.8-3.10		Revisit Phase 2
Week 4.1-4.2		Revisit Phase 3
Week 4.3	t.b.a.	**P4: Formal Process Assessment**
Week 4.3-4.8		Finalize Thesis
Week 4.10	t.b.a.	**P5: Final Assessment**

Table 5.1: General schedule for completing the proposed research.

6 | Preliminary Results

As part of the orientation into the subject and the definition of the research scope, a preliminary set of results have been gathered.

6.1 | Modelling

Figure 6.1 shows a 3D Mesh created on the fly using the Constructor app. As this is one of the Project Tango native applications, this model may be tagged as native output. The model was created in the Geolab at the faculty of Architecture, Urbanism and Building Technology.

As can be seen, a coloured 3D Mesh can be built with the device, in which features are visually recognizable. However, no semantics are added to the model. In the shape of the room, a slight drift can be detected. Though an apparent cause to this would be the trajectory taken when building the model, sufficient drift correction on this scale should be possible.

A cause for the apparent drift could be way the model is built: frame by frame. This method enables on-the-fly processing, but is prone to large error. An example of this can be seen in figure 6.2. Two chairs are distinguishable, with a person sitting on the left one. Though most of the person's shape is captured, only a part is coloured accordingly. This due to the fact that the person sat down on the chair, after the first few frames containing the chair were processed.

Furthermore, the model contains a lot of noise, mainly from people moving around. This raises the question whether the device will be capable of providing for sufficient indoor localization in crowded places, a subject that should be tested more extensively.

6.2 | Building / Following the ADF

To test the functionality of the ADF, some experiments have been performed at the faculty of Electrical Engineering, Mathematics and Computer Sciences. An ADF was built on the first and second floor of the main building and tested at several moments. For all images in this section, the blue line represents the trajectory in respect to the starting point as registered by the device's IMU. The green line represents the trajectory in respect to the ADF.

Figure 6.3 (Left) shows that it may take some time for the device to position itself within the ADF. Usually it takes a while, which might be a down side for the applicability for commercially available indoor applications. More specifics about the initialization time can be tested in a later phase. Figure 6.3 (Right) shows a new trajectory added to an existing ADF. The added trajectory is the thicker one on the bottom, which was followed twice and recorded practically the same. This test suggests a sufficient capability of adding new area to the ADF on the fly.

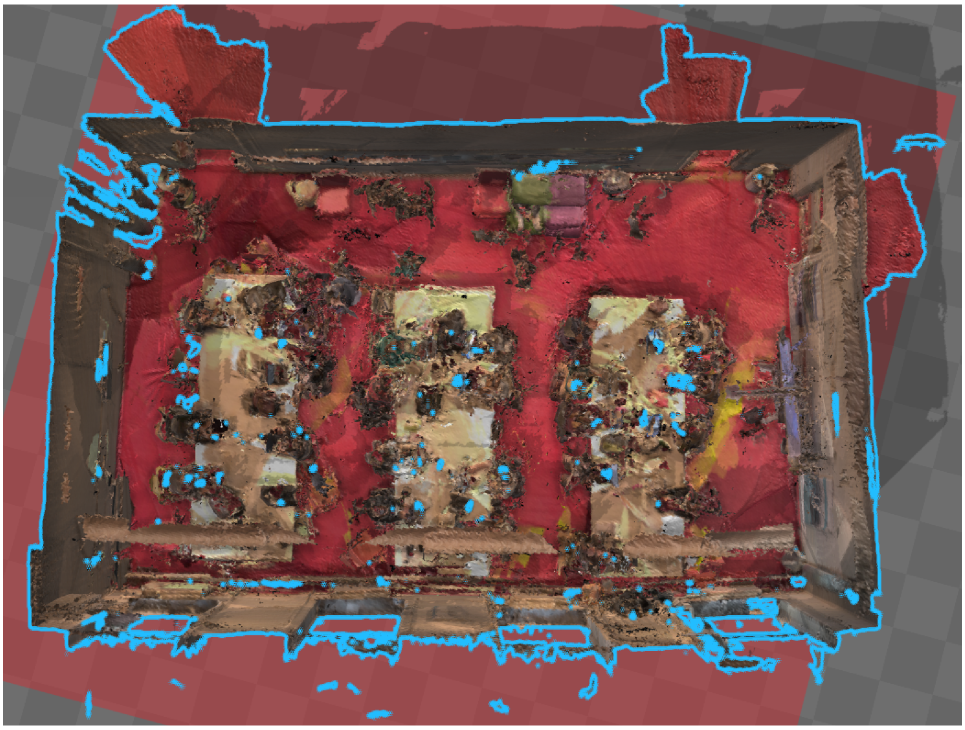

Figure 6.1: 3D Mesh created by the Constructor app.

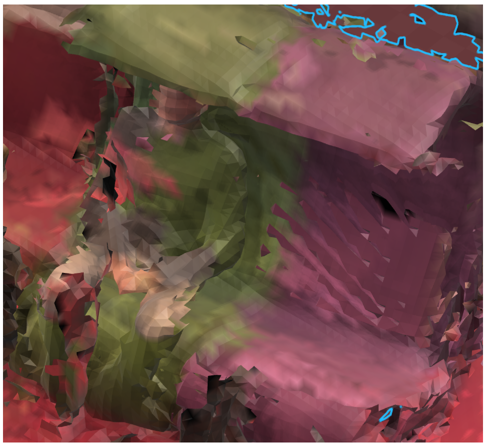

Figure 6.2: Detail of 3D Mesh created by the Constructor app, showing error caused by fast processing.

How well the movement of the device is tracked, can be tested by walking straight lines back and forth (Figure 6.4). Even when walking slowly, a drift is noticeable. However, the third iteration over the same trajectory was plotted over the first one again, proving an application of the drift correction embedded into the device. When walking fast, the drift is much stronger, causing strong jumps in the trajectory in respect to the ADF.

When following a complex route through an area stored in an ADF, the device tracks the route taken quite well (Figure 6.5). Sudden turns and changes over (sub)levels are nicely represented. However, when performing a similar experiment during rush hour (Figure 6.6), the device loses its track whenever the camera cannot track enough landmarks and keyframes from the ADF, as the camera is blocked by humans walking around. This causes jumps in the trajectory, as in figure 6.4. Furthermore, a stronger drift in all direction can be detected in the images.

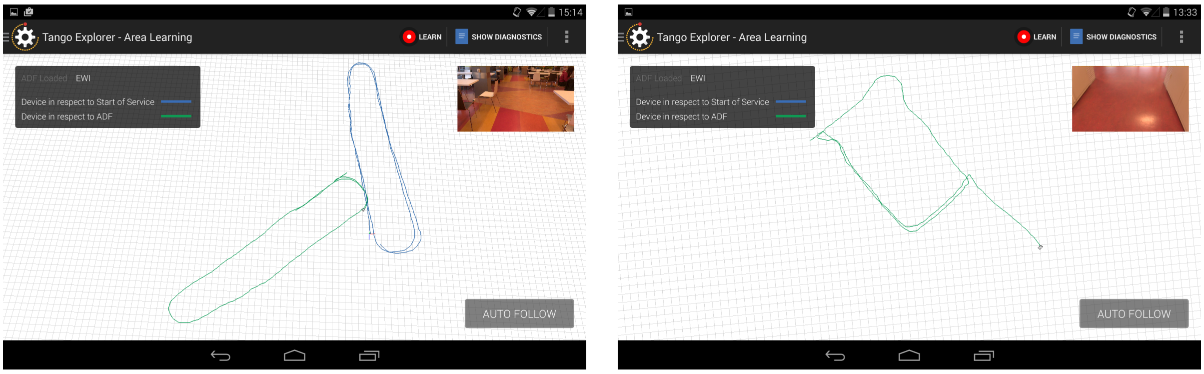

Figure 6.3: Initializing: Time it takes the device to recognize its location (Left), handling a new area within the ADF.

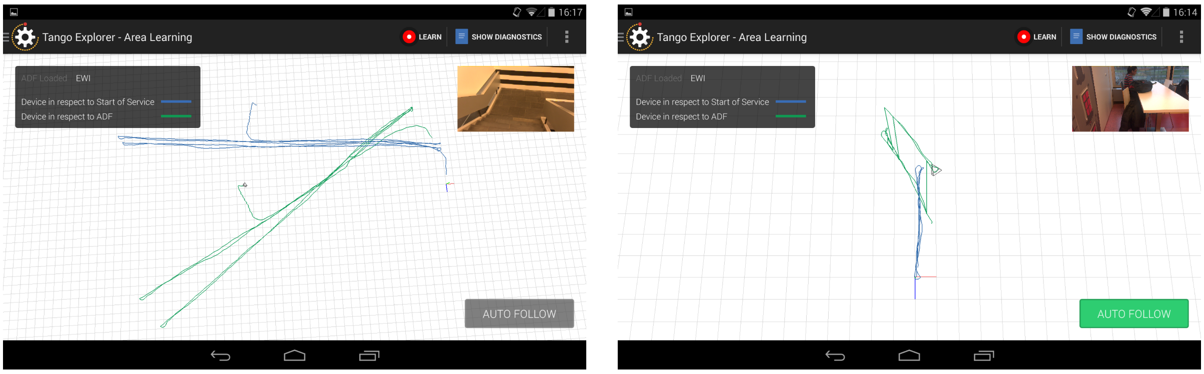

Figure 6.4: Following straight lines while walking slow (Left) and fast (Right).

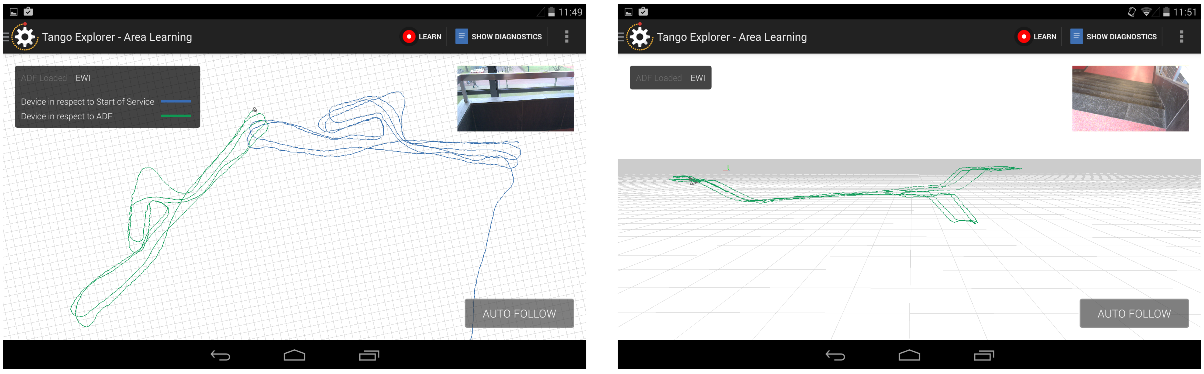

Figure 6.5: Following a complex route in 2D (Left) and 3D (Right).

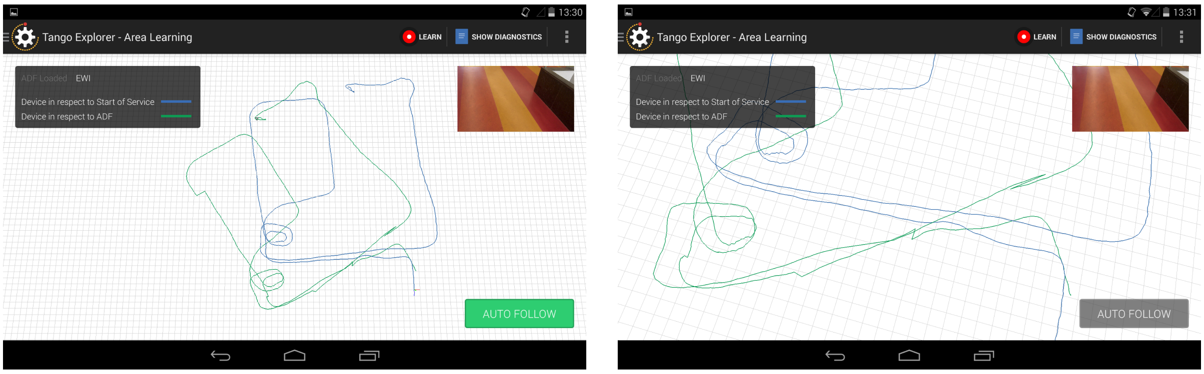

Figure 6.6: Following a complex route during rush hour.

Bibliography

P. Agarwal, W. Burgard, and L. Spinello. Metric Localization using Google Street View. *arXiv preprint arXiv:1503.04287*, pages 3111–3118, 2015. ISSN 21530866. doi: 10.1109/IROS.2015.7353807.

A. A. Diakité and S. Zlatanova. FIRST EXPERIMENTS WITH THE TANGO TABLET FOR INDOOR SCANNING. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume III-4*, pages 67–72, Prague, Czech Republic, 2016. XXIII ISPRS Congress, 12-19 July 2016. doi: 10.5194/isprsannals-III-4-67-2016.

D. R. Dos Santos, M. A. Basso, K. Khoshelham, E. De Oliveira, N. L. Pavan, and G. Vosselman. Mapping Indoor Spaces by Adaptive Coarse-to-Fine Registration of RGB-D Data. *IEEE Geoscience and Remote Sensing Letters*, 13(2):262–266, 2016. ISSN 1545598X. doi: 10.1109/LGRS.2015.2508880.

D. Gálvez-López, M. Salas, J. D. Tardós, and J. Montiel. Real-time monocular object SLAM. *Robotics and Autonomous Systems*, 75:435–449, jan 2016. ISSN 09218890. doi: 10.1016/j.robot.2015.08.009. URL `http://dx.doi.org/10.1016/j.robot.2015.08.009http://linkinghub.elsevier.com/retrieve/pii/S0921889015001864`.

Google. Tango Customers Page, 2017a. URL `https://get.google.com/tango/`.

Google. Tango Developers, 2017b. URL `https://developers.google.com/tango/`.

P. Henry, M. Krainin, E. Herbst, X. Ren, and D. Fox. RGB-D Mapping : Using Depth Cameras for Dense 3D Modeling of Indoor Environments. *RGBD Advanced Reasoning with Depth Cameras Workshop in conjunction with RSS*, 1(c):9–10, 2010. ISSN 0278-3649. doi: 10.1177/0278364911434148. URL `http://ils.intel-research.net/uploads/papers/3d-mapping-iser-10-final.pdf`.

U. Isikdag, S. Zlatanova, and J. Underwood. A BIM-Oriented Model for supporting indoor navigation requirements. *Computers, Environment and Urban Systems*, 41:112–123, 2013. ISSN 01989715. doi: 10.1016/j.compenvurbsys.2013.05.001. URL `http://dx.doi.org/10.1016/j.compenvurbsys.2013.05.001`.

K. Khoshelham and S. Zlatanova. Sensors for Indoor Mapping and Navigation. *Sensors*, 16(5):655, may 2016. ISSN 1424-8220. doi: 10.3390/s16050655. URL `http://www.mdpi.com/1424-8220/16/5/655`.

M. J. P. M. Lemmens. GIS: The Data Problem. Technical report, 1991.

M. Lopez-Antequera, N. Petkov, and J. Gonzalez-Jimenez. Image-based localization using Gaussian Processes. In *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–7. IEEE, oct 2016. ISBN 978-1-5090-2425-4. doi: 10.1109/IPIN.2016.7743697. URL `http://ieeexplore.ieee.org/document/7743697/`.

R. Mautz. *Indoor Positioning Technologies*, volume 85. 2012. ISBN 9783908440314;. doi: 10.3929/ethz-a-007313554. URL http://e-collection.library.ethz.ch/eserv/eth:5659/eth-5659-01.pdf.

Merriam-Webster. Dictionary. URL https://www.merriam-webster.com/.

A. Moteki. Fast and Accurate Relocalization for Keyframe-Based SLAM Using Geometric Model Selection. *IEEE Virtual Reality Conference 2016 19–23*, pages 235–236, 2016.

E. A. Ramirez. An Experimental Study of Mobile Device Localization. Technical report, Massachusetts Institute of Technology, 2015.

T. Sattler, B. Leibe, and L. Kobbelt. Improving image-based localization by active correspondence search. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7572 LNCS(PART 1):752–765, 2012. ISSN 03029743. doi: 10.1007/978-3-642-33718-5_54.

T. Sattler, B. Leibe, and L. Kobbelt. Efficient & Effective Prioritized Matching for Large-Scale Image-Based Localization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, X (1):1–1, 2016. ISSN 0162-8828. doi: 10.1109/TPAMI.2016.2611662. URL http://ieeexplore.ieee.org/document/7572201/.

T. Schöps, T. Sattler, C. Häne, and M. Pollefeys. 3D Modeling on the Go: Interactive 3D Reconstruction of Large-Scale Scenes on Mobile Devices. In *Proceedings - 2015 International Conference on 3D Vision, 3DV 2015*, pages 291–299, 2015. ISBN 9781467383325. doi: 10.1109/3DV.2015.40.

T. Schöps, T. Sattler, C. Häne, and M. Pollefeys. Large-Scale Outdoor 3D Reconstruction on a Mobile Device. *Computer Vision and Image Understanding*, 2016. ISSN 10773142. doi: 10.1016/j.cviu.2016.09.007.

C. Sweeney and M. Turk. Efficient Computation of Absolute Pose for Gravity-Aware Augmented Reality Efficient Computation of Absolute Pose for Gravity-Aware Augmented. (November), 2016. doi: 10.1109/ISMAR.2015.20.

The Economist. Indoor Positioning: Finding the way Inside. *Technology Quarterly*, (December), 2012. URL http://www.economist.com/news/technology-quarterly/21567197-navigation-technology-using-satellites-determine-your-position-only-works.

A. Yassin, Y. Nasser, M. Awad, A. Al-Dubai, R. Liu, C. Yuen, and R. Raulefs. Recent Advances in Indoor Localization: A Survey on Theoretical Approaches and Applications. *IEEE Communications Surveys Tutorials*, PP(99):1, 2016. ISSN 1553-877X. doi: 10.1109/COMST.2016.2632427.

S. Zlatanova, L. Liu, G. Sithole, J. Zhao, and F. Mortari. *Space subdivision for indoor applications*. Number 66. 2014. ISBN 9789077029374.