CREATING 3D INDOOR FIRST RESPONDER SITUATION AWARENESS IN REAL-TIME THROUGH A HEAD-MOUNTED AR DEVICE

Bart-Peter Smit [a], Robert Voûte [a,b], Edward Verbree [b],

[a] CGI Nederland B.V.
George Hintzenweg 89, 3068 AX Rotterdam, The Netherlands
bart-peter.smit@cgi.com, robert.voute@cgi.com

[b] Delft University of Technology, Faculty of Architecture and the Built Environment,
Architectural Engineering and Technology, Digital Technologies
2628 BL Delft, The Netherlands
r.l.voute-1@tudelft.nl, e.verbree@tudelft.nl

TS_DigitalTwins

KEY WORDS: First Responder, Real-Time, Digital Twins, Indoor, Mapping, Tracking, Situation Awareness, HoloLens

ABSTRACT:

Emergency operations are a key example for the need of digital twins in the way it is complex, urgent and uncertain. First, the process is complex, as many organizations are involved. Second, it is urgent, as most damage is done in the first moments of an emergency. Third, it is uncertain, as situational conditions tend to change quickly.

For outdoor operations, spatial information systems help in creating an overview of the situation, for example by displaying positions of first responder units involved with the incident. However, spatial data of indoor environments is scarce. Static information of the building, such as floor plans, are often outdated or non-existent. Dynamic operational data such as positions of first responders within the building are only available in a very limited way as well, and often without visual representation.

To create situation awareness of indoor first responder operation environments, this paper successfully proposes a proof of concept with two objectives. First, the proof of concept will collect spatial environment data in the form of mapping and tracking data by using a Microsoft HoloLens. This means the geometry of the building will be collected, together with traversed routes within the building. Second, the data will be streamed and displayed to a remote first responder coordinator in real-time to create a common operational picture. This enables the coordinator to quickly build situation awareness of the operation environment, enabling the coordinator to improve the quality of decisions, thereby improving first responder performance. The proof of concept showed that situation awareness on all three levels increases with the real-time (live) availability (visualisations) of 3D indoor environments. This concept needs to be tested further on usability and performance.

1. INTRODUCTION

This paper is constructed in 7 chapters. First, we will give context to key concepts in the introduction. Second, we will explain this context in a perspective of related academic works. Third, we will explain how we implemented lessons learned from other academics, into a proof of concept (PoC). This PoC will display the feasibility of real-time 3D spatial data acquisition and presentation in emergency operations. Fourth, we will present results of the PoC regarding mapped environments. Fifth, we will discuss performance of the PoC, in terms of accuracy, precision, robustness and added value. Sixth, we will present the conclusion, answering to what extent of real-time 3D spatial data acquisition and presentation is possible in emergency operations, making use of different levels of situation awareness. Finally, suggestions for future work are presented.

1.1 Emergency response, operations, and spatial data

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 (Prati and Pietrantoni, 2010; Dilo and Zlatanova, 2011). 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).

Emergency operations are complex, urgent and uncertain (Kapucu and Garayev, 2011). First, the process is complex, as many organizations are involved. Second, it is urgent, as most damage is done in the first moments of an emergency (Dilo and Zlatanova, 2011). Third, it is uncertain, as situational conditions tend to change quickly (Dilo and Zlatanova, 2011; Kapucu and Garayev, 2011).

For outdoor operations, spatial information systems help in creating an overview of the situation, for example by displaying positions of first responder units involved with the incident (Seppänen and Virrantaus, 2015). However, spatial data of indoor environments is scarce (Rantakokko et al., 2011; van der Meer et al., 2018). Static information of the building, such as floor plans, are often outdated or non-existent. Dynamic operational data such as positions of first responders within the building are only available in a very limited way as well, and often without visual representation.

1.2 The nature of situation awareness

Situation awareness (SA) is a concept that describes to which extent someone is aware of what is happening in a situation, while using that information for gaining an understanding of what that information means in the present or in the future (Endsley, 2016). Next to this, SA is goal-oriented, meaning the awareness can be described in the added value of the information for a specific goal or operation. This paper adopts the formal definition of SA 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". From the definition, three levels of SA can be distilled: perception of data elements, comprehension of the data elements and projection of their status.

Levels of situation awareness

Having SA of an environment means understanding what is going on in that environment (Endsley, 2016). This awareness can be described in three levels. The first level is the lowest level of SA, while the third level describes the highest extent to which SA can be reached.

The first level of SA (perception) relies on availability of data elements that describe a situation. The second level of SA (comprehension) is the transformation and filtering of raw data elements into information that is useful to the system operator. An aspect of this is the combination of raw data elements into a combined interpretation of the data. Finally, the third level of SA is projection. The third level takes the interpretation process of a situation 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.

Any of the three levels of situational awareness can only be reached if there enough mental resources left to understand the situation (Endsley, 2016). Therefore, a user-centered design is imperative, as this the information needs and system requirements of users to prevent information overload are taken into account (Endsley, 2016).

2. RELATED WORK

For the extent to which SA can be built, reliability of the information sources are of the essence (Endsley, 2016). 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 (Seppänen and Virrantaus, 2015). Incomplete or unreliable data sources do therefore hurt the creation of SA. In many domains, the collection of the data to reach level one SA is therefore already challenging (Endsley, 2016). This holds true for first responder operations, as for example the presence of smoke may obstruct visual data collection.

2.1 First responder context

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, a 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).

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 SA (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 (Zlatanova, 2010; Endsley, 2016).

2.2 First responder spatial information availability

If an incident occurs, a process is initiated to mitigate the effects of the incident. Procedures exist for response to many types of incidents, but we do also know that the way in which these procedures are executed can change a lot per situation (Dilo and Zlatanova, 2011; van der Meer et al., 2018).

Dilo and Zlatanova (2011) discern between two types of information at the base of emergency response operations: dynamic and static information. Static data holds information that is 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 and risks maps of the area (Zlatanova, 2010). Dynamic data is volatile in nature and is 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).

To create SA and prevent information overload in first responder operations, we should be aware of the tasks and data requirements for successful response to an incident. Dilo and Zlatanova (2011) provide a data model of Dutch first responder incident response operations, independent of disaster type.

2.3 Spatial data collection

For indoor operations, first responder organizations are often forced to gather data themselves as spatial information is seldom readily available (van der Meer, Verbree and Oosterom, 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. This information is enriched by an indoor exploration of a 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, Verbree and Oosterom, 2018).

The identification of the data sources by exploration is mainly a manual workflow. A real-time indoor application adds to the way dynamic information is shared: by using updated and accurate information about indoor environments, a command and control unit has a better base to make decisions on (Basilico and Amigoni, 2011). A 3D depiction of an environment has an advantage over 2D visualizations in this sense making step, as users are able to view an indoor environment in one complete model instead of viewing it in separate floor plans. Furthermore, 3D data enables users to zoom in to a point of interest to observe it more closely, while a 2D view enables a user to zoom out and still oversee the whole situation (van der Meer, Verbree and Oosterom, 2018). A 2D/3D switch can thereby help in de-conflicting visualisation of situation environments.

To collect 3D data this research uses depth camera. This is a more low-cost alternative to using LiDAR sensing. By using a depth-sensor, distances to surfaces can be sensed instead in a ranged pixel wise way (Khoshelham, Tran and Acharya, 2019). An infrared light is emitted from the sensor, and reflectivity is

mapped in a pixel matrix. As we can measure the time it takes for the infrared light to reflect, we can measure distance by using the Time of Flight (Hübner et al., 2020). The range of this method is less compared to LiDAR scans, but can be acquired in real-time as will be shown in this research.

2.4 SLAM: Basics

While explaining several mapping methods in the section above, the relation between device pose and the mapped environment is introduced. For this research, the estimation of pose (tracking) and the estimation of geometry of the environment (mapping) is treated as a single problem. In literature, this problem is called Simultaneous Localization and Mapping (SLAM) or Concurrent Mapping and Localization (CML) (Durrant-Whyte and Bailey, 2006). A detailed 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, Tran and Acharya, 2019; Nikoohemat et al., 2020). For a detailed explanation of the concept, (Bailey and Durrant-Whyte, 2006) can be consulted.

2.5 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).

This delay of processing is not a problem for many SLAM use cases, as processes such as 'scan to BIM' (Wang, Cho and Kim, 2015) or generation of high quality navigation graphs (Staats et al., 2017; Flikweert et al., 2019; Nikoohema 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 (Rantakokko et al., 2011; Khoshelham, Tran and Acharya, 2019), to which this research aims to make a contribution.

3. PROOF OF CONCEPT

3.1 Coordinator application

The focus of the PoC will be to facilitate the coordinator of the operation with real-time environment data to create SA. The coordinator application will run on a laptop and present mapping and tracking information of the operation environment to the system simultaneous with the explorer process, described below, with the goal to create SA. The operator is able to change the way in which the data elements are visualized and to interact with the data. An example of such interaction is enabling/disabling specific floor levels of the building. This provides an operational picture for this coordinator which unfolds in real time with the action elsewhere.

3.2 Explorer application

As mapping and tracking data of the operation environment is not readily available, we have to collect reliable data as quickly as possible. The data is collected by a first responder who is sent into the building: the 'explorer'.

The explorer is equipped with a Microsoft HoloLens, thereby capable of collecting accurate spatial information within indoor environments (Hübner et al., 2020). To do this, the Microsoft HoloLens is equipped with a depth camera and 4 tracking cameras. Furthermore, the Microsoft HoloLens can transfer data by using Bluetooth and WLAN connections. Additionally, the Microsoft HoloLens is head mounted, leaving the hands of the explorer free. The explorer is therefore more flexible in climbing over or removing obstacles and rubble, or even help victims of the incident. At last, the holographic screen offers the explorer the opportunity to receive visual instructions the holographic screen is able to give visual feedback of the mapped environment, display menus for interaction with the application and receive instructions from the coordinator.

Two kinds of spatial information will be collected simultaneously. First, mapping information will be collected, meaning geometric measurements of the indoor environment in the form of a spatial mesh. This spatial mesh will represent the environment in a 3D model. Second, tracking information of the explorer will be collected. This information will represent the pose (orientation + position) of the explorer over time within the mapped environment.

3.3 Proof of concept development

The PoC is developed in C# in Unity3D. This is a game engine, making it suitable for fast processing and visualization of 3D features. Furthermore, it enables us to deploy interaction methods more easily compared to developing the middleware ourselves. At last, Microsoft has deployed the second version of the Mixed Reality ToolKit (MRTK_V2), to provide developers for easy interaction with the Microsoft HoloLens hardware with C# code.

3.4 Mapping module

At the explorer side, 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 that environment can be displayed. The objective of the mapping module is to capture the indoor 3D geometry, which is done by using the built in mapping capability of the Microsoft HoloLens.

The Microsoft HoloLens is equipped with a time of flight depth-sensor. Surfaces are scanned by the Microsoft HoloLens depth-sensor and the environment cameras, resulting in a point cloud. The Mixed Reality ToolKit transforms this point cloud into a spatial mesh with a set level of detail (Hübner et al., 2020).

At the coordinator side, the mapping information should be presented in real-time to the system operator in a way that it is easily interpreted. The mapping information is collected in the form of a spatial mesh. To prevent information overload and make interpretation easy, the spatial mapping data is visualized in attribute space, object space, and temporal space (Kraak, Ormeling and Ormeling, 2013). When visualized in attribute space, the 3D model does only take geometric aspects such as the normal and the relative height of a surface into account. It is therefore a very robust method, as there are few rules applied to how the spatial mesh is visualized. When visualized in object space, a higher level of interpretation is applied by the system. If this visualization method is used, the system tries to separate floors, walls, obstructive objects, stairs and ceilings in a visual way by depicting them in different colors. The translation from the raw geometry data into features should help to form a mental model of the situation and through this process, SA (Endsley, 2016). At last, when visualized in temporal space, a time component is added to the attribute visualization. First responder operations are very dynamic, as environment conditions tend to

change quickly. Therefore, it is beneficial to know when a certain part of an environment is scanned. You could say the reliability of a mapped environment 'decays' over time, which is represented by adding other colors to the attribute visualization. Various services can be built upon the collected 3D model, such as indoor navigation. We can ask the system operator to interpret whether a space is navigable. However, this is assumed to be a difficult and time consuming task that can be automated (Rantakokko et al., 2011; Seppänen and Virrantaus, 2015). Therefore, this research will explore if calculation of navigable space is possible in real-time for the spatial mesh created by the Microsoft HoloLens by using 'navigation meshes', first introduced by (Snook, 2000).

3.5 Recognizing floor levels

Next to the visualization of the mapped features, one should consider multi floored 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, Verbree and Oosterom, 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. The application will enable the coordinator to show or hide floors with a single press of a button. Therefore, spatial meshes need to be segregated based on floor level.

The PoC uses a method inspired by (Díaz-Vilariño et al., 2017) to recognize floors, 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 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.

3.6 Tracking module

The explorer will be tracked by the system in terms of a series of 'poses'. The pose is collected by the Microsoft HoloLens and is a combination of a position (x, y and z coordinate in cartesian system) and orientation (roll, pitch and angle of the device). As the Microsoft HoloLens is head mounted, we do not distinguish between the pose of the device and the pose of the head of the explorer.

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. This conceptual model is enhanced by literature of the likely predecessor of the Microsoft HoloLens SLAM algorithm: KinectFusion (Khoshelham et al., 2019). This enables the Microsoft HoloLens to track itself with an accuracy of 2 centimeters (Hübner et al., 2020).

3.7 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. 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. This is identified as a threat for first responder applications, as it would render the system temporarily useless. According to the documentation,

tracking loss can be experienced especially if the following (combination of) aspects are apparent:
- 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 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 (reference points) 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 more low 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.

3.8 Communicating module

The data transfer of spatial data elements from explorer to coordinator device will happen via an Microsoft Azure Queue. By choosing this communication protocol, the data is routed through the Azure service, which acts as a bridge. Therefore, the data is first sent from explorer application to the Azure service. Subsequently, the coordinator application polls for new messages and retrieves them from the server if they are available.

It is assumed a continuous WLAN connection is troublesome in first responder operation environments. To prevent connection problems, a mobile phone connected with a fourth generation mobile network is brought with the explorer. The smartphone is used as a hotspot, thereby connecting the Microsoft HoloLens indirectly to a mobile network.

3.9 Situation Awareness Testing

To discuss the outcome of the PoC, three meetings with first responders have been organised. Both the requirements evaluation and first responder meetings are used to indicate whether the PoC is able to create a certain level of SA as described by (Endsley, 2016): perception (level 1), comprehension (level 2) or projection (level 3).

4. RESULTS

4.1 Mapping, Tracking, and Communicating

Every 0.5 seconds, a depth image is captured and processed into a mesh with a maximum size of 500 triangles. That mesh is subsequently sent to the coordinator application. With this settings, the explorer can walk and look around in an environment, while capturing the environment simultaneously in a spatial mesh. The testing environment used to illustrate the process is an office environment. Office environments are often a blind spot for first responders, as data about the interior is often unavailable or outdated. Furthermore, the environment depicted in this research has been chosen for its distinctive furniture and subspaces. The two left spaces of the room contain two tables, while in the right space there is a number of chairs facing a presentation stand.

Both mapping data and tracked poses were collected without problem. Every time a mapping or tracking data element was created by the explorer application, the data element was

transferred within a second to the coordinator application. Therefore, real-time mapping, tracking and data transfer is found to be possible.

4.2 Capturing indoor points of interest

By capturing indoor geometry, surfaces like walls, floors and stairs are mapped. A highly requested feature of involved stakeholders 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 for use in augmented reality. With this menu, objects can be pinpointed in space with a coloured sphere, which is sent to the coordinator together with the geometry.

4.3 Created situation awareness

The mapping module describes the collection of raw data within the explorer application. The collected data has to be presented in the coordinator application in a way that the coordinator 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 SA levels easier (Endsley, 2016). As stated in the methodology, three visualization perspectives will be used: Geometry focused, Time focused, and Object focused.

4.3.1 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 connectedness 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.

The geometry focused visualization can be seen in Figure 1. 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 1: From left to right: Geometry focused visualization in colour, normalized and normalized + height

As we can observe from the coloured normal visualization, walls have a distinctive colour from the floor. As the walls have a horizontal normal, all walls are coloured distinctive from the vertical normals. However, as the walls are placed almost perpendicular to each other, all walls do also have a different colour 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 colours are harmonized to the verticality of the normal: because of this, the horizontal direction does not matter anymore.

4.3.2 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 colour from a separate variable: the last update time. If a spatial mesh has not been updated for a set amount of time, a colour will be added to current representation of the mesh, see

Figure 2. The colours and time thresholds are set dynamically by the coordinator, as different operators might prefer different settings for different operation environments.
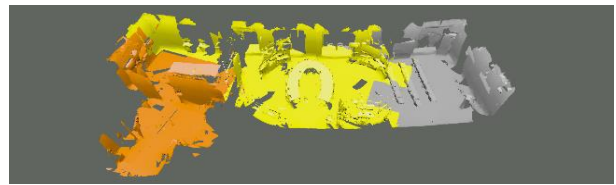

Figure 2: Time focused visualization. From left to right: > 5 minutes = orange; 3-5 minutes = yellow; 0-3 minutes = grey

4.3.3 Object focused

Finally obstructive objects, doors and walls are shown distinctively (Figure 3). Generally, this presentation is received as the best visualization of the spatial mesh by first responders. 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 3: Object oriented visualization. Orange: obstructive objects. Black: walls. Red: floors

4.4 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. Figure 4 shows the extracted navigable space in a Unity3D NavMesh.

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 PoC, 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.

Figure 4: Navigable space represented in a NavMesh

4.5 Tracking Presentation

In Figure 5 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. As our vision

is directed to motion, this makes it easy to follow the explorer while it moves through the scene.

Figure 5: Tracking representation. The left part of the model has been scanned while the right part of the model is not yet scanned.

4.6 User Interface

In
Figure 6 the user interface of the coordinator application is shown. As can be observed, the object oriented spatial mesh visualization is used in this interface. Several items are displayed in the 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).
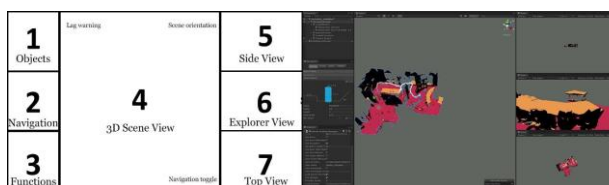

Figure 6: User interface. Left: conceptual overview. Right: the system as it is in use.

5. PERFORMANCE

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 SA: 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.

5.1 Accuracy

The accuracy tests describe errors in the way that objects are represented by the PoC at the right position: it tests whether the measurements are actually there or if the mapping information has drifted a bit. From (Khoshelham, Tran and Acharya, 2019), we do already know the local accuracy of the Microsoft HoloLens spatial meshes is about 5 centimetres from the real world representation. Hübner (2020) shows the local tracking capabilities of the device are around 2 centimeters. Therefore, we will not discuss the accuracy of the Microsoft HoloLens on single

floors further. Both Hübner and Khoshelham (2019) only consider single floors. There is reason to suspect differences between horizontal and vertical accuracy, as the HoloLens SLAM did not perform well on staircases.

5.2 Precision

Precision measures the consistency of the measurements, often described by the resolution of the data (Luhmann et al., 2013). Precision translates to the level of detail of PoC. 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. A visual comparison between the spatial mesh and the point cloud can be observed in Figure 7. We can distinguish desks (orange), chairs and people (black/orange), floors (red), walls (black), and columns (black). The columns are window frames. We can see that the geometry in the images looks the same. However, the geometry of the window frames and the chairs in the PoC scan are often jagged. Therefore, the scan looks to be imprecise. Precision could be increased within the current setup of the PoC at the cost of scan update frequency. The SA tests will cover if low precision of the data is a problem for the feasibility of gaining SA from the data.
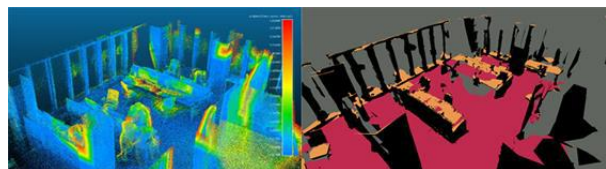

Figure 7: Scanned environment. MLS point cloud (left), indicating distance to the PoC scan (right)

5.3 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 PoC fails for whatever reason to report environment data in real-time to the coordinator application, it is regarded as a failure of the robustness of the system.
In the development and testing phase of the PoC, a couple of aspects have been identified that are important for the robustness of the system. They will be explained in the subsections below.

5.4 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. 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 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.

5.5 Added value as defined by first responders

Three meetings with first responders have been organized to discuss added value of the PoC to indoor first responder operation SA.
The first meeting set the requirements for the PoC. Two officers of duty of a Dutch Safety Region were involved. They agreed with the statement that basic, geometric collection of mapping and tracking data would benefit indoor SA greatly. They stressed the data should be presented in a interpretable way: displaying

geometry was not simple enough. Furthermore, indoor navigation should be a focus of the eventual system. At last, besides catching indoor geometry, important objects such as fire hydrants should be added to the 3D model as well.

The second meeting was organized with two groups of first responders of different organizations at a Dutch event for 3D data use for first responder operations. Here, progress of the PoC was discussed with about twenty first responders. The participants noted they were impressed with the transfer of the mapping and tracking data. Although it was visualized in the basic geometry focused presentation (see section 4.3.1), the first responder stated they could interpret the data fairly easily. They also focused on the importance of time (both in mapping and tracking data) and on the ability to add objects such as exit signs and victims, which was only partly implemented at the time.

Finally, a third meeting was organized at a Dutch Safety Region to discuss the added value of the PoC to the creation of indoor SA. A demo was given to the 13 participants, among who officers of duty and indoor map makers were present. Next to the provision of mapping and tracking data, the potential of extracting navigable space for future navigation and evacuation simulation purposes were highly valued. The data was understood and comprehended into mission critical information, as was stated by the first responders. Of course, this statement is subjective in nature, but it does still support the conclusion that creation of level two SA is possible with the PoC. Furthermore, the temporal component of both mapping and tracking data has been presented, of which one officer of duty stated this enabled him to follow and predict the states of the situation environment.

6. CONCLUSION

This original research creates SA from real-time shared 3D mapping data. It shows that a common operational picture for a first responders commander or control room coordinator can not only be realized, but also helps to improve the SA.

The PoC creates both mapping and tracking data elements in an indoor first responder operation environment and is able to transfer these data elements to a remote coordinator in real-time. The collected data elements fit the spatial data requirements for first responder operations.

As we know it is important for first responders to know to what degree they can trust the system. We have seen the data elements are reliable in a way that they are accurate (below 10 centimetres deviation from TLS results), relatively precise (features are jagged, but can be interpreted) and robust under set circumstances (PoC delivers continuous streams of data elements in real-time if tracking is not lost).

To aid the interpretation process, mapping elements can be visualized in attribute, temporal, and object space as explained in section 3.4. Attribute and temporal space visualization is most reliable as these visualizations are based on few rules. Object space is more easy to interpret, as elements separated into floors, walls, stairs, and obstructive objects, giving more meaning to the 3D model compared to the raw geometry of the maps. For object space, ease of interpretation comes at a cost of more complex visualization rules making the visualization less reliable. Therefore, we recommend to use object space as default visualization method, while using attribute space visualization as fallback option. The mapping information can be extended with named objects, such as fire hydrants. The tracking information is integrated with the mapping information, enabling operators to quickly review the spatial data of the operation environment.

The findings can be attributed with levels of SA. We have seen the PoC fulfils the data requirements of first responders regarding mapping and tracking. This fulfils the requirements for the first level of SA: perception. Furthermore, we have seen the mapping and tracking data can be integrated in one common operational picture, combining elements in a interpretable way. This leads to the second level of SA. Nevertheless, this is a theoretical conclusion based on subjective testing together with first responders.

Concerning projection (as third level of SA) it is concluded that by showing the age of data elements, the model does not only give a comprehensive idea of when a certain area has been scanned, but it also assures the user that always the most current situation will be displayed. This might be used to project change over time.

7. FUTURE WORK

7.1 Connecting multiple explorer devices

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.

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 one distinctive hash. These hashes can be stored in the cloud, for example by using Azure Spatial Anchors. Ideally, both models would be joined in one large model of the operation environment. One large model will likely require less mental resources of a coordinator compared to two smaller models.

7.2 Dividing mapped and unmapped space

We 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 PoC. While the PoC runs, the voxels within the dataset are classified in 'mapped empty', 'mapped surface', and the default state 'unmapped'.

7.3 Improving navigable space extraction

Navigable space is not always continuous in the models generated by the PoC. 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 while it is passed. Careful mapping is therefore required to scan a complete area. Future research could focus on filling such gaps by post-processing the spatial mesh of the model. If combined with a mapped/unmapped space division, such post-processing could be more aggressive for parts that have not been scanned but are likely to be connected to mapped space.

REFERENCES

Bailey, T. and Durrant-Whyte, H. (2006) 'Simultaneous localization and mapping (SLAM): part II', *IEEE Robotics Automation Magazine*, 13(3), pp. 108–117. doi: 10.1109/MRA.2006.1678144.

Basilico, N. and Amigoni, F. (2011) 'Exploration strategies based on multi-criteria decision making for searching environments in rescue operations', *Autonomous Robots*, 31(4), p. 401. doi: 10.1007/s10514-011-9249-9.

Díaz-Vilariño, L., Verbree, E., Zlatanova, S., Diakité, A. (2017) 'Indoor modelling from SLAM-based laser scanner: door detection to envelope reconstruction', in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. ISPRS Geospatial Week 2017*, Copernicus GmbH, pp. 345–352. doi: 10.5194/isprs-archives-XLII-2-W7-345-2017.

Dilo, A. and Zlatanova, S. (2011) 'A data model for operational and situational information in emergency response', *Applied Geomatics*, 3(4), pp. 207–218. doi: 10.1007/s12518-011-0060-2.

Durrant-Whyte, H. and Bailey, T. (2006) 'Simultaneous localization and mapping: part I', *IEEE Robotics Automation Magazine*, 13(2), pp. 99–110. doi: 10.1109/MRA.2006.1638022.

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

Flikweert, P., Peters, R., Díaz-Vilariño, L., Voûte, R. L., Staats, B. R., (2019) 'Automatic extraction of a navigation graph intended for indoorGML from an indoor point cloud', *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-2/W5, pp. 271–278. doi: 10.5194/isprs-annals-IV-2-W5-271-2019.

Hübner, P., Clintworth, K., Liu, Q., Weinmann, M., Wursthorn, Sven. (2020) 'Evaluation of HoloLens Tracking and Depth Sensing for Indoor Mapping Applications', *Sensors*, 20(4), p. 1021. doi: 10.3390/s20041021.

Kapucu, N. and Garayev, V. (2011) 'Collaborative Decision-Making in Emergency and Disaster Management', *International Journal of Public Administration*, 34(6), pp. 366–375. doi: 10.1080/01900692.2011.561477.

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

Kraak, M.-J., Ormeling, F. J. and Ormeling, F. J. (2013) *Cartography : Visualization of Spatial Data*. Routledge. doi: 10.4324/9781315847184.

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

van der Meer, T., Verbree, E. and 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, pp. 655–660. doi: 10.5194/isprs-archives-XLII-4-655-2018.

Nikoohemat, S., Diakité, 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, p. 103109. doi: 10.1016/j.autcon.2020.103109.

Prati, G. and 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(3), pp. 403–417. doi: 10.1002/jcop.20371.

Rantakokko, J., Rydell, J., Strömback, P., Händel, P.(2011) 'Accurate and reliable soldier and first responder indoor positioning: multisensor systems and cooperative localization', *IEEE Wireless Communications*, 18(2), pp. 10–18. doi: 10.1109/MWC.2011.5751291.

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

Snook, G. (2000) 'Simplified 3D Movement and Pathfinding Using Navigation Meshes', in *Game Programming Gems*. Charles River Media, pp. 288–304.

Staats, B. R., Diakité, A. A., Voûte, R. L., Zlatanova, S. (2017) 'Automatic generation of indoor navigable space using a point cloud and its scanner trajectory. *SPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4(2W4), pp. 393–400. doi: 10.5194/isprs-annals-IV-2-W4-393-2017.

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

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*.