Creating Remote Situation Awareness of Indoor First Responder Operations using SLAM

Master's Thesis - July 2020

Author

Bart-Peter Smit

Supervisors

ir. E. Verbree

ir. R. Voûte

Responsible professor

Prof. dr. ir. P.J.M. van Oosterom

University

Delft University of Technology

Colophon

Bart-Peter Smit: *Creating Remote Situation Awareness of Indoor First Responder Operations using SLAM* (2020)

This master thesis is commissioned by the University of Utrecht, University of Wageningen, University of Twente and the Delft University of Technology as part of the Geographical Information Management and Applications (GIMA) Master of Science.

In cooperation with CGI Nederland B.V.
Vital Civil Infrastructures

Supervisor Delft University of Technology:
ir. E. Verbree

Supervisor CGI Nederland B.V.:
ir. R. Voûte

Responsible professor Delft University of Technology:
Prof. dr. ir. P.J.M. van Oosterom

Contact information:

Bart-Peter Smit
BartPeterSmit@Gmail.com

Acknowledgements

For six months, I have been working on this thesis that now see in front of you. Working on this project proved both challenging and rewarding, as it enabled me to work on an interesting topic that I believe has real potential for use in first responder operations. Over the course of this project, I have met many great people who have made it a joy to spend my time on careful construction of code, argumentation, and the acquirement of knowledge.

First of all, I would like to thank my daily supervisors, ir. Edward Verbree and ir. Robert Voûte. When I met Edward for the first time, he was presenting about the added value of direct use of point clouds, thereby opening my eyes to the merits of three-dimensional geo-information. Patiently and critical, he provided and continues to provide many insights into this new and sometimes complex world. Robert on the other hand, stressed the importance of adding even two more dimensions: time and level of detail. As I saw Robert almost every day in the office of CGI, he energized and motivated me to not only reach for the best result possible, but also to keep a clear definition of scope for my thesis. For this, I am thankful to him, but also for his enthusiasm with allowed me to engage in other activities of CGI. Thereby, I could sometimes draw some attention away from my thesis and return to it with new perspectives and ideas.

Second, I wish to thank all other graduate interns with whom I worked at CGI. The discussions in the coffee corner inspired many, whether they were about the 'AtoB-ANWBABB-Busje' or about the climate impact of big oil companies. In particular I wish to thank Sebastiaan den Boer, who was never afraid to share his C# wisdom; Ullas Rajvanshi, with whom I discovered the beauty of table soccer; Hessel Prins, photogrammetry and Lego expert; and Laurens Oostwegel, fellow HoloLens developer and fierce PinBall opponent.

Third, I want to thank all employees of CGI who made me feel very welcome at the office. Among these people were Jafeth van Elten, a noble Viking warrior with outstanding programming skills; the men from the GameFactory, who helped me tackle several nasty errors in Unity3D; and (the team of) Ben van Tricht, always up for good conversation.

Fourth, I want to thank the safety region Rotterdam-Rijnmond, and in particular Huib Fransen, for their insightful input and support.

Fifth, I would like to thank Drs. Maarten Zeylmans Van Emmichoven, who helped me a great deal by providing a Terrestrial Laser Scanner and spending an afternoon to scan the Vening Meineszbuilding together with me.

Finally, but certainly not least, I would like to express my gratitude to my friends and family for encouraging me to reach my goals and comforting me when times were tough.

Acronyms

3D	Three Dimensional
BIM:	Building Information Model
CML:	Concurrent Mapping and Localization
COP:	Common Operational Picture
EBSE:	Evidence-Based Software Engineering
EDM:	Emergency Decision Making
FIFO:	First In, First Out
GNSS:	Global Navigation Satellite systems
LIDAR	Light Detection and Ranging
LIFO:	Last in, First Out
MLS	Mobile Laser Scanner
SA	Situation Awareness
SLAM	Simultaneous Localization And Mapping
SLR:	Systematic Literature Review
SPAM	Simultaneous Posing And Mapping
SSA	Shared Situation Awareness
TLS	Terrestrial Laser Scanner

Definitions

Mapping:	The creation and visualization of spatial data, meaning data that includes a reference to the location or the attributes of objects or phenomena located in a defined space.
Space:	A spot, surface or volume with a spatial definition or spatial context
Addressing:	The capability of referencing to 'where' a space is.
Location:	A space with spatially defined physical borders. Locations have context, meaning it has meaning in relation to objects, phenomena and spaces.
Position:	A point, surface or volume in space, which is always addressed according to a reference frame. Positions do not have context, meaning they do not have meaning in relation to objects, phenomena and spaces.
Cognition:	Mental action or process to acquiring knowledge and understanding through thought, experience, and the senses
Reliability:	A construct consisting of accuracy, precision and robustness
Accuracy:	The extent to which measurements represent the real world situation.
Precision:	The extent to which the measurements are spread over the operational environment.
Robustness:	The extent to which the system performs continuously within different operational environments
Primary keyword:	Keyword that has been documented directly from observed literature.
Secondary keyword:	Keyword that has been documented indirectly from observed literature, which is connected to a domain
Field of study:	A branch of knowledge about a specific research subject.
Pose:	Combination of both the position and the orientation of a device
First responder performance:	Accountability for the measurement of success in executing the designated task of saving human lives and reducing private and public property damages in emergency situations.

List of Figures

Contents

1 Abstract

First responders operate in complex and dynamic environments to mitigate the effects of incidents. To maintain an overview of the operation, they are aided by spatial information to create a Common Operational Picture (COP). Such a COP aids decision making processes by displaying for example unit movements or incident effects, thereby creating situation awareness. Situation awareness exists of three levels: (1) perception, (2) comprehension, and (3) projection. The higher the level is, the better the quality of decisions are, thereby increasing performance. Although spatial data elements prove useful for increasing first responder performance, most services are limited to outdoor uses. Static data sources of buildings such as floor plans or BIM models are often non-existent or unavailable. Furthermore, they do not represent the dynamic nature of first responder operations, in which the situation changes and pathways may be blocked. Therefore, demand exists for dynamic methods of creating indoor situation awareness.

This thesis successfully proposes a novel method to create remote situation awareness in real-time for indoor first responder operation environments. It does so by demonstrating a Proof of Concept (PoC). Required data elements are identified from literature and stakeholder interviews. This directed the PoC in collecting 'mapping' and 'tracking' data elements within indoor first responder operation environments by using Simultaneous Localization and Mapping (SLAM) algorithms. These data elements are transferred from the indoor environment to a remote operation coordinator, where the data elements are visualized in a comprehensible way to facilitate the remote coordinator with situation awareness.

The PoC is evaluated on reliability, meaning the accuracy, precision and robustness of the system. Reliability is important, as system operators need to know to what extent they can trust the system. First, both mapping and tracking data elements are found to be accurate, representing the real-world well with minimal drift. Second, the detail in which this real-world is represented, also called precision, is rather low, as features appear jagged. Third, the system is robust if certain limitations are respected, as the system was able to work continuously and without notable lag in different indoor environments. A major threat to the robustness of the system is tracking loss, which was found to appear if the explorer is navigating narrow spaces or stairs.

Besides reliability, situation awareness is evaluated. The evaluation is based on stakeholder meetings with safety regions, in which the PoC is demonstrated. We found that collecting and visualizing 'mapping' and 'tracking' data elements does indeed build situation awareness in a reliable way. The object-oriented visualization method was perceived best, as it enables first responders to easily identify floors, walls, stairs and obstructive objects. Furthermore, the combined presentation of the mapped environment and the first responder poses (position + orientation) was found meaningful, as it enables first responders to comprehend the situation better. Therefore, the second level of SA (comprehension) is reached, which is the extent to which the PoC contributes to the real-time creation of situation awareness in indoor first responder operation environments.

Future research may focus on mitigating the effects of tracking loss, extending the PoC capabilities to multiple devices or devices with another nature, and focusing on extracting continuous navigable space from the mapping elements.

Keywords: Indoor: Mapping; Tracking; SLAM; Real-Time; 3D; Situation Awareness; Common Operational Picture; HoloLens; Navigable space; Remote; Visualization

2 Research scope

People all around the world flock towards cities (United Nations, Department of Economic and Social Affairs, Population Division, 2019). This process of urbanization calls for densification of urban areas, to reduce the spatial footprint of cities on the overall availability of land (European Commission, 2011; Neill and Schlappa, 2016). The combined ambition to reduce urban sprawl and facilitating urban housing drives construction of high-rise buildings. This can be observed in the recent decision of the Dutch city of Rotterdam, as the city decided to increase the maximum height of buildings from 200 to 250 meters (Kurvers, 2019).

The process of densification is accompanied by a growing supply of urban information flows (Hashem et al., 2016). Urban performance does not only rely on the quality of physical infrastructure, as it is also influenced by availability and reliability of communication of knowledge and infrastructure (Caragliu et al., 2011). Outdoor spaces have been mapped in great detail, providing reliable static spatial information. Location based services, such as Global Navigation Satellite Systems (GNSS) provide navigation and tracking (Rantakokko et al., 2011). Next to this, recent additions of networks of sensors provide dynamic data about the urban environment and are supported in this task by an increasing number of connected devices (Hashem et al., 2016). In part, this raw data is transformed into meaningful information, which creates many opportunities for data driven processes in fields like sustainable urban planning (Geertman et al., 2013), monitoring mobility (Semanjski et al., 2017) and emergency management (Keating, 2016).

In literature, this is described by the concept of 'situation awareness', meaning the predicted status of elements within an operation (Endsley, 2016). Situation awareness was first applied to human factor research related to the domain of aviation. However, the scope of the concept quickly expanded to research fields like air traffic control, military operations, transportation, power systems, law enforcement, health care, space, transportation, education, mining, and oil, gas operations and the focus of this research: emergency management (Endsley, 2015a).

In the case of an emergency, both dynamic and static data sources aid professional first responders in their task to prevent and reduce losses of lives and property in case of an incident (Zlatanova et al., 2004). The data sources enrich the perception of a situation that professional first responders, such as police departments, fire departments and mobile medical personnel (Dilo and Zlatanova, 2011). Situation awareness of an operation is an important factor in first responder decision making (Li et al., 2014). According to Endsley (2016), the extent to which situation awareness is reached depends on the availability and transformation of raw data into insightful information, which drives the quality of decisions and thereby team performance. Therefore, it is argued the (un)availability of insightful data influences first responder team performance, which is why reliable data availability is a key factor in the success of first responder operations.

2.1 Problem statement

Although information sources prove to be beneficial for emergency response operations, spatial data availability is limited in the case of indoor operation environments (van der Meer, 2018). Increased complexity of urban areas due to adoption of high-rise adoption of high-rise in city centers changes the urban fabric and draws attention from outdoor space to indoor space (Tashakkori et al., 2015). As first responders are now getting used to availability of real-time data in outdoor situations, the demand for such data grows for indoor operation environments as well (Rantakokko et al., 2011). There are three main causes for the limited availability of indoor information flows.

First, positioning through GNSS services is unreliable within indoor environments, as buildings tend to block the signals send by GNSS (Rantakokko et al., 2011).

Second, only a few buildings are well mapped, meaning there is limited availability of indoor spatial data. Even if a building is mapped, the information is often limited to (outdated) 2D floor plans of the first floor, which give a poor representation of reality of a building (Staats, 2017; Tashakkori et al., 2015; van der Meer, 2018; Zlatanova et al., 2004).

Third, indoor environments are likely to change, which damages the reliability of the available data (Khoshelham et al., 2019; Rantakokko et al., 2011; Seppänen and Virrantaus, 2015), meaning the extent to which the data represents the real world in a continuous way (Van der Ham, 2015). The dynamic circumstances can introduce chaos in first responder operations, in which different decision makers have a different perception of the situational environment (Kapucu and Garayev, 2011). These limitations force first responders to rely on traditional operation methods for indoor operations with a lower level of situation awareness.

To maintain the same level of situation awareness for indoor operation environments, professional first responders need to adapt to the increasingly complex urban environment (Kurvers, 2019; Tashakkori et al., 2015). 'indoor mapping' and 'indoor tracking' are active fields of research, with the goal to increase situation awareness of indoor operational environments. Indoor mapping provides spatial information of the interior of buildings, for example by providing navigable space, indoor navigation and Building Information Models (BIM) (Bailey and Durrant-Whyte, 2006; Fuentes-Pacheco et al., 2015; Kang et al., 2019; Nikoohemat et al., 2020; Staats, 2017; Zlatanova et al., 2013). Indoor tracking focuses on creating reliable ways of tracking assets and persons within buildings (Rantakokko et al., 2011).

Several methods exist already to create spatial information of indoor environments like LiDAR scans and photogrammetry (Luhmann et al., 2013). These techniques deliver reliable results and some techniques are even mobile and wearable, offering more flexibility (Khoshelham et al., 2019; Rantakokko et al., 2011). These techniques depend on Simultaneous Mapping And Localization (SLAM) (Endres et al., 2012; Karam et al., 2019; van Schouwenburg, 2019).

SLAM systems are also used in first responder context, to create spatial information on an as-need basis (Rantakokko et al., 2011; Tashakkori et al., 2015). However, these methods are often not available in real-time as the raw scanned data is only available after completion of the scan (Luhmann et al., 2013). This unavailability of real-time data generates a large problem for first responders, as the commanding officers need reliable and updated information to be supported by up-to-date situation awareness in a rapidly changing operation environment (Li et al., 2014; Seppänen and Virrantaus, 2015).

2.2 Research objectives

The goal of this thesis is to research to which extent a first responder command-and-control center can be supported with reliable (near) real-time situation awareness of the indoor operation for the purpose of remote decision-making.

A proof of concept will be developed to test the feasibility of real-time collection of spatial data of indoor first responder operations, to test integration of the data elements into a Common Operational Picture (CoP). The spatial data elements will consist of geometric measurements of the environment (mapping) and a list of device poses (position + orientation). All data elements are accompanied with temporal information, namely the time and date of collection.

To create a CoP, all spatial data will be sent from the operation environment to a remote coordinator in real-time, to create a three-dimensional image of the operation environment. Subsequently, this image can be used to build situation awareness and thereby add on existing knowledge, to make better decisions and improve first responder performance.

2.2.1 Central research question

The research question of this thesis is:

To what extent may a head mounted (near) real-time simultaneous localization and mapping system (SLAM) support reliable spatial decision making to increase first responder operation performance in indoor environments by improving situation awareness?

2.2.2 Sub-questions

The central question will be broken down into sub-questions. By breaking down the central question into sub-questions, components of the research can be individually tested on feasibility and added value.

The sub-questions are categorized into four categories:

1. Recognition of the research domain
2. Feasibility to make a proof of concept
3. Testing of the proof of concept
4. Integration of the results

2.2.2.1 Recognizing the research domain

To identify the place of this research in the research domain, a mapping study will be conducted. The preliminary literature review and contact with the supervisors of this thesis have identified the research topic and key concepts in the domain. However, it is unclear how these concepts relate to the research domain. To identify this relation, the following question will drive a mapping study: *To what extent has the relation between 'indoor mapping', 'indoor tracking' and 'first responder decision making' already been established by the academic community?*

2.2.2.2 Feasibility of (near) real-time indoor SLAM system

To research the feasibility a head mounted (near) real-time indoor SLAM system for first responder decision making, a proof of concept of the system will be developed. The proof of concept will be developed with three main components: mapping, tracking, and communicating.

The second sub-question will focus on providing mapping of indoor operation environments by use of a head mounted device. The sub-question is defined as following: *To what extent may a head mounted augmented reality device be used to map first responder indoor operation environments in (near) real time?*

A tracking component will be added to the system to enable officers of duty to track first responder movements within the mapped indoor environment. The sub-question is defined as following: *To what extent may a head mounted augmented reality device be used to track first responders in (near) real time within indoor operation environments?*

To support the decision making process of a first responder coordinator, the system will have to work in (near) real-time, as first responders work in highly dynamic environments (Seppänen and Virrantaus, 2015). Therefore, information has to be simultaneous and up-to-date, or the coordinator will make an outdated decision. Communication is key, as the way in which data is presented to a user creates bias (Endsley, 2016). A way to communicate from the operation environment to a remote coordinator in an interpretable manner will therefore be examined within this sub-question. Therefore, the fourth sub-question will be: *To what extent may a head mounted device be used to communicate in (near) real time about mapping and tracking information of indoor first responder operation environments?*

2.2.2.3 System reliability

First responders rely for safety and performance of their operations on the reliability of their tools (Fischer and Gellersen, 2010). Therefore, the system will be tested on this indicator. Reliability is defined in this research as a construct consisting of the accuracy, precision and robustness of the system. Accuracy is the extent to which measurements represent the real-world situation. Precision is the extent to which the measurements are spread over the operational environment. Robustness is the extent to which the system performs continuously within different operational environments (Van der Ham, 2015). The sub-question will be: *To what extent may the reliability of the mapping, tracking, and communicating capabilities of the proof of concept support first responder decision making for indoor operation environments?*

2.2.2.4 Situation awareness

The sub-questions above research the nature of key aspects, the feasibility of developing a real-time indoor SLAM system and the reliability of such a system. The last sub-question will answer the question how the integrated aspects of the system will work together in providing situation awareness to aid first responder decision making and thereby increase first responder performance in indoor operations. The sub-question is formulated as following: *To what extent may the proof of concept improve first responder indoor performance by providing situation awareness to support first responder decision making?*

Out of scope

Some topics are related to the research, but specifically placed out of scope due to time constraints or other reasons. The topics are listed below.

2.2.2.5 Pre-processing of mapping results

The results of the mapping process will not be processed before pushing the raw results to the officer of duty: only visualisation of the raw data will be considered. This means that clutter in the data will not be removed, nor will the data be segmented by object or classified as particular objects.

2.2.2.6 Indoor localization

As will be explained further in the theoretical framework, there is a large semantic difference between localization and positioning. For this thesis, positioning is regarded as part of SLAM, focusing on the tracking component of the localization part of SLAM. For the proof of concept proposed by this thesis, only the pose (relative x,y,z coordinates + orientation) will be considered for tracking purposes, which is why the system might also be called an (near) real-time indoor Simultaneous Posing And Mapping (SPAM) system. For practical reasons, namely possible confusion with unsolicited message systems, the term SPAM will not be used for the system.

2.2.2.7 Using direct sensor output of the Microsoft Hololens

Since the introduction of 'research mode', the possibility of direct interaction of Microsoft HoloLens sensor output exists. Explorative results indicate the possibility of combining the output of the environment camera's with output of the depth sensors. If the overlay of these sensors is calculated, an RGB-D result could be generated. In this result, a pixel has both RGB values and an angle/distance value. By projecting these values with the local coordinate system, a point cloud could be generated. This point cloud can be communicated to the officer of duty at CoPI and this method offers a lot of flexibility, as the data can be processed outside of the mixed-reality toolkit.

However, using this research mode is complex, as there is few documentation on indoor mapping based on the raw sensor output of the Microsoft HoloLens. The system would have to be developed by using DirectX in combination with C or C++. This is assumed to take too much time. Furthermore, the research mode is only available for research and experimenting developers. The modus cannot be enabled by an application that is downloaded from a database like the windows store. This limits practical application and potential deployment of the system for first responders.

2.2.2.8 Operations where (internet) infrastructure is completely unavailable

The technical question of getting data from the head mounted device to the officer of duty will be out of scope. An operational internet connection (Wi-Fi or a 4G mobile connection) is assumed for operating environments of the proof of concept. Although incidents like fires and floods would damage infrastructure, rendering the real time part of the application useless, there are also a lot of first responder operations for which the incident stays intact. Examples of such situations are gas leaks and hostage situations.

2.2.2.9 Operations where visual view is obstructed

As the SLAM method of the proof of concept depends on the use of cameras capturing visual light, the proof of concept is only expected to work in spaces with unobstructed view. First responder operations in which there is an obstruction of view, such as fires with a lot of smoke generation, will therefore be out of scope.

2.2.2.10 Change detection with known information

The assumption of the proof of concept is to be deployed in an unknown environment. This means there are no previous scans or information sources like BIM models. Change detection with previous scans is therefore out of scope, as there is no information to compare scans with. However, change detection will be within scope for single scans, meaning the 3D model of the environment will be updated as the system runs.

2.2.2.11 Aligning scans with absolute coordinate systems

The proof of concept is based on a SLAM method that defines an arbitrary coordinate system, referenced to the starting position. This arbitrary coordinate system is not be aligned with coordinate systems that have the status of consensus, called absolute coordinate systems (Luhmann et al., 2013). There are advantages of such alignment: one could image integrating outdoor and indoor operating environments in one application. However, this alignment is not within the scope of this thesis. Instead, recommendations for aligning the local 3D model with world space will be given in the discussion part of the thesis.

2.2.2.12 Collaboration of multiple Microsoft HoloLens devices starting from multiple positions

As there is no absolute coordinate system, there is no fixed frame of reference. Therefore, multiple scanning devices cannot combine data if they started at different positions. Therefore, this will be be out of scope. Instead, recommendations for aligning coordinate systems of multiple devices will be given in the discussion part of the thesis.

2.2.2.13 Integration of indoor and outdoor spatial services

An incident location is not never completely isolated, not even if the incident is in an indoor environment. For example, the relation to outdoor water resources or entry/exit points of buildings such as flight paths may be important for emergency management.

The proof of concept will not be capable of estimating its position in a global coordinate system out of the box. There is no built in GPS module and compass. Instead, it does only know where it is by referencing to observed objects and surfaces. Approximation of global position by techniques like outdoor Wi-Fi positioning (Sheng-Cheng Yeh et al., 2009) and visual feature matching (Naseer and Burgard, 2017) are fields of study, but they are still meters away from even GPS reliability. Therefore, they are considered to be too unreliable as they do not match the accuracy of the 3D model of the proof of concept. Spatial data of the indoor operation environment shall therefore be presented in isolation, apart from outdoor information.

3 Literature scope

The research domain of this thesis is a combination of the research fields of simultaneous indoor mapping & indoor tracking, and situation awareness driven decision making in first responder operations. This domain includes a lot of different topics, such as the aforementioned automatic processing of scans into 3D models and Building Information Model (BIM) (Wang et al., 2015), real-time medical asset tracking (Van der Ham, 2015) and decision support systems for first responders (Kapucu and Garayev, 2011).

Indoor mapping is a broad topic. Research bias is inevitable due to the impossibility of exploration of all relevant literature. Therefore, choices must be made regarding the literature to be studied. These choices must be documented to create insight in the bias created by the choices. This idea is not new, as for example Paré et al. (2015) and Kitchenham et al. (2010) describe the problem of research bias in detail. Furthermore, a mapping study has value for the academic field (Endsley, 2015b), as related domains are more easily connected if the same concepts and keywords are used.

The goal of mapping studies is to "*provide an overview of a topic area and identify whether there are sub-topics with sufficient primary studies to conduct conventional Systematic Literature Review (SLRs) and also to identify sub-topics where more primary studies are needed*" (Kitchenham et al., 2011). Mapping studies are useful to software engineering researchers. The search strategy is to provide an overview of the domain and the fields of study, including the research gap, by classifying primary research papers in the domain. The SLR goes beyond this mapping study, as it is answers specific questions based on literature, instead of being topic driven (Kitchenham et al., 2011). Although conducting an SLR after the mapping study might improve the practical usage of the mapping study, it would take too much time and is therefore regarded out-of-scope.

First, an example of the importance of using specific keywords in relation to specific domain is given. Second, the methodology of the mapping study will be illustrated. Third, the results of the mapping study will be presented. Fourth, a conclusion will be deduced from the results, thereby answering the first sub-question: *To what extent has the relation between 'indoor mapping', 'indoor tracking' and 'first responder decision making' already been established by the academic community?*

3.1 Identification of concept and keyword use

Within the context of this research one could distinguish both 'situation' and 'situational' awareness. Both terms are used interchangeably in literature. Nevertheless, the following queries executed in Google Scholar give the following results:

"first responder" <u>AND</u> "situational awareness" :	**4160 results**
"first responder" <u>AND</u> "situation awareness" :	**1420 results**
"first responder" <u>AND</u> "situational awareness" <u>OR</u> "situation awareness":	**4840 results**

We can see that 'situation awareness' returns fewer results compared to 'situational awareness' in combination with 'first responder'. Because of this, we know the term 'first responder' is related to a larger number of publications if used in combination with the term 'situational awareness'. Although this research refers to 'situation awareness' as it refers directly to the 1995 Situation Awareness model of Endsley, we would not want to miss out on 'situational awareness' papers. Therefore, using the third query gives a more complete overview for the first responder domain, as it combines the results of the first two queries. From the combined results, we can see that there are 4840 – 4160 = 680 overlapping results. This overlap contains 47.9 percent of the 'situation awareness' results, meaning that these papers can be found by either using the first or the second query.

1. **"first responder" <u>AND</u> "situational awareness"**

Time	**Results absolute**	**results percentage**
All time:	4.160	100
2015 – 2020:	1.800	43.2
2010 – 2015:	1.280	30.8
Before 2010:	998	24

2. **"first responder" <u>AND</u> "situation awareness"**

Time	**Results absolute**	**results percentage**
All time:	1420	100
2015 – 2020:	659	46.0
2010 – 2014:	430	30.3
Before 2010:	314	22

3.2 Methodology mapping study

To establish the place of this research in the research domain, a systematic method for the literature review will be used. As the thesis relies heavily on the development of a proof of concept to test the added value of a real-time indoor SLAM system for first responder decision making, the literature study will follow the concept of Evidence-Based Software Engineering (EBSE) and uses a 'mapping study' as a methodology (Kitchenham et al., 2011).

According to Kitchenham et al. (2011), the first four phases in the ESBE framework that will be followed for the mapping study:

- Construction of a question
- Identifying evidence to answer the question
- Carefully and systematically appraising the evidence
- Using the evidence to answer the question

A preliminary research identification has been completed for the extended research proposal. This identification has led to the sub-question: *To what extent has the relation between 'indoor mapping', 'indoor tracking' and 'first responder decision making' already been established by the academic community?*

This question will be answered by means of a systematic search strategy, providing evidence to answer the question. The result of the search strategy will be a set of papers that are related to the research domain of this thesis. The frequency of occurrence per research field will be documented. This documentation will be used to analyse the frequency combinations of the research fields to gain insight in the relations of the research field and the size of the research domain. More important however, is the storage of returned papers for further inspection.

To create insight in the connection of important concepts to fields of study, the API of the academic search engine 'Scopus' will be used to combine primary and secondary keywords. A frequency matrix will be created for the (combinations of) primary and the secondary keywords to indicate how often they occur either alone or in pairs of two in academic publications. The papers are subsequently sorted to relevance, and the first 25 listed papers are stored for easy retrieval.

3.2.1 Primary keywords

First, primary keywords are identified to describe associations between fields of research. The primary keywords documented in the preliminary literature review for the research identification will be used for this purpose. A primary keyword is a keyword that has been documented directly from observed literature. These primary keywords can be used to explore the domain in greater detail and are used for the construction of the central research question and the sub-questions. The identified keywords are shown below in Figure 1.

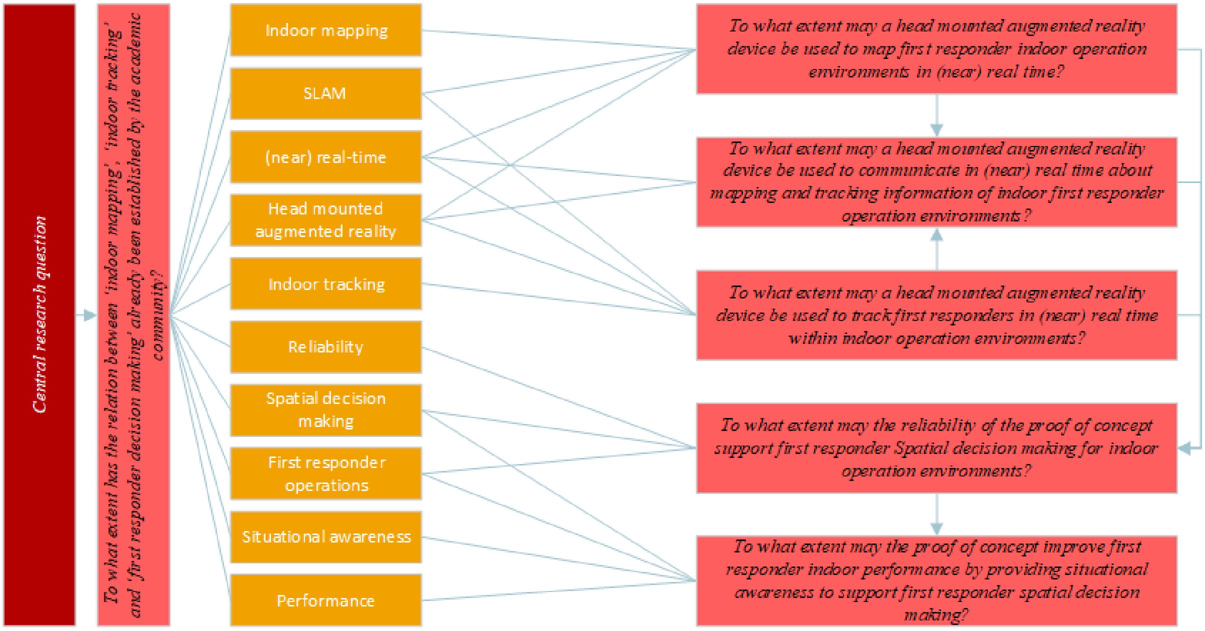


Figure 1: relation between primary keywords and research (sub) questions

3.2.2 Secondary keywords

However, as the primary keywords are arbitrarily documented, care should be taken to ensure unnecessary bias. Different fields of research, and especially different domains, use different keywords to describe the same concept. This is illustrated by Endsley (2011), who describes that the concept of 'situation awareness' is identified differently across domains. This concept has roots in the domains of aviation and cognitive psychology and is primarily used in these domains, while 'situational understanding' is used to describe the same concept within the domain of military operations. Synonyms describing the same concepts as the primary keywords will be documented while reviewing literature and are called secondary keywords. These synonyms are evaluated on their meaning in a qualitative way.

3.3 Results mapping study

Preliminary research resulted in a list of 32 keywords related to the research. These keywords are identified qualitatively. Furthermore, the keywords can be independent or related to other keywords, such as 'situation awareness' and 'situational awareness'. Subsequently, the Scopus database has been queried by using the API and pairs of keywords. The number of returned papers is stored in a two-dimensional matrix. Besides this frequency matrix, an overview of related literature is composed by the association mapper application. This overview stores a maximum of 25 related papers with metadata per set of keywords, such as 'First Responder' AND 'Situation Awareness'. The papers are filtered on relevancy by Scopus. As the literature is categorized to subject in an excel worksheet, the dataset can be used for browsing through subjects and related papers. The literature mapping study therefore formed an important input for creating an overview of the research domain for the theoretic background.

3.3.1 Mapped associations

Associations between keywords are mapped in a quantitative manner. This means that the number of paper results for a given keyword pair is registered. From the frequency matrix below, we can identify research fields that have a large or a low number of publications. Although it mainly serves as a reference, some patterns can be observed. For example, all aspects to do with reliability (evaluation, accuracy, precision, and robustness) are on the right and are very frequently found in combination with all keywords. This makes sense, as evaluation and measurements are assumed to be associated with the scientific field as a whole. Furthermore, the subjects 'mapping' and 'tracking' are marked in green as well, meaning these topics are frequently researched as well. HoloLens', 'reality eyewear' and 'head mounted' show up in red in a vertical direction, meaning these terms are not often used in combination with the other keywords. So do 'situational understanding' and 'depth-map'.

	indoor	head mounted	hololens	reality eyewear	simultaneous localiz	SLAM	first responder*	emergency responde	disaster management	emergency manage	situation awareness	situational awarenes	situational understan	information overload	decision making	mapping
indoor	425648															
head mounted	1051	18267														
hololens	186	578	1851													
reality eyewear	3	7	4	19												
simultaneous localization and m	5894	136	55	0	13940											
SLAM	10438	268	95	3	10573	49887										
first responder*	1123	74	8	0	91	172	12785									
emergency responder*	474	33	1	0	48	69	726	4003								
disaster management	1195	44	6	0	40	175	1007	483	41858							
emergency management	844	53	6	0	13	112	1248	683	8372	50091						
situation awareness	760	394	33	0	69	133	268	170	844	1165	24603					
situational awareness	598	216	30	0	86	212	354	276	939	1192	4473	19266				
situational understanding	9	3	1	0	0	1	7	2	6	22	112	74	503			
information overload	86	35	3	0	3	20	33	20	93	218	332	224	6	15048		
decision making	11538	912	93	0	380	2049	1428	519	8125	10336	8300	4037	163	4488	1605794	
mapping	36629	2172	296	3	13940	21251	1082	343	9124	6102	2442	2264	60	1309	135676	2898370
tracking	54251	5873	547	8	7301	14637	1334	468	2387	1926	3898	3903	100	1044	55443	157154
depth*map	48	1	0	0	10	18	0	0	3	1	1	0	0	0	44	138
point cloud*	5701	226	87	1	1580	3799	46	11	398	114	76	160	5	8	1166	17501
RGB-D	6751	290	85	2	2101	5002	28	7	44	20	80	67	2	4	405	7739
LiDAR	4807	94	34	1	1645	3423	109	29	1113	523	165	292	3	7	3345	40390
Photogrammetry	5287	254	71	1	925	2251	122	61	2647	880	208	301	5	36	5710	65976
positioning	40887	899	116	3	3667	6503	968	400	1848	1192	877	962	18	212	22262	73003
localization	53118	1629	239	4	13940	21012	1050	402	1147	966	1138	1120	28	159	35756	411525
depth sensing	262	93	25	0	32	101	2	1	0	0	5	2	0	1	54	1090
real*time	2773	444	42	0	551	1283	73	39	251	238	278	332	6	51	2500	7879
reliability	22166	1129	90	1	602	2481	1173	393	3767	5986	4382	2386	40	1012	158404	137923
evaluation	155209	7221	721	10	3733	13381	5102	1493	14809	20576	11623	7078	210	5914	639943	787656
performance	203135	8365	664	7	5325	15289	4268	1304	10730	13972	14715	8548	240	6047	531263	679285
accuracy	50257	3043	342	3	4253	9413	1487	469	3207	3473	2682	2384	37	1806	151514	319263
precision	17311	1464	150	2	1651	3767	332	140	1015	760	868	833	9	436	37242	125471
robust	47596	1891	201	7	7513	15316	896	288	2795	2233	2036	2032	54	613	82276	207011

	tracking	depth*map	point cloud*	RGB-D	LiDAR	Photogrammetry	positioning	localization	depth sensing	real*time	reliability	evaluation	performance	accuracy	precision	robust
tracking	1359116															
depth*map	77	658														
point cloud*	8689	30	45485													
RGB-D	9455	16	4701	19251												
LiDAR	13068	9	14863	1837	124792											
Photogrammetry	15261	33	16518	1577	25581	139302										
positioning	90785	33	3354	1668	6455	10353	553302									
localization	140543	37	6801	6500	8224	7497	90403	2704021								
depth sensing	797	3	146	274	89	108	163	515	9570							
real*time	14630	15	1025	1369	802	908	2799	5256	110	71245						
reliability	82093	18	2192	834	4726	7781	36461	66071	681	4872	2249215					
evaluation	393546	223	15397	7856	41852	53231	152020	553724	3232	23107	937170	1.3E+07				
performance	533981	175	14448	8455	38022	40129	170689	387327	2687	29997	833705	3708350	1.1E+07			
accuracy	220466	87	16551	6677	33464	55666	120735	198950	1036	11623	309821	1214736	1047748	3143185		
precision	102009	17	7523	2006	13693	20172	69687	69354	667	4250	90276	384920	432405	376767	1356185	
robust	282361	104	15844	10243	16214	21179	60915	155623	781	12135	143280	536520	821915	280099	120633	2023459

Figure 2: Absolute association between keywords. Red: percentile 0-30. Yellow: percentile 30-80. Green: percentile 80-100

In the figure above, we can see which research fields are well described in literature. This is an interesting overview, which especially allows for identification of research areas that are less well covered. For example, 'LiDAR' and 'photogrammetry' have over 100.000 related papers, while RGB-D has less than 20.000. We can also distinguish that 'first responder' is more often used in literature compared to 'emergency responder', with over 10.000 findings.

Although the absolute associations are interesting for identifying research gaps, we are also interested in the relative associations: how often do keywords appear together? This is calculated by the percentage of overlap between the returned results of a keyword pair (for example 'Mapping' and 'Localization') and the returned results of one of the keywords (for example 'Mapping'). In this specific case, this would be 411.525/2898370*100 = 14.199%.

The percentage in itself does not say a lot. However, some patterns can be observed. For example, of all publications mentioning 'first responders' in the title, abstract or keywords, 8.46 percent is related to ‘mapping’. Compared to the mapping results, the association between ‘tracking’ and ‘first responders’ is a little higher with 10.43 percent. Based on this finding, in combination with the finding of 12785 papers for ‘first responders’ alone, we say there is indeed interest in the domain of mapping and tracking applications for first responders. We do also see there that ‘indoor’ and ‘first responders’ only overlap for 0.26 percent, which is rather low. This corroborates the idea that there is little research done at first responders in indoor environments, let alone in combination with mapping and tracking.

	indoor	head mounted	hololens	reality eyewear	simultaneous localization ai	SLAM	first responder*	emergency responder*	disaster management	emergency management	situation awareness	situational awareness	situational understanding	information overload	decision making	mapping
indoor	100.00															
head mounted	0.25	100.00														
hololens	0.04	3.16	100.00													
reality eyewear	0.00	0.04	0.22	100.00												
simultaneous localization a	1.38	0.74	2.97	0.00	100.00											
SLAM	2.45	1.47	5.13	15.79	75.85	100.00										
first responder*	0.26	0.41	0.43	0.00	0.65	0.34	100.00									
emergency responder*	0.11	0.18	0.05	0.00	0.34	0.14	5.68	100.00								
disaster management	0.28	0.24	0.32	0.00	0.29	0.35	7.88	12.07	100.00							
emergency management	0.20	0.29	0.32	0.00	0.09	0.22	9.76	17.06	20.00	100.00						
situation awareness	0.18	2.16	1.78	0.00	0.49	0.27	2.10	4.25	2.02	2.33	100.00					
situational awareness	0.14	1.18	1.62	0.00	0.62	0.42	2.77	6.89	2.24	2.38	18.18	100.00				
situational understanding	0.00	0.02	0.05	0.00	0.00	0.00	0.05	0.05	0.01	0.04	0.46	0.38	100.00			
information overload	0.02	0.19	0.16	0.00	0.02	0.04	0.26	0.50	0.22	0.44	1.35	1.16	1.19	100.00		
decision making	2.71	4.99	5.02	0.00	2.73	4.11	11.17	12.97	19.41	20.63	33.74	20.95	32.41	29.82	100.00	
mapping	8.61	11.89	15.99	15.79	100.00	42.60	8.46	8.57	21.80	12.18	9.93	11.75	11.93	8.70	8.45	100.00
tracking	12.75	32.15	29.55	42.11	52.37	29.34	10.43	11.69	5.70	3.85	15.84	20.26	19.88	6.94	3.45	5.42
depth*map	0.01	0.01	0.00	0.00	0.07	0.04	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
point cloud*	1.34	1.24	4.70	5.26	11.33	7.62	0.36	0.27	0.95	0.23	0.31	0.83	0.99	0.05	0.07	0.60
RGB-D	1.59	1.59	4.59	10.53	15.07	10.03	0.22	0.17	0.11	0.04	0.33	0.35	0.40	0.03	0.03	0.27
LiDAR	1.13	0.51	1.84	5.26	11.80	6.86	0.85	0.72	2.66	1.04	0.67	1.52	0.60	0.05	0.21	1.39
Photogrammetry	1.24	1.39	3.84	5.26	6.64	4.51	0.95	1.52	6.32	1.76	0.85	1.56	0.99	0.24	0.36	2.28
positioning	9.61	4.92	6.27	15.79	26.31	13.04	7.57	9.99	4.41	2.38	3.56	4.99	3.58	1.41	1.39	2.52
localization	12.48	8.92	12.91	21.05	100.00	42.12	8.21	10.04	2.74	1.93	4.63	5.81	5.57	1.06	2.23	14.20
depth sensing	0.06	0.51	1.35	0.00	0.23	0.20	0.02	0.02	0.00	0.00	0.02	0.01	0.00	0.01	0.00	0.04
real*time	0.65	2.43	2.27	0.00	3.95	2.57	0.57	0.97	0.60	0.48	1.13	1.72	1.19	0.34	0.16	0.27
reliability	5.21	6.18	4.86	5.26	4.32	4.97	9.17	9.82	9.00	11.95	17.81	12.38	7.95	6.73	9.86	4.76
evaluation	36.46	39.53	38.95	52.63	26.78	26.82	39.91	37.30	35.38	41.08	47.24	36.74	41.75	39.30	39.85	27.18
performance	47.72	45.79	35.87	36.84	38.20	30.65	33.38	32.58	25.63	27.89	59.81	44.37	47.71	40.18	33.08	23.44
accuracy	11.81	16.66	18.48	15.79	30.51	18.87	11.63	11.72	7.66	6.93	10.90	12.37	7.36	12.00	9.44	11.02
precision	4.07	8.01	8.10	10.53	11.84	7.55	2.60	3.50	2.42	1.52	3.53	4.32	1.79	2.90	2.32	4.33
robust	11.18	10.35	10.86	36.84	53.90	30.70	7.01	7.19	6.68	4.46	8.28	10.55	10.74	4.07	5.12	7.14
Legend	0%	5	10	15	20	25	30+	percents								

	tracking	depth*map	point cloud*	RGB-D	LiDAR	Photogrammetry	positioning	localization	depth sensing	real*time	reliability	evaluation	performance	accuracy	precision	robust
tracking	100.00															
depth*map	0.01	100.00														
point cloud*	0.64	4.56	100.00													
RGB-D	0.70	2.43	10.34	100.00												
LiDAR	0.96	1.37	32.68	9.54	100.00											
Photogrammetry	1.12	5.02	36.32	8.19	20.50	100.00										
positioning	6.68	5.02	7.37	8.66	5.17	7.43	100.00									
localization	10.34	5.62	14.95	33.76	6.59	5.38	16.34	100.00								
depth sensing	0.06	0.46	0.32	1.42	0.07	0.08	0.03	0.02	100.00							
real*time	1.08	2.28	2.25	7.11	0.64	0.65	0.51	0.19	1.15	100.00						
reliability	6.04	2.74	4.82	4.33	3.79	5.59	6.59	2.44	7.12	6.84	100.00					
evaluation	28.96	33.89	33.85	40.81	33.54	38.21	27.48	20.48	33.77	32.43	41.67	100.00				
performance	39.29	26.60	31.76	43.92	30.47	28.81	30.85	14.32	28.08	42.10	37.07	29.26	100.00			
accuracy	16.22	13.22	36.39	34.68	26.82	39.96	21.82	7.36	10.83	16.31	13.77	9.58	9.19	100.00		
precision	7.51	2.58	16.54	10.42	10.97	14.48	12.59	2.56	6.97	5.97	4.01	3.04	3.79	11.99	100.00	
robust	20.78	15.81	34.83	53.21	12.99	15.20	11.01	5.76	8.16	17.03	6.37	4.23	7.21	8.91	8.90	100.00

Figure 3: Association between keywords. The legend is given at the bottom of the table.

3.3.2 Trend analysis

To get a better view of the development of the research domains, we mapped the associations over time. We assume we can display research trends by employing this approach. An example of this is the association between the ‘mapping’ keywords and ‘LiDAR’, ‘photogrammetry’ and ‘RGB-D’: if the association changes over time, it is assumed we can distinguish trend changes in research focus within the indoor mapping domain. This association can be seen in Figure 4.

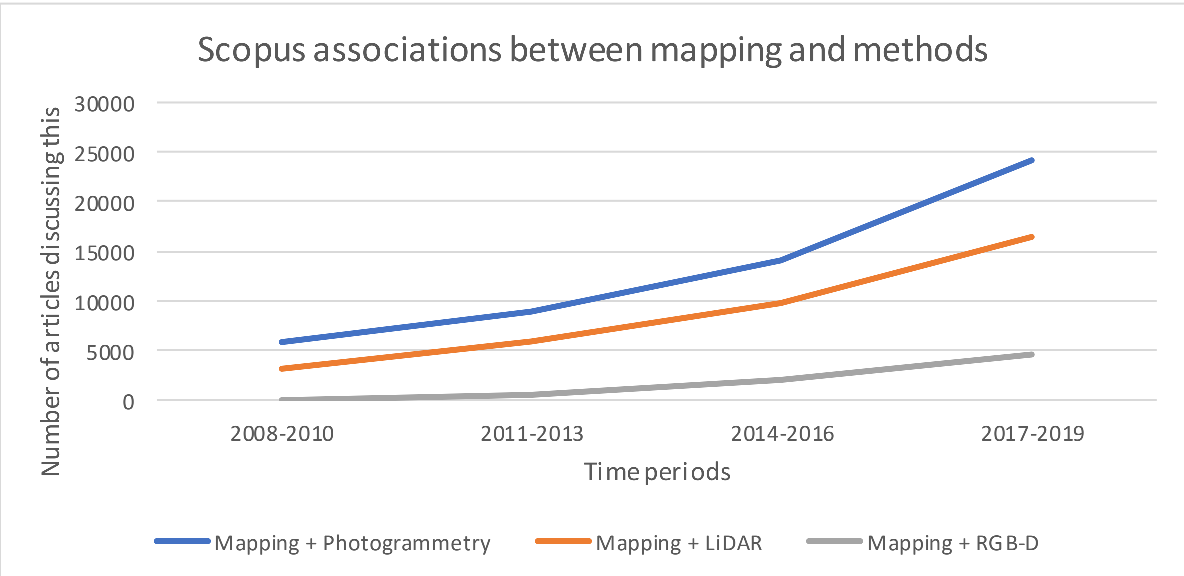


Figure 4:Scopus associations between mapping and scan methods

Graphs such as seen in Figure 4 can be made for every mapped combination of keywords. However, there is a catch. From the raw data it appears that every keyword has more records in the Scopus results compared to previous years. A reason for this has not been found, but it thought that Scopus has made more connections to other research groups. This is thought to create bias in the results. Furthermore, there are very large differences with other academic search engines such as Google Scholar. To compare the registered Scopus associations with scholar associations, the parameters of the Scopus search are used for a similar, manual search in Google Scholar. Below, the approximate results are shown.

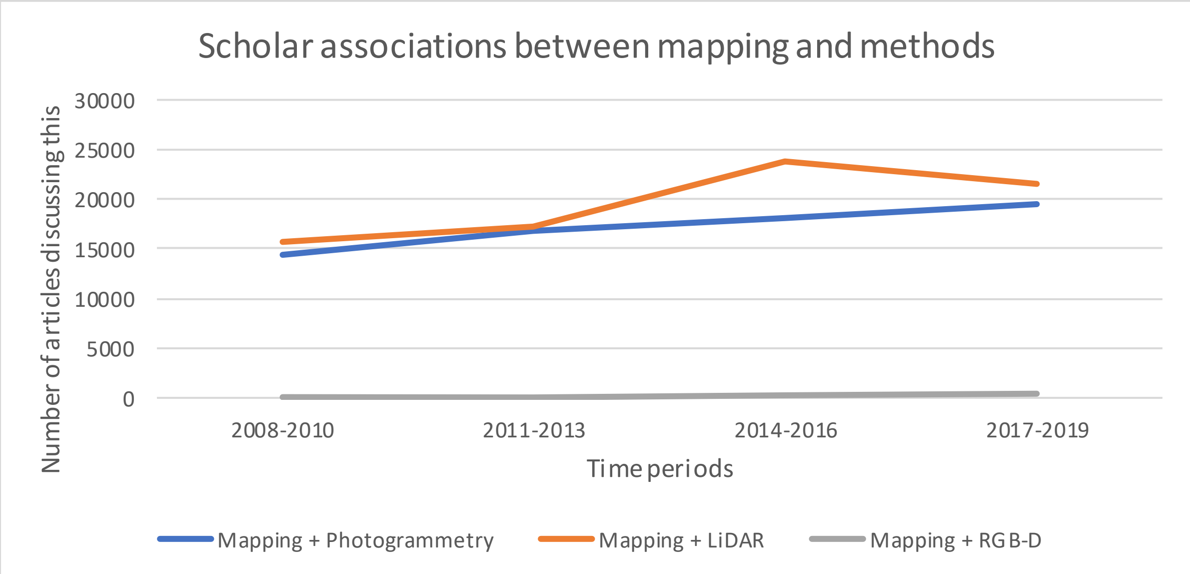


Figure 5: Scholar associations between mapping and scan methods

As can be seen, there is still an increase in results over time. However, the results are very different from the Scopus results. For now, it cannot be said which results are better. Therefore, more research is needed to give a definite answer about research trends, for example by combining multiple search engines as described in (Endsley, 2019).

4 Theoretic background

This section will provide a foundation of academic literature, giving academic context to the central research question. Part of the relevant body of knowledge will be presented here serving two main goals. First, to explore the presence of the research gap. Second, to provide a justification for the research methodology.

A proof of concept will be developed to provide answers to the central research question. Because of this, a lot of time within this thesis will go to development of the real-time SLAM proof of concept. A 'technology-centered design' perspective is therefore lurking. Traditionally, a lot of systems were developed in a technology-centered design perspective, meaning that systems are built from a perspective of what is possible (Endsley, 2016). Although such systems can be impressive, it does not necessarily mean that the system does also provide a solution for what is needed by the user. Because of this mismatch, the added value of a system might get lost. Therefore, many engineering domains have switched from a technical perspective to a user-centric based perspective (Endsley, 2016; Hennig and Belgiu, 2012).

The literature study itself gives context to three topics: (1) the context of situation awareness, (2) the way in which first responders use data in their operations to make better decisions, and (3) collection and visualization of indoor spatial data elements. Insights from the literature study will be used to create a system design that does not only provide a system for indoor mapping and tracking of first responders in indoor environments, but does also facilitate the workflow of first responders in a meaningful way.

4.1 Situation awareness

Emergency management are exemplary in the way it is complex, urgent and uncertain (Kapucu and Garayev, 2011). First, the process is complex, as many organizations are involved. Second, the process is urgent, as most damage is done in the first moments of an emergency (Dilo and Zlatanova, 2011). Third, the process is uncertain, as situational conditions tend to change quickly (Dilo and Zlatanova, 2011; Kapucu and Garayev, 2011). This appearance of uncertainty is even worse for indoor operations. Static information of the building, such as floor plans, are often outdated or non-existent (Tashakkori et al., 2015). Dynamic operational data such as positions of first responders within the building and situational data such as blocked pathways are available in a very limited way as well (Tashakkori et al., 2015).

Emergency management entails a continuous cycle of activities, which can be distinguished in the phases 'preparation', 'response', 'recovery' and 'mitigation'. First responders use these phases to prevent emergencies from happening (mitigation), prepare for the occurrence of an incident (preparation), react to an occurring incident (response), and providing aid after an incident has occurred (recovery) (Keating, 2016). This research focuses on the response phase of first responder operations, which has a reactive nature. An incident of some sort happens, triggering a response of first responder organizations. Quickly, an overview of the situation needs to be created, for example by assessing the scale and nature of the incident.

For complex and dynamic situations, decision makers rely on the constantly evolving picture of the state of an environment, also called 'situation awareness' (Endsley, 2016). Dilo and Zlatanova (2011) identify 'situation awareness' to be a key factor within the context of information use within emergency response operations. Other authors from various domains within the academic community do also state that the availability of situation awareness can make or break first responder operations (Heard et al., 2014; Li et al., 2014; Seppänen and Virrantaus, 2015; Tashakkori et al., 2015), making it a main influencer for first responder performance. However, situation awareness and its role in first responder operations is still undefined.

As the hypothesis exists that an indoor slam may support decision making of first responder officers of duty by creating (shared) situation awareness, the next section will explore the concept of situation awareness and define its role within first responder decision making.

4.1.1 First responder performance

Situation awareness is goal oriented, as the importance of elements that describe an environment in which decision makers operate rely on the goals that the decision maker aims to complete (Endsley, 2016). Therefore, to understand first responder operations, the core or the 'why' of first responder operations should be discussed first. Performance is understood as accountability for a measurement of success in executing a designated task, as derived from the Oxford dictionary. To apply this definition to first responders is not straight forward due to the complexity of the first responder operational environment: As has been defined above, first responders are heterogeneous groups, each group designated with their own focus and tasks. However, all first responder organizations are designated with the tasks to save human lives and reduce public and private property damages (Liu et al., 2014; Prati and Pietrantoni, 2010; Zhang et al., 2018).'

4.1.2 Data driven decision making

Although situation awareness is a relatively new concept in the domain of cognition that was first introduced by (Endsley, 1988), the idea that situation awareness describes is not new. According to (Flach, 2015), a junction between the state of mind (awareness) and the state of matter (situations) has been made by by describing situation awareness. This means that humans tend to base their decisions on knowledge of matter: a mental model of what is going on and what their previous experiences with such a situation were (Endsley, 2016). Therefore, (Endsley, 2016) argues that if operators are able to achieve a higher level of situation awareness, better decisions can be made that lead to better system performance (*Figure 6*).

Figure 6: Situation awareness as a driver for decision making, leading to better performance (Endsley, 2011)

4.1.3 The nature of situation awareness

Now we have explored the importance of situation awareness for decision making and operation performance, we will define what situation awareness means. The concept of situation awareness, sometimes also named situational understanding (Endsley, 2016), situational assessment (Klein, 1993) or sensemaking (Weick, 1995), is used in many domains with a human factors perspective, such as aviation, education and medicine (Endsley, 2016, 1988; Vassallo et al., 2017).

Mica R. Endsley is regarded as an expert on situation awareness with over two hundred scientific publications about this topic. She coined the concept (Endsley, 1988) and describes situation awareness as being aware of what is happening around a person, while using that information for gaining an understanding of what that information means to the person in the present or in the future (Endsley, 2016). Next to this, situation awareness is goal-oriented, meaning the awareness can be described in the added value of the information for a specific goal or operation. This thesis adopts the formal definition of situation awareness as stated by Endsley (1988), being:

'The perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future'.

Although Endsley states this formal definition of situation awareness in the context of military aviation back in 1988, she argues the definition is still up-to-date and applicable in many domains. This is built by the argument that individual elements of the situation awareness may differ from domain to domain, but the essence of using situation awareness as the basis of decision-making remains (Endsley, 2016). Seppänen and Virrantaus (2015) add on this by stating that the elements can be regarded as data and information. The difference between data and information is herein stated by following English (1999), in which data can be described as the raw input product, while information is a final product. Within this thesis, this difference is understood as the conversion from descriptive, raw data of a situation to insight about that situation.

The use of the definition of situation awareness within the context of first responder operations is illustrated for example by Dilo and Zlatanova (2011), Heard et al. (2014); Li et al. (2014), and Seppänen and Virrantaus (2015). Seppänen and Virrantaus (2015) state that in disaster management, situation awareness is commonly understood as availability of situational information, such as evidence of what events are happening and where these events take place. As can readily be derived from this description, the information is both spatially and temporally oriented (Seppänen and Virrantaus, 2015).

4.1.4 Sharing Situation Awareness

Furthermore, Endsley (2011) makes a distinction between situation awareness (SA) and shared situation awareness (SSA). This definition is stated as:

The degree to which team members have the same SA on Shared SA requirements

In this definition, the idea of shared situation awareness can be understood as situation awareness that is shared between team members. Our objective is to share spatial data elements from indoor first responder operation environments with a remote coordinator to create situation awareness. This involves team members with different roles: someone who explores the situation (called 'explorer') and a team member who coordinates the situation (called 'navigator') (van der Meer, 2018). Although this objective might seem similar to creating shared SA, we want to stress that it is not the same process. The definition of shared SA specifically names the degree to which team members have the same SA. The proof of concept proposed in this paper does only send the data elements from the indoor environment to the remote coordinator and builds situation awareness only at the receiving side of the system.

Although the topic of the situation awareness is the operation environment for both sides of the system, we do not intend to build situation awareness for both team members. Therefore, we introduce a new term by stating the proof of concept creates 'remote situation awareness' instead of 'shared situation awareness', as we create situation awareness in real-time at a remote location without necessarily building situation awareness within the operation environment itself. Although not within scope of this research, remote situation awareness does also allow for robot explorers as illustrated by (Chellali and Baizid, 2011).

4.1.5 Levels of situation awareness

Having situation awareness of an environment means understanding what is going on in that environment (Endsley, 2016). However, such understandings can be very basic or very advanced. As a car driver, basic situation awareness of what the car is doing is sufficient to succeed in the goal of driving around. To know the stiffness of the suspension of the car is unnecessary to reach this goal. As a formula 1 driver however, one needs to know exactly what is going on within the vehicle as the goal is different: not only driving around but actually driving with the best performance. Different levels of situation awareness are therefore not a bad thing, but depend on the goal of the operator.

To understand to which extent situation awareness can be reached, three levels of situation awareness as defined by Endsley (2016) will be shortly described. The first level is the lowest level of situation awareness, while the third level describes the highest extent to which situation awareness can be reached.

4.1.5.1 Level 1: perception

Basic understanding of a situation relies on the status, attributes, and availability of relevant data elements in the environment. The requirements can be perceived by a multitude of senses, for example visual, auditory, or tactile. A flight controller may depend on displays showing him information about the airplanes near an airfield, while a heart surgeon will rely much more on his tactile senses. Both verbal and non-verbal communication are also a part of the information sources.

For the extent to which situation awareness can be built, reliability of the information sources are of the essence. The confidence that an operator has in the different information inputs plays a large part in the trust that he has for depending on the data, which is assessed by personal experience of working with data sources and defined metadata specifications. Incomplete or unreliable data sources do therefore hurt the creation of situation awareness. In many domains, the collection of the data to reach level one situation awareness is therefore already challenging (Endsley, 2016).

4.1.5.2 Level 2: comprehension

The second level entails transformation of raw data into insightful information. This level can only be reached if the 'perception' elements are available, and therefore the second level can only be reached if the first is already reached. Data elements that are necessary for level one 'perception' may have a separated nature. An aspect of level two situation awareness is to join these data elements together into new insights. An example within the scope of this research is the combined visualization of 'mapping' and 'tracking' data elements.

Besides system design, operator experience is a large influencer for the extent to which separated data elements may be integrated into insightful information. A mental model of the situation is important, which is easier developed by experienced operators. To this mental model, new information can be added into a comprehensive overview of the situation (Endsley, 2016).

4.1.5.3 Level 3: projection

Creating a comprehensive overview of a situation is already a challenging objective. The third level takes this process one step further, by stating that not only current, but also future states of the situation should be able to be perceived and comprehended. Therefore, it is necessary that the system is enabled for temporal component and use these components to project to future states of the situation. Such a projection requires a highly developed mental model of the situation and requires significant mental resources.

Although the requirements for level three situation awareness are high, the benefits are large as well. By projecting current states into future states, system operators are enabled not only to react to ongoing events but also to prepare for future potential problems. Due to the high mental cost of projecting current states into future states, information overload should be limited to a minimum by the system. Furthermore, level three situation awareness requires training and experience (Endsley, 2016).

4.1.6 Testing for situation awareness

The central research question tests the extent to which a head mounted (near) real-time simultaneous localization and mapping system may support reliable spatial decision making to increase first responder operation performance in indoor environments by improving situation awareness. As we wonder to what extent situation awareness can be increased to increase performance, we have to be able to measure situation awareness.

Critics on the 1995 model on situation awareness (Endsley, 1995) are often placed in the context of measuring situation awareness (Endsley, 2015b). Endsley has published a systematic review in which several ways to measure the extent to which situation awareness is created by using certain systems (Endsley, 2019). The paper has a broad scope in the domains that are addressed and not specifically for testing situation awareness for first responder operations. However, the broad nature of the paper makes it possible to weigh several aspects of the methods that are used to measure situation awareness and apply which testing methods can be used for this research. This section will describe the publication from a first responder perspective. Of course, the publication itself can be consulted for the full review.

4.1.6.1 Process testing

Many authors have used observation of processes to measure situation awareness. This method uses a behavioral approach by observing reactions of test participants to stimuli aimed on creating situation awareness. Reactions are observed for example by eye tracking, communication measurements and physiological measurements while a stimuli may be a certain state of a situation. The goal is to capture attention of a test participant of certain aspects of the situation and to measure response times and error (Endsley, 2019).

By observing the decisions made and the behavior displayed by test participants, an understanding of how situation awareness is created can be developed by the researcher. The measurements can be boiled down into conclusions of the level of situation awareness. However, the results can only be used to indirectly infer the quality and completeness of the created situation awareness (Endsley, 2015b). Furthermore, behavior is not only explained as a reaction on given stimuli: response is guided by a combination of the present stimuli and earlier, external, experiences. This means that two people who are tested in the same environment and who are exposed to the same stimuli are likely to arrive at a different extent of situation awareness. For emergency events these consequences are even worse, as they are very different from normal events in terms of workload and stress (Wickens, 2000). Therefore, one should always remember that inferences about situation awareness are highly constrained by the testing environments that are used (Endsley, 2019).

She concludes that process measurement may provide insight in the creation of situation awareness. However, this insight is only partial in the way in which it describes the information that is attended to and how the information is processed into a certain level of situation awareness. Therefore, Endsley (2019) claims isolated process measurements are inadequate for measuring situation awareness clearly and objectively.

4.1.6.2 Direct testing

Although process measures may give insight in the way in which situation awareness is developed, it does not tell us precisely to what extent situation awareness is created. Besides process measurements, there are direct measurements. The goal of taking direct measurements is to determine the situation awareness as a state of knowledge (Endsley, 2019). These measurements can be taken either subjective or objective.

Subjective assessment is easy, as researchers can ask participants about their mental model of the situation. This easy determination comes with a catch: participants may not be aware of the knowledge that they do not possess. Therefore, it is reasoned that instead of situation awareness levels, subjective assessments are rather measuring confidence levels.

To counter subjective measurements, a researcher can ask test participants about the state of the environment. The answers to these questions can be objectively scored into being false or true, ruling out subjective answers (Endsley, 2019). The results of these questions, of which an example question list is placed below, can be used to categorize the situation awareness into one of the three situation awareness levels (perception, comprehension, projection). Nevertheless Endsley (2015) agrees with Wickens (2000) that level three situation awareness testing is still underdeveloped compared to level one and level two tests.

Level 1 SA
- Mark all aircraft on the attached sector map.
  - (Completed map provided for all subsequent questions)
- What is the airspeed of the indicated aircraft?
- What is the heading of the indicated aircraft?
- What is the type of the indicated aircraft?
- Is the indicated aircraft currently level, climbing, or descending?
- Which aircraft are currently experiencing an emergency?

Level 2 SA
- Which aircraft have been issued assignments (clearances) that have not yet been completed?
- Which aircraft are not conforming to their clearance?
- Which aircraft are not in communication with you?
- Which aircraft are currently being affected by weather?
- Which aircraft will violate special airspace separation standards if they stay on their path?

Level 3 SA
- What is the next sector of the indicated aircraft?
- Which pairs of aircraft will lose separation if they stay on their current (assigned) course?
- Which aircraft must be handed off to another sector/facility in the next 2 min?

Figure 7: Example from Endsley (2019) of questions used in a direct objective testing method

4.2 First responder context

First responders, or emergency responders, are defined as the organizations and individuals who are responsible for protection and preservation of life, property and the environment in the early stages of an accident or disaster (Prati and Pietrantoni, 2010). The nature of these organizations can differ across publications, although a general understanding of first responders seems to be a combination of fire departments, paramedics and police departments (Dilo and Zlatanova, 2011; Prati and Pietrantoni, 2010). If the scale of disaster increases, other organizations such as (paramilitary) defense units might be recognized as first responders as well (Dilo and Zlatanova, 2011). Because of the Dutch context of this thesis, first responders are defined as all organizations that are alarmed and coordinated by the control room. These are defined in the Dutch law 'wet veiligheidregio's': the fire department, police department, ambulance services and the medical assistance teams. The research will focus on facilitating safety regions with the proof of concept, specifically fire departments.

The modus operandi of first responder organizations rely to a large extend on well-established procedures (Zlatanova, 2010). As the organizational structure for emergency response differ across organizations and between countries, depending on the vulnerability and preparedness of an organization for disasters, the procedures are tailored to the specific needs of an organization. For example, country that deals with frequent earthquakes may have different disaster procedures compared to a country in which earthquakes are rare. This causes differences in the way in which disasters are handled by different organizations and in different countries (Zlatanova, 2010). These differences in procedures ask for a general understanding of the Dutch first responder operational context, as this will affect the requirements of good system design for the proof of concept.

4.2.1 Coordination in repressive first responder operations

Within first responder operations, two kinds of coordination can be distinguished: low level and high-level coordination. Low level coordination means coordination at a local scale, such as deployment of a fire department repressive team in a building. These units are coordinated by a local commander, who may direct one or more teams while being at the place of the incident (van der Meer, 2018). High level coordination is aimed on using shared situation awareness to reduce the impact of an incident, by collaboration between different organizations. These meetings are often not held at the place of the incident but at a central place of command.

4.2.1.1 Low level coordination

Local teams at the place of the incident are coordinated by a local commander. For firefighters, one commander is normally coordinating a team of five firefighters and a multi-purpose vehicle. A detailed overview regarding low level coordination of on-site preparation and execution of Dutch firefighters can be found in (van der Meer, 2018).

4.2.1.2 High level coordination

Every safety region has a main structure of disaster management and crisis mitigation. Collaboration and communication procedures between decision makers of different first responder organizations are described within the GRIP structure. The main structured entails the units as stated below, translated to English by the author of this thesis from the law 'Besluit veiligheidsregio's'.

	English:	Dutch:	Abbreviation:
•	Control Room	Meldkamer	
•	Command Place Incident	Commando Plaats Incident	CoPI
•	Regional Operational Team	Regionaal Operationeel Team	ROT
•	Municipal Policy Team	Gemeentelijk BeleidsTeam	
•	Regional Policy Team	Regionaal BeleidsTeam	

The control room alerts all involved first responder departments and supports their efforts in a global way. CoPI and ROT are both multidisciplinary teams, involving at least a member of each involved first responder department. The individual team members are responsible for the processes of their own departments and communication about these processes with the other multidisciplinary team members. The difference between CoPI and ROT is the scale: CoPI are focused at the immediate threat or the source area, while ROT's are focusing on the indirect threat outside of the source area. Therefore, CoPI is often situated in proximity of the incident in a mobile command and control meeting place, or a building close by. The ROT's operate at a larger scale and do often meet at a predefined place somewhere in the safety region, such as the head quarter of the safety region (Cools et al., 2017).

While the CoPI and ROT are both concerned with the first responder deployment directly, the municipal policy team and the regional policy team have a task that is aimed at advising the responsible supervisors of the incident. This is either the mayor (phase 1 – 3) or the safety region chairman(s). Although not being a legal requirement, first responder organizations are often represented in the policy teams as well (Cools et al., 2017).

4.2.2 First responder spatial information availability

If an incident occurs, a process is initiated to mitigate the effects of the incident. Information sharing is part of this process, to inform organizations and individuals of the current situation of the incident (van der Meer, 2018). In Endsley (2016) we have seen that performance is influenced by the quality of decisions, which is dependent on the level of situation awareness. Situation awareness is created by translating raw data into insightful information which is necessary for successful completion of tasks. For this, we must identify what information would be meaningful for first responders. Dilo and Zlatanova (2011) provide a data model of Dutch first responder incident response operations. The formal model is created based on work of (Xu et al., 2008) and (Scholten et al., 2008) and used for development of system for data management and sharing in emergency situations (Zlatanova, 2010), similar to the role that the proof of concept produced by this research will have.

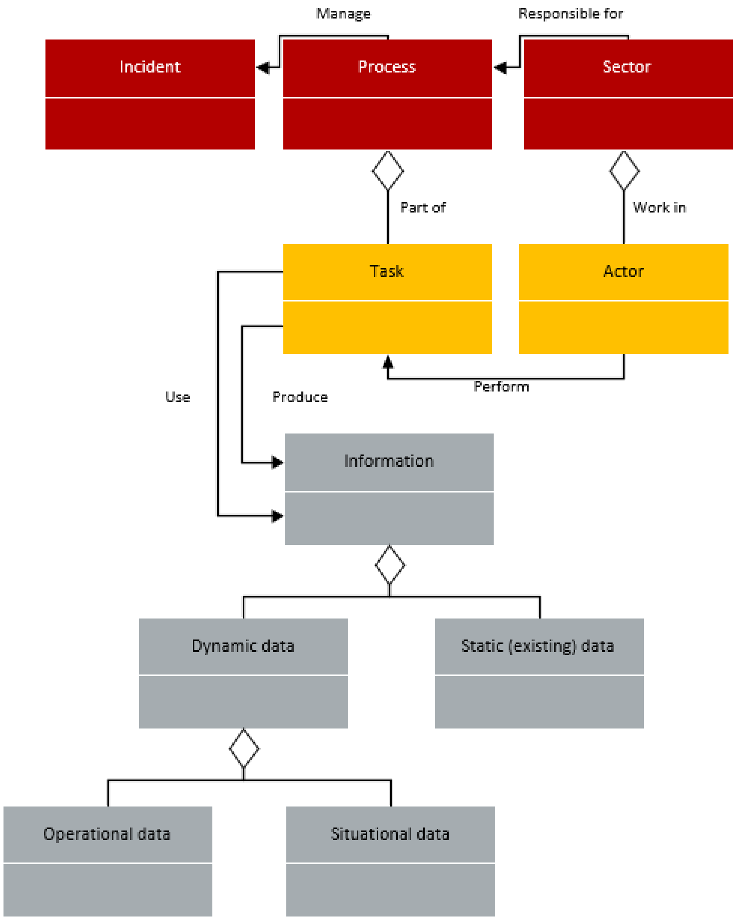


Figure 8: Conceptual schema of first response incident response (Zlatanova, 2010)

In the model (Figure 8), we see a clear link between tasks and information, which is what we would expect based on the goal oriented model of Endsley (1995). Furthermore, we see that the actor performs the tasks, and is therefore the ultimate decision maker in this process. As the actor uses information to perform the task, we conclude that the quality of the information and the quality of the decisions made are indeed linked.

Dilo and Zlatanova (2011) discern between two types of information at the base of emergency response operations (Figure 8): dynamic and static information.

Static data hold information that are not likely to change during an incident, such as managerial and administrative data, and risk maps. An officer of duty of a fire brigade for example, request several pieces of information such as topographic maps, a map of water resources, optimal route information an d risks maps of the area (Zlatanova, 2010). Dynamic data are volatile in nature and are collected during an emergency operation. Within this category, operational and situational data are identified.

Operational data describes data about the operation, including information about the ongoing processes such as responsible departments and persons, together with their roles. Situational data describes the incident itself and the impact of the incident on its environment, such as the type of the incident, the affected area and the number of trapped, missing or injured people (Dilo and Zlatanova, 2011).

4.2.3 Spatial data dissemination

A multitude of Spatial Data Infrastructures (SDI) provide first responders with access to the collected spatial data sources, in the form of command & control systems and early warning systems (Zlatanova, 2010). (van der Meer, 2018) explains how fire brigades make use of Mobile Operational Information Systems. These systems are mobile geographical information systems allowing the user to use different scales of data. For example, units can request/receive navigation instruction to move from the fire station to an incident location. While being on route, officers of duty can subsequently zoom to a local scale, showing building specific information if available (Figure 9).

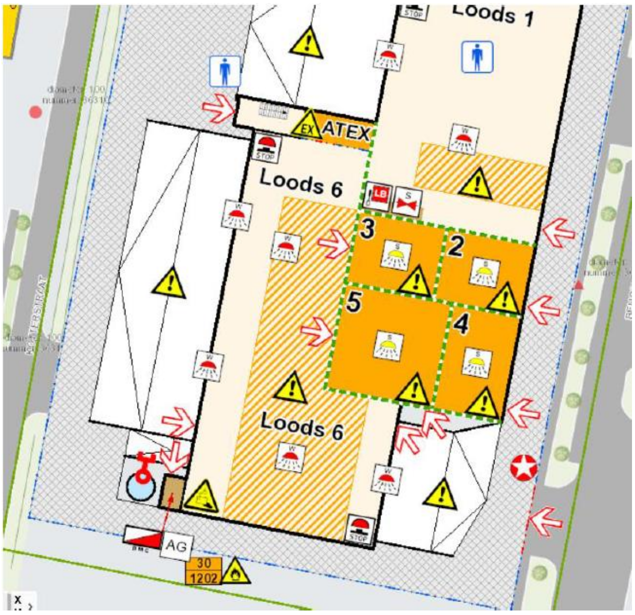


Figure 9: Digital attack map of the Mobile Operational Information System of safety region Rotterdam-Rijnmond (Van der Meer, 2016)

Although the technical embedment of indoor spatial data sources resulting from this research within an SDI is out of scope for this research, the research will consider data dissemination on an operational level. This means placement of the data within the operational format of first responders with the aim to provide the data to the right people, meaning the people who have to make decisions based on the resulting information to improve execution of tasks and processes for emergency response.

(Keating, 2016) offers an overview of the governance of geo-information flows and coordination of emergency responses applied from an SDI perspective in a Dutch context. Two user groups are distinguished, one at operational level (e.g. emergency responders) and one at a tactical level (e.g. safety region) to demonstrate the information needs of emergency managers. A technical implementation of data sources within an SDI for emergency services can be found in (Dilo and Zlatanova, 2011), (Keating, 2016) and (Scholten et al., 2008).

4.2.4 Spatial data collection

Data is at the heart of spatial data infrastructures (Keating, 2016). The static data that is used to inform first responders is often stored in datasets with a nation-wide coverage. As has been illustrated already, these datasets comprise for example topographic maps, water resources, route information and risk maps (Zlatanova, 2010). In line with the ambition of the Dutch government to collect data sources once and use them many times, most of these datasets are not collected by the first responders organizations but provided by secondary sources such as municipalities, Kadaster or Rijkswaterstaat (Dilo and Zlatanova, 2011; Van Capelleveen et al., 2008).

For indoor operations however, the first responder organizations are often forced to gather data themselves as spatial information is seldom readily available (van der Meer et al., 2018). Risk maps are drawn up by the safety regions for vulnerable buildings in the preparation phase, in which the basic geometry of ground floors are depicted (Van Capelleveen et al., 2008). This information is enriched by an indoor exploration of an repressive team, to identify the fire source, to determine attack routes and to check if the fire can be extinguished with resources present in the building (van der Meer et al., 2018).

The identification of the data sources by exploration is mainly a manual workflow. A real-time indoor SLAM application could add to the way dynamic information is shared. By using updated information about indoor environments a command and control unit is better equipped with a more accurate base to make decisions on (Basilico and Amigoni, 2011). The next section will explore current mapping and positioning practices and connect these practices to first responder operations.

4.3 Collecting indoor spatial data: Simultaneous Localization and Mapping (SLAM)

The previous section states that availability of information creates situation awareness, which improves first responder decision making processes and thereby first responder performance. It has also been stated that indoor spatial data is often missing or sparsely available, thereby hurting first responder indoor performance. There are automated ways to create spatial information of indoor environments, as will be illustrated in this section.
First, an explanation of the importance of Simultaneous Localization and Mapping will be explained in concept and application, opposed to collection of mapping and tracking data separately. Second, the concept of mapping will be further explained. Third, methods of collecting spatial mapping information will be introduced. Fourth, the concept of tracking will be further explained. Fifth, methods of tracking will be introduced. At last, a summary of first responder application for SLAM use will be given.

4.3.1 SLAM: Introduction

Simultaneous Localization and Mapping (SLAM) algorithms refer to a variety of algorithms that enable mobile simultaneous mapping and tracking for a wide range of mobile devices (Rantakokko et al., 2011). Among these devices are for example backpack systems, handheld sensors, trolley systems and head mounted devices (Khoshelham et al., 2019; Nikoohemat et al., 2020). This research makes a distinction for mobile devices that are carried by hand and devices that are wearable, meaning they can operate without being held. Instead, these wearable devices are placed on the body of the user. Wearable devices offer opportunities to first responders, as they have their hands free to handle equipment, remove rubble or transport victims (Khoshelham et al., 2019).

Next to TLS, mobile laser scan (MLS) systems are used for indoor mapping processes (Lehtola et al., 2017). These systems have a higher flexibility compared to the TLS systems, as they do not require several set up sequences. MLS systems do also work with LIDAR and require therefore a line of sight to measure distances to surfaces. However, as these systems are mobile, their line of sight can be changed dynamically. To calculate the distance to a surface, the MLS should know its own position at the moment of sending and receiving the laser pulse. For this, the systems make use of Simultaneous Localization and Mapping (SLAM).

4.3.2 SLAM context: a brief history

The problem of Simultaneous Localization and Mapping has roots in robotics, where consistent mapping and positioning has an important place in the field as it is regarded to be a key functionality for fully autonomous robots (Durrant-Whyte and Bailey, 2006). The need for consistent mapping and continuous positioning emerged in the 19'eighties, when probabilistic methods had only made their first steps in robotics and artificial intelligence (AI). The idea emerged that one could never be sure about one's position in an environment, as the environment may be changing rapidly and because measurements are incomplete and not fully continuous. Therefore, spatial reasoning should be based on the probability of being at an estimated pose in an estimated mapped environment (Durrant-Whyte and Bailey, 2006; Smith et al., 1990).

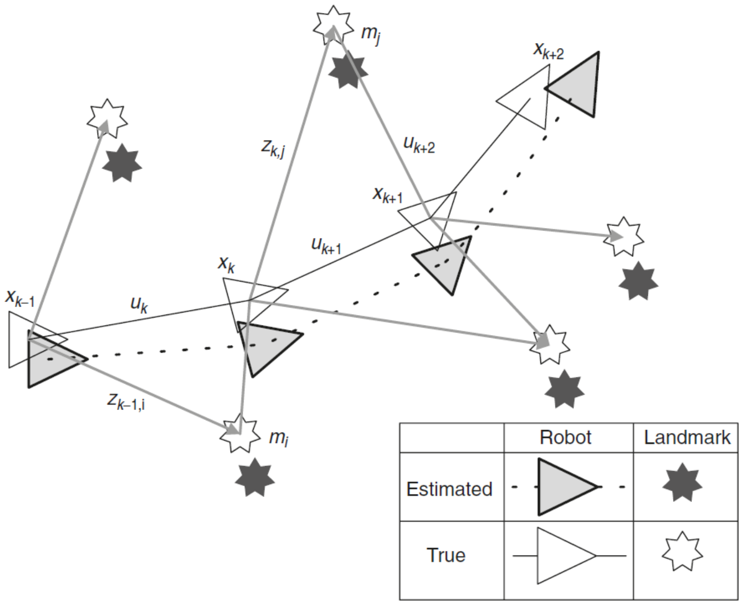


Figure 10: estimated and true robot and landmark poses, describing spatial relationships between robot and correlated landmarks (Durrant-Whyte and Bailey, 2006)

From this idea, rapid progression was made both for the conceptual and computational issues raised by SLAM (Durrant-Whyte and Bailey, 2006). A key element in this progression was the finding of a statistical basis for describing the relationship between a mapped object and the manipulation of geometric uncertainty (Smith and Cheeseman, 1987). This means enabling a robot to reason and demonstrate its knowledge of the approximated relationship between itself and objects in its spatial proximity by using sensor information to reduce both positional and orientational (pose) uncertainty. The objects are also addressed to as 'landmarks', as they serve as reference objects.

At the same time, advancements in visual navigation were made, for example on sonar-based navigation. Visual navigation and probabilistic reasoning about a spatial relationship in a set environment came together in a paper by (Smith et al., 1990), also dubbed 'the landmark' paper (Durrant-Whyte and Bailey, 2006). The main message from this fundamental paper was the proposition that as a robot moves through an unknown spatial environment, it has to keep track of the positional estimation of landmarks in its surrounding spatial environment. However, it should not observe the landmarks in isolation, but it should recognize the estimated spatial relationships in a holistic way as the estimates are fundamentally correlated due to the cumulated error (drift) in the estimated position and orientation of the device (Smith et al., 1990).

The consequences of this realization were profound, as it meant that a device could only tackle the combined SLAM problem by keeping track of both the pose of the device and the position of all of the mapped landmarks. These relations would have to be updated for every (new) mapping of a landmark, requiring a state vector with complexity related to the size of the maintained landmarks. This inhabits a computational scaling of O(n^2), in which n is the size of the mapped landmarks (Smith et al., 1990). As this is a quadratic function, computation time increases very rapidly as the mapped area (and therefore the mapped landmarks) increase, making the use of SLAM algorithms on mobile devices in large areas almost impossible. After this, scientific focus shifted to either using assumptions to limit or even eliminate spatial relationships between landmarks or by treating pose estimation and mapping as separate problems (Durrant-Whyte and Bailey, 2006).
The tide for SLAM was turned by another conceptual breakthrough, namely by treating pose estimation and landmark mapping estimations as one single estimation problem. This realization causes the computational complexity of the problems to become convergent instead of divergent. This causes the time complexity of the algorithm to be minimized to O(n Log n) which is a significant improvement. However, maybe even more importantly, it was recognized that the spatial relationship between landmarks should be treated as more meaningful if the correlation grew, the quality of the solution would also grow (Durrant-Whyte and Bailey, 2006).

The term SLAM was presented at a robotics conference together with the single estimation problem in 1995 (Durrant-Whyte et al., 1996). It was a renewed kick-off for many institutions to do research at SLAM algorithms, which were also named Concurrent Mapping and Localization (CML) algorithms (Durrant-Whyte and Bailey, 2006). The algorithms were applied to indoor, outdoor and even subsea environments. Two main SLAM solutions emerged from these developments, either using Extended Kalman Filter (EKF-SLAM) or Rao-BlackWellized particle filters (FastSLAM) (Montemerlo et al., 2002) (Durrant-Whyte and Bailey, 2006).

Research of the last decennium has extended the computational time, reliability and flexibility of SLAM even more (Mur-Artal and Tardós, 2017; Taketomi et al., 2017; Tateno et al., 2017). The perspective of the field widened from robotics to applications such as computer vision-based online 3D modeling, augmented reality visualization and self-driving cars (Taketomi et al., 2017). As mentioned before, mapping components of the early SLAM systems made use of sonar-based sensors. Later on, SLAM algorithms were adapted to work with a wide range of sensors, for example Mobile Laser Scanners (MLS), Inertial Movement Sensors (IMU), rotary encoders, GNSS and camera systems (Taketomi et al., 2017). The adaptation of a wider range of input sensors caused adaptation of the existing SLAM algorithms such as Visual SLAM (vSLAM), which have made impressive progress in the 2010's in both reliability and affordability (Taketomi et al., 2017).

4.3.3 SLAM basics

As there are various technical implementations of SLAM algorithms that are already explained in detail in existing literature, this research will only cover the basics of the concept. We have already seen the importance of spatial correlation between the estimated pose of the device and the estimated position of landmarks, as it is described in the short history of SLAM. We will now explore these concepts further.

4.3.3.1 Landmarks

SLAM systems estimate their pose and build a map of their environment in a so called 'online' manner. This implies that there is no need for former knowledge of the environment (Durrant-Whyte and Bailey, 2006). Furthermore, there is a distinction between global and local map matching: global map matching uses all available landmarks, while To estimate the pose of landmarks in space, finding and tracking the landmarks is a key function from SLAM algorithms. The nature of a landmark may be confusing, as landmarks are often used to describe places or routes. In mental mapping for example, complete objects are often used as spatial anchors to describe spatial relationships (Epstein and Vass, 2014; Xia et al., 2008). A church tower, a statue or a specific plant can all be described as important landmarks and may serve as orientation points to make navigating through a space easier. In SLAM systems, landmarks are not necessarily individual objects (Durrant-Whyte and Bailey, 2006). However, an analogy between the mental mapping landmarks and the SLAM landmarks can be found in the statement that the landmarks are distinctive features in the spatial environment.

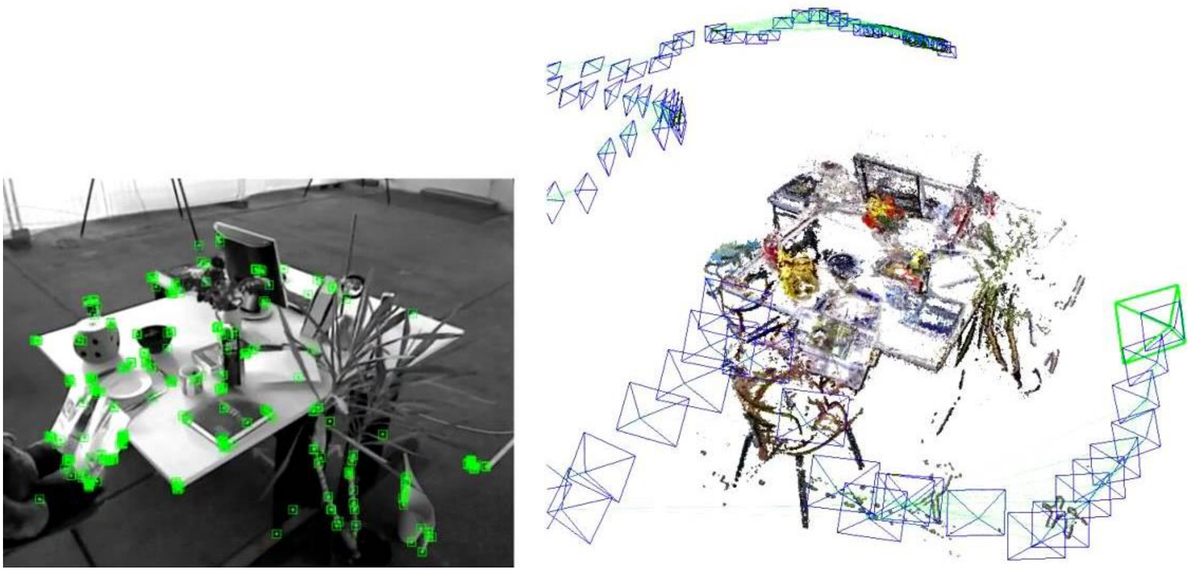


Figure 11: Landmarks identified by a vSLAM method in real-world representation (left) and as abstracted reality with observer poses (right) (Mur-Artal et al., 2015)

Examples of distinctive features are color based distinctions (like a red spot on a white wall), geometry distinctions (points, lines or a shape like a pillar in an empty office environment) or even invisible signal receptions (like distinctive Wi-Fi signals). SLAM algorithms are likely to use a combination of sensors, a process called sensor fusion, to receive a combination of distinctive features related to a specific position into one landmark (Coppens, 2017). These signals can be active or passive, in a way that devices make use of the readily available infrastructure (passive) or by the placement of beacons sending out a distinctive signal (active). The combination of features makes landmarks even more unique, while the redundancy does also offer reliability in case some features are lost. Coppens (2017) can be consulted for a more detailed overview of motion tracking and related computer vision techniques.

4.3.3.2 Loop closing

Another crucial element of SLAM is loop closing, which should happen if a specific place is revisited (Newman and Kin Ho, 2005). This is an especially difficult problem for SLAM algorithms as landmarks are scanned again, after which their position and correlated spatial relationships in the SLAM model are updated (Durrant-Whyte and Bailey, 2006). The first problem of loop closing is recognizing a revisited spot. If an area is not recognized as revisited it will be added as a new area with new landmarks, with a different global location and orientation. Error accumulates dramatically because of such an event, making it possible that the device is lost for the entirety of the remaining scan time. Therefore, not only asserting a loop but even recognizing small possibilities of loop closure is of the essence (Newman and Kin Ho, 2005). The hypothesis that a device is unlikely to have the same orientation or heading once it revisits an area increases the chance of missing a loop closure, which makes robust recognition of landmarks even more important.

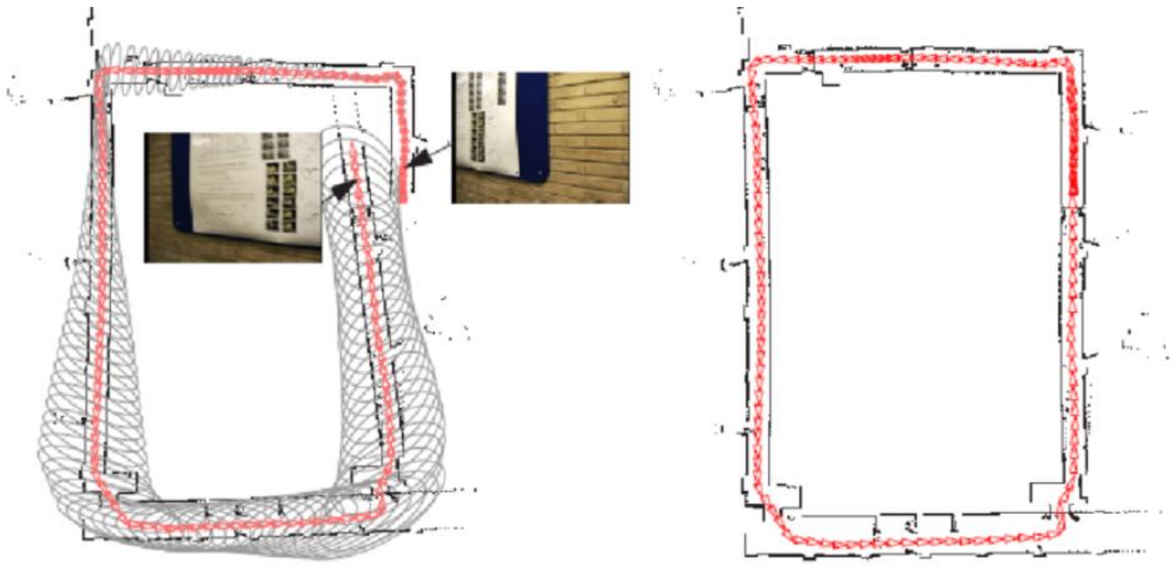


Figure 12: Advantage of loop closing: left: a SLAM snapshot just before loop closure. Global uncertainty is depicted with grey circles, for which a larger circle means larger pose uncertainty. Right: result after loop closure.

However, there are large merits to loop closing as well. Correct recognition of loop closure gives the opportunity to re-align measurements over a large space, thereby increasing reliability of the results. Many technical aspects of loop closure, including recognition of loop closure and re-aligning landmarks after a loop closure can be found in literature. Newman and Kin Ho (2005), (Williams et al., 2009) and (Labbé and Michaud, 2014) give detailed technical explanations in these topics.

4.3.4 The concept of tracking

Tracking is a key aspect of a SLAM process, as consistent mapping can only be acquired by knowing the spatial relationships from scanner to its spatial environment (Durrant-Whyte and Bailey, 2006). A mobile scanner can hardly effectively scan environments if it does not know its where it is in space (Luhmann et al., 2013). Next to this, many applications benefit from knowing where objects or phenomena are in space. For example, it enables indoor navigation (Sithole and Zlatanova, 2016), finding first responders of a rescue team in a building (Nuaimi and Kamel, 2011) and to track assets for the purpose of asset management (Nuaimi and Kamel, 2011).

4.3.4.1 Location, pose and position ambiguity

Localization is a term frequently used in SLAM literature, which makes sense as it is one of the words in the abbreviation of SLAM (Durrant-Whyte and Bailey, 2006). Its meaning, however, appears to be ambiguous. Position, location and pose are often used with similar meaning. In (early) SLAM literature, location is often referred to as a reference to a spatial construct with six degrees of freedom: x, y, z coordinates and roll, yaw and pitch angles (Durrant-Whyte and Bailey, 2006; Nüchter et al., 2005; Smith and Cheeseman, 1987).

Besides the SLAM use of location, location is also embedded in natural language. Here, terms such as 'position', 'location', 'place' and 'area' are often used interchangeably (Sithole and Zlatanova, 2016). However, subtle differences between these words exist according to recent research. According to Sithole and Zlatanova (2016), ambiguous words for spaces are frequently used in indoor navigation systems. To prevent ambiguous perception of the functions of the system, the terms 'localization', 'positioning', and 'posing' will be defined.

To explain the relevance of this distinction, first the word 'Space' will be defined. 'Space' means a spot, surface or volume with a spatial definition or spatial context. Consequently, the word 'addressing' will be used for the capability of referencing a space. The way in which spaces are addressed can differ per person and per situation (Sithole and Zlatanova, 2016). The local knowledge and perception of a space is important to this capability. For example, someone who has a good mental model of a building might address location within the building differently from someone who does only have a blueprint of the building (Sithole and Zlatanova, 2016). The first one might address a certain space in the building as 'the room next to the coffee machine on the second floor, in wing B', while the latter one might address this same location as 'room 213B'. They are speaking about the same space but use different words to describe it.

4.3.4.2 Localization

According to Sithole and Zlatanova (2016), a location is absolute, meaning it is defined by physical borders. Furthermore, locations have context, meaning it has meaning in relation to objects, phenomena, spaces and areas. Locations can contain places and can be contained in areas, which are both contextual relative spaces without a defined spatial border. This spatial relationship is illustrated below in a floor plan of a first floor of a building.

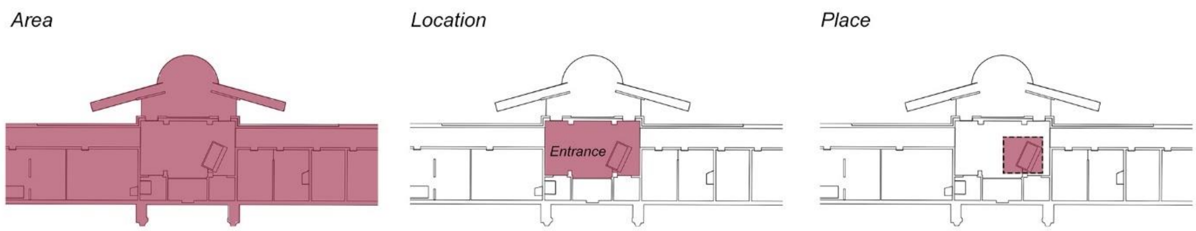


Figure 13: Difference between Area (I am in the middle part of the building); Location (I am in the entrance hall); and Place (I am at the reception desk)

4.3.4.3 Positioning

In contrast to a location, a position does not contain context and is therefore unambiguous. This makes it easier for computer systems to reason about space. Instead of areas, locations and places, a position does not have a surface but is a point in space. To address a position, a reference frame (for example cartesian) and therefore a position is always relative (Sithole and Zlatanova, 2016). For example, a point, surface, or volume can be positioned in space by using a coordinate system. The reference frame can either be arbitrary or absolute. While absolute frames of reference are agreed upon by consensus, such as the WGS 1984 coordinate system, arbitrary frame of reference are locally defined and not agreed upon. A non-georeferenced scanner may use such an arbitrary frame of reference, in which positions are always relative to the origin of the scan (Luhmann et al., 2013).

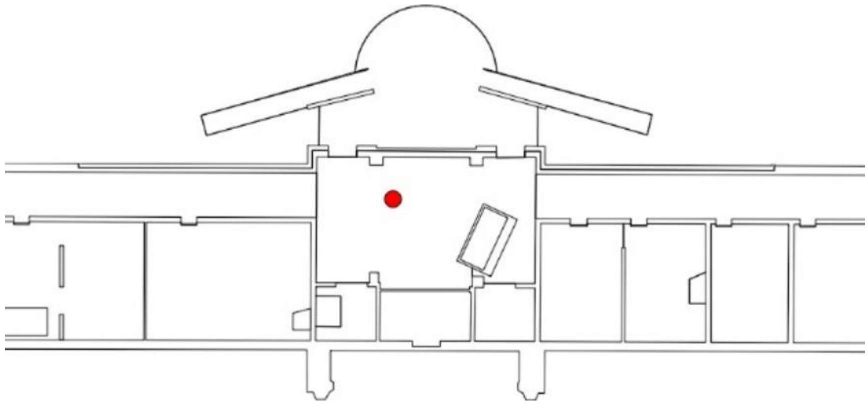


Figure 14: Position illustrated by red dot: my position is at x = 52.0056702, y = 4.3705018

4.3.4.4 Pose estimation

For SLAM systems, the estimation of a position in 3D space is an important component of the localization part of the system, as it is used to determine spatial relationships between the observing device and correlated landmarks. As can be seen in the explanation of the SLAM use of 'localization', not only the position in 3D space is important for these spatial relationships, but also the angles from which a landmark is observed. Although the combination of position and orientation are referred to in SLAM literature as 'location' (Smith and Cheeseman, 1987), this might cause ambiguity by recent definitions of location and positions (Sithole and Zlatanova, 2016).

Therefore, this research sees the localization component of SLAM rather purely as an estimated combination of x, y, z coordinates and the roll, yaw and pitch angle of the device. This research will use 'pose estimation' to refer to the combination of these six degrees of freedom in space.

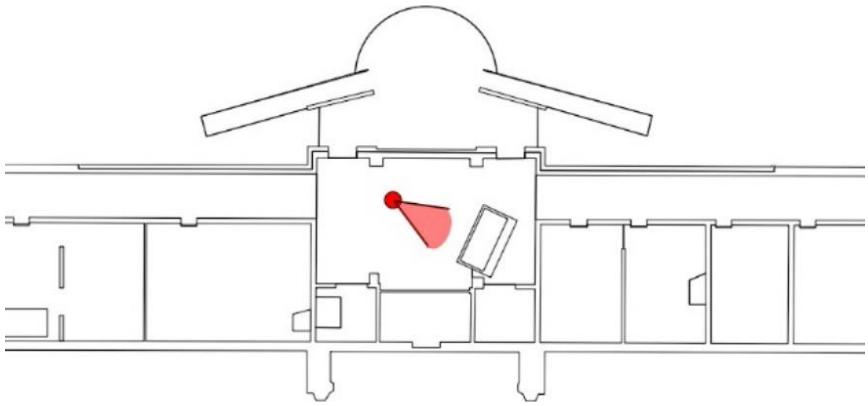


Figure 15: 2D Pose illustrated by arbitrary position + orientation: my position is at x = 52, y = 4.3 and my yaw angle is 295°

4.3.5 The concept of mapping

The word mapping is ambiguous, as it can have many meanings. If the English oxford dictionary is consulted, mapping is defined in two ways ("Oxford Dictionaries, s.v," n.d.). One definition is the process of discovery or the presentation of information about the structure of something. An example of this is 'gene mapping', but one can also think about 'mind mapping' or a 'literature mapping study'. However, 'maps' are also ambiguous in meaning, as maps can be presented in many different ways (Kraak et al., 2013).

Although maps can have different nature, they are always used to visualize geospatial data, meaning data that includes a reference to the location or the attributes of objects or phenomena located on earth (Kraak et al., 2013). The transformation from paper to digital maps has driven a evolution from static information containers to dynamic interfaces, enabling the presentation of multi-dimensional subjects. This enables for example (interactive) presentation of 3D spatial objects and differences over time (Kraak et al., 2013).

It is assumed that the creation part of 'mapping', as covered in the oxford dictionary definition, is left out by Kraak et al. (2013) as they describe 'maps' as a product. However, because 'mapping' is a verb, both the creation and visualization part of maps are considered important for the word. Moreover, mapping is not restricted to objects bound to earth. Therefore, the definition of 'mapping' has been adapted from both sources into the creation and visualization of spatial data. First, raw environment data is collected. Second, the data organized for meaningful visualization. These two processes will run in parallel if mapping is to be used in real-time. Because of the scope of this thesis, the defined space of the mapping process of this thesis is restricted to indoor environments.

4.3.6 Methods of mapping indoor environments

As it has now been defined what it means to map an indoor environment, the next step is to explore ways to capture and visualize locations or attributes of objects or phenomena within indoor environments. Such objects or phenomena can range from floor plans to positions or attributes of lamps, people and furniture (Khoshelham et al., 2019). These objects and phenomena tend to change over time, ideally paying attention not only to registration at one point of time but to continuous or frequently updated registration (Khoshelham et al., 2019).

Although 2D indoor environment products like floor plans exist, these products are often derived from 3D data (Wang et al., 2015). There are multiple ways to automatically capture indoor environments, each with their own advantages and disadvantages. Some methods are specifically good at representing the real world in a very accurate way, while others might lack accuracy but provide a higher efficiency or flexibility (Khoshelham et al., 2019). For in-depth information of optical indoor mapping and positioning, Mautz and Tilch (2011) provide a very detailed overview, including scanner specifications and local coordinate transformation methods. However, this thesis provides a more general overview of the difference between LIDAR and photographic scanning methods.

A trade-off can be identified between three mapping performance attributes, as named in Khoshelham et al. (2019). First, efficiency is defined as the extent of space which can be mapped in a certain period. Second, accuracy is defined as the extent to which the scanned environment represents the real-world environment, with scan resolution as a main indicator. Third, flexibility is defined as the extent to which the method offers opportunities to alter parameters or use gathered data during the scan. Mapping methods are usually only performing well on one or two of the attributes in comparison to other methods. This trade-off will be used to illustrate and compare the methods.

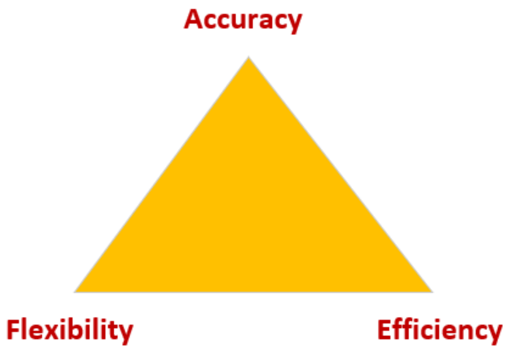


Figure 16: mapping tradeoff: accuracy, efficiency, flexibility

4.3.6.1 LiDAR

For indoor scans, LiDAR (LIght Detection And Ranging of Laser Imaging Detection And Ranging) technology is often used (Nikoohemat et al., 2020, 2019; Orthuber and Avbelj, 2015; Staats, 2017). LiDAR is an active method, in a way that it transmits beams of light and calculates distances by calculating the time it takes light to reflect from surfaces back to the sensor (Luhmann et al., 2013). If the pulse encounters a surface, it scatters, giving detailed information about the distance traveled from the scanner to the surface and back. As the distance and the return rates of the laser beam is captured, a point with x, y and z coordinates can be recorded at the spot in which the surface is detected. If combined with a high-speed scanning system, the laser system can capture very dense measurements of a 3D space (Walsh et al., 2018). This results in a 3d data structure that is called a 'point cloud'. A point cloud can consist out of millions or even billions of points, representing the real world very well (Richter and Döllner, 2014). Such point clouds are frequently used for automated indoor mapping purposes (Khoshelham et al., 2019; Koeva et al., 2019; Nikoohemat et al., 2019; Rusu et al., 2008; Wang et al., 2015; Zlatanova et al., 2013).

Point clouds do generally only provide geometric information and per-point attributes (Richter and Döllner, 2014). The points are therefore placed in an unstructured, unorganized and hierarchy lacking 3D space (Verbree et al., 2019). Point clouds can be rendered in different ways, e.g. photorealistic, non-photorealistic and rendering based on per-point attributes or point density (Richter and Döllner, 2014). It should be noted that RGB values are not captured by LIDAR, but points can be colored by using an overlay of a RGB camera. We distinguish two types of LiDAR sensors: The Terrestrial Laser Scanner (TLS) and the Mobile Laser Scanner (MLS).

TLS are mounted and immobile, which enables them to calculate distances with millimeter levels of accuracy (Lehtola et al., 2017). A TLS assumes the sensor remains at a fixed position while scanning. This means the coverage of a TLS is limited by the viewpoint: only surfaces with a direct line of sight will be mapped by a TLS. Therefore, a complete indoor scan will often require multiple scans from different positions. This makes mapping an indoor space with a TLS an accurate, flexible but inefficient solution.

MLS systems offer more efficiency compared to the TLS systems, as they do not require several set up sequences. This makes them more suitable for quickly scanning indoor environments (Lehtola et al., 2017). MLS systems do also work with LiDAR and require therefore a line of sight to measure distances to surfaces. However, as these systems are mobile, their line of sight can be changed dynamically. To calculate the distance to a surface, the MLS should know its own position at the moment of sending and receiving the laser pulse. This requires some sort of SLAM algorithm to estimate the position of the device while scanning.

As we have seen, the spatial relationship between landmarks and device are not fixed, meaning the SLAM algorithm changes the relation during the scan. This is troublesome for direct, real-time use of the data as the final data is only available after the scan is complete. This makes MLS an accurate, efficient, but inflexible solution.

4.3.6.2 Photogrammetry

Besides using LIDAR, photogrammetry is used for mapping indoor environments (Khoshelham et al., 2019). Photogrammetry has been defined as: "*the method that encompasses image measurement and interpretation in order to derive the shape and location of an object from one or more photographs of that object (Luhmann et al., 2013).*".

This historical understanding means the use of two dimensional (2D) imaging to deliver 3D object data (Luhmann et al., 2013). This means that 3D data can be derived from multiple 2D images. This is done by calculating the coordinates of the pixels within a local coordinate system by computing overlap of different images (Luhmann et al., 2013). The result of these calculations are first pixels that do also have positional information. By combining these 3D pixel images, a point cloud of the environment can be created (Luhmann et al., 2013). In this description, the 2D images are captured in a passive way as it captures a signal not produced by itself.

Photogrammetry can be applied to all sorts of camera output. The most common application is to use mainstream RGB cameras because of the low cost of these cameras (Khoshelham et al., 2019; Taketomi et al., 2017). For accurate results, a high image resolution is beneficial for the 3D mapping process (Luhmann et al., 2013). Today, even small cameras like the ones on smartphones can sport sensors able to capture millions of pixels (megapixels) in one image. This enables efficient scanning of environments, while systems are still able to deliver accurate results with centimeter resolution. Furthermore, the scanning process does not have to be completed to present intermediate results, making it usable for real-time processing on (part of) the data. However, calculating overlap of images is a resource intensive task, therefore requiring high quality hardware or cloud computations to get real-time results. This makes photogrammetry an accurate, efficient but inflexible solution.

4.3.6.3 Depth-imaging

As an alternative to LiDAR and photogrammetry, time of flight range cameras are used to measure distance to objects. Like LiDAR, range cameras are active and use the reflectance of infrared light. The difference with LiDAR however is the width of the infrared beam, of which the reflectance is measured in a pixel wise manner. Although range is less compared to LiDAR, it enables for quick accurate scanning of objects which are only a couple of meters away. Furthermore, it requires less computational resources compared to photogrammetry as depth is sensed and not calculated. This makes it a good option for constantly moving devices such as a head mounted Microsoft HoloLens, as (Hübner et al., 2020) demonstrates in detail. Therefore, if one takes the limited range into account, a range camera is a reasonably accurate, efficient and flexible solution.

4.3.7 SLAM research gap for first responders

Although SLAM algorithms enable systems to map and track efficiently within indoor environments, most systems are limited to delivering their results until completion of the scan. Sometimes, even additional postprocessing is needed to generate a reliable 3D representation of the environment (Luhmann et al., 2013). The reason of this is the nature of the SLAM algorithm, which enables the system to change the mapping and its track as the scan progresses. This minimizes drift, needed to correct for internal positional errors that are for example generated by the Integrated Mobility Unity (IMU) of the device (Luhmann et al., 2013). This delay of processing is not a problem for many SLAM use cases, as processes such as 'scan to BIM' (Wang et al., 2015) or generation of high quality navigation graphs (Nikoohemat et al., 2020) do not require instant accessibility. First responders however, need the resulting data as soon as possible due to the dynamic environment of first responder operations (Kapucu and Garayev, 2011; Seppänen and Virrantaus, 2015). This asks for more research to the added value of real-time SLAM systems to the first responder context (Khoshelham et al., 2019; Rantakokko et al., 2011), to which this research aims to make a contribution.

4.4 Designing a Common Operational Picture

In the three levels of situation awareness, we have seen that although raw data is a start for the creation of situation awareness, data does not guarantee the creation of situation awareness. First, data has to be transformed into information, after it may be used in decision making processes to perform tasks in a qualitative way (Endsley, 2016). A system that takes or creates raw data and communicates this data in an informative way to a user might be perceived as a 'decision support system', as it aids the user in the decision making processes and may even advise or take certain decisions based on the input data.

Furthermore, a user oriented design is preferred as this means taking into account the information needs and system requirements of users to prevent information overload (Endsley, 2016; Wickens et al., 2013). This reduces mental stress which makes it easier to reach the 'comprehension' and 'projection' levels of situation awareness. Besides this, the effects of using user-centric design opposed to technology-centric design are reduced chances on errors, improved safety, improved user acceptance & satisfaction and finally, improved productivity (Endsley, 2016).

4.4.1 Identifying user requirements

From a situation awareness perspective, dynamic and existing data form the elements needed for creation of the first level of situation awareness: perception. The data elements are transformed into information. This information forms the input for the execution of tasks in a certain manner: the decision-making process. If all required data elements are there for solving the task and if the data elements can be comprehended, the decision should be of better quality compared to a situation in which (part of) the data elements are missing. The better decision leads to increased performance and thereby to the second level of situation awareness: comprehension. Furthermore, if the data elements also provide temporal information and if information overload can be reduced, even the third level of situation awareness is possible: projection of current states of the operation into future states of the operation. Next to qualitative data elements and comprehension, reaching this level of situation awareness requires an experienced user.

However, first responders are usually no geo-spatial specialists and therefore they do often lack deep understanding of terminology and structures used for spatial data (Zlatanova, 2010). Proper filter techniques should be applied to (spatial) data support systems, as they tend to create an information overflow rather easily instead of increasing situation awareness (Endsley, 2016). Information overflow is often described in a first responder context, especially in the case of using spatial information under pressure and within stressful conditions (Zlatanova, 2010). A solution that can be applied to battle information overflow is to offer information in different levels of detail and customized to different tasks or user groups (Endsley, 2016; Zlatanova, 2010). This stems from the idea that an officer of duty who directs one team within an indoor operation needs different information compared to an officer of duty who is present at a ROT meeting.

To identify data requirements of first responders, Zlatanova (2010) follows a two-step approach. First, a complete overview is created of all the data sources that are needed to respond to a specific incident. Second, specific data sources are narrowed down for specific tasks within the emergency response formulation. Close collaboration with first responders is recommended to identify user requirements for the data. Zlatanova (2010) hints to several methods of involving first responders, such as organizing interviews, filling out questionnaires, open discussions, trainings and studying organizational instructions.

Both Endsley (2016) and Wickens (2013) support the involvement of users to create situation awareness and provide useful recommendations on doing so. Nevertheless, although consultation of users should be a key process within user-centric design, it does not mean that every design advice of a user is a good one. Endsley (2016) explains how system designs have failed by implementation of just the functionality that was asked for by the users. Users are experts in their field, which is also the case for the first responders who would be using the proposed proof of concept of this research. However, most first responders are no experts in design principles such as effective presentation of information and human interactions with complex information systems. The risk of an endless list of ideas and requirements emerges, while other users might ask for completely different things. Users must be regarded as a valuable source of information, which should be combined with design principles and user requirements from literature (Endsley, 2016).

4.4.2 Support, not limit decision makers

In the context of designing systems for the creation of situation awareness, Endsley (2016) states that "*User-centered design does not mean systems that make decisions for the user*". What is meant here is that a decision support system may act as a black box, meaning that the system is not transparent in the way it comes to an advice to the user. Such systems do not necessarily end up in systems that provide best performance because of several reasons (Endsley, 2016). Especially ambiguous results tend to slow human decision making processes down and lower the quality of the decision (Endsley, 2016). Speed is of the essence for first responder operations, while the safety of first responders and potential casualties depend on the quality of the decisions made at operational level. Ambiguity of the system is therefore regarded as a threat, and transparency of the system is deemed important for the indoor SLAM system.

Furthermore, although a system that provides advices on the decisions that have to be made may provide wrong advices. This can lead to 'decision biasing', where a decision support system guides a decision maker in the wrong way, thereby leading to a wrong decision (Endsley, 2016). In these cases, the user would be better off without the decision provided by the system, as he/she would not have been misdirected without it. Professionals themselves are also prone to making wrong decisions, but their experience in the field is an important factor in preventing such mistakes (Wickens et al., 2013).

4.4.3 Reducing information overload

Processing information implies a cost, both for processing information by humans and computers (Endsley, 2016; Wickens et al., 2013). A human operator is only capable of processing a certain amount of data into meaningful information and to act on the gained information. Very dynamic and stressful environments decrease this limitation further (Endsley, 2016). The notion that more data does not equal more information is becoming more widespread, which asks for solutions to filter data and to try and only present valuable data to decision makers (Endsley, 2016). The gap between produced data and needed information is called the 'information gap' (Endsley, 2000).

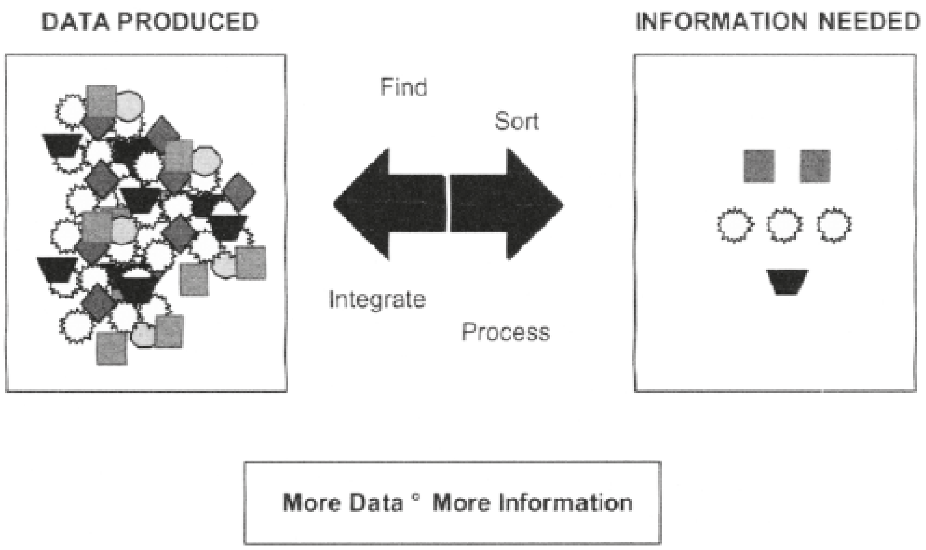


Figure 17: The information gap, as illustrated by Endsley (2000)

A user can only oversee a certain amount of information. Consequently, if a user is supplied with unnecessary information, crucial information might be overlooked leading to error. Although these errors are often attributed as 'human error', design-induced error might be a more accurate term to describe them as they are caused by bad system design (Endsley, 2016). Not all first responders need the same information, as the information needed is task related (Zlatanova, 2010). An officer of duty who is attending a ROT and who oversees deployment of dozens of people needs different information compared to a command present at a local CoPI commander who supervises a small indoor attack team. Therefore, applying a correct level of detail for information products is crucial to prevent information overload.

Although supplying users with just the information that they need for completing a certain task within a process sounds ideal, this is not what user-centric design means according to Endsley (2016). It has proven to be a very difficult task to instruct computers and information systems to present just the right information. The risk of system error is consequently high, as wrong information may be presented or removed at a given time. Even if we would be able to instruct computers to present only the right data, it is known people need some time to perceive the information and translate it to situation awareness (Endsley, 2016). Operators must quickly switch between tasks and goals, which is why some consistency in the data provision is needed to prevent confusion. This confusion is also caused in highly dynamic environments, such as first responder operations, as appearance of data might be lost by constantly switching data types on and off. This hurts the ability of operators to predict future states of a situation (Endsley, 2016), which is a key aspect of situation awareness. Therefore, it is better to present a little more information that might be useful regardless of the current task.

4.4.4 Visualization of data elements in CoP

To create an extent of situation awareness, users need to interpret data elements from the system (Endsley, 2016). If first responders use a real-time mapping and tracking system within indoor environments, raw data must be converted into information in real-time as well. Furthermore, we know first responder rely on the reliability of the data elements as they need to trust the data to be able to make good decisions. Determining exact questions that first responders will have during an operation is complex, especially as the CoP may be used to facilitate different user roles. As Endsley (2016) stated, making decisions for system operators may introduce decision biasing. Therefore, it may be better to create a generic knowledge base that system operators interpret or query themselves compared to deciding which knowledge system operators need.

Kraak (2013) states the presentation of spatial data elements can have three main characteristics: location space, attribute space and temporal space. These characteristics can be used to place emphasis on generic questions, namely to query 'where' something is, 'what' something is and 'when' something is (*Figure 18*). These characteristics are usable for reducing the information gap between the information presented and the information needed for making a decision.

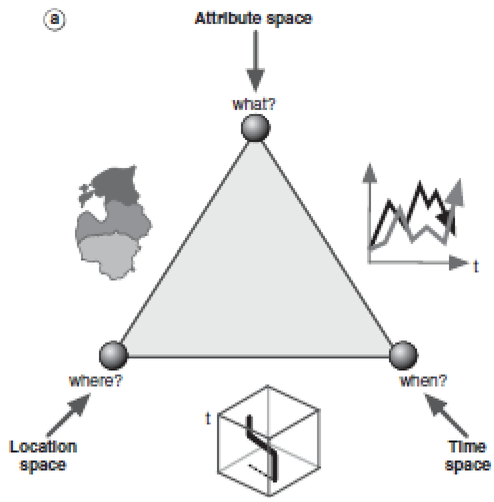


Figure 18: mapping visualization based on three main characteristics: location space, attribute space and temporal space (Kraak et al., 2013)

Van der Meer et al. (2018) describe visualization of indoor models for fire brigades in detail. There are some take home messages which will be briefly described here. First, Van der Meer et al emphasize the importance of both 2D and 3D visualization of indoor environments. 2D visualizations allow for a quick overview, while 3D visualizations allow for more accurate measurements and detailed inspection. Although the 2D and 3D visualization are often treated as separate environments, the publication states that a combined view could be beneficial for fire fighters. Furthermore, the publication explains different symbols and data requirements.

4.5 Conceptual model

Below, the conceptual model of this thesis is presented. To the left, the most important input requirements for situation awareness are stated. They are deduced from literature and the requirements identification. Mapping and tracking data elements are found necessary to create a level of situation awareness, needed to complete a certain task in an indoor first responder operation environment. These data elements need to be available in real-time, because of the dynamic and complex conditions of first responder operations. If data elements are successfully transformed into information, while reducing information overload and guaranteeing sufficient system reliability, the second level of situation awareness can be reached. On top of that, if the system is operated by an experienced user and if mental stress remains low, the third level of situation awareness can be reached.

The quality of the decision-making process is influenced by the level of situation awareness, and subsequently influences the performance of first responders for indoor operation environments. The research gap of this thesis is to what extent remote situation awareness can be created by using indoor situation awareness, described in one of the levels of situation awareness.

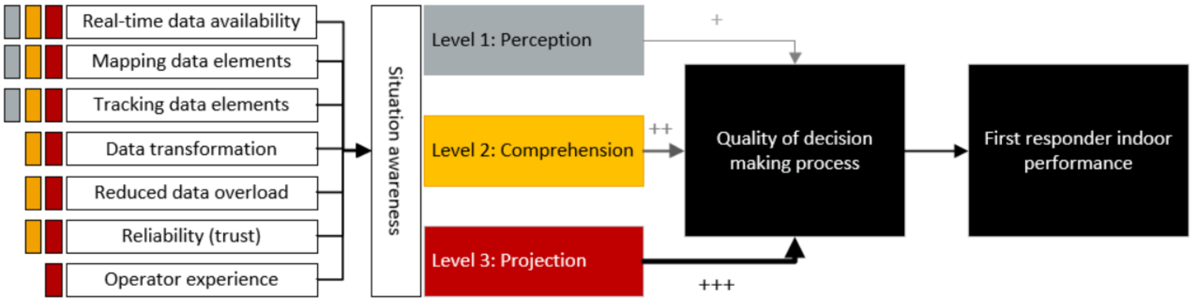


Figure 19: Conceptual model, describing requirements (left), situation awareness levels (middle) and influences on first responder indoor performance (right)

5 Methodology

This research will provide an answer to the central research question. This will be done by first providing answers to the individual sub-questions as defined in section 0. The methodology regarding the domain recognition has been covered already in section 3.2. This has been done so because the methodology of the domain recognition influences the theoretical framework.

The methodology is aimed to test the feasibility of a (near) real-time indoor mapping and tracking system for the creation of first responder situation awareness. For this, a proof of concept will be developed. The system will communicate the SLAM results from an indoor operation environment with a remote coordinator to create situation awareness.

The extent to which a system is trusted by the system operator is a large influencer for the creation of situation awareness. Therefore, a methodology for evaluating reliability in terms of accuracy, precision, and robustness of the developed PoC will be presented next.

Finally, a methodology for evaluating the extent of reached situation awareness will be proposed. This methodology is aimed at classifying the extent of situation awareness in one of three levels: (1) perception, (2) comprehension, or (3) projection.

The process is illustrated below. As it is mostly a linear research process, some results from previous sub-questions form the input for latter sub-questions.

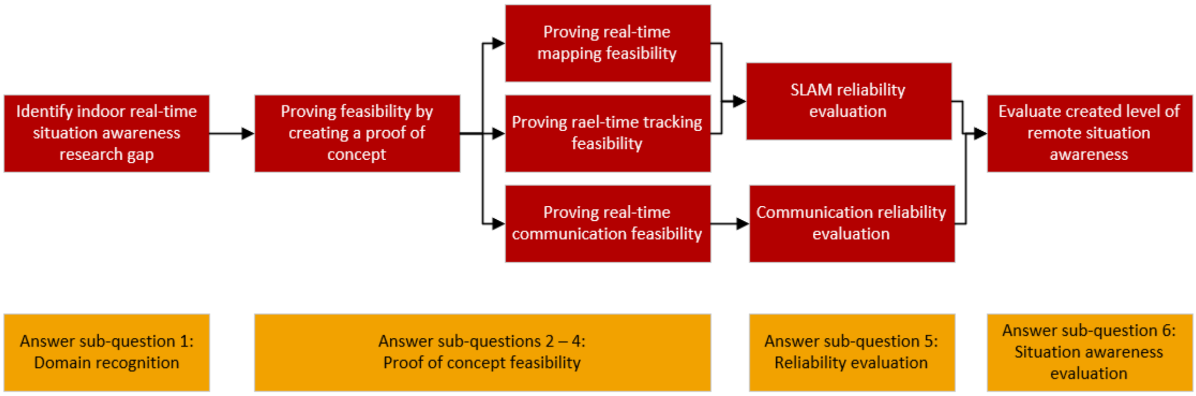


Figure 20: Step wise research methodology

5.1.1 Stakeholder involvement

As explained by Endsley (2016) and Wickens (2013), creating a user-centered system without stakeholder involvement is almost impossible. It is essential to identify tasks within incident management and to identify the data requirements related to these tasks. First, this is important to facilitate users with the data they need to accomplish the task with good performance. Second, the users should not be exposed to an abundance of data as this causes an information overload, exposing users to mental stress that hurts the extent to which situational awareness can be required (Endsley, 2016).

To involve stakeholders and design the system in a user-oriented way, three stakeholder meetings will be organized. The first meeting identifies user requirements. The second meeting will focus on evaluation of reliability. The third meeting will focus on evaluation of the integrated proof of concept in terms of situation awareness. The domain identification, reliability evaluation, and situation awareness evaluation can be regarded as phased objectives related to answering specific sub-questions. This phasing can be seen below. Also, a more detailed explanation of the set-up of the meeting will be given.

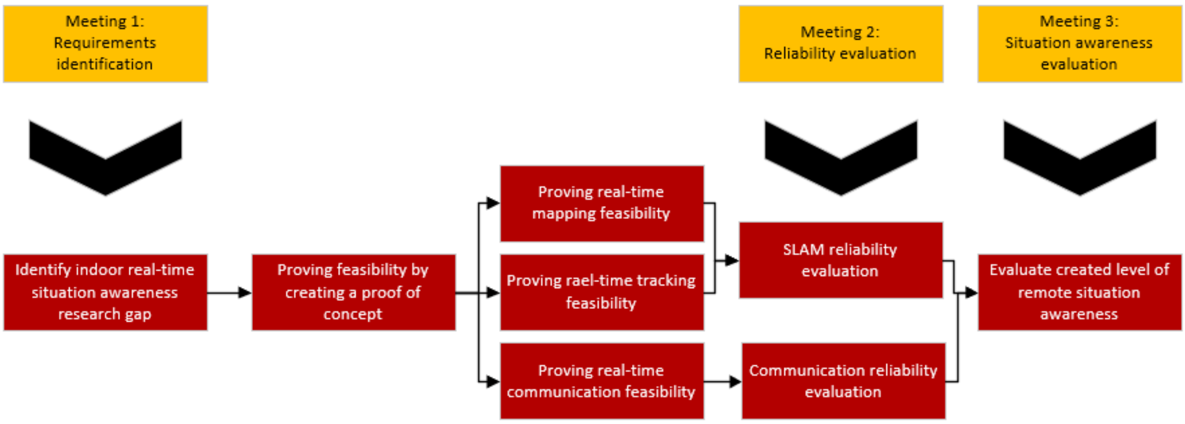


Figure 21: Phased stakeholder involvement

The first meeting will be a semi-structured interview with multiple officers of duty. The goal oriented formal identification of the first responder incident framework of (Zlatanova, 2010) will be used as basis to identify data requirements.

The second meeting will be a demo of the system, which will be planned halfway the research progress with a diverse group of first responders from different fields (firefighters, police, medical care). This meeting evaluates added value of current functionality and explores potentially missing functionality. This meeting will also be used to evaluate the level of reliability that the system should offer. Users will be asked to take on the role of the coordinator. This means they interpret the data presented by the PoC. Simultaneously, an environment is scanned, or such a process will be simulated. This meeting will test the perceived added value in terms of situation awareness in a subjective way, by observing the process and by discussing the added value of the proof of concept for indoor first responder operations.

At last, a conclusive meeting with a system demo will be planned to discuss the extent to which the system could offer situation awareness if deployed in the field. This meeting will be used as input for the final evaluation of the system in terms of added value to answer the central research question. For this test, an indoor first responder operation environment will be simulated as well, just as in the second meeting. However, this meeting will be organized with fewer first responders and is aimed on in-depth evaluation of the PoC. During the test, the officer of duty will be asked to give objective answers in according to direct testing methods. An example of such a question is to query the current floor level of the exploring first responder in the indoor environment.

5.2 Proof of concept Development

The proof of concept has two components: a component to be used within the indoor situation environment and a component that will be used to facilitate a remote coordinator.

A first responder equipped with a head mounted scan device will enter the indoor operation environment. This first responder is called the 'explorer'. This explorer collects spatial data elements regarding the environment, while the explorer is simultaneously tracked. All collected data elements are immediately transferred to the remote coordinator.

The remote coordinator is called the 'navigator'. This first responder has a coordinating function. For example, the navigator can be a commander of one or more first responder attack teams. The navigator application facilitates the visualization of an indoor 3D model of the operation environment. This view is combined with a track of the explorer within the environment.

5.3 System development

All modules of the proof of concept will be developed in the same development environment, which will be explained below.

5.3.1 Development device: Microsoft HoloLens

The development platform should at least be capable of mapping environments, tracking the device in the scanned environment and communicate these mapping and tracking results in (near) real-time in some wireless manner. As the Microsoft HoloLens fulfils these requirements it has been chosen as development platform.

- First, the Microsoft HoloLens has a built in SLAM system by utilizing its range camera combined with four tracking cameras. The mapping and tracking data elements can be processed in real-time by the onboard computer. (Hübner et al., 2019; Khoshelham et al., 2019).
- Second, the Microsoft HoloLens is capable of wireless data transmission by means of Wi-Fi and Bluetooth connections (Microsoft, 2020).

Besides the necessary requirements for mapping, tracking and communicating environment data, the Microsoft Hololens has other capabilities that are nice to have for first responder operation use.

- First, the Microsoft HoloLens is head mounted. This leaves the hands of the explorer free. This increases the mobility of the first responder, as the first responder is less restrained in removing rubble or climbing over obstructions.
- Second, the Microsoft HoloLens is a mixed reality device, meaning it has a screen to overlay virtual projected holograms with the real environment. This enables the device to provide the explorer with visual support. Applications of such visual support are indication of what has been mapped, menu

interaction, highlighting of important objects or features and even visual navigation aid by following a projected lifeline.

- Third, Microsoft facilitates developers with a Mixed Reality ToolKit (MRTK_V2). This toolkit provides basic building blocks for a wide range of platforms, including Microsoft HoloLens (both generations), Windows Mixed Reality Headsets and OpenVR headsets such as the HTC Vive and Oculus Rift. One of the building blocks is a basic SLAM implementation for the Microsoft HoloLens and an input configurator. The toolkit is well documented and it's functionality is accessible by use of different technology paths. Both the frequently used game engines Unity and Unreal are supported, as well as browser based WebVR experiences. Besides these options, it is possible to build own frameworks or middleware to access low-level device functionality by utilizing a DirectX implementation.

5.3.2 Development platform: Unity3D

The system will be developed by using the game engine 'Unity3D'. This game engine is frequently used for HoloLens development and 3D visualization of objects (Jana et al., 2017). The MRTK does provide basic functionality which the proof of concept will extent upon. The MRTK is written in the object-oriented C# programming language (pronounced as 'See Sharp'). This programming language will also be used for the proof of concept.

The language has been intended to be efficient in terms of memory and processing power requirements, although it is not competing directly in terms of performance with the familiar C language. However, it does support functionality such as automatic garbage collection, strong type checking, array bounds checking and the detection of attempts to use uninitialized variables. Therefore, it is generally perceived as a language that has a good tradeoff between simplicity and raw performance (Fourment and Gillings, 2008; Hejlsberg et al., 2003).

5.3.3 Simulator

No incident is the same, but for research and training purposes it should be possible to replay scans collected and transferred by the proof of concept. After each scan, all data collected by the explorer will be stored in a datafile. From this datafile, the data can be read and displayed appropriately in the navigator application. The timestamp created to depict any data transfer delay will be utilized by the incident simulator, making it possible to replay scans. The simulator will make it possible to alter the replay time by a dynamic multiplier or instantly show all the data. This will make it possible to test the added situation awareness of the system by showing the same situation to multiple navigators, with different visual effects or functions.

5.4 Mapping module

This section of the methodology aims to provide answers for the second sub-question: *To what extent may a head mounted augmented reality device be used to map first responder indoor operation environments in (near) real time?*

Priority is given to the spatial mapping capability of the system. As there is generally little information about indoor geometry, an environment should be scanned before the position of the explorer within the environment can be depicted. Below, the requirements for the mapping module are stated.

First, objective of the mapping module is to capture the indoor 3D geometry in real-time. This requires the system to continue scanning without hick-ups due to data processing. These hick-ups can be observed if the scanning application is unable to run at 60 frames per second.

Second, the stakeholders who have been involved in the requirements identification stated that mapping information would be beneficial, if they were able to quickly get an overview of the environment. Therefore, perceiving features in the mapping information is regarded more important compared to actual precision of the results.

Third, the stakeholders stressed the importance of being able to navigate within first responder environments. Navigation within the model by finding the shortest route from a certain point in the model to another point in operation environments is an important task for first responders (Fischer and Gellersen, 2010; Rantakokko et al., 2011; Seppänen and Virrantaus, 2015). For this, the concept of 'navigable space' is important, which is space that can be traversed normally within a building. We want to experiment with facilitating first responders in this task. Therefore, presentation of visual data will be combined with calculated navigable space if possible.

Last, the stakeholders asked for the option to add the positions of certain objects such as fire hoses and exit signs to the mapping information.

5.4.1 Mapping indoor environments in real-time

As (Khoshelham et al., 2019) show, the Microsoft HoloLens is able to capture its spatial environment. This mapping process is embedded in a closed SLAM system. As the SLAM system is closed, it serves to a certain extent as a black box. However, we can interact with the SLAM system by using the second version of the Mixed Reality ToolKit (MRTK_V2), which is a cross platform toolkit developed by Microsoft (Microsoft, 2020). Khoshelham et al also show this SLAM algorithm captures multi-room indoor environments without apparent deformation. Furthermore, accuracy deviates approximately five centimetres from a TLS scan and is therefore considered accurate.

However, mapping an indoor environment by using the depth-camera of the Microsoft HoloLens is a resource intensive process. Surfaces in front and up to about three meters away of the device are scanned, resulting in a point cloud (Hübner et al., 2020).

It should be noted that direct processing of the point cloud holds advantages in terms of resource efficiency (Linsen, 2001; Preiner et al., 2012; Richter and Döllner, 2014). Often, detail is lost by the transformation from point cloud to mesh representation. Also, updating a small part of a 3D mesh model requires reprocessing of the complete model instead of just removing/adding points from a point cloud, which is a resource intensive process (Richter and Döllner, 2014). Despite these point cloud advantages, the MRTK converts the point cloud into a spatial mesh. A mesh can be seen as a connected point cloud, with vertices which are connected by triangles (Richter and Döllner, 2014).

Of course, there are also advantages of a mesh representation opposed to a point cloud representation. A point cloud is unstructured by default, while a mesh is structured. This makes indexation and performing spatial queries less resource demanding (Richter and Döllner, 2014). Furthermore, meshes do in general provide a better way of perceiving, interpreting and interaction of 3D data compared to point clouds (Richter and Döllner, 2014).

The spatial 3D environment model is divided in surfaces covering about 1 to 1.5 cubic meter. This eliminates the need to rebuild the complete model upon every spatial awareness update. Each separate mesh has a unique identification number (ID). The interval of scanning for new surfaces can be set and ranges generally between 0.5 and 5 seconds for good performance (Microsoft, 2020). As only portions of the model have to be recalculated if a mesh is updated, resource needs stay low enough to maintain a high framerate of 60 frames per second.

The amount of points within the point cloud can be reduced by the MRTK compared to the number of vertices in the mesh. Developers can choose how aggressive this minimization of points should be: a higher number of vertices per cubic meter may represent reality better, but processing takes significantly more resources. This is because time complexity of mesh generation by Delaunay tetrahedrizations is O(n^n), with 'n' being the number of input points (Linsen, 2001). Therefore, the explorer application should consider a tradeoff between application performance, mapping precision and mapping update frequency (Microsoft, 2020).

The MRTK provides options to set the precision of the generated mesh to 'fine', 'medium', 'low' or a custom value of triangles per qubic meter (Figure 22). Quick mesh processing is important for the real-time application. Furthermore, the high quality setting has large implications for the processing time of mesh generation (Borycki, 2018). To account for real-time use, this research uses the 'low' setting with an update frequency of 0.5 seconds. These settings enable users to look around and scan the environment without having to wait for the depth sensor to update its view.

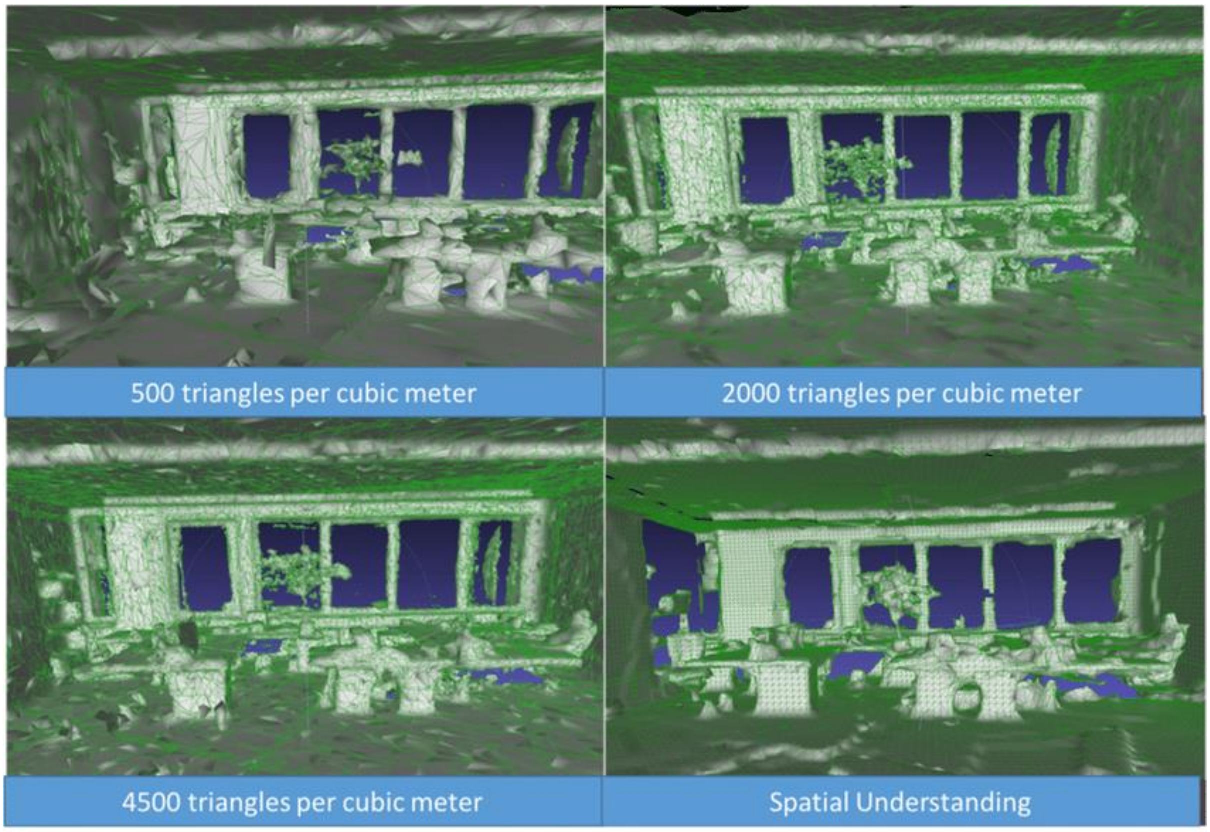


Figure 22: Spatial mapping of the hololens in different resolutions (Bondar et al., 2018). Spatial understanding is a post processing step of the 'HoloToolKit', unavailable for use during run-time.

The MRTK lets developers subscribe to events that relate to the creation, updating or removal of specific meshes. It does so by providing an unique identification (ID) number together with a 'spatial awareness mesh' every time environment data is changed (Microsoft, 2020). If a mesh is observed for the first time, the mesh is created by the MRTK. The event created by the MRTK is captured by the proof of concept by registering the object ID in a dictionary, creating a reference to the object. The proof of concept uses these references to address specific meshes. The coordinates of the mesh are transformed from local coordinates attached to a spatial anchor into a global coordinate system, after which the mesh is sent to the navigator application.

Several variables, such as the create time of a mesh, are not available in the spatial awareness mesh object. Therefore, a class 'SharedMeshObject' will be constructed for shared mesh representations between the explorer and the navigator applications. As the Microsoft HoloLens is a device with limited processing and memory resources, the explorer application will reference rather than clone objects of the MRTK to minimize the use of resources.

5.4.2 Mapping objects

As a result of the first user-requirement meeting with two first responder officers of duty, the importance of reference objects came to light. When navigating in an indoor environment, first responders use positions of landmarks such as lightswitches and exit signs as a reference in space (van der Meer, 2018). It would be a great help, so was said in the requirement identification meeting, if such objects could also be mapped. This would enable the positions to be stored as holograms, guiding first responders visually in a way that they already knew. Furthermore, positions of items like fire hydrants are important both for explorer and navigator, as such objects can be used to mitigate the effect of incidents (van der Meer et al., 2018).

As it makes sense to let the explorer add (object)information to the mapped environment besides the spatial awareness meshes, this functionality will be implemented. The explorer can make a gesture while pointing at a position on a mapped mesh in 3D space, after which a menu appears. From this menu, items such as a fire hydrant, exit sign or even victim positions can be selected and are added to the spatial mapping model with the specified position. Per object, a colored sphere appears in the 3D model of both the explorer and the navigator. This menu can also be used for other user input, for example to stop the system.

5.4.3 Privacy considerations

Basic privacy will be considered for the explorer application, as the mapping module involves scanning of the environment. Geometry is perceived to be the most important factor for situation awareness building. Therefore, the application will not save or send any RGB camera images or pixel output. Furthermore, due to the low mapping resolution, a sparse mesh will be created only showing basic geometry. These precautions should make identification of individual persons more difficult (ENISA, 2017). Furthermore, it hides sensitive information from observed documents and computer monitors which could be left in plain sight in case of an emergency.

5.5 Tracking module

This section of the methodology aims to provide a method to give an answer for the third sub-question: *To what extent may a head mounted augmented reality device be used to track first responders in (near) real-time within indoor operation environments?*

By using a world-scale coordinate system, the position and orientation of the explorer can be retrieved. The position of the device will be stored and sent every half a second, to be able to follow the explorer accurately. Furthermore, the orientation of the device is stored with every position. By using both the position and orientation of the explorer, the navigator can see where the explorer is and in which direction the explorer is looking. The position (x, y, z) and the orientation (roll, yaw and pitch angles) are stored in a list of poses. Depending on a user defined update time, the registered pose is sent to the navigator every so many seconds.

Within this research, the update time is 0.5 seconds. This means if the explorer application can store and send its pose continuously every 0.5 seconds, the explorer application is regarded successful in tracking a first responder in real-time within an indoor environment.

5.5.1 Explorer implementation

The Microsoft HoloLens renders its environment based on a so called 'stage' (Microsoft, 2020). The stage could be perceived as a cartesian coordinate system, in which the device and holograms positions are estimated. There are different stages for different use cases, depending on the scale of the use case. For room sized stages, it is assumed the positional drift of the device is small enough to relate all positions to a single reference point in the coordinate system: the origin position (Microsoft, 2020). However, the proof of concept will be built will be used in large multi-room or even multi-floor indoor environments. Because of the size of the stage, a single reference point in the coordinate system would make pose estimation of the device difficult resulting in pose drift. A SLAM system with landmark recognition can reduce the error of pose estimation as explained in section 4.3.3: SLAM basics.

Unsurprisingly, the Microsoft HoloLens uses a SLAM algorithm to correct for pose drift (Khoshelham et al., 2019). Although the HoloLens SLAM algorithm itself is largely unpublished due to the proprietary rights of Microsoft, the mixed reality documentation enables researchers to use it and to estimate what is going on.

5.5.2 Spatial anchors

At the centre of the world-scale experience performance, there are spatial anchors. These anchors represent SLAM landmarks as discussed in the literature review (section 4.3.3.1) in the way that they recognize distinctive features in space. From these spatial anchors, local coordinate systems define where holograms should be placed if they are in close proximity (smaller than a couple of meters) of the anchor. The anchors relate to each other, to the device pose, and to the origin of the coordinate system in a spatial way. By measuring distance to these spatial anchors, the device can know its location relative to the spatial anchors and correct for estimated drift caused by the IMU. Furthermore, because the spatial anchors are connected to the origin of the coordinate system, the device knows its relative position to the origin. Just as SLAM landmarks, the relative estimated positions of the anchors are constantly updated.

Microsoft does provide information about the nature of features used to create spatial anchors and to recognize them again in space to close loops. Microsoft provides this information to create transparency about privacy considerations, but it can be used to create insight in how spatial anchors are captured, processed and stored. According to the documentation, both geometric and visual characteristics are captured and processed into a sparse point cloud of unique points. Per point, the characteristic data of these points are processed subsequently stored as a unique hashed string. As can be seen below, the representation of a spatial anchor is rather abstract.

Figure 23: Left: real world scenario. Right: derived spatial anchor (Microsoft, 2020)

5.5.3 Tracking loss

Although a SLAM algorithm is implemented by the MRTK to estimate poses in the real-world, it is possible the device loses its reference system (Microsoft, 2020). After tracking loss, all content becomes pose locked instead of world-locked and all spatial meshes will be removed from view if tracking is regained. According to the documentation, such tracking loss can be experienced especially if the following (combination of) aspects are apparent (scooley, n.d.):

- Lighting conditions are too bright, too dark or lighting conditions change too sudden
- A room with strongly reflective surfaces
- Landmark poor environments, such as a hall without a lot of distinctive features
- Places that look very similar, such as office spaces with the same interior for every floor
- Movement in place, for example in crowded areas
- Rooms without Wi-Fi connections, as Wi-Fi fingerprinting enables the device to recognize spatial anchors more quickly

Preventing tracking loss on a device level would require improving the SLAM algorithm of the Microsoft HoloLens. This is out of scope, as it would require a lower approach to the device. Therefore, the limitations as stated above should be considered when scanning. Furthermore, tracking loss will be an important aspect within the reliability tests which will be discussed later in this research.

5.6 Communicating module: Data transfer

This section of the methodology aims to provide a method to give an answer for the fifth subquestion: "*To what extent may a head mounted augmented reality device be used to communicate in (near) real time about mapping and tracking information of indoor first responder operation environments?*"

As explained in the previous section, a method for dynamically capturing indoor first responder operation environments with a Microsoft HoloLens will be developed. This data collection will be taken care of by the 'explorer' application of the system. The 'navigator' application will receive data from the explorer application and present it to a first responder with a coordinating role. Although there is relevance in the spatial mapping process in itself, this research aims specifically to facilitate a first responder coordinator with situation awareness. The following section will describe the receiving, processing and presentation process in more detail.

5.6.1 Communication protocol

To provide the navigator with up-to-date data, the communication protocol is of upmost importance. From literature, it can be derived that first responder operations are highly dynamic (Kapucu and Garayev, 2011). Furthermore, operation coordinators have to trust a system to deliver reliable data in order to create situation awareness (Endsley, 2016). Therefore, the following requirements have been set up for the communication protocol:

- First, the location of the navigator should not matter. Therefore, a data transfer connection via the internet is due to its global availability preferred.
- Second, the communication protocol should be quick, transferring collected data within 10 seconds if connected.
- Third, the package transfer from explorer to navigator application should not be blocking the scanning process of the explorer application: the explorer app should continue with data collection even if there is a momentarily disconnection of data transfer capabilities.
- Fourth, if reconnected after a disconnection, the explorer application should send the data it could not send before to keep the model as complete as possible.
- Fifth, package loss should be minimized.

The HoloLens has the availability of two kinds of wireless data transfer: via Bluetooth or via Wi-Fi (2.4 and 5 GHZ bands). A direct connection through Wi-Fi makes most sense as a requirement is an internet connection. However, Wi-Fi may not be available or may not have full-building coverage. Therefore, a secondary device with a mobile internet connection (preferably 4^{th} or 5^{th} generation) shall be carried by the explorer. This device sets up a 5 GHZ hotspot connection to which the Microsoft HoloLens can connect to gain access to the internet. The navigator will connect via a Wi-Fi or mobile internet connection as well, for example through a SIM enabled laptop. Different network configurations will be tried until a suitable connection is found. First, a cost-free web socket will be tried.

5.6.2 Navigator device

The test device for the navigator application will be a laptop with an Intel core I7 processor, 4G SIM mobile internet connection, Nvidia Quadro p1000 graphics card and 16 GB of RAM. The data will be received in the editor of Unity. This offers more flexibility compared to building a native windows application, as we do not have to worry about building interactive elements. The touchscreen will enable the laptop to lay flat on a table with the screen pointing upwards, making it well visible from all viewing angles. This is done to make collaborative decision making easier with the aim of creating shared situation awareness (Kapucu and Garayev, 2011). Due to the increased processing power of the navigator device, the navigator is able to process the environment data much quicker compared to the Explorer device.

5.6.3 FIFO queue versus LIFO stack data structures

The data packages send from explorer application to the navigator application will be addressed to as 'messages'. A message might consist out of a single spatial mesh, a single pose or a single command. To reduce the size of the message, objects are first serialized into byte strings before they are sent.

As multiple messages may be generated almost simultaneously, the networking protocol might need some time to process the transfer of the messages. All messages will be sent, but some message transfers will have to be prioritized over others. A choice has to be made between prioritizing messages. Meshes, poses and commands are treated equally: only the time of creation matters for the message priority. A list of items can be ordered to creation time in two ways: a queue, with a 'first in, first out' (FIFO) structure or a 'stack', with a 'Last in, First Out' (LIFO) structure.

Both message priorities make sense for first responder use. FIFO would prioritize messages that have been created most recently, therefore showing the direct environment of the explorer and its most recent pose first. This might make communication about that environment easier, for example by a telephonic communication device. On the other hand, a FIFO priority could create gaps in the data that are only filled up later on. Therefore, a way out of the building might not be directly visible and important commands like victim identifications might be missed. Furthermore, the growing of the 3D model seems more natural if it is continuous, which might it make more easier to interpret and might thereby reduce information overflow. Because of the wish to show the model as complete as possible, the system will prioritize the messages in a FIFO manner, showing messages that are sent first also first on the navigator application.

However, if using a queue, the navigator should be made aware of any delays in receiving messages. This is why a timer will be shown displaying the time in seconds that the last received message was sent by the explorer. For this purpose, a timestamp will be added to each message.

5.7 Mapping presentation

As soon as a mesh, pose or command message is received by the navigator application it will be visualized for the navigator in the growing 3D model, to create a Common Operational Picture. The goal of this presentation is to inform the navigator about the state of the situation. This should require minimal mental effort, with the ultimate goal to create situation awareness.

It is assumed the navigator application runs on a device with a more computing resources compared to the HoloLens. The aim of the communication module is to make the indoor 3D model well interpretable. (van der Meer, 2018) gives design principles for designing for first responder indoor models.

The 3D model is collected by the explorer, sent to the navigator, and immediately visualized. This means there is no manual step between collecting and visualizing the results. Therefore, all data presented to the screen will have to be generated by code. This is for example different from the concept used by Van der Meer (2018), who could select and change elements within a BIM model. Because of this, basic universal rules will have to be applied to visualize the model by code.

In computer graphics, screen elements are visualized pixel by pixel by a so called 'shader'. Shaders are computer programs used for programming visual effects such as scene lighting, darkness and color in a rendered image and are usu. The basic rules that will be defined for visualizing elements in the 3D scene will be implemented within the shader. There is little academic relevance in explaining in the way shaders work, for example (Zucconi and Lammers, 2016) give a good overview of the possibilities of shaders. However, there is academic relevance in the way shaders are used to visualize dynamically created 3D models for creating situational awareness, as there is little research on this topic.

Other indoor mapping research that are using the Microsoft HoloLens as a 3D capturing device do not pay attention to the visualization of the model (Hübner et al., 2019; Khoshelham et al., 2019). This research will extent on their research in a way that geometric features are used for visualization of the model to make interpretation easier and to reduce information overload. The foci of the mapping characteristics are taken from (Kraak et al., 2013) as explained in section 4.4.4. These characteristics are visualization of spatial information in location space, attribute space, and temporal space.

5.7.1 Geometry focused

Location space-based visualization is focusing on spatial representation of an object. It answers the 'where is something' question (Kraak et al., 2013). However, we have seen location is ambiguous (see section 4.3.4.1). Therefore, we rather use the mathematical term geometry to describe the spatial attributes of the model, saying our location space visualization is geometry focused. In this context, geometry is explained as the mathematician of space. In other words, it is concerned with relative position, shape, size, and spatial properties such as the surface normal of an object (Risi, 2015).

The proof of concept does not collect any color information from its environment. This is done to reduce required processing resources, minimize the amount of data transferred and to consider potential privacy violations. This means only geometry is captured by the proof of concept. Within the geometry, there is the element of orientation of a surface. The orientation of a surface is a large influencer for the way light reflects on it and therefore, a large influencer in how the surface is perceived. In the 3D model, a normal vector representing the orientation of a surface can be calculated for every triangle within the spatial mesh. The first visualization model will present the mesh geometry as raw as possible in a scaled black-white representation. The normal of a triangle will be used to color horizontal triangles as white and vertical triangles as black. Any value in between will be interpolated, therefore depicting a grey scale color. Other spatial properties, such as relative height, can be used to further distinguish features.

The geometry focused visualization is created with the idea to make little assumptions about the environment. This means the system lets the system operator interpret the model almost entirely on its own, without support of the system. This makes it a reliable way of visualization, as there is little risk of information bias created by the system. However, it is likely to require a lot of mental resources of the operator since the system does little to support the system operator and minimize information overload. Therefore, reaching a higher extent of situation awareness (second or third level of situation awareness) may be more difficult by using this way of visualization.

5.7.2 Time focused

Temporal space focuses on the question: 'when is something' (Kraak et al., 2013). This question is important for first responders because first responder operations are dynamic: circumstances change over time. As circumstances tend to change, the reliability of perceived data elements decreases as time progresses. Mapping a building takes time, as the whole building cannot be perceived from a single viewpoint. This causes a difference in age of specific spatial mapping data elements. Although the aging of spatial information is inevitable due to the scanning capacities of the device, a navigator should be able to request the age of a mesh. A visualization focused on depicting spatial mesh age will be developed to take care of this capability.

Scan time is linear and therefore there is always an order available (Kraak et al., 2013, p. 157). This order will be used to display time combined with the spatial mesh. The time-based visualization will take the geometry focused visualization as basis. To this basis representation, time can be added by adding color. Color can be added either be applied stepwise or in a scaled gradient, representing the variable continuously or ordered.

A continuous scaled gradient can depict time in a precise way, for example changing from white to black over an hour with an interpolated grey scale for each second. However, although the color depicts the variable very precisely, it is unlikely this precision is also perceived by a user due to the minimal differences between color values. Furthermore, a scaled representation would cause the model to be continuously changing, asking more attention and therefore mental resources of the operator (Endsley, 2016; Kraak et al., 2013; Wickens et al., 2013).

As a user-centered system design should prevent information overload (Endsley, 2016) and because perceiving small changes in time are deemed unimportant for this use-case, an ordered, step-wise visualization of time will be implemented in the system. The spatial meshes will change color at the arbitrary chosen threshold values of 3, 5 and 10 minutes on a color scale from white to red.

Another use of a time focused design would be to depict change: for example, a transition from one state into another if a spatial mesh is updated. This has been declared out of scope (section 2.2.2.10).

5.7.3 Object focused

Attribute space is focused on answering the question 'what is something' (Kraak et al., 2013). By identification of doors, walls, stairs, floors, and obstructive objects, some of the raw geometric data is translated into a representation with more meaning. By doing so, the system interprets some of the data with the aim of relieving the system operator of mental stress. An object-oriented visualization method will be developed to focus on distinction between objects and navigable space.

Van der Meer (2018) describes the importance of removing visual clutter from indoor models for first responder use: not all information is of importance for a navigator. This notion can be linked to the statement of Endsley (2016) who says system operators should be protected from information overload. The main reason for which first responders need the mapping functionality of the system is for navigation purposes, for example moving via the quickest route to an objective or evacuating a space by using the quickest route to the exit (van der Meer, 2018). To know how first responders can operate within an indoor environment, the concept of navigable space is therefore important. The navigator application will be built in a way that is focused on the user need of quickly perceiving (in)navigable space.

From Van der Meer (2018) it was derived that geometric classification of floor levels, walls and stairs/ramps are important for first responder operations. Within the object class, no further distinction will be made. At last, closed doors will be identified to prevent them from obstructing the navigable space. This results into the following list:

- Floors
- Walls
- Stairs/ramps
- (closed) doors
- Obstruction objects

Walls, floors, objects, stairs and doors will be identified by using the preprocessing steps as described in the sections below. The indoor features of walls, obstructive objects, ramps and floors will be visualized by using a single shader. The mesh objects of these features will not be classified themselves, as they can contain multiple features.

An exception on this rule is the representation of stairs/ramp meshes. They are not identified by using the mesh geometry alone, but by also using the track.

5.7.4 Recognizing floors

First, one should consider multifloored buildings. Although a complete 3D model of a building has its purposes, one may want to zoom in to an overview of a separate level (van der Meer et al., 2018). This may be useful for viewing the position of the explorer without clutter of other floors and may also be used to follow the explorer on a (2D) map (van der Meer, 2018). The application will enable the navigator to show or hide floors with a single press of a button. Therefore, spatial meshes need to be segregated based on floor level.

To divide a 3D model into floors, one could observe the vertical surface histogram of a model. As floors and ceilings are usually horizontal, they provide a large surface for a specified height in the model. From this, floors and ceilings can be extracted (Okorn et al., 2010). However, such a method might not be robust enough for non-Manhattan buildings, in which floors and ceilings might not be consistent or completely horizontal (Nikoohemat et al., 2020).

The proof of concept will use a method inspired by (Díaz-Vilariño et al., 2017), stating that the scan trajectory and scanned surfaces are related to each other. This research will use timestamped explorer positions to relate surfaces to a floor level, utilizing the position of the explorer device (Figure 24) at the time of observing a spatial mesh. If a spatial mesh is created or updated, it is always observed from a certain point in space: the position of the explorer device. As the explorer moves around space, the explorer is always standing on navigable space when observing a spatial mesh. This means that an offset between the floor height and the height of the explorer device can be determined.

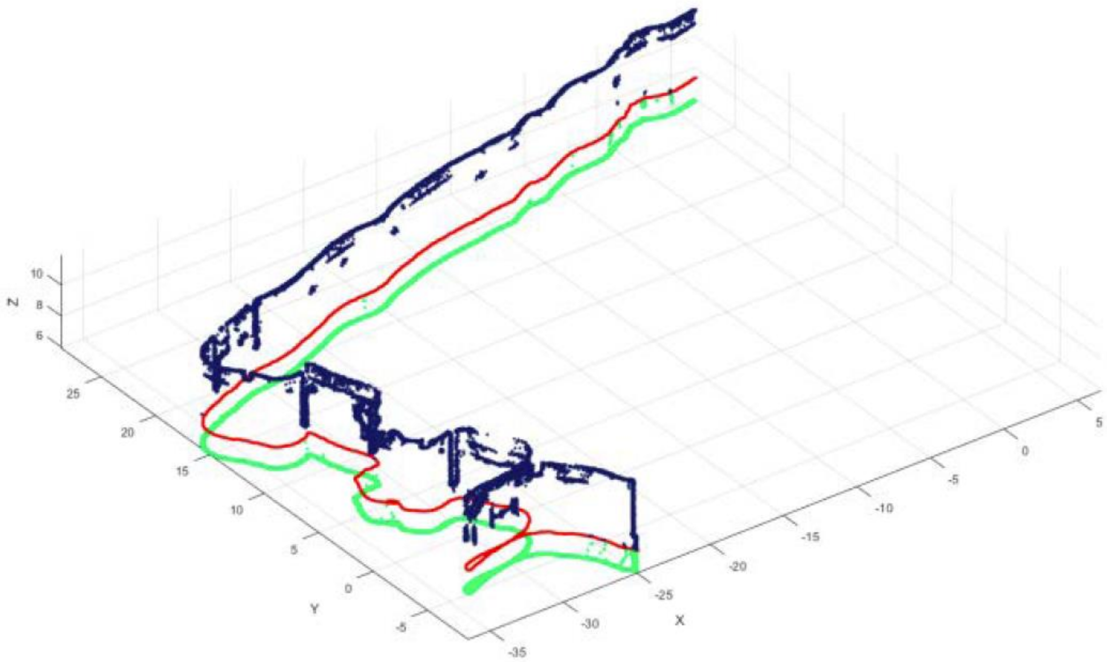


Figure 24: MLS trajectory (red) with floor height (green) and ceiling height (blue) directly below and above the MLS track (Díaz-Vilariño et al., 2017)

Therefore, two assumptions are made. The first assumption is that navigable space is almost always a floor and not an object like a table or a chair. The second assumption is that the floor surface directly beneath the explorer is mapped just before it is traversed, as the scan direction is usually forward. Following these two assumptions, a virtual ray can be cast downwards from the 3D explorer position every time that a position is received. The ray length is limited to two meters: the maximum length of the explorer. If a surface is found directly beneath the explorer, the relative height of the intersection point between surface and explorer position is taken to calculate the height of the intersection point relative to the center of the global coordinate system of the model. The resulting height value is stored as the last observed floor level and kept until a new floor height is observed. In the meantime, the last observed floor height is added to all observed spatial meshes. The variable be used to cluster meshes on observed floor level by looking for significant height differences between observed floor surface levels.

Logic: floor = surface AND below explorer AND vertical surface normal AND slope < 25 degrees

5.7.5 Recognizing walls

Although it is recognized that buildings are not always Manhattan shaped (Nikoohemat et al., 2020), the assumption is made that most walls are at least placed vertical in space. As walls are vertical surfaces, the normal vector (the perpendicular vector of the surface) of the wall surfaces are horizontal. Surfaces with a horizontal normal are therefore classified as walls. It should be noted, of course, that most walls have a horizontal normal vector but not all surfaces with a horizontal normal vector are walls. For example, a closet does often also have a large surface with a horizontal vector. This is a distinction which will not be solved, as the assumption is that a navigator will be able to interpret the walls separately from other vertical obstructions.

Logic: wall = surface AND NOT below explorer AND horizontal surface normal

5.7.6 Recognizing stairs and ramps

The assumption is made that the only way to move from one floor to another is by use of a stair or a ramp, which makes these floor transition paths of great importance for first responders (van der Meer, 2018). Therefore, they are recognized in the navigator application. Ramps will be assigned to a floor, as it a gradual transition from one floor to another. A significant height breakoff point can be used to divide two floors separated by a ramp. Stairs are more common compared to ramps to separate floors in buildings. These surfaces will be recognized to aggregate them to one floor (either the bottom or the top floor) or to a separate object. The slope of the explorer path is used to detect stairs: if the slope has an absolute value 25 or more degrees for the arbitrary duration of 1.5 seconds, surfaces observed from the start to the end of the sloped track are classified as 'stairs'. These objects can be connected to a floor or treated and visualized as separate objects.

Logic: stair/ramp = surface AND below explorer AND track slope >= 25 degrees

5.7.7 Recognizing obstructive objects

In indoor environments, there are all sorts of obstructive objects. An office area has for example a lot of desks and chairs, while a factory contains large machines that obstruct movement. An abstraction of the object is made in which the exact nature of the object does not matter: the objects are assumed to be immovable and therefore only their obstructive nature matters. This is not a real-world scenario, as chairs are for example more easily moved compared to a desk. However, object classification is not within the scope of this research as it is assumed that the classification would be considered too unreliable. Although objects will not be classified by The proof of concept as a supportive function, operators can perceive the object themselves on a visual basis to estimate what the nature of the object is and whether the object could be easily moved.

As objects are considered to be immovable for the calculation of navigable space, all surfaces that stand on the floor are determined as being an obstructive object.

Logic: obstructive object = surface AND above floor height AND NOT horizontal surface normal

5.7.8 Closed door detection

Space is navigable if it is below the explorer and the surface is not obstructed. An open door does not obstruct the navigable space. However, due to the way walls are recognized, there is no way of making a distinction between a wall and a closed door based purely on the geometry. Just as (Díaz-Vilariño et al., 2017) did, doors will be recognized by using both mapped geometry and the track of the explorer. If the geometry and the track intersect, a door will be placed at the floor at the point of intersection. The intersection of track and surface proves that the door can indeed be opened instead of being locked, which holds an information advantage over recognition of doors with computer vision.

Logic: closed door = surface AND above floor height AND horizontal surface normal AND track intersection

5.7.9 Calculating navigable space

First responders benefit from goal oriented navigation, in which the system guides first responders through indoor spaces (Fischer and Gellersen, 2010; van der Meer, 2018). To calculate routes through space, navigable space needs to be defined and calculated (Flikweert, 2019). From this navigable space, navigation graphs can be extracted that can be used for quick routing. Extracting navigation graphs from the 3D model is out of scope for this research, but calculating and visualizing navigable space is within the scope of the research. Navigable space is taken within scope due to the opportunity it offers for indoor routing and running evacuation simulations, but also as it is estimated it could be a good proxy for estimating scale, room clutter and thereby perception of suitable escape paths.

By using the definitions of floors, walls, stairs, obstructive objects and closed doors we can calculate space that is navigable or not. For a space to be navigable, it should have a certain surface size. Furthermore, it must not be obstructed by objects, walls, or ceilings. Ceilings are not classified, but a check will be done at potential navigable space if the space above the floor surface is free for at least two meters: the height of the explorer. Closed doors are navigable, as proven by the track intersecting with the door surface.

Logic: navigable space = Floor surface OR Stair OR Ramp OR Door AND NOT Wall AND NOT Object

5.8 Tracking presentation

The latest received pose will be visualized with a blue sphere in three-dimensional space (or blue dot in 2D representation), as is a usual default representation of position in frequently used mapping applications such as Google Maps and Esri applications. A trail will run through former poses of the explorer, changing colors in a gradient based on a temporal scale and indicating both speed and direction. Both colors and temporal scale should be easily changeable by the user, as different users might prefer different representations within different operation environments.

5.8.1 Pose estimation versus positioning versus localization

As discussed in section 4.3.4.1, pose estimation is often confused with localization. The proof of concept could work with both, either presenting the pose (position + orientation) of the explorer or the place (location) of the explorer. The pose would be presented with a dot (or sphere, in 3D space) together with an optional viewing direction or viewing frustrum. Location would be represented by a space indication, for example by coloring the room in which the explorer is present. There are advantages to both presentations of the track: the pose would present the first responder tracking seemingly more precise, while coloring the room in which the explorer is present may be quicker to be observed by the navigator. A hybrid solution, showing both pose and location could also be presented.

Due to practical reasons, only a series of explorer poses will be presented to the navigator. A list of poses is easily composed, while determining room boundaries require the system to make assumptions about the difference between obstructions and room borders. The latter solution is more prone to error, as the pose estimation is guaranteed by a certain reliability due to the 'tracking loss' functionality within MRTK (as described in section 5.5.3: Tracking loss).

5.9 User interface

According to (van der Meer, 2018): "The firefighters also demanded interaction with the 3D model. They wanted to zoom, rotate and pan to certain locations.". The proof of concept will provide an operational dashboard in which data is presented and in which interaction with the model is possible. The dashboard will be created within the unity editor.

5.9.1 Scene view

The heart of the dashboard will be the 'scene' view. This view enables users to move around the 3D model by changing viewing positions dynamically. It does also enable users to zoom in onto features and spatially select and pan to features, such as a specific spatial element such as a mapped body or to the explorer game object. It will also be possible to (de)activate visualization of screen elements, such as specific floor levels.

5.9.2 Hierarchy

All elements of the 3D model, being spatial meshes, the explorer pose and mapped objects like victims, will be listed in the hierarchy section of the dashboard. Here, elements can be selected, and elements can be activated or deactivated, either showing or hiding them.

5.9.3 Inspector

Some modeled elements are enabled for user input, like changing the visualization method of a spatial mesh. Such options will be available in the inspector screen of the dashboard.

5.9.4 Cameras

A navigator may not be fully engaged with the proof of concept. A situation may require the navigator to focus on other aspects of emergency management like sharing knowledge with other first responder decision makers. This may distract the navigator from the proof of concept. To enable a quick overview of the model, three cameras will offer a generic view of the situation environment.

First, a top down view of the environment will be presented in a separate screen on the dashboard. As this is a top down view, the three-dimensional nature of the model will be lost. Instead, a two-dimensional overview of the environment will be visible, enabling a navigator to quickly glance and observe the environment. The importance of such a 2D overview is further explained in (van der Meer, 2018).

Second, a 'first person view' of the explorer will be visualized in another camera screen within the dashboard. Bodycams are frequently used in first responder operations, streaming RGB pictures or videos from a explorer to a coordinator. For correct transfer of these images, a reasonably fast connection from explorer to navigator is needed. As the last pose of the explorer is known and the environment is being mapped, it is possible to visualize a view from the pose of the explorer with data that is already known to the navigator. If this visualization appears to be adding situational awareness, it could be used as an alternative for a bodycam.

Third, a side view will be displayed in a third separate camera screen within the dashboard. This is presumed to add multiple floors of a situational awareness in a good perspective.

All cameras will move to 'follow' the explorer by keeping the last explorer pose in the center of the camera screen. This requires no further interaction of the navigator and aims to offer a low-cost overview alternative to scene navigation.

5.10 Reliability evaluation

This section of the methodology aims to provide a method to give an answer for the fourth sub-question: *To what extent may the reliability of the mapping, tracking and communicating capabilities of the proof of concept support first responder decision making for indoor operation environments?*

System operators need to know the capabilities and limitations of a system in order to trust the system (Endsley, 2016). The quality of a scan depends on a very large scala of factors (Luhmann et al., 2013). Reliability is one of the factors that are important for first responders (Seppänen and Virrantaus, 2015). For this thesis, reliability is understood as a construct consisting out of accuracy, precision, and robustness. These factors will be first tested individually. After this individual test, the results will be integrated in one conclusion. Stakeholders will be involved in the evaluation of the reliability results as illustrated in section 0:

Stakeholder involvement. These subjective measurements will be integrated in the conclusion as well. Finally, the conclusion will answer the fourth sub-question.

5.10.1.1 Accuracy

In practice, accuracy is often used as a term indicating overall quality. However, this thesis defines accuracy as the extent to which a scan resembles the real-world (Van der Ham, 2015) and therefore accuracy can only be compared with reference data of higher accuracy (Luhmann et al., 2013, p. 92). As a SLAM algorithm is used, the accuracy of mapping and tracking data are assumed to be highly correlated: if tracking is inaccurate, the mapping data will also be inaccurate. Therefore, only the accuracy of the mapping module is evaluated.

To test the accuracy, the scan results of the proof of concept will be compared with a 'ground truth' of the environment. The proof of concept scan results will be compared with a high resolution 3D point cloud, in a way derived from Hübner et al. (2019) and Khoshelham et al. (2019). The evaluation is repeated to check whether any drift is added by using a real-time implementation of the Microsoft HoloLens mapping capabilities. Also, buildings with multiple floors will be considered. To our knowledge, such a vertical evaluation of Microsoft HoloLens mapping accuracy has not yet been published.

Of course, to calculate distances between our measurements and the real world, a ground truth is necessary. This ground truth is offered in the form of a LiDAR point cloud, scanned at the same moment as the proof of concept scan. Below, processing steps are described to compare the proof of concept scan with the ground truth scan.

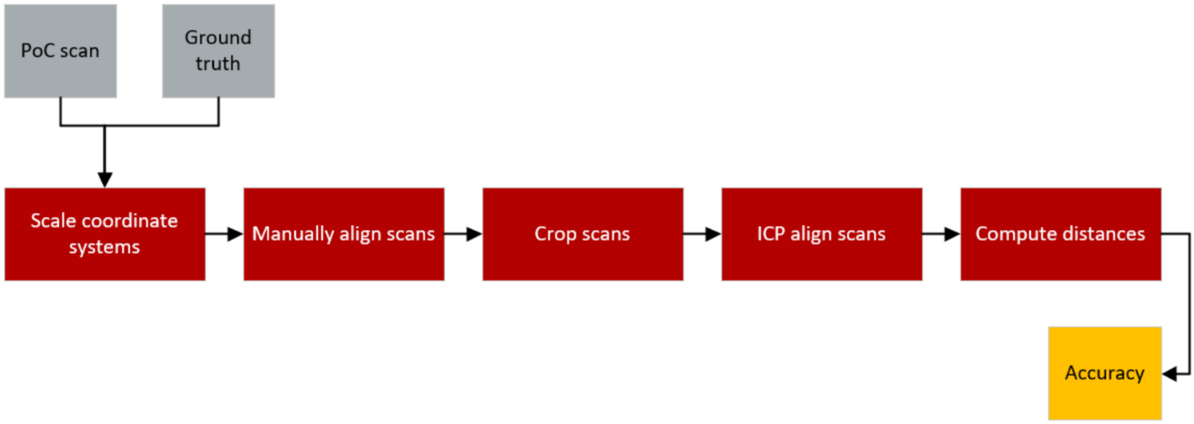


Figure 25: Comparing accuracy of PoC by ground truth distance comparison

To compare the scans, the proof of concept scan results will first be scaled to be 100 times smaller, to match the centimetre measurements of the MLS scan. Both scans have a metric coordinate system. However, the MLS used for the ground truth scan uses centimetres as units and the proof of concept uses meters.

Next, both scans are overlayed in the program 'CloudCompare', version 2.10.2 (2019). The scans are manually matched based on visual interpretation. Matching will be performed in six degrees of freedom, meaning the rotation and the orientation of one of the scans will be altered to overlay the other scan. The scans are not transformed, as scale is left untouched.

If the coverage (area that has been scanned) of the ground truth and the proof of concept scan is different, the largest scan extend will be cropped to the smallest scan extend. Otherwise, measurements of the largest scan would be matched to boundary points of the smallest scan. This would result in unrealistic results: namely an error that is far to large.

After manually roughly aligning the scans, the Iterative Closest Point (ICP) algorithm as implemented in CloudCompare will be used to closely align the scans. (Pomerleau et al., 2013) describe the ICP matching algorithm in detail. Again, this matching will be performed in six degrees of freedom, leaving the scale of the scans intact.

Finally, the (mean) distance between the proof of concept scan and the ground truth scan will be calculated. For this, the 'cloud/mesh distance' function of CloudCompare will be used.

5.10.1.2 *Precision*

Visual inspection of the results will be used to report on the precision of the proof of concept. This evaluation will be done together with stakeholders, to research to which extent the results are interpretable. This interpretation will be discussed with in the stakeholder involvement sessions.

5.10.1.3 *Robustness*

Robustness is understood as the way in which the proof of concept performs continuously in different operating environments. For this reason, different types of environments will be scanned and visually compared with each other. Furthermore, findings of tracking loss and connection failures will be described in the results.

5.11 Situation awareness testing

The sixth sub-question: "*To what extent may the proof of concept improve first responder indoor performance by providing situational awareness to support first responder decision making?*" will be answered as last.

The central research question has been segmented into five parts, as the components are expected to give a reasonably well idea of how the proof of concept would perform in first responder indoor operations. However, the performance of a system is more than the added value of the individual components. Therefore, the proof of concept will be evaluated as a whole, resulting in an answer to the sixth sub-question.

Because of this, the results of the domain recognition, feasibility, and reliability related sub-questions will be considered. As illustrated in the conceptual model (section 0), we regard 7 main elements to create situation awareness in indoor first responder operations, related to mapping and tracking:

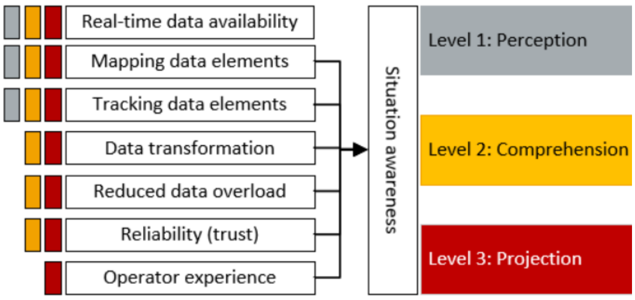


Figure 26: Key requirements for creating situation awareness, for different levels

Based on these aspects, we will provide argumentation whether the creation of remote situation awareness succeeded and to which extent. The aspects will also be used in a stakeholder session. This stakeholder session will demo the proof of concept, to discuss the added value of the PoC. Both the listed evaluation and first responder meetings will be used indicate whether the proof of concept is able to create level 1, 2 or 3 situation awareness as described by (Endsley, 2016).

6 Results

This chapter will provide results as input for answering the central research question. The sub-questions have been grouped in section 0:

Research objectives.

6.1 Domain recognition

For readability, the results related to the domain recognition have been provided in section 3. They will not be repeated in this section, but will be used to answer the first sub-question: *To what extent has the relation between 'indoor mapping', 'indoor tracking' and 'first responder decision making' already been established by the academic community?*

Furthermore, a meeting with stakeholders has been organized. Two officers of duty were willing to give context to findings in theoretical framework. Furthermore, they were asked to help in identifying requirements for the methodology. The results of the meeting are stated below.

6.1.1 Stakeholder meeting: requirements identification

The first meeting was held with two officers of duty of the fire department of the safety regions. The officers of duty can perform coordination tasks on different GRIP levels if an incident occurs, which is why they can provide context and feedback on different scales of emergency management. One of the officers has been closely involved with the work of (van der Meer, 2018), which means the officer was already familiar with the concept of mapping indoor environments from a spatial information perspective.

6.1.1.1 Current practice

First, current practices of indoor emergency management were discussed. Van der Meer (2018) describes these practices in greater detail. The most important finding of the discussion considers the current availability of spatial data: even for vital, public buildings there is little indoor information available. The safety regions have employed people who are tasked with making 'availability maps' of indoor environments. These maps contain indoor geometry and important objects like fire separators. Often, only the first floor of a building is available in such a way to the first responders. If an incident occurs, first responder use this information to get into the building and find their way across the first floor. If they go up or downstairs, they tear evacuation maps from the walls that are available in the building itself. Of course, these information sources are drawn up in two-dimensional space while a three dimensional model would present reality better and adds more value, according to the officers of duty. Furthermore, the collection of spatial data of indoor environments is time consuming and difficult to keep up to date.

6.1.1.2 System requirements

Second, the requirements of the proof of concept were discussed. A very early version of the proof of concept was presented to the officers of duty, providing (local) spatial mapping with a Microsoft HoloLens and visualization of the mapped 3D model on a laptop. The demo had three objectives:

- Presenting the possibility of using head mounted devices to map indoor environments
- Discussing the added value of mapping environments in 3D by using a head mounted device
- Discussing requirements of the proof of concept

The first responders reacted positively to the first version of the proof of concept. They said the ability to map indoor spaces without having to hold a device would be beneficial, as first responders would be able to clear rubble with their hands without dropping the mapping device.

The presented 3D model needed quite a lot of explanation, as it was presented in a default grey 'blob'. Improvements were needed in the visualization part to make interpretation of the model more intuitive. Furthermore, the position of the explorer was depicted in the model by a camera symbol. The first responders were fond of the tracking capability. However, they stated they wanted to know not only the position, but the track of the explorer as well. They suggested to present the track by a fading line with a color gradient, to add a temporal component to the mapping and tracking information. This request was anticipated and was already planned for. However, the request does indicate the significance of having 'time' as an additional variable next to indoor geometry.

At last, a new requirement was stated. The officers of duty stated that fire fighters depend on traditional methods to find their way within an indoor environment as vision is often limited. Methods to do so are also covered in (van der Meer, 2018), but include for example the following of walls for navigation purposes. Landmarks on the walls, such as light switches, aid fire fighters in knowing where they are by serving as reference points. If buildings increase in complexity, wall following may not be sufficient anymore and finding landmarks is more time consuming. Therefore, the officers of duty stated, visual aids might help to identify and recognize landmarks to increase speed and safety of indoor operations. The holographic screen of the Microsoft HoloLens offers the ability to show landmarks within the indoor environment. Following the request of the officers of duty, the visualization of landmarks has been taken within the scope of this research, as could be seen in the methodology. The effects of the implementation of these requirements will be discussed in the proof of concept evaluation sections.

6.2 Feasibility of real-time mapping & tracking communication

The sections below will describe the final functionality of the proof of concept and compare it to the requirements set for the methodology chapter of this research. This comparison will answer the research questions related to the feasibility of developing the proof of concept. This section is aimed on describing whether the technical capabilities of the proof of concept fulfil the technical requirements that have been specified in the methodology chapter. The results of the mapping, tracking and communicating modules will be described separately.

The mapping and tracking modules collect data in real-time. If the requirements of the mapping and tracking modules are met, the data elements are available for creating the first level of situation awareness: perception (Endsley, 2016). However, before this situation awareness level can be reached, the data needs to be transferred from explorer to navigator.

The communicating module is needed for real-time data transfer and real-time data presentation. Both aspects are requirements for the second level of situation awareness: comprehension. This statement needs some explanation:

- First, the data needs to be transferred from the explorer to the navigator. If this transfer fails, the navigator is obstructed from gaining any data elements, therefore creation of situation awareness will fail.

- Second, if the data is transferred, it needs to be interpreted correctly by the navigator. This means that the raw data of the mapping and tracking data needs to be transformed into information.

6.3 Mapping

This section aims to provide elements to answer the sub-question: *To what extent may a head mounted augmented reality device be used to map first responder indoor operation environments in (near) real time?*

6.3.1 Capturing indoor geometry

Hübner et al. (2019) and Khoshelham et al. (2019) have already shown the capabilities of the Microsoft HoloLens to capture indoor geometry. As in their research, the SLAM functions of the Mixed Reality ToolKit (MRTK) are used to enable the explorer application to capture indoor environments. The result of the mapping process is a spatial mesh. As the mapping process has been described and applied to indoor environments already, there is little scientific value in describing the mapping process in itself.

The focus of this research has been on real-time capturing operational environments. In the methodology section a tradeoff between mapping frequency and perceived detail is described. Mapping frequency is herein the time the application waits for processing data from the RGB-D sensor, while the perceived detail is an indicator for the number of triangles per cubic meter. By using a mapping frequency of 0.5 seconds, it was found users do not have to wait for the application while looking around. Furthermore, if the level of detail is set to 'low' (translating into 500 triangles per cubic meter), the application was able to run by rendering at 60 frames per second: the highest possible result. Due to these settings, the explorer can walk and look around in an environment, while capturing the environment simultaneously in a spatial mesh.

Figure 27: visualization of the mapping process: explorer view

The explorer is aided in the mapping process by the augmented reality display of the Microsoft HoloLens. If a real-world surface is mapped, the spatial mesh is projected over it. In that way, the explorer can observe which parts of the indoor environments have been mapped already and which parts should still be mapped.

This process is pictured in Figure 27: the left part of the displayed table has been mapped, while the right part is not mapped yet. From this picture, we can also observe that complex surfaces like the meshed framework at the left-back side of the scene is captured in more detail compared to the simple flat surfaces of the table.

6.3.2 Capturing indoor objects

By capturing indoor geometry, surfaces like walls, floors and stairs are mapped. A highly requested feature of first responders was to add objects such as exit signs, victims and light switches to the spatial mapping mesh as well. For this purpose, an explorer menu has been developed, shown in Figure 28.

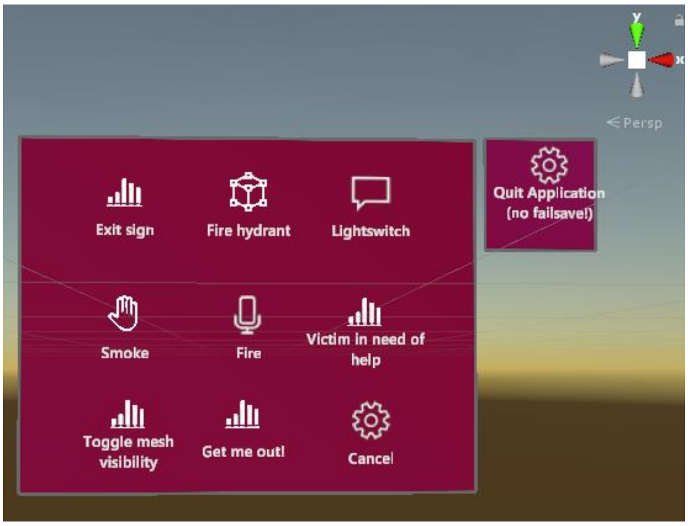


Figure 28: Indication of a HoloLens menu in the unity editor. The menu will appear in front of the explorer if triggered. The icons are chosen random

A user of the explorer application can choose to show this menu by clicking with a head-directed cursor on the spatial mesh. Then, the explorer can choose an option from the menu. If this option is related to a spatial object, like an exit sign, a colored sphere is both placed within the mapped environment (Figure 29) and sent to the navigator application. The sphere changes color specified for every mapped object.

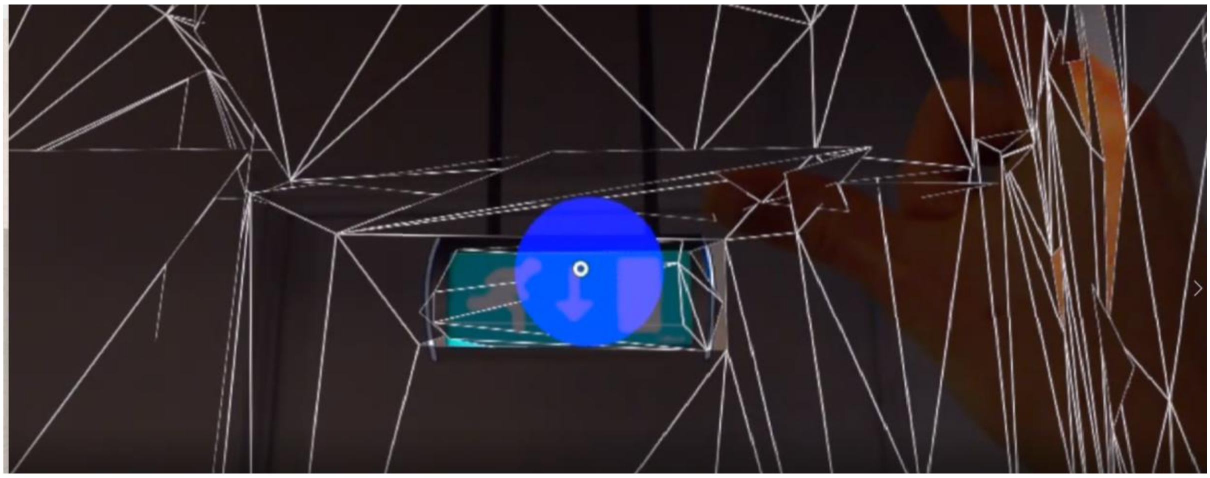


Figure 29: adding a visual indicating sphere in 3D space to an exit sign in the explorer application. The exit sign is there in real life, and the blue sphere represents the exit sign in the mapped environment.

6.3.3 Coverage

Although coverage is not identified as a component of reliability, we still want to draw some attention to it. Coverage means the extent that can be mapped by the proof of concept device. For the Microsoft HoloLens, the distance for which it can detect surfaces is limited to approximately three meters. In general, the device maps the viewpoint of the explorer and a bit below the viewpoint. This means that surfaces that are gazed at by the explorer are mapped, but if the explorer is walking straight, the floor is also mapped. As the ceiling is not mapped at the same time, the HoloLens depth sensor appears to be face slightly downward.

The images of the reliability tests look very accurate and precise because they are scanned with great attention. This takes more time compared to walking (or even running) through a building, which may be needed for first responder operations.

6.4 Tracking

This section aims to provide elements to answer the sub-question: *To what extent may a head mounted augmented reality device be used to track first responders in (near) real time within indoor operation environments?*

Like the mapping module, there is little academic relevance to describing the technical feasibility results of the tracking component. The tracking module has been implemented according to the described method (section 5.5.1) without any problems. This means that a series of poses can be stored in a list data structure at a given interval. Lists in C# can be 2 gigabytes large in memory. If a pose would be saved every frame (60 times per second), the size of the list is still within the bounds before the battery of the Microsoft HoloLens gives out. Therefore, it is concluded that the tracking module of the explorer application is able to fully fulfil the need for spatial tracking data elements to create situation awareness in real-time.

6.5 Communication

This section aims to provide elements to answer the sub-question: *To what extent may a head mounted augmented reality device be used to communicate in (near) real-time about mapping and tracking information of indoor first responder operation environments?*

6.5.1 Communication Protocol

The following requirements have been set for the communication protocol in the methodology:

- The data should be transferred via internet
- The data should be transferred within 10 seconds, if connected
- The data transfer should not block the explorer application
- If reconnecting after a disconnection, all unsent data should be sent in a first in, first out manner
- Package loss should be minimized

Three connection methods were used to transfer serialized data elements, also called messages, from explorer to navigator:

- Web socket
- Microsoft OneDrive
- Microsoft Azure

The Azure connection proved to be most suitable for the proof of concept. Below, short explanations of design choices will be explained.

6.5.2 Web socket

The first connection protocol that has been tested was a web socket. Web sockets are network protocols making use of a connection via the Transmission Control Protocol (TCP). In this case, the explorer device and the navigator device are connected via this protocol. Streams of messages can be sent and received over the internet from both ways via the web socket.

The web socket has been set up with the 'NetMQ' library. This library offers a .Net implementation of the ZeroMQ library, which is needed for integration with the code written for the Microsoft HoloLens. Furthermore, ZeroMQ works without a data broker. This means the devices are connected directly, without a server rerouting the data. This enables ZeroMQ to set up a web socket with minimal latency, no administration and without charging any fees. Technical aspects are available on NetMQ is available in the ZeroMQ documentation.

The web socket was implemented and tested. In a controlled environment with a stable Wi-Fi connection the web socket appeared to work well. All messages were sent within ten seconds and there was no observed package loss. However, the application would block if the connection was broken. This means that both applications would stop functioning until the connection was re-established. This does not comply with the requirements for the connection. Furthermore, as both the explorer and the navigator application rely on a mobile internet connection, a direct connection is likely to disconnect more often compared to an indirect connection with a broker. Because of the blocking behavior of the web-socket and because the connection could not guarantee to be stable, development with the web socket protocol was stopped.

6.5.3 Microsoft OneDrive

A second attempt to set up a reliable communication protocol has been performed with Microsoft OneDrive. This service enables users to store files in the cloud. These files can be accessed from multiple devices. By using this service, the explorer can save messages to text files which will be uploaded to the cloud if an internet connection is available. The files are subsequently pulled from the OneDrive server and processed in the navigator application. The advantage of such a method is that there does not have to be a stable connection between the explorer and navigator device: the explorer connects to OneDrive and the navigator does also connect to OneDrive. The connection to OneDrive is assumed to be always stable. Because of this assumption, the explorer/navigator device only need to have a connection to the internet, not to each other.

The method has been tested and was able to transfer data. However, the OneDrive connection was not able to keep up with the steam of messages and would begin to lag. Therefore, the data transfer threshold of ten seconds would be exceeded for large scans. Because of this limitation, development with this data transfer protocol was halted.

6.5.4 Microsoft Azure

At last, a Microsoft Azure implementation was used to send the messages. Azure is a cloud service with various options, such as saving very large files (so called 'Azure blobs') and sending messages in a so called 'Azure Queue'. From the Microsoft OneDrive implementation, we learned that sending data through a broker would me more reliable compared to a direct connection. In this case, Azure services as the broker. If a command, spatial mesh or explorer pose is created, it is transformed into a Microsoft Azure Queue Message.

Subsequently, it is put into the Azure Queue with an asynchronous function, meaning that the application will not block while doing this. Once pushed to the Queue, the navigator can retrieve the message from the in a FIFO manner. This data transfer protocol complies to all set requirements.

A disadvantage of this method is that the service is not free to use like the web socket and the Microsoft OneDrive implementation. However, total costs for using the Azure Queue for this research remained below one euro in total for the entire duration of the research.

6.5.5 3D model geometry presentation

The mapping module describes the collection of raw data within the explorer application. The collected data must be presented in the navigator application in a way that the navigator can make sense of the 3D model. This interpretation from raw data into information should require a minimal amount of mental resources, to reach comprehension and projection situation awareness levels easier (Endsley, 2016). As stated in the methodology, three visualization perspectives will be used:

- Geometry focused
- Time focused
- Object focused

The results of these perspectives will be discussed below.

6.5.6 Description of the space used for the 3D model

The visualization will be applied to a 3D model that has been created from a scan of the explorer application. To aid interpretation of the visualizations, a short description of the space will be given. The 3D model represents a multi-purpose space. It is multi-purpose in a way that it can be used for demos, meetings, and presentations. The space consists of three areas:

- The most left area, which is connected to the main entrance and which is dominated by a large table. Several monitors are placed on the walls, together with a collection of objects.
- The middle area contains a large, oval table and is secluded by four arced metal frames
- The right area is used for presentations and contains many chairs

The model will be represented from an angle of approximately 60 degrees. The back wall contains windows which are not scanned. Furthermore, the walls of the structure have a height of approximately 3 meters. Some parts of the ceilings are scanned. They are not visualized however, as surfaces are single faced: they are only visualized if they are viewed from the side from which they are scanned.

6.5.7 Geometry focused

The first objective was to make a geometry focused representation of the mesh. This means that the visualization takes only the geometric aspects of the mesh, such as connectiveness of the mesh vertices, global height and the normals within the spatial mesh. Interpretation of this model is done solely based on geometric features of the spatial mesh collected by the Microsoft HoloLens, which has a good accuracy according to Hübner et al. (2019) and Khoshelham et al. (2019).

First, normals were used to color the meshes. If a full RGB scale is used for 360-degree normals, this results in an image like the one below. What we can see is that horizontal surfaces, such as floors and table surfaces, have a vertical normal and are colored in a light green color. We can also observe vertical surfaces are colored distinctively from the horizontal surfaces. This is beneficial, as it enables observers to separate floors and walls from each other. Ceilings are collected for some parts of the structure, but as the model is observed from above the ceilings are fully transparent. The same applies for the walls on the side of the observer, at the southern side of the model.

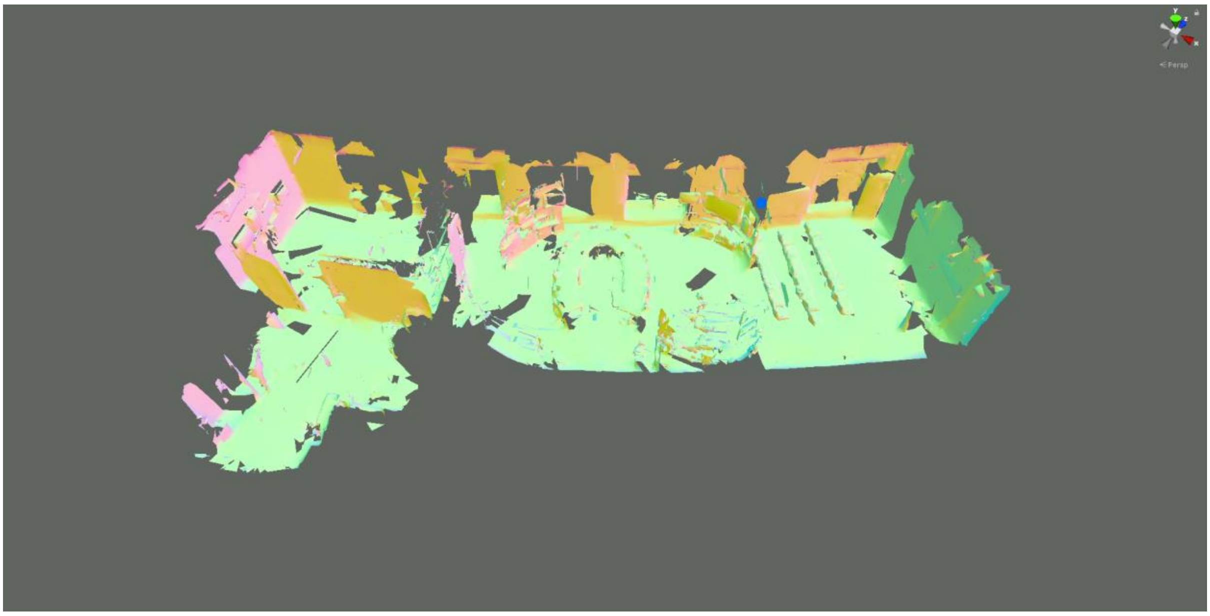


Figure 30: Geometry focused visualization: colored to normal

As we can observe from the colored normal visualization, walls have a distinctive color from the floor. As the walls have a horizontal normal, all walls are colored distinctive from the vertical normals. However, as the walls are placed almost perpendicular to each other, all walls do also have a different color from the other walls. This is an unnecessary overload of information, as we are only interested in the information if a surface is a wall or not. Therefore, the colors are harmonized to the verticality of the normal: because of this, the horizontal direction does not matter anymore.

In the visualization below, a completely vertical normals (horizontal surface, such as floors) are visualized in white. Completely horizontal normals (vertical surfaces, such as walls) are visualized in black. All normals between complete horizontality and complete verticality are visualized in a scaled tone of grey. We see that all large vertical surfaces (the walls) are colored in the same color. Furthermore, this depiction is a distinction in contrast rather than color. Therefore, this visualization is also suitable for people who are color blind.

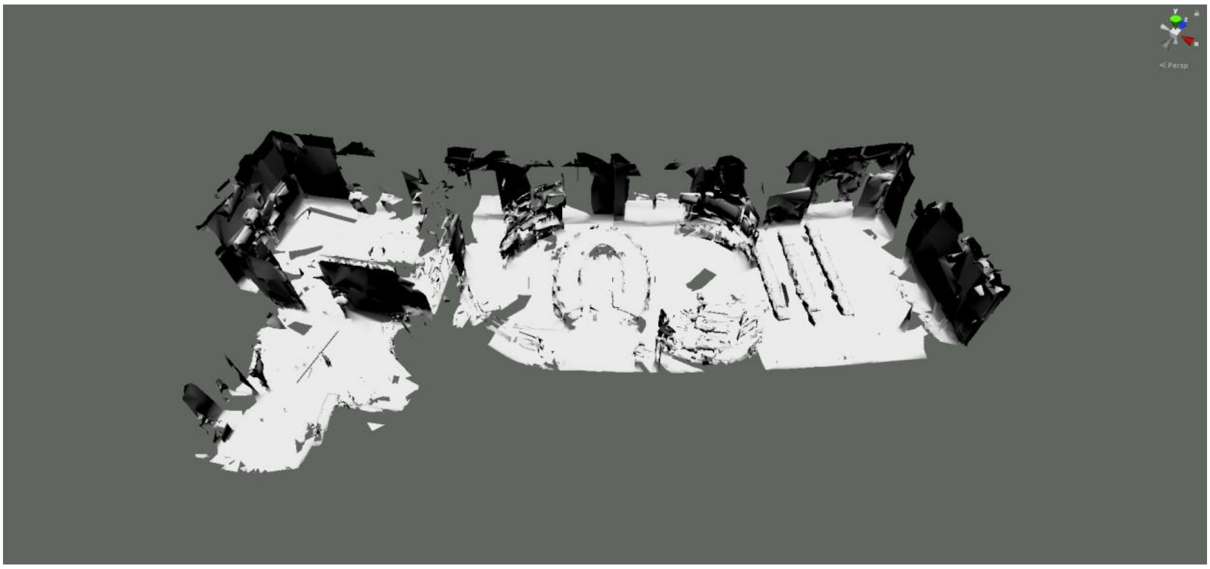


Figure 31: Geometry focused visualization: colored to normal, harmonized into grey scale

The distinction between floors and walls is well interpretable in the image above. However, there is no distinction between objects with horizontal surfaces and the floors. The oval conference table for example is hard to distinguish from the floor. This makes it hard to interpret the clutterness of a space and to determine whether a first responder could walk naturally on a surface or not.

A solution for this misinterpretation lays in utilizing another geometric feature of the model: relative height. By adding a black to white height based gradient overlay to the image above, we get the image below. Walls can still be distinguished, although the contrast between walls and floors have deteriorated. The big advantage of this visualization is the distinction between floor and horizontal object surfaces, such as the table in the middle of the image. This table is now colored in a different shade of grey compared to the floor.

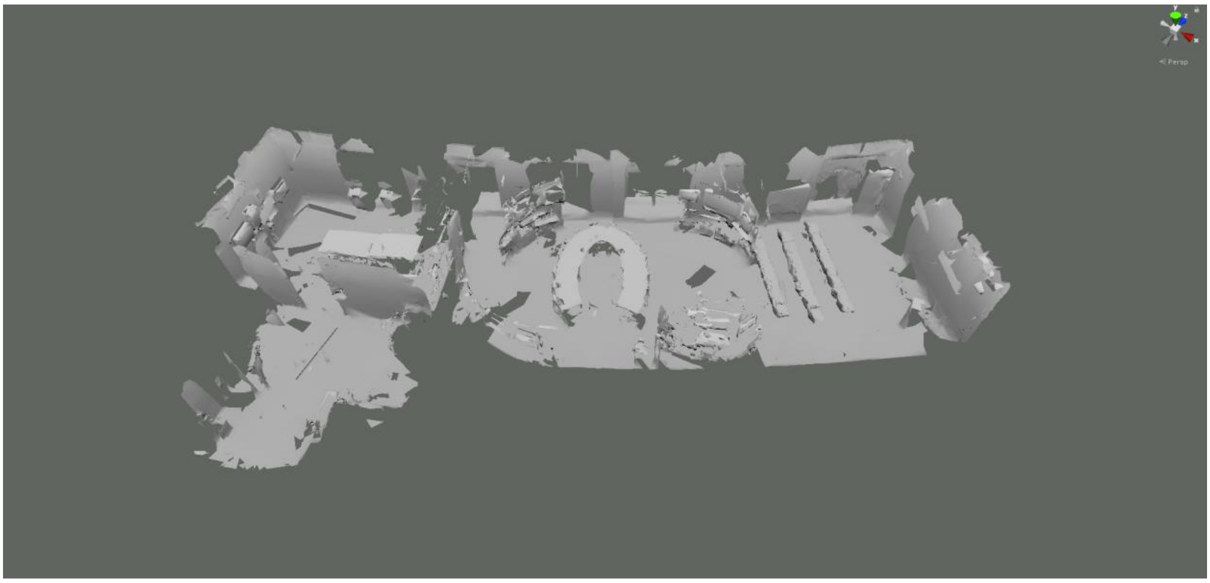


Figure 32: Geometry focused visualization: colored to normal, harmonized into grey scale and added relative height

6.5.8 Time focused

As first responder operation environments tend to change quickly, the way in which spatial information represents the true state of an environment is bound to time. The reasoning here is that relatively old data is less reliable compared to newer data.

This temporal component is added to the basic geometry by adding color from a separate variable: the last update time. If a spatial mesh has not been updated for a set amount of time, a color will be added to current representation of the mesh. Currently, the color scheme used is:

- 0 to 3 minutes: use the geometry based representation (right part of image below)
- 3 - 5 minutes: change to yellow (middle part of image below)
- 5 - 10 minutes: change to orange (left part of image below)
- > 10 minutes: change to red (not shown)

The colors and time thresholds are set dynamically by the navigator, as different operators might prefer different settings for different operation environments.

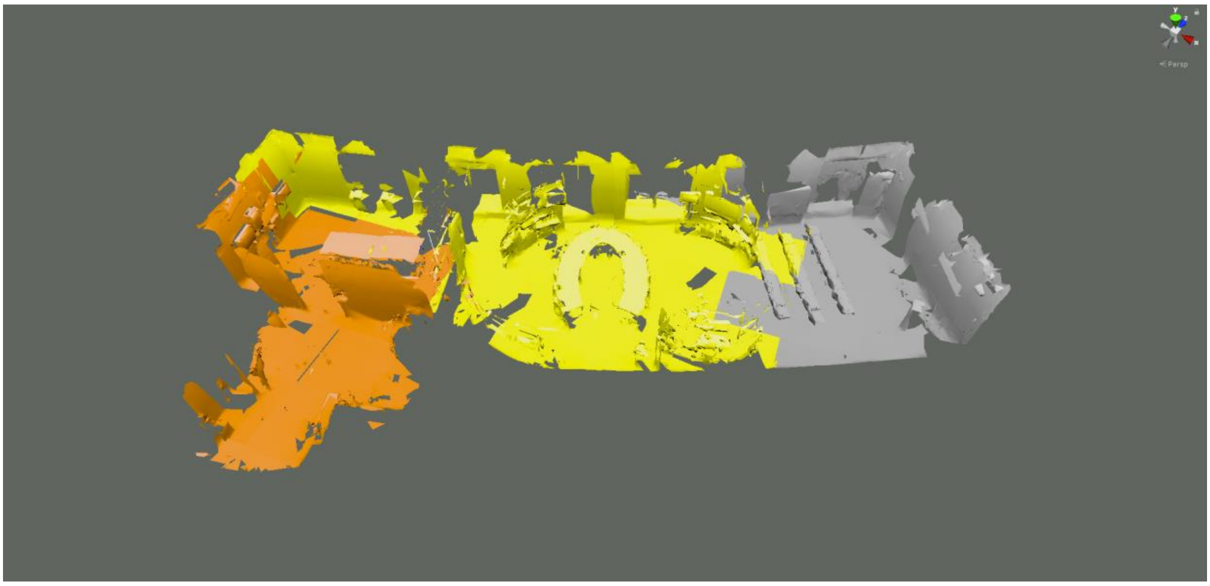


Figure 33: Time focused visualization: from left to right: orange (5-10 minutes old), yellow (3-5 minutes old), geometry based visualization (0-3 minutes old)

6.5.9 Object focused

Finally, the rules described in section 5.4.2: Mapping objects are applied to the 3D model. Because of the visualization rules, stairs, obstructive objects, doors and walls are shown distinctively. Generally, this presentation is received as the best visualization of the spatial mesh. Interpretation of objects, floors and walls is easy. However, due to the many rules applied to the spatial mesh, it is also easy to make mistakes in the classification of spatial features. Therefore, it is important that system operators are able to switch quickly from an object focused representation to a geometry based representation. The geometry based representation is more reliable compared to the object focused representation, as the geometry based representation depends on less and more robust rules.

Figure 34: Object oriented visualization. Red: floors; orange: obstructive objects; black: walls

6.5.10 Extracting navigable space

From the mapping information, navigable space can be extracted by fitting the shape of an 'agent' in the model. If the agent fits in the model at a certain space, that space is navigable. This navigable space can be extracted for different agents with different specifications. For example, an agent with a height of two meters and a width of 0.3 meters (the specifications of the navigable space below) can make different moves compared to an agent with a width of 0.5 meters. Variables such as 'step height', the maximum threshold an agent can step up, can also be defined.

By using the navigation mesh, agents are able get a route from one position to another position. Practical use of this functionality is not implemented in the proof of concept, as the navigable space that is extracted is deemed to be too rough: not all navigable spaces are connected to each other while they should be. As can be observed in Figure 35, the navigable space is not entirely continuous. This hinders using the Nav Mesh for future navigation purposes.

Figure 35: extracted navigable space (blue)

6.5.11 Tracking presentation

In the figure below, the tracking component is visualized within the object focused spatial mesh representation. The explorer is represented by a blue dot (in the middle of the figure). From this blue dot, a blue track follows the explorer, changing from blue to white to black over one second of time. Because of the time based gradient, the track is continually moving over time, making it easier to distinguish the track from the background. Furthermore, our vision is directed to motion, which makes it easy to follow the explorer while it moves through the scene.

As we can also see in the figure, the left part of the model has been scanned while the right part of the model is not yet scanned.

Figure 36: Scan in progress: in blue, we see the position and the track of the explorer

6.5.12 User interface

In the figure below, the user interface of the navigator application is shown. As can be observed, the object oriented spatial mesh visualization is used in this interface.

(1): Object menu, displaying all data elements. Spatial mesh elements are categorized to floor, making it possible to enable or disable visualization of floor levels. (2): Navigation menu, where navigation setting such as step height and agent size can be altered. (3): Function menu, where functionality such as a specific visualization space can be selected. (4): The scene view, which is the main screen of the coordinator. This is a full 3D view of the environment in which the coordinator can zoom, select, and rotate the contents. On the right, three virtual cameras move along with the explorer, either displaying a side view (5), a first-person view (6), and a top-down view (7).

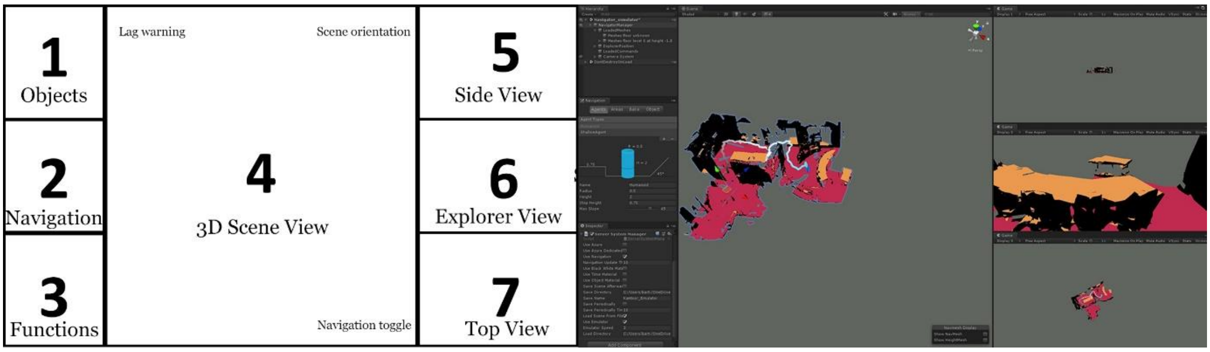


Figure 37: User interface. The left side represents a conceptual overview, while the right side displays the system as it is in use.

6.6 Reliability testing

Reliability has been defined in the theoretical framework as being a construct, consisting out of accuracy, precision, and robustness. The construct is used to describe to what extent measurements reflect the 'real world' situation. Data elements must be reliable to be used in the creation of situation awareness: the system operator must be able to trust the data that is received from the system. Therefore, this section will provide test results of the reliability of the data elements.

This section aims to provide elements to answer the sub-question: *To what extent may the reliability of the mapping, tracking, and communicating capabilities of the proof of concept support first responder decision making for indoor operation environments?*

6.6.1 Test environments

Three test environments will be used test the reliability of the proof of concept. The test environments have each been chosen for specific reasons, which will be explained below. Furthermore, the tests have been performed in different stages of proof of concept development. Therefore, the test conditions will be shortly described to give more context. The most important test conditions are specified on top of the description, namely:

1. the specific test characteristic
2. The ground truth reference
3. The communication protocol used

Below, the test environments are listed.

1. The Vening Meineszbuilding of the University of Utrecht
2. A large office building
3. The architecture building of TU-Delft

The second scan of the large office building is leading. The other scans will be used to extend on the results of the second scan.

6.6.1.1 Vening Meineszbuilding of University of Utrecht

Specific characteristic: multi-floor building
Ground truth: Terrestrial Laser Scanner point cloud
Communication protocol used: Microsoft OneDrive

Hübner et al. (2019 and Khoshelham et al. (2019) describe the accuracy of the Microsoft HoloLens SLAM algorithm in a single floor environment. However, first responders are faced with increasingly complex, multi-floor buildings. Therefore, reliability of multi-floor buildings should be considered. Because of this, accuracy and robustness have been tested while exploring multiple floors of the Vening Meineszbuilding of Utrecht University.

Simultaneously, a point cloud of the environment is collected with a TLS. This scanner gives a ground truth with an accuracy of 3 millimeter.

6.6.1.2 Large office building

Specific characteristic: Very large single floor office space
Ground truth: Mobile Laser Scanner point cloud
Communication protocol used: Microsoft Azure

At the end of the development phase of the proof of concept, a second test was scheduled. The test has been completed together with a officer of duty of a safety region. The test environment was a large office floor of approximately 2500 m2. By scanning the environment with both the explorer application and a MLS, a large area of the office space could be covered. The MLS has a slightly lower accuracy compared to the TLS of the first test, but as the coverage of this test is larger, this test will be used as baseline for the reliability tests.

6.6.1.3 Architecture building of TU-Delft

Specific characteristic: Little distinctive visual features due to industrial look
Ground truth: CAD drawing
Communication protocol used: Microsoft Azure

Finally, the Architecture building of TU-Delft has been scanned with the explorer applications. Here, a large space is present with little visual features. Therefore, a large probability of tracking loss was assumed. A CAD drawing of the building is available to overlay the scan with a ground truth.

6.6.2 Accuracy

First, the way in which measurements reflect the real world will be tested: accuracy. Accuracy is of importance for both the mapping and the tracking modules, as will be explained below. The first part of accuracy will describe the results of the office test environment, as this environment has been scanned with large coverage by an MLS and the explorer application.

As has been described in the results of the feasibility of the mapping module of the proof of concept, the result of mapping an indoor operation environment is a spatial mesh. These spatial meshes can be represented in a variance of manners, such as geometry focused, time focused, and object focused. However, we do not know how well this spatial mesh represents the real world. For example, it is unknown if the space between two objects is indeed large enough to be able to walk there. Furthermore, as the navigator receives the mapping information at a remote location, it is assumed to be impossible to validate the results by visual inspection. Therefore, the navigator has to rely on the assumption that the mapping information is correct.

This subsection will describe to what extent the assumption of correct mapping measurements is true by calculating distance variety of the proof of concept mesh and the real world. This will result in a global mean distance offset of the proof of concept scan in centimeters and ‘distance maps’, depicting local distance offsets.

6.6.2.1 General description: floor of office building

The test that will be described now has been performed in the latest phase of testing with an almost completed proof of concept. The test environment is a very large, multi-story office building. The building was in use at the moment of scanning, meaning there were people working in the building. The test describes a scan of a part of the seventh floor of the building. The Azure communication protocol is used to transfer data from the explorer to the navigator application. A connection to the internet is made by using a wireless mobile connection. Due to unforeseen circumstances, only a part of the floor has been saved. These circumstances will be reflected upon in subsection 6.6.4: Robustness. However, a large part of the scan remained and will be used for the evaluation of the accuracy results of mapping in a horizontal space (one floor of a building, without stairs).

6.6.2.1.1 Overview of the environment

Below, an overview of the environment as mapped by the HoloLens can be observed in an object oriented manner.

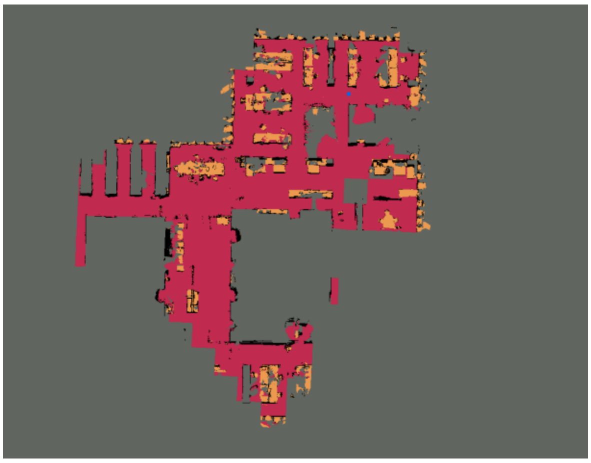


Figure 38: Top down view of scanned office environment

Below, a side view of the mapped environment can be observed.

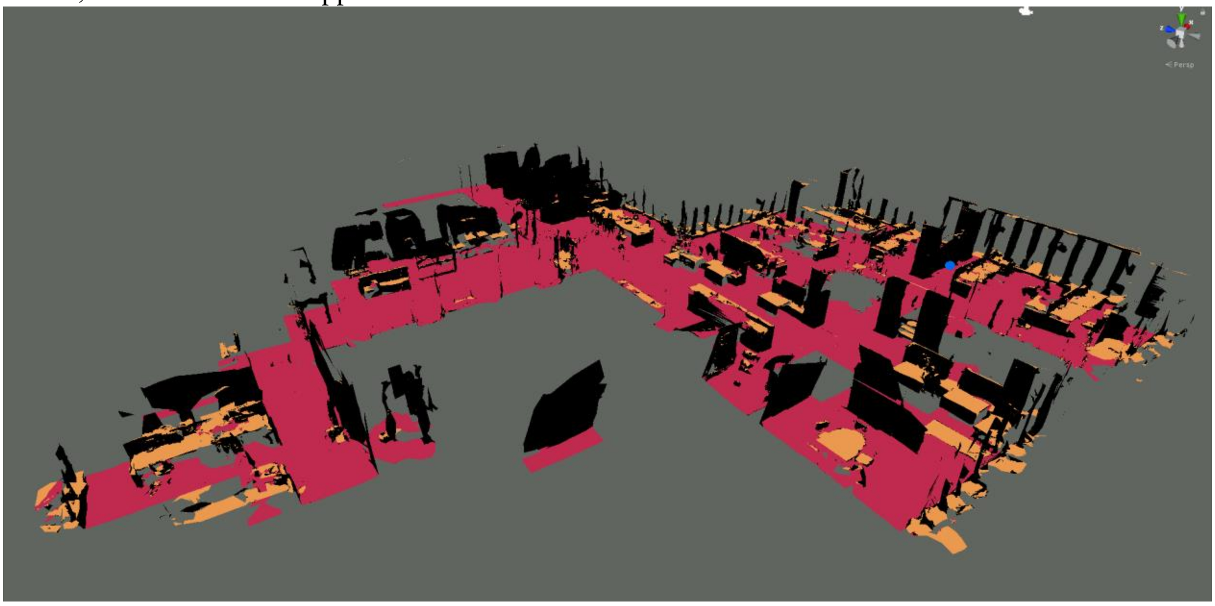


Figure 39: Birds eye view of scanned office environment. Surfaces with a normal with a direction from the viewer are fully transparent

The environment contains a lot of desks and chairs, which can be observed in the images. Furthermore, all red information is floor space.

6.6.2.1.2 Overlay

For the manual overlay, the ceiling of the point cloud has been removed (Figure 40). In red, the floor is depicted. In green, objects on the floor are depicted. In white, the spatial mesh of the proof of concept is depicted. As can be seen, the overlay is already well positioned and rotated to match the reference point cloud.

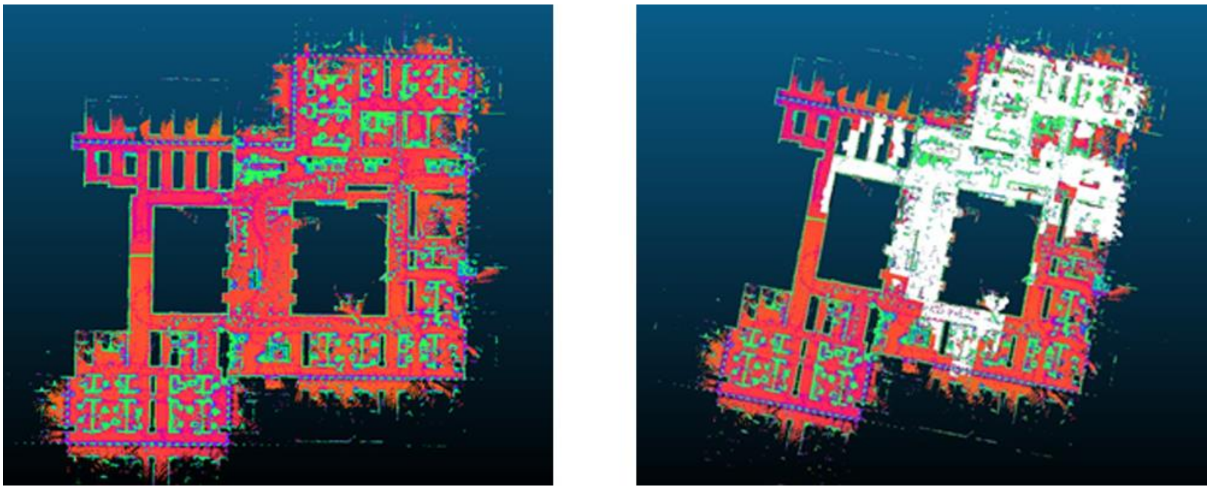


Figure 40: Left: ground truth. Right: Overlay of ground truth and PoC scan in white

6.6.2.1.3 Cropping

The point cloud will be minimized in size to ensure a better fit to the spatial mesh. Two methods are applied to ensure a good fit:

1. Only points that are within a buffer of 0.5 meter within from the spatial mesh will be used for the distance calculation. The value of 0.5 meter has been chosen after visual inspection of the overlay and by using the approximate mean distance of the point cloud to the mesh, which is 0.78 meter.
2. The ceiling will be removed from the point cloud by using a histogram distribution describing the height of the points. By extracting the points above the height with the most points, which is identified to be the ceiling, only the points representing the floor and objects close to the spatial mesh remain. This is done as the ceiling is not actively scanned by the proof of concept.

The removed points are depicted in grey in Figure 41, while the green points remain. After cropping, the point cloud is finely registered to the mesh by using the ICP algorithm.

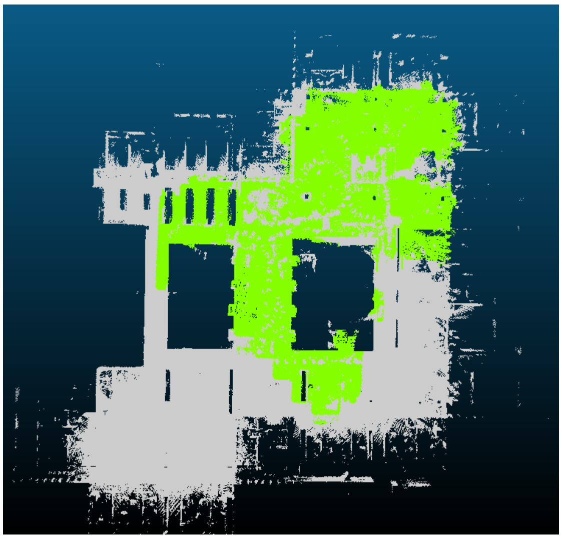


Figure 41: Maintained points (green) and removed points (grey) after cropping

6.6.2.1.4 Distance maps

After the crop has been performed, the distance from the point cloud to the mesh has been calculated. This means that from every point in the cropped point cloud, the closest distance to a mesh vertex will be calculated. From this calculation, a mean distance error of 0.0359 meters, or 3.6 centimeters has been identified between all of the points of the point cloud and the spatial mesh. Figure 42 presents the distribution of the point distance to the mesh. From this graph, we can observe that more than 72 percent of the MLS points has a distance to the spatial mesh smaller than 0.1 meters.

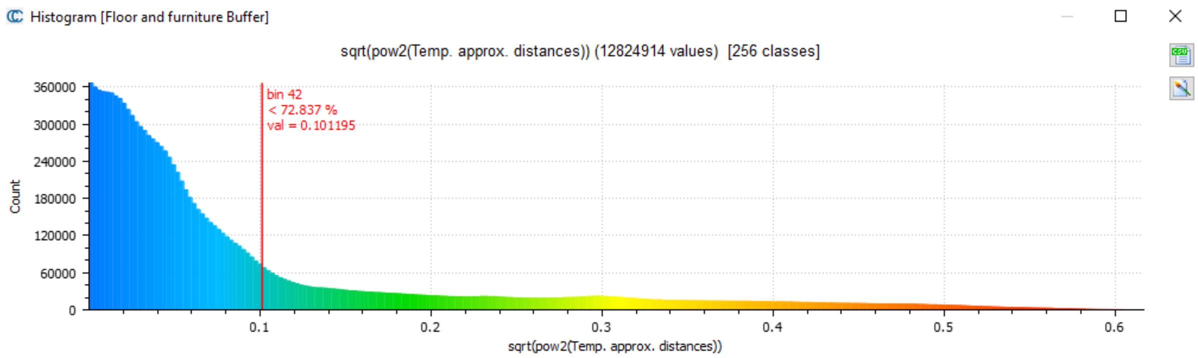


Figure 42: Histogram of all point distances to the mesh. Distance units are in meters.

Although the distribution of error is valuable to get insight in global accuracy, we know little about local accuracy. This means the representation of specific features within the scan. In Figure 43, a visual representation of the error is given for the entire the proof of concept scan. There are a couple of red spots, but most points are very green. This means they are captured well with a maximum offset of about 0.3 meters, according to the scale bar in the right part of the picture. Movable objects are less well represented, which makes sense as there is some time between the MLS and the proof of concept scan.

Figure 50 presents the same data as Figure 43, but from a different angle. Now, we can also see that the top of immovable objects such as walls and window frames are often not scanned by the proof of concept, which creates a relatively large error. Moreover, the buffer technique that is used to crop the MLS point cloud to the proof of concept scan extent is visible on the edges of the scan.

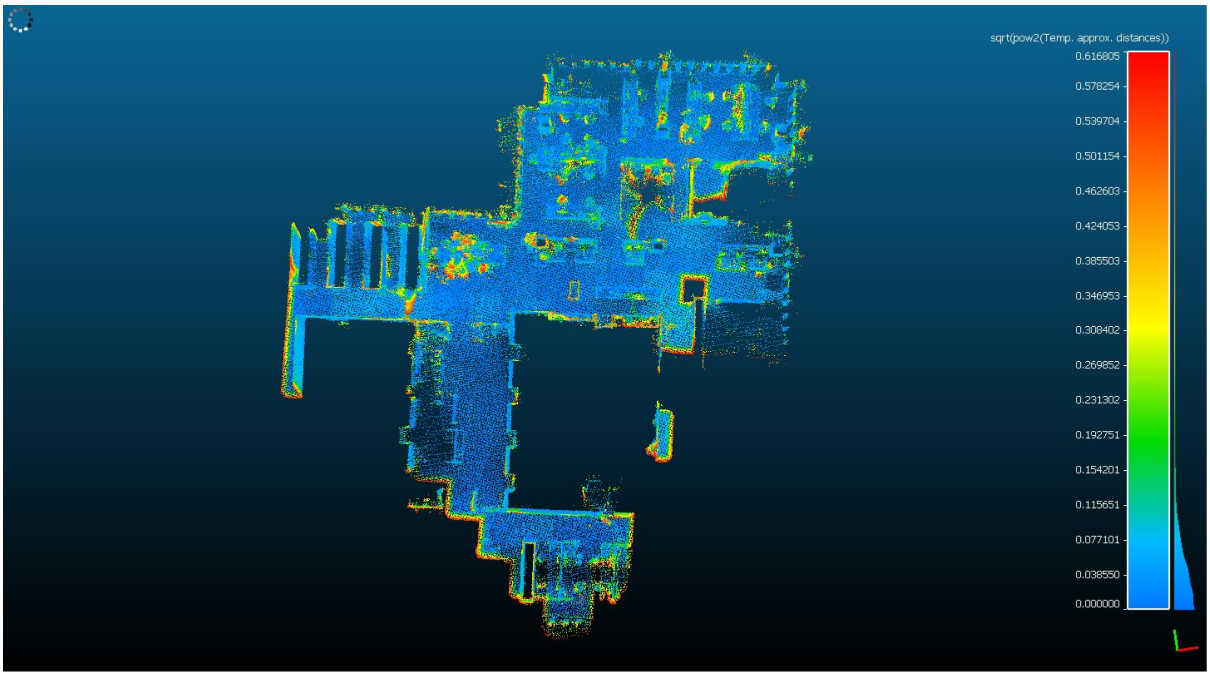


Figure 43: top down view of distance map of office building

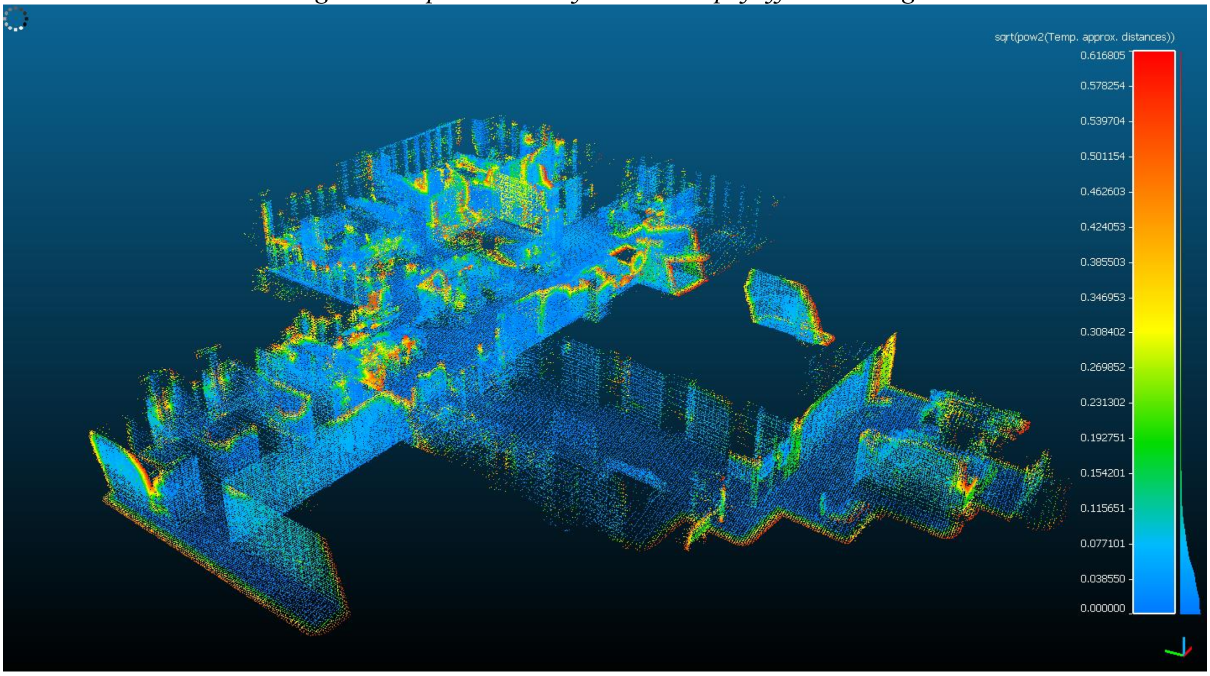


Figure 44: Birds eye view of distance map of office building

6.6.2.2 Vertical accuracy: multi-floor scan

Most of the accuracy of the office floor scan stayed within ten centimeters. This was expected also based on the results of other academic publications (Hübner et al., 2019; Khoshelham et al., 2019). These publications did not cover vertical accuracy that is needed for scanning complex, multifloor buildings. There is reason to suspect differences between horizontal and vertical accuracy, as the HoloLens SLAM did not perform well on staircases. This finding will be elaborated upon in the robustness tests.

Figure 45: TLS set up in Vening Meineszbuilding. Left: Scan position 1. Right: Scan position 2.

This test has been performed in the Vening Meineszbuilding of Utrecht University. This building has several floors with a gap in the middle of the floors. This gap makes simultaneous scanning of multiple floors possible with a TLS, without the need to change positions. The positions and field of view of the scanner are shown in Figure 45. Below, the proof of concept scan of the environment is shown. The TLS positions are indicated with the red circles. The left red circle should be on a lower floor, which is indicated by the downwards arrow.

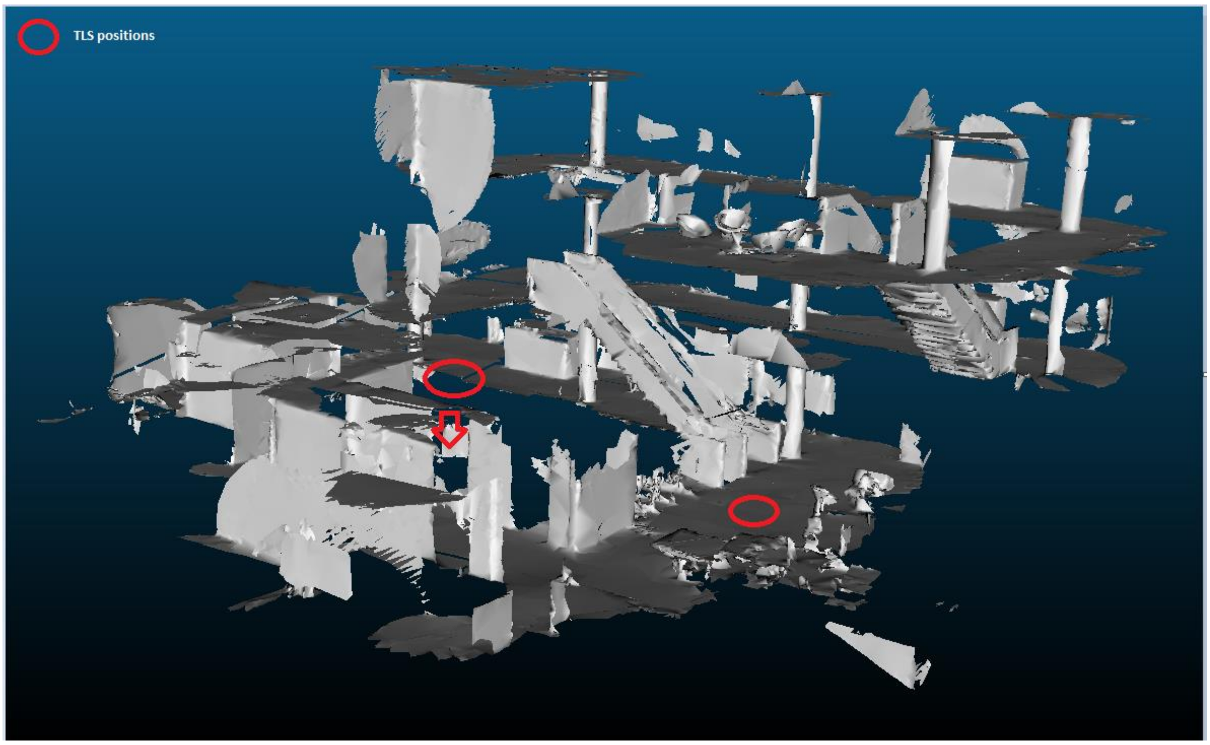


Figure 46: The proof of concept scan of Vening Meineszbuilding

On visual inspection, the scan looks to be quite reliable: the pillars supporting the floor have been scanned neatly on top of each other and the stairs are also intact. A geometric evaluation will be done by comparing the TLS scan and the proof of concept scan, with a method similar to the one described for the office space scan. However, the results should be interpreted differently compared to the interpretation of the office scan results. The TLS scans its environment from a stationary position. From this point, the scanner can only cover a limited area of its environment, as its viewpoint is limited. Therefore, the back of a pillar or the floor of a higher floor level is invisible to the scanner, unless the scanner is moved. Calculating the mean distance does therefore not make a lot of sense, as many points of the mesh are not visible to the scanner anyway. Furthermore, the extends of both scans do only partially overlap, instead of the full overlap of the office environment scan. Therefore, points are plotted on the mesh surface with a density of 200 points per square meter. This allows us to use the TLS scan as reference scan.

Below, a visualization presents the calculated distances between the proof of concept scan and the TLS scan. The point size has been enlarged to make the structure more visible. Although the picture looks worse compared to the office scan, we should take the TLS limitations into account. For example, the bottom of the stairs in the foreground is primarily green. This is likely as the TLS scans the bottom of the stairs, while the explorer walked over them and scanned the top part of the stairs. Just as with the office scan, only points that are within 0.5 meters of the reference point cloud are considered, as the scans have a different extend. This is the reason for the red outline of the image.

To validate the accuracy of the proof of concept scan we look at two features: the ceilings and the pillars. Both are stationary objects. The ceiling is used to measure vertical drift, while the pillars are used as a reference for horizontal drift.

- The ceilings are primarily blue, meaning that they have been mapped with high accuracy. This can be observed for example in the middle top part of the image. If there had been a lot of vertical drift, this accuracy would have been lower compared to the accuracy observed in the office scan. Therefore, it is concluded there is little vertical drift.
- The pillars are also colored blue in the picture, which means they are also captured with good accuracy. This means there is little horizontal drift, as was already proven by the office scan.
-

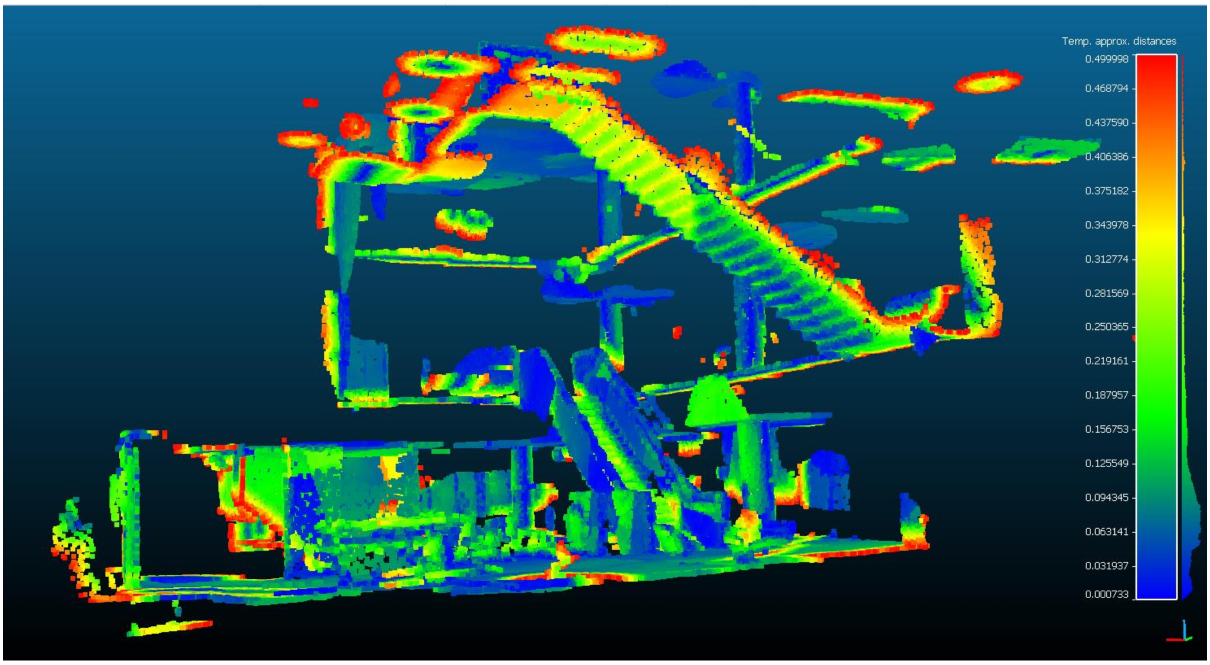


Figure 47: visualisation of distance between mesh and ground truth. The viewpoint is behind scan position 1. The scale is in meters. Although there are many red lines visible around the model, we can observe that the percentage of red points is very low in the total amount of points if we look at the histogram to the right.

6.6.3 Precision

The accuracy tests describe errors in the way that objects are represented by the proof of concept at the right position: it tests whether the measurements are actually there or if the mapping information has drifted a bit. Precision measures the consistency of the measurements, often described by the resolution of the data (Law, 2007).

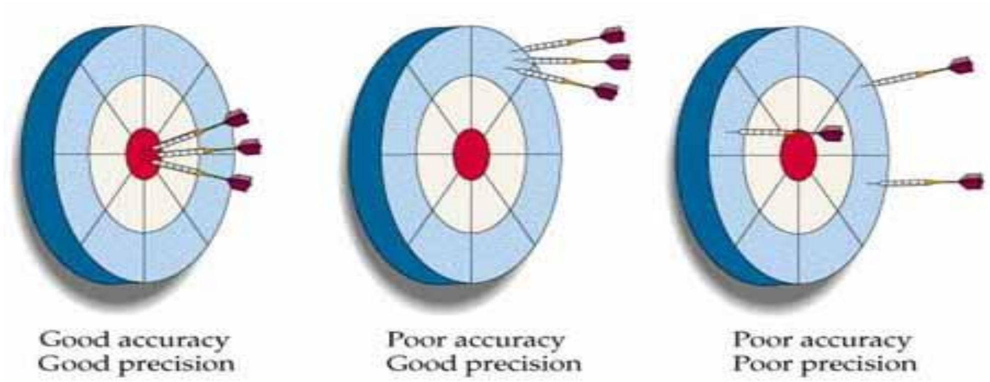


Figure 48: accuracy versus precision (Law, 2007)

Precision translates to the level of detail of the proof of concept. We know that the level of detail of the spatial mapping system has been set to 'low', which does not sound very promising for our precision measurements. To evaluate precision of the spatial mapping, two scans will be compared visually. We will use the point cloud and the spatial mesh of the office space discussed in in subsection 6.6.2.1: General description: floor of office building. A visual comparison between the spatial mesh and the point cloud can be observed in Figure 49 and Figure 50.

In Figure 49 we can distinguish desks (orange), chairs and people (black/orange), floors (red), walls (black), and columns (black). The columns are window frames. If we compare *Figure 49* with *Figure 50*, we see that the geometry in the images looks the same. However, the geometry of the window frames and the chairs in the proof of concept scan are often jagged. Therefore, the scan looks to be imprecise. Precision could be increased within the current setup of the proof of concept at the cost of scan update frequency. The situation awareness tests will cover if low precision of the data is a problem for the feasibility of gaining situation awareness from the data.

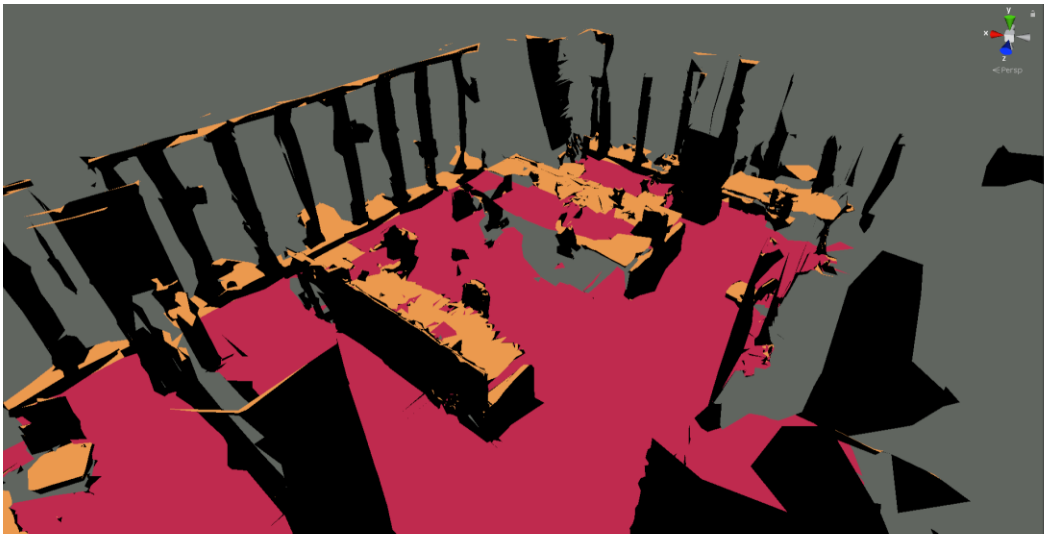


Figure 49: Part of the office building as scanned by the proof of concept

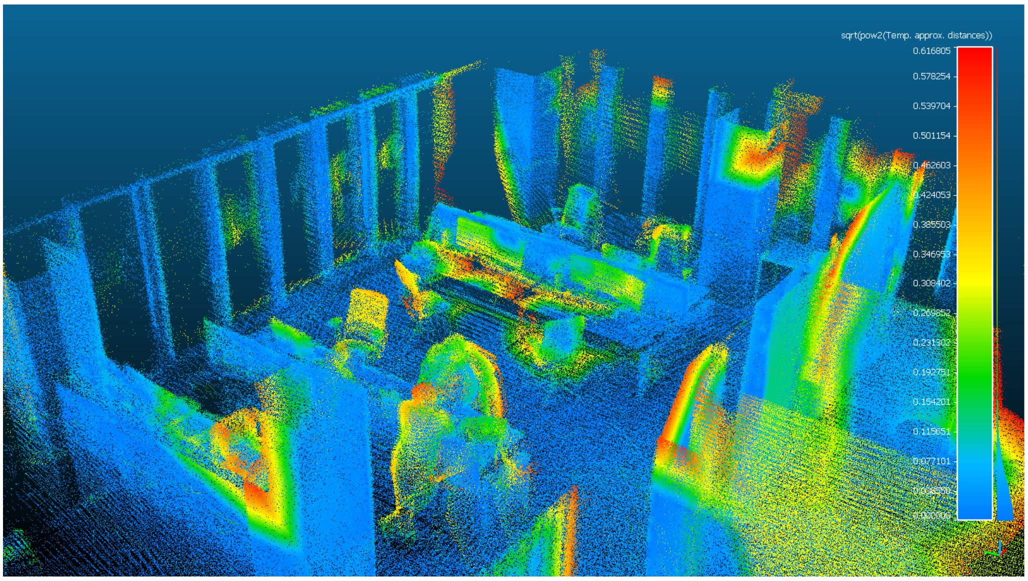


Figure 50: Part of the office building as scanned by the MLS. Points are colored to distance from the spatial mesh (Figure 49)

6.6.4 Robustness

Robustness is the third component of reliability. While accuracy and precision give us information about the quality of the measurements, we do not have information about the continuity of the measurements. If the proof of concept fails for whatever reason to report environment data in real-time to the navigator application, it is regarded as a failure of the robustness of the system.

In the development and testing phase of the proof of concept, a couple of aspects have been identified that are important for the robustness of the system. They will be explained in the subsections below.

6.6.4.1 Tracking loss

As has been identified in the methodology section, tracking loss of the Microsoft HoloLens SLAM algorithm was to be expected. From the documentation we know that the algorithm may for example fail if the environment is badly lighted, landmark scarce or if there are a lot of black/reflective surfaces. This subsection will describe the consequences of tracking loss, how often it was encountered and what was done to mitigate the effects.

6.6.4.1.1 Consequences

If tracking is lost, the spatial mapping module of the system is disabled. Therefore, no new environment data is collected. A symbol is shown on the explorer screen, indicating tracking has been lost. The system will try to regain tracking capability. Regaining the tracking capability is done quicker if the explorer walks back, as spatial anchors may be recognized if a space is revisited. Of course, both the disabled spatial mapping module and the need to walk back in a certain direction is unfavorable in an first responder operation where time is of the essence.

If tracking is restored, the system defines a new coordinate system for the whole environment. The origin of this coordinate system is the position at which tracking was restored. As this origin is different from the previous origin, all current mapped data should be shifted to match this new origin. Sometimes, this shift is done correctly, especially if the to be shifted environment is revisited. However, more often the translation fails and the navigator ends up with unreliable results. An example of this will be shown below.

6.6.4.1.2 When does tracking loss occur: validation

The methodology lists a couple of conditions for which tracking loss is expected to occur. They are repeated below:

1. Lighting conditions are too bright, too dark or lighting conditions change too sudden
2. Places that look very similar, such as office spaces with the same interior for every floor
3. A room with strongly reflective surfaces
4. Landmark poor environments, such as a hall without a lot of distinctive features
5. Movement in place, for example in crowded areas
6. Rooms without Wi-Fi connections, as Wi-Fi fingerprinting enables the device to recognize spatial anchors more quickly

Although all of these conditions are tested and recognized by this research, some conditions were worse than others. The conditions above have therefore been listed on likelihood of losing tracking capability. Condition 1, 2, and 3 have a very high likelihood of causing tracking to fail. Tests in such environments did almost always cause tracking loss. Condition 4 and 5 did only cause tracking loss a couple of times. For condition 6, no difference in tracking loss was perceived. This was tested by using the application in the same environment with enabled and disabled Wi-Fi connectivity. However, condition 6 could influence the robustness of the algorithm in environments where one or more of the other conditions are present. Although it does not influence the robustness of tracking loss, spatial anchors of environments were recognized more quickly if Wi-Fi was enabled.

6.6.4.1.3 When does tracking loss occur: new finding

Next to validation of expected tracking loss conditions, this research found that the traversal of stairs is also found to be very likely to cause loss of tracking. This loss of tracking was experienced early in the development phase of the proof of concept while the proof of concept was first tested in a stairwell of an office building.

Multiple reliability tests have been deployed in the Vening Meineszbuilding, as described in section 6.6.2.2: Vertical accuracy: multi-floor scan. The goal of these tests was not only to test if there is a difference in accuracy when scanning multi-floor, but also to test if there is a difference in tracking loss.

Five attempts were made to move to another floor by the stairwell of the building, of which 4 attempts failed causing tracking loss. Furthermore, the application failed almost all the time to shift known environment data back in place if tracking was recovered.

We have seen that tracking loss is likely in spaces that look similar. This could be the reason for tracking loss in stairwells. Therefore, next to stairwell tests, multiple tests were taken on stairs outside of stairwells. The system was able to maintain tracking better on 'normal' stairs, but still failed about half of the time. The results shown in section 6.6.2.2 display a successful test. The image below displays the spatial mapping results of a failed test. Five floors were traversed for this test, traversing stairs outside of stairwells. When traversing from the 4th to the 5th floor, tracking was lost and eventually regained. The mapping however was not restored, resulting in the 3D model as shown below. As can be seen, spatial meshes are often disconnected and displaced. The model is therefore very hard, if not impossible, to interpret by an observer.

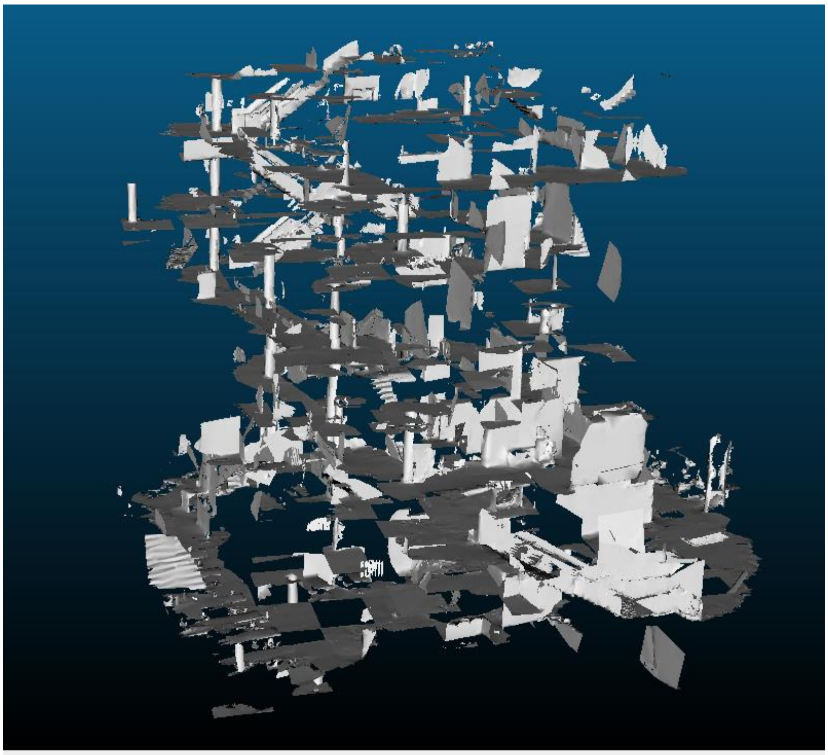


Figure 51: tracking loss on stairs in the Vening Meineszbuilding

6.6.4.1.4 Overall reliability of tracking

In general, the Microsoft HoloLens is able to track its position within an environment well, especially if a user respects the limitations of the system as described in sub-sub-sections 6.6.4.1.2 and 6.6.4.1.3. For example, tracking was only lost once while tracking the visual feature rich environment of the large office space. In a visual feature sparse environment such as the architecture environment (Figure 52), tracking was lost twice. If tracking is not lost at a staircase, the SLAM algorithm is almost always able to restore a consistent mapped image if tracking is restored. This can be seen in the image below, where tracking loss is also indicated.

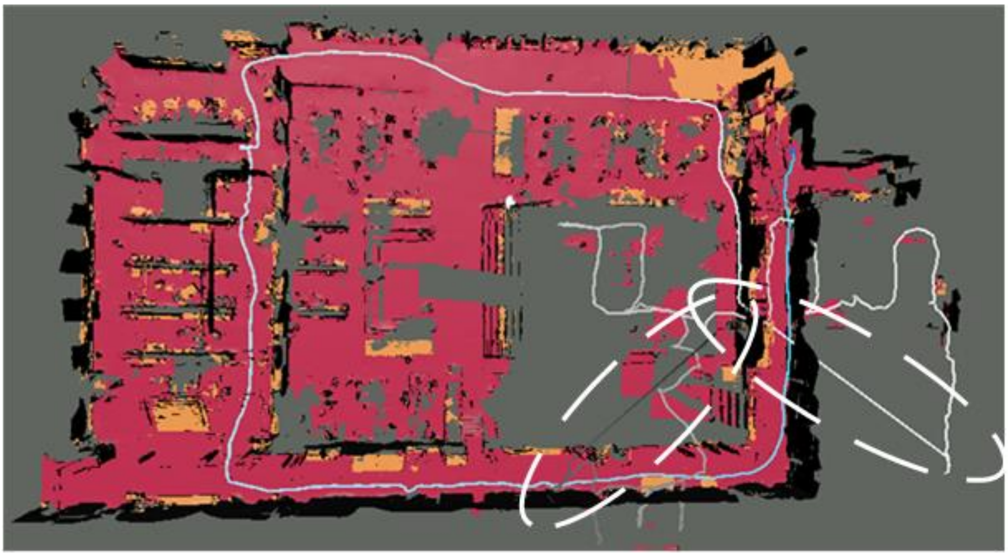


Figure 52: indoor environment of the building of Architecture of Delft University of Technology. This environment is regarded to be visual feature sparse, due to the large open space in the middle of the scanned environment. The right bottom corner shows two cases of tracking loss, displayed by the completely straight tracks.

6.6.5 Communication protocol

Besides real-time data collection, the communication protocol is of importance to facilitate the navigator with up-to-date data. From the results of the third sub-question (section 6.5), we know that real-time data transfer is possible. Data is transferred within ten seconds if the azure communication protocol is used, as data is transferred within a second if both the explorer and navigator devices are connected to the internet.

The requirement of an internet connection is a condition that may fail and therefore cause trouble in the continuity of the data stream. To give insight in the reliability of the communication protocol, the proof of concept has been tested both in an environment with and without a stable Wi-Fi connection. Only the Azure connection is tested on reliability, as the Web Socket and OneDrive communication protocol were found to be infeasible according to the requirements of real-time data transfer.

To test any lag of data transfer, the explorer application adds a timestamp to all collected data. This is subsequently sent to the navigator application, that is able to calculate time difference between creation time and the time at which data is received. If this data transfer takes more than 1 second, a warning is given to the navigator.

6.6.5.1 Wi-Fi connection

The reliability of the Wi-Fi connection has been tested in the architecture building of Delft University of Technology. The resulting scan is shown below. The duration of the scan was 9 minutes and 3 seconds. Within the building, the Wi-Fi network of Eduroam is available. As the environment is rather large (the scan covers 1600 square meters), the network is transmitted by multiple access points. This means that the explorer device has to switch between these access points to continue wireless data transfer. Although such a switch of access point was expected to cause some lag, no warnings were issued by the navigator application. This means there was no observed lag of over 1 second for the entire duration of the scan. Therefore, it is concluded the communication protocol of the proof of concept works reliable while using a Wi-Fi connection.

6.6.5.2 Mobile connection

As Wi-Fi is not always available to first responders, for example because of bad signal coverage or missing Wi-Fi credentials, the reliability of proof of concept data transfer has also been tested by using a mobile internet connection. In this case, the explorer device is connected to a Wi-Fi hotspot of a mobile phone, which is connected to the internet via a 4G signal. The mobile connection has been tested in the large office environment. As this is a large indoor environment, signal interruptions were expected. This was especially expected in the elevator area of the building at the heart of the building (Figure 53), as it was expected mobile broadband internet signals would have a hard time penetrating this part of the building. However, no lag was observed.

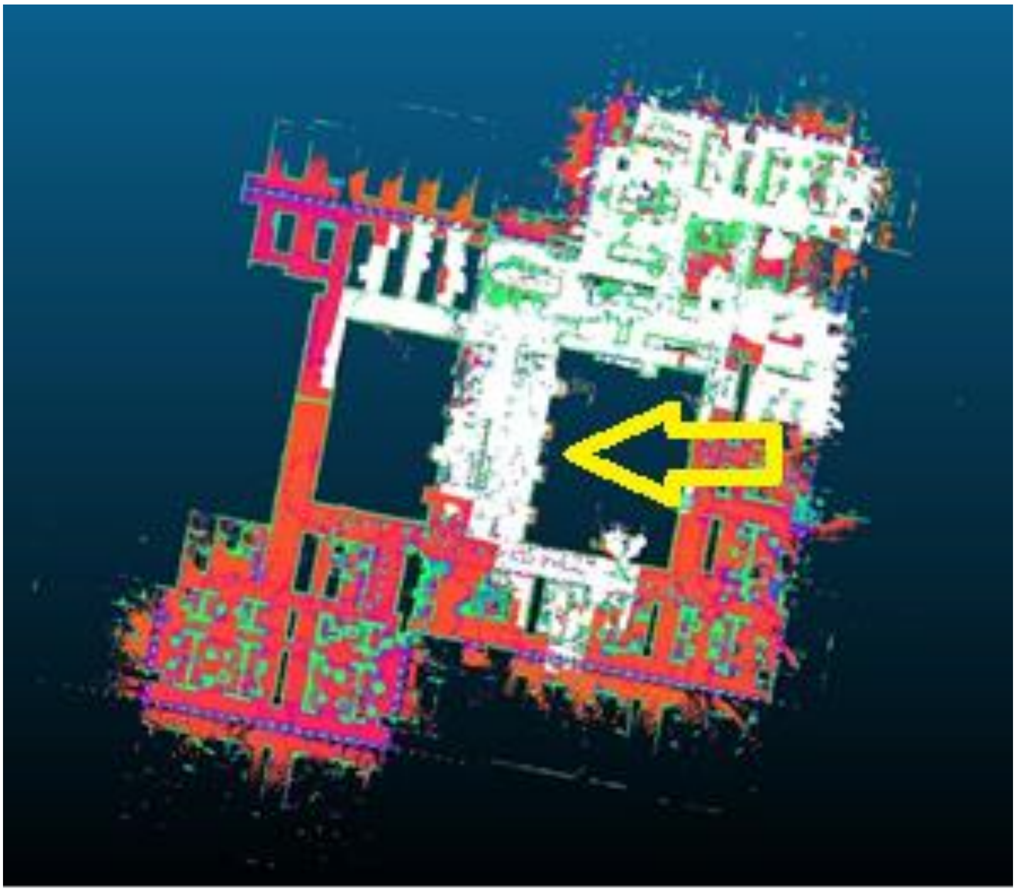


Figure 53: map of office environment. The yellow arrow points to the elevator area

6.6.5.3 Interrupted connection

Although no disconnections were noticed ether in the Wi-Fi test or the mobile internet test, we tested what would happen if the connection was lost. The explorer application should still collect data if the connection to the internet was lost. If the explorer device reconnects to the internet, all data that was collected during the absence of the internet connection should still be sent in a First In, First Out structure.
This condition was tested by connecting the explorer application to a hotspot of the mobile phone. Other Wi-Fi signals were disabled in the system settings of the Microsoft HoloLens. Therefore, the explorer application would fully depend on the hotspot connection. This concept was used in the following manner to test what would happen if the internet connection was interrupted:

1. At the begin of the test, data was sent via the hotspot as usual.
2. After one minute, the hotspot connection of the phone was disabled. The explorer application kept functioning, meaning that it continued with data collection. However, the data was not received at the navigator application any longer.
3. After one minute of connectivity, the hotspot connection of the phone was restored. Within 5 seconds, all data that was collected during the internet absence was sent by the explorer application and received by the navigator application in a FIFO structure. After this, both applications continued to work as usual.

As the data was quickly sent and received in a FIFO structure after the period of interrupted internet availability, the test is successful. Therefore, the communication protocol works reliable even if internet availability is interrupted.

6.6.5.4 Application stability and performance

Both within and outside of the reliability tests, the proof of concept has not shown any unresolved blocking errors. Therefore, both the navigator and the explorer application are regarded to be stable and robust.

Furthermore, the explorer device (the Microsoft HoloLens) has limited computational resources. Within the explorer application, an indicator gives information about the performance of the application. Results of this indicator are summarized below, as bad performance would give a bad user experience and more importantly, it might cause errors and stop continuation of the system.

6.6.5.5 Frames per second

As a performance indicator, frames per second is used. Every frame, the application will have to make calculations and render images to present on the display. Sixty frames per second are preferred, as this is an indication that computations can be made without delaying the system. If the framerate drops, this is an indicator that the application has trouble to process the data quick enough.

It was found the application does not always maintain a framerate of sixty frames per second, which is preferred for a good user experience. Sometimes, the framerate drops to about 30 frames per second, for example after reconnecting to the internet if the connection was interrupted. However, this drop in frame rate is not perceived to be bothersome as the mapping, tracking and communicating processes are continued without visible interruption. Also, the application has always quickly restored from frame drops.

6.6.5.6 Memory usage

All spatial mapping data is saved first in the working memory of the explorer device. This data is maintained in working memory for the entire duration of the scan. Memory management is handled mostly by the Mixed Reality Toolkit. In the development process, maintaining all spatial data was treated as a threat: if the working memory would overflow the explorer application might crash. However, the device indicated that the working memory has never exceeded 0.6 GigaBytes of the available 2.0 Gigabytes while The proof of concept was running. This was even the case for continuous scans of over 20 minutes. Therefore, it is concluded memory management does not pose a threat to the robustness of the proof of concept.

6.6.5.7 Battery level

The Microsoft HoloLens is powered by a built-in battery. On the test device, which is a couple of years old, it was observed that running the proof of concept does not drain the battery very quickly. The drainage was not continuous, but as a rule of thumb we used the rule that one minute of scanning would take half a percent of battery power. Therefore, scans with a duration of over one hour should not be a problem.

6.6.6 Reliability stakeholder meeting

The second meeting was organized at a first responder conference about the use of 3D data in first responder operations. CGI was invited to organize a workshop with the proof of concept, which was scheduled to have a length of one hour in total. At the day of the event, the workshop was split into two smaller workshops, each with a length of approximately 45 minutes. The concept of situation awareness in combination with the concept of (real-time) indoor mapping and tracking was introduced in a short presentation, after which the participants could test the proof of concept themselves. During the demo, the participants were asked to evaluate the reliability of the system and give feedback about the added value of the proof of concept. Each group contained approximately twelve participants, of who most were working for fire departments.

The first workshop was used mostly as a try out for the concept. The presentation took a bit longer than expected (25 minutes). The last 20 minutes of the meeting was used to demo the proof of concept. During the demo, the first responders were asked to give input to improve the proof of concept, which was written down on a white paper (see appendix 10.1).

The second workshop was improved with the experience from the first one. One of the improvements was better time management: the presentation was limited to 15 minutes. Then, a demo of 10 minutes was given after which a separate discussion 20 minutes was undertaken with the participants. The results from the discussion were again written down on a white paper (see appendix 10.2).

6.6.6.1 Perceived reliability

The topic of reliability of the navigator application was touched upon, discussing the way in which real-world objects and indoor geometry were represented in the mapped 3D model. The level of detail was deemed sufficient by the participants, as they declared to be able to distinguish geometric features of the indoor environment easily. The tracking component was also perceived to be reliable, as the explorer positions were presented without visible drift.

The reliability of the communication module was put to doubt. Due to the way the communication protocol was set up the transfer of messages got more latency if scan duration increased. For the small demo, the transfer time stayed within the near real-time threshold of ten seconds. Although this duration should not be that big of a problem for creating situation awareness, first responders stated it would be more troublesome if a larger area had to be scanned.

6.6.6.2 Mapping interpretation

The 3D model was presented to the first responders in a simple normal based grey to black visualization. Participants could recognize the walls, doors, stairs and mesh representations of participants easily. Besides humans, there were almost no objects within the demo environment. It should be noted however that the participants were in the same room as the scanning device, which makes interpretation of the geometry easier as it can be compared with real-world observations. It is assumed interpretation of unknown environments requires more mental resources, as comparison with the real-world is not possible which makes interpretation more abstract.

Besides the three dimensional representation as discussed above, the question was posed whether the proof of concept should also present a two dimensional presentation as proposed by (van der Meer et al., 2018). The participants recognized the statement that a two-dimensional overview can give a quick overview of a situation.

6.6.6.3 Tracking interpretation

Just as in the first meeting, only the last known position of the explorer was shown in the 3D modelled environment, although it was now represented with a blue dot instead of a camera symbol. Participants could easily interpret the blue dot as the 3D position of the explorer. However, they repeated the request of the first meeting to visualize the track of the explorer in a time dependent gradient.

6.6.6.4 First meeting conclusion

The added value of the system was recognized by all participants, as it was deemed to offer a useful addition on current indoor operation resources. The first responders deemed the level of detail high enough to distinguish individual features in the 3D model and could easily recognize explorer positions within the 3D model. The participants suggested to add functionality to mark visitors with the explorer application and send that information to the navigator as well. Finally, the communication protocol as presented was recognized to transfer data from explorer to navigator quickly enough for small operation environments.

The input data could therefore successfully be interpreted by the participants, who stated it would be beneficial mostly for indoor navigation and victim extraction. When they were asked which first responder role should be facilitated with the data of the 'navigator application', the answer was that either the local commander should receive the data or a completely new role should be set up. As the data elements for these tasks are provided and can be translated into useful information for the completion of the tasks. Projection of the current state into future states has not been tested. Therefore, it was concluded the proof of concept enabled the participants to gain at least the second level of situation awareness.

The participants recognized the potential of a system like The proof of concept for mitigating the effects of an incident in the response phase. Nevertheless, they stated such a system might be even more valuable in the preparation phase. The explorer application could be used to scan buildings completely and add valuable information such as fire separators to the building. This would automate parts of the process of making availability maps of indoor environments, thereby enabling first responders to map more buildings within the same time span. Real-time transfer of data would not be needed in such a scenario, as the data can be processed after the completion of the test. A test for such deployment of the system had been scheduled but was cancelled due to the COVID-19 crisis of 2020. The test was scheduled with a team leader of team that makes first responder availability maps of crucial buildings. The aim of the test would have been to test which indoor objects would be important for preparation phase scans and how indoor objects could be best visualized/presented to observers and first responder coordinators.

6.7 Situation awareness testing

In the conceptual model (section 0), key requirements for creating situation awareness were presented from literature. In the methodology for testing situation awareness (section 0), these requirements are repeated as stakeholder involvement did not change them. These requirements will now be evaluated by the researchers. After the researcher evaluation, the results of the last stakeholder meeting will be presented, in which situation awareness was specifically evaluated.

The first level of situation awareness, perception, requires mapping and tracking elements to be available in real-time. The PoC has proven these requirements to be feasible, as mapping and tracking data elements could be collected without delay due to limited computing resources. Furthermore, the data elements could be sent in real-time both by using Wi-Fi and mobile internet connections. Delay between sending and receiving data elements has remained below 1 second, which is well within the requirement of a maximum of 10 seconds delay for (near) real-time transfer.

Based on the results found for the previous sub-questions, we see the creation of situation awareness is possible by using the PoC in indoor first responder operations. Identified requirements for collection and presentation of data elements have been proven to be feasible in real-time. Furthermore, we see that data elements can be successfully transformed into interpretable information. This interpretation is aided by the geometry, time, and especially the object oriented visualization techniques that reduce information overload. Last, the reliability of the PoC is enough to put reasonable trust in the system. This would lead to the second level of situation awareness: comprehension. Projection, the third level of situation awareness, requires more experience in the field to be judged by the researchers.

As has been stated by both (Endsley, 2016) and (Wickens et al., 2013), a system has to be designed with the user in mind instead of technique. (Dilo and Zlatanova, 2011) enforces this idea from a first responder perspective, as systems used within first responder operations should meet varying first responder requirements. Three stakeholder meetings have been organized to involve first responder in setting system requirements and evaluating the results. All meetings have been planned with employees of the Dutch safety regions. Most of the involved first responders were working in the fire department of the safety region. This is why the proof of concept has been designed mostly with fire departments in mind, building forth on the work of (van der Meer, 2018). A final stakeholder meeting was organized to discuss the overall results of the PoC in terms of added situation awareness. The results are stated below.

6.7.1 Third stakeholder meeting: discussing added value of PoC

Finally, a third meeting has been organized at a safety region office. Despite the intention to organize an in-depth interview with an officer of duty, the meeting was transformed in an open group discussion. A diverse group of first responders were present, of whom most were working directly for the fire department. Among them were multiple officers of duty and commanders. Also, a creator of indoor availability maps was present.

Figure 54: demonstrating the navigator application to the safety region participants

The goal of the meeting was to present a completed version of the proof of concept in the form of a live demo. A volunteer of the safety region was asked to perform the role of the explorer by putting on the

Microsoft HoloLens and to scan the office environment. Meanwhile, the author stayed with the other participants of the meeting to discuss the transfer of explorer data to the navigator application.

6.7.1.1 Perceived reliability

During the demonstration, the mapping and tracking information was transferred in real-time thanks to the new Microsoft Azure communication protocol. A delay in data transfer was not noticeable.

During the test, the explorer application lost it position: it was not able to maintain tracking. This event was perceived while the explorer walked down the stairs in the stairways of the office. The duration of this tracking loss was a couple of seconds, after which tracking was regained. The model did not suffer too much from this tracking loss, but was transformed to meet the new coordinate system of the explorer application. The first responders did not think tracking loss to be a large problem as the perception of the model remained almost completely intact.

6.7.1.2 Mapping interpretation

Interpretation of the model is possible, but the group thought identification of obstructive objects would add more value compared to the normal based grey-black visualization. After the meeting, this object oriented visualization method has been implemented which was well received by the officer of duty who had proposed this way of visualization.

6.7.1.3 Tracking interpretation

For this meeting, the time dependent visualization of the explorer track was completed. This feature was well received and “added a new dimension” to the proof of concept, according to one first responder.

6.7.1.4 Meeting conclusion

Although the mapping visualization was not finalized, a fruitful discussion about the added value of the proof of concept was still possible. The first responders identified real-time indoor navigation of the most important objective of the proof of concept. Figure 54 shows the author of this research demonstrating the navigable space component (depicted in blue) of the mapping module. Extracting navigable space is the basis of navigation, which was recognized by the first responders.

The tracking component was deemed sufficient and the time dependent track visualization was well received. One first responder made the remark that this would enable coordinators to predict future scenarios. Others agreed with the first responder. As predicting future scenarios is a proxy for projection of future states, the conclusion has been drawn that level three situation awareness was reached by the demonstration.

7 Conclusions

Before the central research question will be answered, all sub-questions will be answered.

Most research questions evaluate the proof of concept, in which a first responder 'explorer' uses a Microsoft HoloLens to collect indoor environment mapping and tracking data elements. These data elements are then sent and presented in real time to a first responder 'navigator', who coordinates the first responder operation from a remote location.

7.1 Domain recognition

This conclusion will answer the sub-question: *To what extent has the relation between 'indoor mapping', 'indoor tracking' and 'first responder decision making' already been established by the academic community?*

First, the quantitative results of the literature mapping study (section 3.3) are evaluated. As there is about 10 percent overlap in publications, we state there is a basis to conclude that first responder researches are often associated with both mapping, tracking and decision making. Overlap between these domains are not tested. Furthermore, the results of the trend analysis however are too unreliable to say for sure whether the domain is growing in relative importance, as the Scopus and Google Scholar queries provide ambiguous results.

Second, the theoretical framework is the result of the qualitative analysis of the domain recognition. Here, it was found there is a research gap between indoor first responder operations and real-time mapping and tracking applications. At last, no other research was found providing an in-depth explanation of the relation between real-time mapping and tracking for creating remote situation awareness for indoor first responder operations. Thereby it is concluded this research does indeed add to the academic body of knowledge. The answer to the sub-question is:

The relation between 'indoor mapping', 'indoor tracking', and 'first responder decision making' has been established by the academic community. However, this is not the case for real-time mapping, tracking, and communicating within indoor first responder operation environments.

7.2 Mapping feasibility

This conclusion will answer the sub-question: *To what extent may a head mounted augmented reality device be used to map first responder indoor operation environments in (near) real time?*

The explorer application is able to map indoor geometry in real-time by a naturally moving first responder. Furthermore, objects can be added to this geometric representation of the indoor operation environment. These aspects cover the system requirements as identified by the methodology.

As the explorer can map first responder operation environments in real-time while moving naturally, the answer to the sub-question is:

The mapping component of the explorer application can fully fulfil the need for indoor spatial mapping elements to create situation awareness in real-time.

7.3 Tracking feasibility

This conclusion will answer the sub-question: *To what extent may a head mounted augmented reality device be used to track first responders in (near) real time within indoor operation environments?*

The tracking functionality has been implemented according to plan. Therefore, the answer to the sub-question is:

The tracking module of the explorer application can fully fulfil the need for spatial tracking data elements to create situation awareness in real-time.

7.4 Communication feasibility

This conclusion will answer the sub-question: t*o what extent may a head mounted augmented reality device be used to communicate in (near) real time about mapping and tracking information of indoor first responder operation environments?*

First, this section has proved the possibility of real-time data transfer of indoor mapping and tracking data in first responder operations, as data is received by the navigator application within a second after creation of the data. Second, this section has proposed meaningful ways of presenting both real-time collected mapping and tracking data to an observer. Third, the observing navigator is facilitated with a user interface that enables both interaction with the data and automated views of the data.

Mapping data is presented in a geometry, time, and object focused ways. Geometry focused presentation is the most robust presentation as it depends on very basic rules. Time focused presentation enables observers to project future scenarios of the environment and enables observers to determine what trust they should put in the dynamic data. Object focused presentation enables observers to quickly interpret floors, obstructive objects, walls, doors and stairs from the 3D model, thereby reducing information overload. However, object focused presentation comes at a cost: as there are many rules that are used to visualize the 3D model, the presentation is less robust compared to geometry and time focused presentation. The three visualizations complement each other. Therefore, it is possible to switch between presentations for a more complete overview of the operation environment. At last, navigable space is calculated and presented to the observer.

Tracking data is presented by a blue dot, just like in many other navigation applications. A time dependent track follows the explorer, indicating the path that the explorer takes. This compliments the temporal component of the time focused mapping presentation.

As data can be transferred and presented to a first responder coordinator in real-time, the answer to the sub-question is:

A head mounted augmented reality device may be used to communicate mapping and tracking information of indoor first responder operation environments to a full extent.

7.5 Reliability testing

This conclusion will answer the sub-question: *To what extent may the reliability of the mapping, tracking and communicating capabilities of the proof of concept support first responder decision making for indoor operation environments?* As reliability is a construct of accuracy, precision, and robustness, all of these concepts will be used for answering the sub-question.

First, the proof of concept proved to be accurate in capturing indoor environments. By comparing the mapping results of the proof of concept with a mobile laser scanner, the proof of concept mapping data was found to be able to collect over 70 percent of the data with an accuracy within ten centimeters of the ground truth. This accuracy was found for both horizontal and vertical measurements. The results enable first responder coordinators to make accurate estimations of distances and navigable space.

Second, the proof of concept proved to be rather imprecise when capturing indoor environments. The mapping data provides a 3D model that contains holes and is often very jagged. A more precise model could be generated at assumed cost of scan update frequency.

Third, the proof of concept proved to be robust under certain limitations. The communication protocol works well, offering real-time data transfer both by Wi-Fi and mobile connections. If the explorer device is temporarily disconnected, the application continuous to collect environment data and will send the data in a First In, First Out structure to the navigator. Loss of tracking does pose threats to the robustness of the proof of concept. Although tracking may be lost under several known conditions, such as bad lightning and similar looking places, the Microsoft HoloLens SLAM algorithm is often able to restore its spatial model of the environment if tracking is restored. However, if tracking is lost on stairs, the 3D model of the environment is often permanently damaged. This applies both to tracking loss in stairwells and on stairs outside of stairwells, although the first is worse. Therefore, for a robust experience, it is recommended to avoid stairwells.

Because of the good accuracy, mediocre precision and limited robustness, the conclusion of the sub-question is:

The reliability of the proof of concept enables first responders to trust the data elements provided by the proof of concept to a reasonable extent, if certain limitations are respected. If navigators are trained to use the system they will know when to trust the system and when not to trust the system.

7.6 Situation awareness testing

This conclusion will answer the sub-question: *To what extent may the reliability of the mapping, tracking and communicating capabilities of the proof of concept support first responder decision making for indoor operation environments?*

First, the proof of concept proves that a real-time indoor system that communicates mapping and tracking data from the indoor first responder operation environments to a coordinator is feasible. Second, the data is accurate and precise, while the system is robust if actions are undertaken to reduce the change of tracking loss. Third, proof of concept functionality is regarded to be successful in fulfilling data requirements of indoor operation environments. Fourth, first responders can trust the data that is presented to them. Fifth, first responders were able to translate the data into information about the operation environment, as they could interpret the data with ease.

As first responders are able to quickly interpret the raw data elements into information to solve tasks, the requirements for reaching the second level of situation awareness as defined by (Endsley, 2016) are fulfilled. There are even indications that the third and last level of situation awareness, called projection, may be reached by the proof of concept. Therefore, the conclusion of the sub-question is:

The proof of concept improves first responder indoor performance by providing at least the second level of situation awareness: comprehension, to support first responder decision making

7.7 Central research question

Finally, the following central research question will be answered: *To what extent may a head mounted augmented reality based (near) real-time simultaneous localization and mapping system (SLAM) support reliable spatial decision making to increase first responder operation performance in indoor environments by improving situation awareness?*

First, as we now know, the feasibility of creating an head mounted augmented reality based real-time SLAM system is feasible. Augmented reality aids the mapping process, while the head mounted device captures both mapping and tracking data. These data elements are transferable via the internet with less than 1 second of latency to a remote observer.

Second, the data can be trusted by the observer, as it found to be reasonably reliable in terms of accuracy, precision, and robustness. Training should help to distinguish between reliable data elements and unreliable data elements. Furthermore, first responders are found to be able to be able to comprehend the data elements that are provided by the system. They are aided in this task by the way in which the data elements are presented to the observer, reducing information overload. This means the second level of situation awareness (comprehension) is reached by using the system.

Third, first responders have suggested the proof of concept can be used to predict future states of an environment in a data driven way. This hints to achievement of the third level of situation awareness, projection.

Finally, it can be concluded the combination of elements that are provided by the proof of concept present useful information that would not have been available without the system. By using this information, first responders are aided in making decisions of better quality for the execution of tasks, thereby improving first responder performance. However, this boost in performance has not yet been tested in real-life situations. Therefore, this research concludes that a system like the proof of concept proves to be a valuable additional asset for indoor first responder operations. The answer to the central research question is:

The extent to which the proof of concept may contribute to first responder performance by improving indoor situation awareness is 'promising', as the second level of situation awareness was created by use of the proof of concept in simulated indoor first responder operation environments.

8 Discussion and future work

Writing this thesis was both interesting and challenging. The research offers insight in how first responders operate in indoor environments and in what way spatial information aids them in doing so. Furthermore, added value of an indoor mapping, tracking and communicating system is proven. Nevertheless, some aspects of the research could have been improved.

8.1 'promising' is a vague result

To start with one of the most important points: the central research question is answered by stating the extent to which the proof of concept may contribute to first responder indoor performance is 'promising'. The term may feel like an unsatisfactory result, as it is a somewhat vague term. The use of this term was not anticipated when the research was set up, but feels fitting for the result that we got. The most important reason for the use of this term is that the proof of concept has not been tested in a real-world first responder operation, as will be explained below.

First responder operations are complex, meaning that it is difficult to estimate how a first responder operation environment will develop over time. From the demonstrations to first responders, the result of the proof of concept was that at least the second level of situation awareness (comprehension) could be gained. Some first responders even claimed they were able to project the current state of the situation into the future, thereby hinting to the third and last level of situation awareness: projection.

However, the tests used in the demonstrations were simulated, meaning the proof of concept was not deployed in a real-world first responder scenario. The first responders observing the performance of the proof of concept could fully focus their attention to the proof of concept. In a real-world situation, it would be likely that first responders, especially decision makers, would also have to respond to other factors (Wickens, 2000). This increases the information load that a decision maker would have to process. Also, it is likely that an increased level of stress would be involved in a real-world scenario. Both increased information load and the increased stress level are likely to hurt the capability of creating situation awareness (Endsley, 2016). Therefore, a one to one translation between our test results and real-world deployment of the system would be a stretch.

Furthermore, a subjective test method was used, meaning the test participants were asked to what extent the proof of concept would add value to the creation of situation awareness. This is an inherently biased method (Endsley, 2019) that does not deliver the most clear results. However, it did provide insight in the way that the proof of concept could aid the creation of situation awareness for indoor first responder operations, by identifying important situation awareness factors. Direct testing is assumed to deliver clearer results compared to the subjective test results that are presented in this research. The methodology describes both subjective and objective testing. Objective testing however has been placed out of scope during the development phase of this research. The reason for this change is that time was scheduled for direct testing with a very limited amount of test participants (one to three participants) who were already a stakeholder in the set-up of the research. Direct testing with such a user group is assumed to be unrepresentative for first responders as a whole, while results would be likely to be biased by the different roles of the user group.

Therefore, it is reasoned that there can be no definite conclusion about the level of situation awareness that would be available to first responders by using the proof of concept. However, methods have been developed for taking direct and objective measurements of situation awareness, as is for example explained in (Endsley, 2019). Future work can take the results of this thesis and continue with direct testing of the proof of concept, thereby researching usability in a more solid way.

8.2 Stairs confuse HoloLens

As has been shown in this research, stairs can confuse the SLAM algorithm of the HoloLens. This causes the Hololens to lose track of the device, stopping the mapping and the tracking functionality of the device. This has a bad effect on the continuity of the proof of concept results, thereby preventing successful deployment in first responder operations. Within this thesis, no effort has been put on improving the SLAM algorithm as this was regarded out of scope. For successful deployment, the reliability of the proof of concept will have to be improved. This might be done by relying more on the integrated mobility unit (IMU) instead of the visual slam algorithm, as the track was still displayed correctly after the mapping information was repositioned when tracking was lost.

Also, the previous coordinate system could be reinstated after tracking has been found again. Now, if position is lost the device resets its position to the 0,0,0 state once tracking is regained. However, holograms and mapping are replaced once tracking is regained, meaning they are stored and converted from the old coordinate system to the new one. In theory, our application could make use of that knowledge by transforming known meshes with the 4x4 matrix so that the scan results are at least continuous. Spatial anchors could be a good help with this.

8.3 One to one relationship is a limitation

The research problem of this topic has a different nature: different explorers might explore different parts of operation environments. Of course, the resulting models are collected separately from each other and therefore they remain isolated (Figure 55).

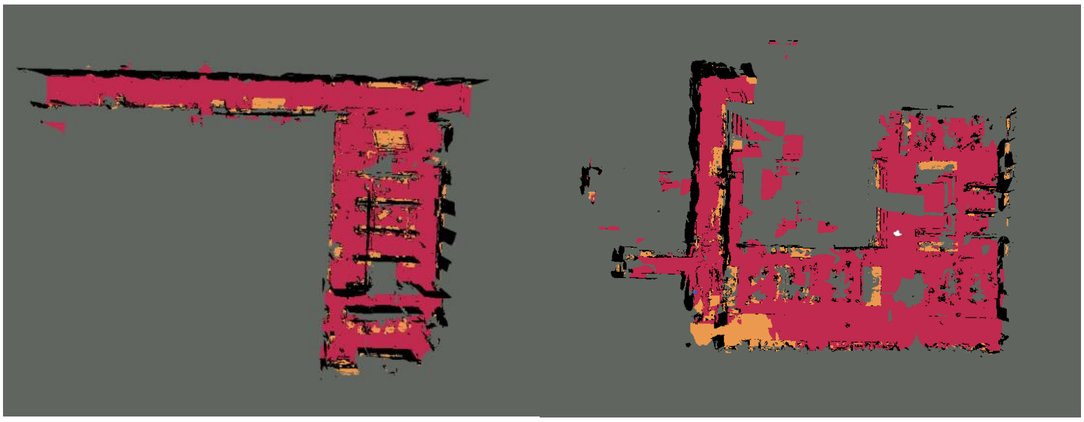


Figure 55: Both left and right are separate scans from separate explorers. However, the spaces connect to each other

However, if two explorers have mapped the same space, this should be recognized by the application. Spatial anchors could offer a solution to recognizing surfaces and spaces. A spatial anchor is a fingerprint that combines visual features, geometry and radio signals into distinctive hashes. These hashes can be stored in the cloud, for example by using Azure Spatial Anchors.

8.4 Divide mapped and unmapped space

As was pointed out by Edward Verbree, only know which surfaces within an indoor environment has been scanned. Therefore, we do not know which surfaces have not been scanned. This seems obvious, but has large consequences. As the Time of Flight scanner of the Microsoft Hololens does only have a range of approximately 3 meters, surfaces are easily missed in the mapping process. Future research could map this unknown space, by combining the known device poses with the mapping information. From all poses, a spatial model can be created by taking the space that should have been mapped from every pose. A frustrum can be calculated from each pose, indicating the space that should have been seen. For example, we know the HoloLens can map for 3 meters in a certain angle of view. If no object is in the way, we can regard this whole frustrum as 'mapped'. Creating an overview of mapped voxel dataset can be done by creating an empty voxel around the initial position of the proof of concept. While the proof of concept runs, the voxels within the dataset are classified in 'mapped empty', 'mapped surface', and the default state 'unmapped'.

8.5 Navigable space

In Figure 35 we can see that the navigable space is not continuous. I.e. a hole has emerged as the floor is not scanned properly. The cause for this is that the Microsoft HoloLens can only scan for 3 meters forward. If a user looks up, a part of the floor might be missed. Careful mapping is required to scan a complete area. Furthermore, the mesh has not been cleaned by post-processing it. Therefore, there has been no attempt to close holes in the mesh automatically. For a more reliable representation of space, such holes need to be handled. At last, we see that navigable space is also drawn upon tables. Such space might be navigable but is not preferred.

Future research could focus on filling the gaps. This would have two main purposes:

- First, successful extraction of navigable space can be used for goal oriented indoor navigation. The proof of concept collects both the geometry of a building and the pose of the explorer. This data is transferred in real-time to a navigator. At the navigator application, navigable space is extracted. The navigator can subsequently pick a positional goal to which the explorer should move. By using the explorer position, the navigator goal, and the extracted navigable space, the system has all elements to calculate a preferred route from the explorer position to the goal. As soon as the route is calculated, the route can be send back to the explorer and visualized in augmented reality.

- Second, agent based simulations can be run in the model if navigable space is known. If a building has been scanned, the extracted navigable space of the model can be used for running relatively low cost agent based simulations. These simulations can for example be used to test evacuation plans. Route calculations are necessary to provide agents with a route to an exit, in case of testing evacuation plans.

For navigable space, we assume floors are most important. Therefore, if the floor is not scanned properly, a hole is likely to emerge. However, we assume floors are generally continuous, making it possible to fill gaps by extending the floor to a certain degree. This process can be made more robust by using a space division in mapped and unmapped space, as described in section 8.4. Then, one is able to make better assumptions about continuous navigable space. For example, in the case of the hole it would have appeared that the space of the floor would have classified as 'unmapped'. It makes sense to think the floor is continuous in this case. However, if the space of the hole was indeed mapped, we should not make such an assumption.

8.6 Extending the method to other devices

The Microsoft HoloLens has proven to be a capable device for mapping, tracking, and communication purposes within indoor environments. However, this does not mean it is the best option for first responder operations. Because the MRTK is multi-platform oriented, the PoC implementation can be ported to different devices. Some of these devices might not be limited by smoke, like the Microsoft HoloLens is. Other devices are able to operate without a first responder being present to wear the device, such as drones or other robots. By using such devices, safety and efficiency of indoor exploration might become even more safe.

8.7 Updating indoor models in preparation phase

An increasing demand for mapping buildings exist, for example for first responder operations and building asset management. Scan techniques are used more frequently to map building automatically, for example by using LiDAR or RGB-D. By mapping the building, the current state of the geometry within a building is saved. However, buildings tend to change over time. This causes the scan reliability to deteriorate, meaning the extent to which a scan represents the real world.

To maintain reliable scan results, a scan should be frequently updated. However, to make a complete new scan may be a resource intensive undertaking. More and more low-cost devices have sensors on-board that can be used for mapping geometry, such as simple camera's, IMU's and depth sensors. These sensors are not very reliable in terms of accuracy and precision, which makes them unsuitable for scanning a whole building.

Although they are too unreliable to make a proper 3D model of an environment, the low-cost devices can be used to detect change. If a reliable scan of the building is available, the low-cost devices can be used to report on changes in parts the environment. This requires a lot of scanning devices over a period of time.

For public buildings, visitors could request navigation from one point in the model to another point in the model as a service. They get a path from origin to goal from a cloud processed request to the 'reliable' 3D model, together with spatial anchors that should be found on the way to the destination. On the way, the device reports about the quality of the spatial anchors that are found. If a spatial anchor is missing, the device reports new information to the cloud infrastructure, thereby detecting a change. If enough reports of the same nature are sent to the infrastructure, an official change is detected and the input of many reports can be used to change the 3D model in the cloud infrastructure in a reliable way. This cloud infrastructure can subsequently be used for other purposes, such as indoor asset management.

As a note: users do <u>only</u> get the information that is <u>needed</u> for them. Often, they don't need the whole 3D model from the cloud infrastructure, as they only need routing information. This routing information is the result of a calculation within the navigable space of the cloud infrastructure model and can be completed in the cloud. The user does only get routing information (a track) and hashed spatial anchors that should be on the path of the user.

9 References

Bailey, T., Durrant-Whyte, H., 2006. Simultaneous localization and mapping (SLAM): part II. IEEE Robotics Automation Magazine 13, 108–117. https://doi.org/10.1109/MRA.2006.1678144

Basilico, N., Amigoni, F., 2011. Exploration strategies based on multi-criteria decision making for searching environments in rescue operations. Auton Robot 31, 401. https://doi.org/10.1007/s10514-011-9249-9

Bondar, S., Salem, B., Stjepandic, J., 2018. Indoor Object Reconstruction Based on Acquisition by Low-Cost Devices. https://doi.org/10.3233/978-1-61499-898-3-113

Borycki, D., 2018. Programming for Mixed Reality with Windows 10, Unity, Vuforia, and UrhoSharp, 1 edition. ed. Microsoft Press, Upper Saddle River, N.J.

Caragliu, A., Bo, C.D., Nijkamp, P., 2011. Smart Cities in Europe. Journal of Urban Technology 18, 65–82. https://doi.org/10.1080/10630732.2011.601117

Chellali, R., Baizid, K., 2011. What Maps and What Displays for Remote Situation Awareness and ROV Localization?, in: Salvendy, G., Smith, M.J. (Eds.), Human Interface and the Management of Information. Interacting with Information, Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 364–372. https://doi.org/10.1007/978-3-642-21669-5_43

Cools, F., Van Duin, M., Wijkhuis, V., 2017. GRIP en de flexibele toepassing ervan. Intstituut Fysieke veiligheid, Arnhem.

Coppens, A., 2017. Merging real and virtual worlds: An analysis of the state of the art and practical evaluation of Microsoft Hololens. arXiv:1706.08096 [cs].

Díaz-Vilariño, L., Verbree, E., Zlatanova, S., Diakité, A.A., 2017. Indoor modelling from slam-based laser scanner: door detection to envelope reconstruction. https://doi.org/10.5194/isprs-archives-XLII-2-W7-345-2017

Dilo, A., Zlatanova, S., 2011. A data model for operational and situational information in emergency response. Appl Geomat 3, 207–218. https://doi.org/10.1007/s12518-011-0060-2

Durrant-Whyte, H., Bailey, T., 2006. Simultaneous localization and mapping: part I. IEEE Robotics Automation Magazine 13, 99–110. https://doi.org/10.1109/MRA.2006.1638022

Durrant-Whyte, H., Rye, D., Nebot, E., 1996. Localization of Autonomous Guided Vehicles, in: Giralt, G., Hirzinger, G. (Eds.), Robotics Research. Springer, London, pp. 613–625. https://doi.org/10.1007/978-1-4471-0765-1_69

Endres, F., Hess, J., Engelhard, N., Sturm, J., Cremers, D., Burgard, W., 2012. An evaluation of the RGB-D SLAM system, in: 2012 IEEE International Conference on Robotics and Automation. Presented at the 2012 IEEE International Conference on Robotics and Automation (ICRA), IEEE, St Paul, MN, USA, pp. 1691–1696. https://doi.org/10.1109/ICRA.2012.6225199

Endsley, M., 2000. Theoretical underpinnings of situation awareness: A critical review. Situation awareness analysis and measurement 3–32.

Endsley, M., 1995. Endsley, M.R.: Toward a Theory of Situation Awareness in Dynamic Systems. Human Factors Journal 37(1), 32-64. Human Factors: The Journal of the Human Factors and Ergonomics Society 37, 32–64. https://doi.org/10.1518/001872095779049543

Endsley, M.R., 2019. A Systematic Review and Meta-Analysis of Direct Objective Measures of Situation Awareness: A Comparison of SAGAT and SPAM. Hum Factors 0018720819875376. https://doi.org/10.1177/0018720819875376

Endsley, M.R., 2016. Designing for Situation Awareness: An Approach to User-Centered Design, 2nd ed. CRC Press, Inc., Boca Raton, FL, USA.

Endsley, M.R., 2015a. Situation Awareness Misconceptions and Misunderstandings. Journal of Cognitive Engineering and Decision Making 9, 4–32. https://doi.org/10.1177/1555343415572631

Endsley, M.R., 2015b. Final Reflections: Situation Awareness Models and Measures. Journal of Cognitive Engineering and Decision Making 9, 101–111. https://doi.org/10.1177/1555343415573911

Endsley, M.R., 1988. Design and Evaluation for Situation Awareness Enhancement. Proceedings of the Human Factors Society Annual Meeting 32, 97–101. https://doi.org/10.1177/154193128803200221

English, L.P., 1999. Improving data warehouse and business information quality: methods for reducing costs and increasing profits. J. Wiley & Sons, New York, NY.

ENISA, 2017. Privacy and data protection in mobile applications.

Epstein, R.A., Vass, L.K., 2014. Neural systems for landmark-based wayfinding in humans. Philosophical Transactions of the Royal Society B: Biological Sciences 369, 20120533. https://doi.org/10.1098/rstb.2012.0533

European Commission (Ed.), 2011. Cities of tomorrow: challenges, visions, ways forward, Oct. 2011. ed, European Union - Regional Policy. Publ. Office of the European Office, Luxembourg.

Fischer, C., Gellersen, H., 2010. Location and Navigation Support for Emergency Responders: A Survey. IEEE Pervasive Computing 9, 38–47. https://doi.org/10.1109/MPRV.2009.91

Flach, J., 2015. Situation Awareness: Context Matters! A Commentary on Endsley. Journal of Cognitive Engineering and Decision Making 9, 59–72. https://doi.org/10.1177/1555343414561087

Flikweert, P., 2019. Automatic Extraction of an IndoorGML Navigation Graph from an Indoor Point Cloud. Delft University of Technology.

Fourment, M., Gillings, M.R., 2008. A comparison of common programming languages used in bioinformatics. BMC Bioinformatics 9, 82. https://doi.org/10.1186/1471-2105-9-82

Fuentes-Pacheco, J., Ruiz-Ascencio, J., Rendón-Mancha, J.M., 2015. Visual simultaneous localization and mapping: a survey. Artif Intell Rev 43, 55–81. https://doi.org/10.1007/s10462-012-9365-8

Geertman, S., Stillwell, J., Toppen, F., 2013. Introduction to 'Planning Support Systems for Sustainable Urban Development,' in: Geertman, S., Toppen, F., Stillwell, J. (Eds.), Planning Support Systems for Sustainable Urban Development, Lecture Notes in Geoinformation and Cartography. Springer, Berlin, Heidelberg, pp. 1–15. https://doi.org/10.1007/978-3-642-37533-0_1

Hashem, I.A.T., Chang, V., Anuar, N.B., Adewole, K., Yaqoob, I., Gani, A., Ahmed, E., Chiroma, H., 2016. The role of big data in smart city. International Journal of Information Management 36, 748–758. https://doi.org/10.1016/j.ijinfomgt.2016.05.002

Heard, J., Thakur, S., Losego, J., Galluppi, K., 2014. Big Board: Teleconferencing Over Maps for Shared Situational Awareness. Comput Supported Coop Work 23, 51–74. https://doi.org/10.1007/s10606-013-9191-9

Hejlsberg, A., Wiltamuth, S., Golde, P., 2003. C# Language Specification. Addison-Wesley Longman Publishing Co., Inc., USA.

Hennig, S., Belgiu, M., 2012. User-centric SDI: Addressing Users Requirements in Third-Generation SDI. The Example of Nature-SDIplus. Geoforum Perspektiv 10.

Hübner, P., Clintworth, K., Liu, Q., Weinmann, M., Wursthorn, S., 2020. Evaluation of HoloLens Tracking and Depth Sensing for Indoor Mapping Applications. Sensors 20, 1021. https://doi.org/10.3390/s20041021

Hübner, P., Landgraf, S., Weinmann, M., Wursthorn, S., 2019. Evaluation of the Microsoft HoloLens for the Mapping of Indoor Building Environments.

Jana, A., Sharma, M., Rao, M., 2017. HoloLens blueprints: experience the virtual and real worlds coming together with HoloLens. Packt Publishing, Birmingham, UK.

Kang, X., Li, J., Fan, X., Wan, W., 2019. Real-Time RGB-D Simultaneous Localization and Mapping Guided by Terrestrial LiDAR Point Cloud for Indoor 3-D Reconstruction and Camera Pose Estimation. Applied Sciences 9, 3264. https://doi.org/10.3390/app9163264

Kapucu, N., Garayev, V., 2011. Collaborative Decision-Making in Emergency and Disaster Management. International Journal of Public Administration 34, 366–375. https://doi.org/10.1080/01900692.2011.561477

Karam, S., Vosselman, G., Peter, M., Hosseinyalamdary, S., Lehtola, V., 2019. Design, calibration, and evaluation of a backpack indoor mobile mapping system. Remote Sensing 11. https://doi.org/10.3390/rs11080978

Keating, E., 2016. Using the SDI framework for the evaluation of geo-information flow & coordination in emergency response. Utrecht University.

Khoshelham, K., Tran, H., Acharya, D., 2019. Indoor mapping eyewear: geometric evaluation of spatial mapping capability of hololens. pp. 805–810. https://doi.org/10.5194/isprs-archives-XLII-2-W13-805-2019

Kitchenham, B., Pretorius, R., Budgen, D., Pearl Brereton, O., Turner, M., Niazi, M., Linkman, S., 2010. Systematic literature reviews in software engineering – A tertiary study. Information and Software Technology 52, 792–805. https://doi.org/10.1016/j.infsof.2010.03.006

Kitchenham, B.A., Budgen, D., Pearl Brereton, O., 2011. Using mapping studies as the basis for further research – A participant-observer case study. Information and Software Technology, Special Section: Best papers from the APSEC 53, 638–651. https://doi.org/10.1016/j.infsof.2010.12.011

Klein, G., 1993. A Recognition Primed Decision (RPD) Model of Rapid Decision Making, in: Decision Making in Action: Models and Methods.

Koeva, M.N., Nikoohemat, S., Elberink, S.J.O., Guarin, J.M.M., Lemmen, C.H.J., Zevenbergen, J.A., 2019. Towards 3D Indoor Cadastre Based on Change Detection from Point Clouds. Remote sensing 11. https://doi.org/10.3390/rs11171972

Kraak, M.-J., Ormeling, F.J., Ormeling, F.J., 2013. Cartography : Visualization of Spatial Data. Routledge. https://doi.org/10.4324/9781315847184

Kurvers, B., 2019. Hoogbouwvisie 2019. Municipality of Rotterdam.

Labbé, M., Michaud, F., 2014. Online global loop closure detection for large-scale multi-session graph-based SLAM, in: 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems. Presented at the 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2661–2666. https://doi.org/10.1109/IROS.2014.6942926

Law, D., 2007. Geodatabases That Are Easier to Set Up and Manage [WWW Document]. esri.com. URL https://www.esri.com/news/arcuser/0807/gdb_precision.html (accessed 3.25.20).

Lehtola, V.V., Kaartinen, H., Nüchter, A., Kaijaluoto, R., Kukko, A., Litkey, P., Honkavaara, E., Rosnell, T., Vaaja, M.T., Virtanen, J.-P., Kurkela, M., Issaoui, A.E., Zhu, L., Jaakkola, A., Hyyppä, H., 2017. Comparison of the Selected State-Of-The-Art 3D Indoor Scanning and Point Cloud Generation Methods. Remote Sensing 9, 796. https://doi.org/10.3390/rs9080796

Li, N., Yang, Z., Ghahramani, A., Becerik-Gerber, B., Soibelman, L., 2014. Situational awareness for supporting building fire emergency response: Information needs, information sources, and implementation requirements. Fire Safety Journal 63, 17–28. https://doi.org/10.1016/j.firesaf.2013.11.010

Linsen, L., 2001. Point cloud representation. Univ., Fak. für Informatik, Bibliothek Technical Report, Faculty of Computer ….

Liu, Y., Fan, Z.-P., Zhang, Y., 2014. Risk decision analysis in emergency response: A method based on cumulative prospect theory. Computers & Operations Research, Multiple Criteria Decision Making in Emergency Management 42, 75–82. https://doi.org/10.1016/j.cor.2012.08.008

Luhmann, T., Robson, S., Kyle, S., Boehm, J., 2013. Close-Range Photogrammetry and 3D Imaging, 2nd ed. ed. De Gruyter, Berlin, Boston. https://doi.org/10.1515/9783110302783

Mautz, R., Tilch, S., 2011. Survey of optical indoor positioning systems, in: 2011 International Conference on Indoor Positioning and Indoor Navigation. Presented at the 2011 International Conference on Indoor Positioning and Indoor Navigation, pp. 1–7. https://doi.org/10.1109/IPIN.2011.6071925

Microsoft, 2020. Windows Mixed Reality developer documentation.

Montemerlo, M., Thrun, S., Koller, D., Wegbreit, B., 2002. FastSLAM: A Factored Solution to the Simultaneous Localization and Mapping Problem. Presented at the Proceedings of the National Conference on Artificial Intelligence.

Mur-Artal, R., Montiel, J.M.M., Tardos, J.D., 2015. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. IEEE Trans. Robot. 31, 1147–1163. https://doi.org/10.1109/TRO.2015.2463671

Mur-Artal, R., Tardós, J.D., 2017. Visual-Inertial Monocular SLAM With Map Reuse. IEEE Robotics and Automation Letters 2, 796–803. https://doi.org/10.1109/LRA.2017.2653359

Naseer, T., Burgard, W., 2017. Deep regression for monocular camera-based 6-DoF global localization in outdoor environments, in: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Presented at the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 1525–1530. https://doi.org/10.1109/IROS.2017.8205957

Neill, W.J.V., Schlappa, H., 2016. Future Directions for the European Shrinking City. Routledge.

Newman, P., Kin Ho, 2005. SLAM-Loop Closing with Visually Salient Features, in: Proceedings of the 2005 IEEE International Conference on Robotics and Automation. Presented at the Proceedings of the 2005 IEEE International Conference on Robotics and Automation, pp. 635–642. https://doi.org/10.1109/ROBOT.2005.1570189

Nikoohemat, S., Diakité, A., Zlatanova, S., Vosselman, G., 2019. Indoor 3d modeling and flexible space subdivision from point clouds, in: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences: ISPRS Geospatial Week 2019. Presented at the 4th ISPRS Geospatial Week 2019, International Society for Photogrammetry and Remote Sensing (ISPRS), pp. 285–292. https://doi.org/10.5194/isprs-annals-IV-2-W5-285-2019

Nikoohemat, S., Diakité, A.A., Zlatanova, S., Vosselman, G., 2020. Indoor 3D reconstruction from point clouds for optimal routing in complex buildings to support disaster management. Automation in Construction 113, 103109. https://doi.org/10.1016/j.autcon.2020.103109

Nuaimi, K.A., Kamel, H., 2011. A survey of indoor positioning systems and algorithms, in: 2011 International Conference on Innovations in Information Technology. Presented at the 2011 International Conference on Innovations in Information Technology, pp. 185–190. https://doi.org/10.1109/INNOVATIONS.2011.5893813

Nüchter, A., Lingemann, K., Hertzberg, J., Surmann, H., 2005. Heuristic-Based Laser Scan Matching for Outdoor 6D SLAM, in: Furbach, U. (Ed.), KI 2005: Advances in Artificial Intelligence, Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 304–319. https://doi.org/10.1007/11551263_25

Okorn, B., Xiong, X., Akinci, B., Huber, D., 2010. Toward automated modeling of floor plans, in: Proceedings of the Symposium on 3D Data Processing, Visualization and Transmission.

Orthuber, E., Avbelj, J., 2015. 3D BUILDING RECONSTRUCTION FROM LIDAR POINT CLOUDS BY ADAPTIVE DUAL CONTOURING. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences II-3/W4, 157–164. https://doi.org/10.5194/isprsannals-II-3-W4-157-2015

Oxford Dictionaries, s.v, n.d.. mapping.

Paré, G., Trudel, M.-C., Jaana, M., Kitsiou, S., 2015. Synthesizing information systems knowledge: A typology of literature reviews. Information & Management 52, 183–199. https://doi.org/10.1016/j.im.2014.08.008

Pomerleau, F., Colas, F., Siegwart, R., Magnenat, S., 2013. Comparing ICP variants on real-world data sets. Auton Robot 34, 133–148. https://doi.org/10.1007/s10514-013-9327-2

Prati, G., Pietrantoni, L., 2010. The relation of perceived and received social support to mental health among first responders: a meta-analytic review. Journal of Community Psychology 38, 403–417. https://doi.org/10.1002/jcop.20371

Preiner, R., Jeschke, S., Wimmer, M., 2012. Auto Splats: Dynamic Point Cloud Visualization on the GPU., in: EGPGV. Presented at the Eurographics Symposium on Parallel Graphics and Visualization, Institute of Computer Graphics and Algorithms, Austria, pp. 139–148.

Rantakokko, J., Rydell, J., Strömbäck, P., Händel, P., Callmer, J., Törnqvist, D., Gustafsson, F., Jobs, M., Grudén, M., 2011. Accurate and reliable soldier and first responder indoor positioning: multisensor systems and cooperative localization. IEEE Wireless Communications 18, 10–18. https://doi.org/10.1109/MWC.2011.5751291

Richter, R., Döllner, J., 2014. Concepts and techniques for integration, analysis and visualization of massive 3D point clouds. Computers, Environment and Urban Systems 45, 114–124. https://doi.org/10.1016/j.compenvurbsys.2013.07.004

Risi, V.D., 2015. Mathematizing Space: The Objects of Geometry from Antiquity to the Early Modern Age. Birkhäuser.

Rusu, R.B., Marton, Z.C., Blodow, N., Dolha, M., Beetz, M., 2008. Towards 3D Point cloud based object maps for household environments. Robotics and Autonomous Systems, Semantic Knowledge in Robotics 56, 927–941. https://doi.org/10.1016/j.robot.2008.08.005

Scholten, H., Fruijter, S., Dilo, A., Borkulo, E. van, 2008. Spatial Data Infrastructure for Emergency Response in Netherlands. Remote Sensing and GIS Technologies for Monitoring and Prediction of Disasters 179–197. https://doi.org/10.1007/978-3-540-79259-8_11

scooley, n.d. Microsoft HoloLens [WWW Document]. URL https://docs.microsoft.com/en-us/hololens/ (accessed 3.5.20).

Semanjski, I., Gautama, S., Ahas, R., Witlox, F., 2017. Spatial context mining approach for transport mode recognition from mobile sensed big data. Computers, Environment and Urban Systems 66, 38–52. https://doi.org/10.1016/j.compenvurbsys.2017.07.004

Seppänen, H., Virrantaus, K., 2015. Shared situational awareness and information quality in disaster management. Safety Science 77, 112–122. https://doi.org/10.1016/j.ssci.2015.03.018

Sheng-Cheng Yeh, Wu-Hsiao Hsu, Ming-Yang Su, Ching-Hui Chen, Ko-Hung Liu, 2009. A study on outdoor positioning technology using GPS and WiFi networks, in: 2009 International Conference on Networking, Sensing and Control. Presented at the 2009 International Conference on Networking, Sensing and Control, pp. 597–601. https://doi.org/10.1109/ICNSC.2009.4919345

Sithole, G., Zlatanova, S., 2016. Position, location, place and area: an indoor perspective, in: ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences. Presented at the XXIII ISPRS Congress, Commission IV (Volume III-4) - 12–19 July 2016, Prague, Czech Republic, Copernicus GmbH, pp. 89–96. https://doi.org/10.5194/isprs-annals-III-4-89-2016

Smith, R., Cheeseman, P., 1987. On the Representation and Estimation of Spatial Uncertainty. The International Journal of Robotics Research 5. https://doi.org/10.1177/027836498600500404

Smith, R., Self, M., Cheeseman, P., 1990. Estimating Uncertain Spatial Relationships in Robotics, in: Cox, I.J., Wilfong, G.T. (Eds.), Autonomous Robot Vehicles. Springer, New York, NY, pp. 167–193. https://doi.org/10.1007/978-1-4613-8997-2_14

Staats, B., 2017. Identification of walkable space in a voxel model, derived from a point cloud and its corresponding trajectory. Delft University of Technology.

Taketomi, T., Uchiyama, H., Ikeda, S., 2017. Visual SLAM algorithms: a survey from 2010 to 2016. IPSJ T Comput Vis Appl 9, 16. https://doi.org/10.1186/s41074-017-0027-2

Tashakkori, H., Rajabifard, A., Kalantari, M., 2015. A new 3D indoor/outdoor spatial model for indoor emergency response facilitation. Building and Environment 89, 170–182. https://doi.org/10.1016/j.buildenv.2015.02.036

Tateno, K., Tombari, F., Laina, I., Navab, N., 2017. CNN-SLAM: Real-Time Dense Monocular SLAM With Learned Depth Prediction. Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 6243–6252.

United Nations, Department of Economic and Social Affairs, Population Division, 2019. World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420). United Nations, New York.

Van Capelleveen, E., Van Duijn, A., Smit, J., Broekhaar, M., Van Wanrooij, M., Zijlstra, G., Geurts, P., 2008. NVBR Digitale Bereikbaarheidskaart. Nederlandse Vereniging voor Brandweerzorg en Rampenbestrijding (NVBR).

Van der Ham, M., 2015. Real Time Localization of Assets in Hospitals using Quuppa Indoor Positioning Technology. Delft University of Technology.

van der Meer, T., 2018. Geovisualization for the Dutch fire brigade. Utrecht University.

Van der Meer, T., 2016. 3D-indoormodellen voor de brandweer – een verkennend onderzoek naar de toepassingsmogelijkheden binnen de veiligheidsketen. TU Delft.

van der Meer, T., Verbree, E., Oosterom, P., 2018. Effective cartographic methods for assisting tactics choice and indoor deployments during building fires – a casestudy the dutch fire brigade. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLII–4, 655–660. https://doi.org/10.5194/isprs-archives-XLII-4-655-2018

van Schouwenburg, S., 2019. Evaluating SLAM in an urban dynamic environment. Delft University of Technology.

Vassallo, R., Rankin, A., Chen, E.C.S., Peters, T.M., 2017. Hologram stability evaluation for Microsoft HoloLens, in: Medical Imaging 2017: Image Perception, Observer Performance, and Technology Assessment. Presented at the Medical Imaging 2017: Image Perception, Observer Performance, and Technology Assessment, International Society for Optics and Photonics, p. 1013614. https://doi.org/10.1117/12.2255831

Verbree, E., Meijers, M., Van Oosterom, P., 2019. GIMA M6 Point Clouds: direct and explorative use.

Walsh, S.. J., Page, P.. H., Brewington, L.., Bradley, J.. R., Mena, C.. F., 2018. 9.13 - A Beach Vulnerability Framework for the Galapagos Islands: Fusion of WorldView 2 Imagery, 3-D Laser Scanner Data, and Unmanned Aerial Vehicles, in: Liang, S. (Ed.), Comprehensive Remote Sensing. Elsevier, Oxford, pp. 159–176. https://doi.org/10.1016/B978-0-12-409548-9.10438-5

Wang, C., Cho, Y.K., Kim, C., 2015. Automatic BIM component extraction from point clouds of existing buildings for sustainability applications. Automation in Construction 56, 1–13. https://doi.org/10.1016/j.autcon.2015.04.001

Weick, K.E., 1995. Sensemaking in organizations. Sage Publications, Thousand Oaks, CA :

Wickens, C.D., 2000. The tradeoff of design for routine and unexpected performance: Implications of situation awareness, in: Situation Awareness Analysis and Measurement. Lawrence Erlbaum, Mahway New Jersey, pp. 211–226.

Wickens, C.D., Lee, J., Liu, Y.D., Gordon-Becker, S., 2013. Introduction to Human Factors Engineering: Pearson New International Edition. Pearson Education Limited.

Williams, B., Cummins, M., Neira, J., Newman, P., Reid, I., Tardós, J., 2009. A comparison of loop closing techniques in monocular SLAM. Robotics and Autonomous Systems, Inside Data Association 57, 1188–1197. https://doi.org/10.1016/j.robot.2009.06.010

Xia, J. (Cecilia), Arrowsmith, C., Jackson, M., Cartwright, W., 2008. The wayfinding process relationships between decision-making and landmark utility. Tourism Management 29, 445–457. https://doi.org/10.1016/j.tourman.2007.05.010

Xu, W., Zlatanova, S., Dilo, A., Van Oosterom, P., 2008. Modelling emergency response processes: Comparative study on OWL and UML, in: Joint ISCRAM-CHINA and GI4DM Conference. Harbin, China.

Zhang, Z.-X., Wang, L., Wang, Y.-M., 2018. An Emergency Decision Making Method Based on Prospect Theory for Different Emergency Situations. Int J Disaster Risk Sci 9, 407–420. https://doi.org/10.1007/s13753-018-0173-x

Zlatanova, S., 2010. Formal modelling of processes and tasks to support use and search of geo-information in emergency response. Proceedings of the 13th annual international conference and exhibition on geospatial information technology and applications - map India 2010 - defining geospatial vision for India, 19-21 January 2010.

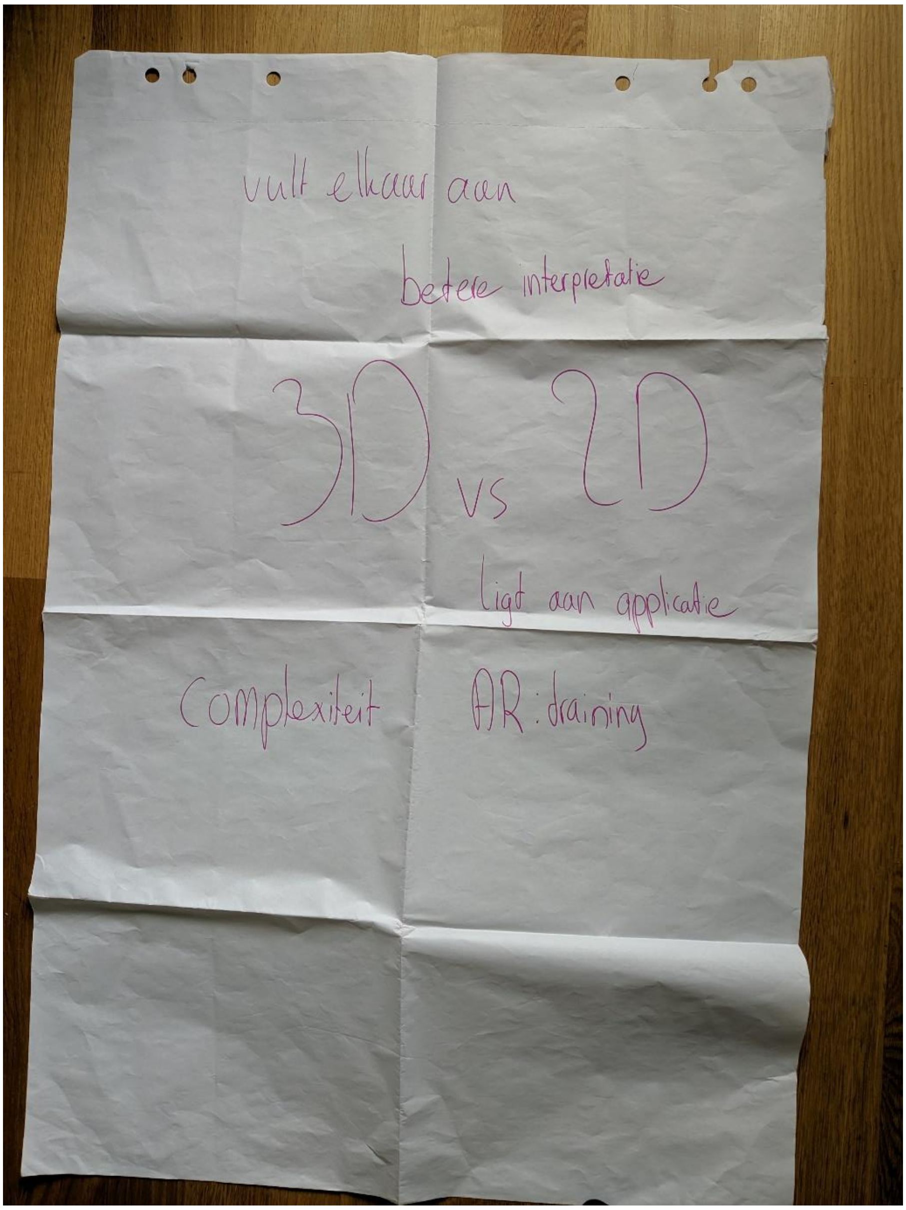
Zlatanova, S., Oosterom, P., Verbree, E., 2004. 3D TECHNOLOGY FOR IMPROVING DISASTER MANAGEMENT: GEO-DBMS AND POSITIONING. Proceedings of The IEEE - PIEEE.

Zlatanova, S., Sithole, G., Nakagawa, M., Zhu, Q., Gist, A., 2013. Problems In Indoor Mapping and Modelling. Presented at the ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. https://doi.org/10.5194/isprsarchives-XL-4-W4-63-2013

Zucconi, A., Lammers, K., 2016. Unity 5.x Shaders and Effects Cookbook. Packt Publishing Ltd.

10 Appendices

10.1 Whiteboard result of first workshop



10.2 Whiteboard result of second workshop