Using Video for Indoor Navigation Guidance

Achmad Maulana Malikul Hakim

University of Tampere
Faculty of Communication Sciences
Human-Technology Interaction
M.Sc. thesis
Supervisor: Jaakko Hakulinen
June 2017

University of Tampere
Faculty of Communication Sciences
Human-Technology Interaction
Achmad M.M. Hakim: Using Video for Indoor Navigation Guidance
M.Sc. thesis, 56 pages, 8 appendix pages
June 2017

Buildings such as shopping malls or university campuses are complex and have always been a challenge to walk through. While map is the commonly used navigation aid, it has several limitations. On the other hand, there are studies about how visual navigation elements can help people to navigate better.

This thesis would like to explore if video can be used as an alternative method to provide navigation information. The findings of visual navigation elements from other studies were implemented in videos.

A user study was performed to see the feasibility of video to be used as an alternative medium to present navigation information. There were 10 participants who joined the user study. In the user study, the participants were asked to understand routes by reading a map and watching a navigation video and then walk to the specified rooms. Questionnaires were provided to the participants to record their experience in using the map and the video to understand the routes. Interviews were conducted at the end of each user study to get more comments and feedback from the participants.

The user study showed that watching navigation videos allow participants to perform navigation in a more efficient manner with fewer error compared to reading maps. The experience of using videos to receive navigation information is also better compared to maps.

The result of the user study suggests that providing navigational information in the form of video can be considered as an alternative to the traditional map. Furthermore, this thesis work also compiled recommendations on how to produce navigation videos.

Keywords: navigation, indoor navigation, video guidance, map, guideline

Table of Contents

Table of Figures

1 Introduction

Buildings such as shopping malls or university campuses are getting bigger, have more rooms, and have always been a challenge to walk through. Looking for a specific location in a complex building is not easy, especially for those who have never been to the building. Furthermore, buildings obstruct GPS signal which makes devices that use GPS technology not work properly. To overcome this, a building usually has a map of the building or provides an information desk for visitors to ask for help. The information that is given from those sources usually comes in the form of a blueprint of the building (map), text with direction information, or spoken directions given by the person at the information desk.

A map is one of the widely-used solutions to help people to reach their destination in a building (Rooke, Tzortzopoulos, Koskela, & Rooke, 2009). However, there are reasons why a map is not the best solution. First, a map requires additional mental effort from its reader (Schaik, Mayouf, & Aranyi, 2015). The nature of a map is *allocentric* which is a way to encode spatial information with respect to elements and features of the environment (Ruggiero, Iachini, & Ruotolo, 2009). As an example of allocentric perspective, a map tells that the location of a computer laboratory is on the second floor of the building regardless where the map reader is currently standing. Another way to encode spatial information is *egocentric* which specifies location and orientation with respect to the person (Ruggiero, Iachini, & Ruotolo, 2009). An example of information with egocentric perspective would be telling that a computer laboratory is 500 meters to the right of the person who was looking for the computer laboratory. Using a map requires a person to process the change of perspective from allocentric perspective (when the person reads the map) to egocentric perspective (when the person starts walking) (Münzer & Stahl, 2011).

Another existing tool to help people navigate indoors is to use Augmented Reality (AR). People can walk and be given the navigation instruction by looking through a device that superimposes additional information to the real world (e.g., Mulloni, et al., 2011). However, this solution requires people to hold such device and bring it with them. This could be a problem as not everyone has a suitable device. If the building would like to provide the device, it would cost more than just providing a better way to inform people about the building.

Until now, only few studies have discussed using video as a media to transfer information about navigation. On the other hand, previous studies have discovered different visual elements that can be useful in helping people navigate such as You-Are-Here symbol (Klippel, Freksa, & Winter, 2006), landmark highlighting (Möller, et al., 2014), path indication (Darken & Sibert, 1996), and presenting information in egocentric perspective (Münzer & Stahl, 2011). These elements could be incorporated into a single video and could allow the video to be a good navigation aid. This possibility allows video to be considered as an alternative media to transfer navigation knowledge to a person.

This thesis aims to explore if a navigation video can be an effective method to give direction information inside buildings. Such navigation videos could then be included in public displays such as information kiosks. For example, when someone is looking for a specific room, he/she can perform a search for that room on an information kiosk and then a navigation video could be presented to tell him/her how to reach the room. This method would not require a person to have special devices as the information kiosks can be provided by the owner of the building. Furthermore, this thesis also provides a set of guidelines of what needs to be considered when making a navigation video based on findings from literature and a user evaluation studying navigation videos and maps. For this thesis work, navigation videos were designed, recorded, and edited based on findings from literature and then evaluated in a user study to see if the navigation videos can help people to navigate in a building.

In summary, this thesis work suggests that video can be considered as an alternative medium to present navigation information. The user study showed that watching navigation videos resulted in better performance in navigating compared to maps including fewer errors and faster time to reach a certain place due to more efficient navigation. The interviews during the user study also revealed that navigation videos have benefits, for example they were considered easy to understand and they allowed the user study participants to remember landmarks to be used during navigation.

The rest of the thesis is following. Chapter 2 explains navigation and its elements, previous studies of indoor navigation, as well as brief explanations of video and user experience. Chapter 3 details how the navigation videos for this thesis work were designed and made. Chapter 4 describes how the user study was conducted. Chapter 5 reports the results from the user study. Chapter 6 discusses the analysis of the results. Chapter 7 presents a set of guidelines to produce navigation video. Lastly, chapter 8 concludes the findings.

2 Literature Review

This chapter describes previous studies that can support the creation of navigation video. The chapter includes the basic understanding of navigation and information that might help a person to navigate better, existing indoor navigation studies, the basics of video, and information about what user experience is.

2.1 Navigation

Per Oxford Dictionary, navigation is “the process of activity of accurately ascertaining one’s position and planning and following a route” (Oxford University Press, n.d.). Based on this definition, there are three main activities of navigation: identifying the position of ourselves in a certain location, coming up with a route plan to reach the destination we would like to go to, and finally executing the plan by following the route.

Navigation consists of two different tasks: motion and wayfinding (Darken & Peterson, 2001). Motion is the motoric task of navigation. It involves physically traveling to the destination a person would like to reach. Wayfinding is the tactical and strategic task of navigation. Wayfinding is more than just an act of planning a route. It also includes the process of constructing as well as using a *mental map*, a representation of physical environment in our head, to find our way in an environment (Darken & Peterson, 2001). A wayfinding system can be deployed to help people to build their mental map. Such a system can employ signage, color codes, or highlighting a specific part of the building that can help people navigate in the building (Rooke, Tzortzopoulos, Koskela, & Rooke, 2009).

Information inside a mental map is called spatial knowledge (Tolman, 1948). Spatial knowledge can be described in three different types of knowledge using LRS model (Siegel & White, 1975): Landmark, Route, and Survey. *Landmarks* are the prominent and static objects in an environment that can help people orient themselves. Once a person knows the landmarks in an environment, *route* will help the person to know how to go from a landmark to another landmark. However, knowing the route from A to B, and from B to C, does not necessarily allow a person to know a direct route from A to C. *Survey* is the complete knowledge of an environment. Once a person acquires the survey knowledge of an environment, the person knows the direct route from A to B, B to C, as well as A to C. Additionally, the way a person acquires survey knowledge also affects their performance in navigation. Survey knowledge acquired from a map requires the person to translate the knowledge from map representation into first-person view representation which might result in a lower performance in navigation.

Related information about navigation such as orienting oneself, landmarks, path, information perspective, map and navigation instruction are described on the next subchapters.

2.1.1 Orient Oneself

When someone uses a navigation instruction such as a map, it is a prerequisite to understand their location before planning a route to take and deciding where to go (Liben & Downs, 1993).

To allow someone to relate a map that they are reading with the environment that they are standing in, there are aspects that should be considered (Klippel, Freksa, & Winter, 2006):

- Alignment

A map should be aligned with the environment in which it is placed. Having the same Left and Right between the map and the environment enhances the formation of correspondence (e.g., Shepard & Hurwitz, 1984; Warren & Scott, 1993). Furthermore, most people do expect maps to be aligned with the environment (Shepard & Hurwitz, 1984).

- Architectural cues

A map should include architectural cues and natural landmarks. Furthermore, the cues should represent the actual shape of the landmarks.

- You-Are-Here (YAH) symbol

YAH symbol should fulfill two functions. The first one is to locate user's location. The second one is to indicate user's orientation. A YAH symbol that fulfils those two functions is referred as a *complex YAH symbol* (Levine 1982).

2.1.2 Landmarks

Landmarks are the easily noticeable objects in an environment. Identifying landmarks is an important part of navigation as they can be used to indicate whether a person needs to make a turn or to keep walking straight. Knowing enough landmarks can also be helpful for someone to assure that they are traveling on the right path (Darken & Peterson, 2001; Hile, et al., 2009). Additionally, finding a landmark can help to locate another landmark if a person already know the route between them (Striegnitz & Majda, 2009; Hile, et al., 2009).

There are several ways to choose objects as landmarks. One way is to evaluate objects by their attraction values (Kallioniemi, et al., 2013). These attraction values are based on their visual, semantic, and structural properties. An object has a high visual value if it has its own specific characteristic, such as its location in the middle of an open area or its form that is unique from the other objects around it. The second property, the semantic value of a landmark is related to the historical or cultural values of the landmark. Lastly, a landmark is structurally attractive if it has importance for the surrounding area, such as a big road intersection, a bridge, or a big plaza.

Another way to choose objects as landmarks is based on their positions in relation to the bend a person needs to take a turn into (Hile, et al., 2009). The landmark is badly positioned if it is located on the opposite of the turn a person needs to take. For example, choosing a statue to the left of the road when the person needs to make a right turn. Next, the acceptable position of a landmark would be in front of the bend a person needs to go into. For example, telling a person to make a turn before a shopping mall. Lastly, the best position of a landmark is inside the bend. For instance, telling a person to make a turn right after a school on the right side of the road.

The next step after choosing landmarks is to help the user notice these landmarks by highlighting them. A previous study on navigation for augmented reality (Möller et al., 2014) used two techniques to highlight landmarks, frame-based highlight and soft highlight. Frame-based highlight would show outer lines around a landmark. The study showed that frame-based highlight technique attracts user's attention greatly which sometimes distracts user during navigation. The second technique, soft highlight, will overlay a landmark with a transparent layer of color. This technique does not distract user too much but gives worse readability if the landmark contains text.

2.1.3 Path and Animation

Path is a way or track that is made for traveling (Oxford University Press, n.d.), such as street or canal. Beyond that definition, path can also be a natural feature of an environment such as riverbank or coastline which people may walk along as if it was a path. Having path indicated in navigation instruction is helpful for the reader. Because following a path is a natural way of people to travel to their destination (Darken & Sibert, 1996). Kevin Lynch also mentioned Path as one of the types of elements of people's cognitive maps. Hence, having a path in the navigation instruction helps reader to plan their route.

Signifying the path in navigation that needs to be taken can be done by putting markers on top of the path, for example arrows or lines. Furthermore, presenting the path with animation also helps people to understand their route better. This combination of marker and animation has two benefits. First, the information of making a turn is coded twice, both symbolically by the marker and implicitly by the movement in animation. Second, the animation bears a resemblance to the movement that the person will do when the person goes along the route (Münzer & Stahl, 2011). Therefore, combining marker and animation to signify path can be one way to improve navigation instructions.

2.1.4 Perspective, Allocentric VS Egocentric

In spatial cognition, there are two ways of encoding spatial information, Allocentric and Egocentric (Ruggiero, Iachini, & Ruotolo, 2009). Allocentric way encodes spatial information with respect of element and features of the environment (Ruggiero, Iachini, & Ruotolo, 2009). For example, saying that classroom ABC is located 300 meters north from the canteen is an allocentric way to explain classroom ABC's location. Map is one of the examples of navigation instruction that has allocentric perspective. On the other hand, Egocentric way specifies location and orientation with respect of the person (Ruggiero, Iachini, & Ruotolo, 2009). Egocentric perspective is the eye-level perspective which is how people see the environment in the actual navigation (Münzer & Stahl, 2011). For instance, telling someone that classroom ABC is located 300 meters to the right of the person, is an egocentric way of locating classroom ABC. Google street view is one of the existing products that uses egocentric perspective.

Previous studies have indicated that telling spatial information in egocentric perspective has more benefits than in allocentric perspective. First, information in egocentric perspective is plausible to be easier to understand and to learn than in allocentric perspective (Münzer & Stahl, 2011). The reason is because information that is given in allocentric perspective requires people to shift the information from allocentric perspective to egocentric perspective. Second, it is possible that having an egocentric navigation instruction allows both men and women to learn their route better (Münzer & Stahl, 2011). Furthermore, previous study also mentioned that female can acquire better understanding from egocentric perspective direction while male can learn both from allocentric and egocentric perspective (Münzer & Stahl, 2011). Lastly, processing information that is in egocentric perspective is more accurate than the allocentric perspective one (Ruggiero, Iachini, & Ruotolo, 2009).

2.1.5 Map

Map is an illustration of a certain area that shows the features of the depicted area (Oxford University Press, n.d.). Map comes in a two-dimensional form and can be used to represent the layout of a specific area, the network of railroads, the roads in a city, and many other things.

For navigating through an area, maps are the commonly-used aid. Big buildings such as shopping malls, hospitals, and university buildings usually provide a map to help their visitors to navigate through the buildings. The provided map might come on a paper brochure, displayed on a wall, or put in an information kiosk.

A map represents spatial information in allocentric perspective. A map tells people that canteen is at the back of the building, computer lab is on the second floor of the building, and there is an exit door on the west part of the building.

Despite the popularity, maps have some disadvantages. Firstly, the nature of maps that presents information of a location based on other location requires more efforts for people to understand it (Schaik, Mayouf, & Aranyi, 2015). People need to translate what they see from the map into what they see with their eyes when walking to their destination. Secondly, maps are relatively abstract representations of the real thing (Schaik, Mayouf, & Aranyi, 2015) which might not provide complete enough set of information to help people to navigate.

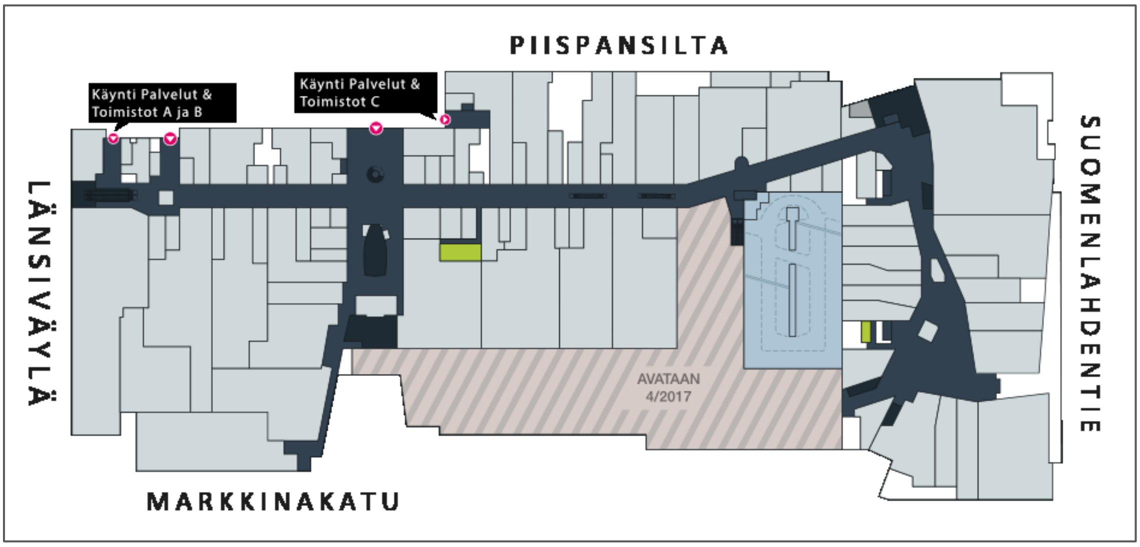


Figure 1. Map of a shopping mall in Finland

2.1.6 Instructions

There are situations where a map or any other visual navigation aid is not available. For example, a person may ask another person on a street how to reach a place the person needs to go. In such situation, the answer might be given verbally as an instruction.

There are a couple of ways to give guidance as verbal or written instruction. First, a guidance can be given as distance-and-turn based instructions. Such instruction will tell a person how far he/she needs to keep going until he/she should make a turn. For example, saying "after walking north for 500 meters, turn right" is a distance-and-turn based instruction. This kind of instruction usually can be found in GPS software in vehicles.

Similarly, a guidance can also be given as activity-based instructions (Mulloni, et al., 2011; Brush, Hammil, & Levi, 2010). An activity-based instruction would consist of a list of user activities to allow them to reach their destination. For instance, "walk to the south for 10 steps", or "walk down the stairs". With this kind of instruction, it is required to know user's exact position as the instruction lack of information about the surrounding area. One study by Mulloni et al. (2011) attempted to use activity-based instruction in AR. In a way, the study complimented the nature of activity-based instruction that is lacking the information of surrounding area by superimposing arrows onto the floor using AR. The study also used checkpoints where the user can see a map augmented onto the floor that shows the location where the user is currently standing and where to go.

Another way to give navigation instruction guidance is to use landmarks. An example of landmark-based instruction would be “walk to the lobby” or “walk to the first intersection and make a right turn”. Previous studies show that landmark-based instruction provides more benefits compared to distance-and-turn based instruction (Hile, et al., 2009). Landmark-based instruction is easier to follow, confirm the correctness of navigation that user has performed, and it is more preferred. Utilizing landmarks also make the navigation more effective as the navigation time is shorter and mental demand to navigate is lower (Goodman, Gray, Khammampad, & Brewster, 2004).

2.2 Previous indoor navigation studies

There have been studies that aimed to provide a better way to aid people to navigate inside buildings. One approach to aid indoor navigation is to use handheld devices to assist people during their journey. A work by Barberis et al. (Barberis, Bottino, Malnati, & Montuschi, 2014) listed several methods to locate a person inside buildings and communicate navigation information which can be performed on mobile devices. One way to locate a person relative to the building is to sense signals. These signals can either be a natural signal such as earth magnetic field signature inside a building or artificial signals such as cellular phone signal, Bluetooth, Wi-Fi, or infrared. However, such signal sensing methods usually require hardware installations that might not be cheap to acquire and maintain. Another way to perform localization inside buildings is to use vision; either by analyzing natural image features or artificial markers. Analyzing natural image requires big computational task such as recognizing edges, colors, and other properties of the environment. On the other hand, a visual marker based system relies on identifying visual markers that are previously placed on specific locations in the environment to recognize the location.

After locating the position of a person, the next step is to inform the route that a person needs to take to reach his/her destination from his/her position. The study listed several methods such as 2D map, 3D map, augmented reality, and instruction. It is also possible to combine these methods to improve the effectiveness of transferring navigation information to a person. One example of combining methods to inform routes is a study by Mulloni et al. (2011) that combined the use of 2D map, 3D map, augmented reality and instruction. Mulloni et al. used augmented reality to show arrows on journey from one point to another point. Visual marker based Augmented reality was also used to show map (both in 2D and 3D map). And text-based guidance was given to the user at all time.

An alternative approach to help people to navigate indoor without requiring them to hold a device is to give the navigation information prior they walk the route. A study by Schaik, Mayouf, & Aranyi (2015) provided walkthrough of a route using a 3-D digital representation of a real building to be used as an aid for route planning. The study aimed to see if providing a walkthrough in 3-D environment could improve route planning and the experience of wayfinding as well as to find out the level of acceptance of such method. The walkthrough itself was presented as a virtual walking in first person (egocentric) perspective.

In addition to the use of 3-D digital environment, the study also put artificial landmarks both in the digital and real building to facilitate wayfinding. The artificial landmarks were big images (1 m (height) × 0.8 m (width)) and small images (0.6 m (height) × 0.4 m (width)) strategically hang on the walls of the building, both the digital and the real building.

The study showed that there was improvement in people's navigation performance (faster task time), less disorientation, and less anxiety during navigation. The results of the study suggested that providing navigation information in the form of virtual walking allows people to have a more positive navigation experience. However, this study used artificial landmarks heavily which makes it difficult to see how big was the role of virtual walking in improving navigation performance.

2.3 Video

A video is a recording of moving pictures recorded digitally or on videotape (Oxford University Press, n.d.). The history of video started from tools such as magic lantern to project pictures which dated back to the 17th century and continued to develop into illusion toys that was famous in 1800s (Walters, 2002). Nowadays, video is widely available and can be displayed with different devices such as TV, mobile phone, or public display.

A digital video has several properties including field of view, resolution, frame rate, and others. Field of view means the area that is visible through an optical instrument (Merriam-Webster, n.d.). In video context, it is about the visible area that is recorded trough a video camera or after a video editing process. The wider the field of view is, the more area is covered in a single frame of a video. Another property of a video is resolution which is the number of pixels that a video has and measured in height and width (Media College.com, n.d.). A high-definition video is a video that has at least 720 pixels on its vertical resolutions (Apple Inc., n.d.). Having a high resolution will allow a video to have more information inside it as well as sharper look. Next, frame rate is the number of images that a video has in one second (Apple Inc., n.d.). Frame rate is measured in frame-per-second (FPS). Higher number of FPS results in a smoother video.

There are number of reasons why this thesis would like to explore the possibility to use video as a tool to provide navigation information. First, the content of a video can be anything including simulation of how a person would see his surroundings. As it has been mentioned in the previous chapter, presenting navigation information in egocentric perspective might be beneficial. Second, a recorded video can be edited further which allows us to add navigation information such as arrows or lines which can help user to understand the information better. The speed of the video can also be adjusted to allow the most effective way to consume the information. Third, previous studies also pointed out that there are benefits of having an egocentric visualization as animation rather than static (Schaik, Mayouf, & Aranyi, 2015; Münzer & Stahl, 2011). The animation of a video in first person point of view is similar to the experience when a person navigates in the real world.

2.4 User Experience

The term 'User Experience' has many definitions. User Experience, often abbreviated as 'UX', may refer to all aspect of user's interaction with products and services of a company (Norman & Nielsen, n.d.). In a bit more elaborate definition, ISO 9241-210 defines UX as person's perceptions and responses from using or anticipated use of a product, system, or service (Pedro, 2013). UX can also be defined as the quality of the experience that a user gets when the user was interacting with a design (Knemeyer & Svoboda, n.d.). Another definition of UX would be a combination of user's internal state, the design of a system, and the context when the user uses the system (Hassenzahl & Tractinsky, 2006).

2.4.1 Properties of Experience

To understand more about experience, Hassenzahl defined five important properties of experience: subjective, holistic, situated, dynamic, and worthwhile (Hassenzahl, 2010). Experience is subjective rather than objective. Experience depends on the person who receives and processes the situation that he/she is in. The subjectivity of experience was previously not covered in the world of human computer interaction (HCI). HCI studies usually focus on objective measures such as time completion, error rate, and gather data by observing people. To relate the subjectivity of experience and objectivity of HCI studies, one needs to focus on achieving certain specific experiences when designing a product. For example, to measure if the experience of being efficient is achieved, the time completion can be used as the measurement.

The second property of experience is holistic which means experience considers everything that a person aims to achieve. A person's goal can be classified into 3 hierarchical levels: motor-goal, do-goal, and be-goal (Carver & Scheier, 1989). Motor-goals are about the basic activities that a person wants to do, such as how to press a button, how to read a text on a screen, and other basic activities. Do-goals are about something that is more concrete such as to watch a movie, to send a text message, or to listen to music. On the highest level is the be-goal. Be-goal is about the reason someone performed do-goals. For instance, when a person is sending a text message, the be-goal might be to keep in touch with another person. Or when a person is listening to music, the be-goal could be to lighten up their mood. Designing for experience should consider all the three levels of goal.

Being situated is the third property of experience. Situated means that there are no two experiences that are identical. Even though the person and the involved tools are the same, the experience will always be different. A simple example can be about sending a text message. A person will experience different feelings between sending a text message to his/her colleague or to his/her lover even though the person uses the same mobile phone.

This situated property of experience makes it difficult to design a product that can achieve specific experience. However, experience can be categorized (Schmitt, 2000) and experience has pattern (Hassenzahl, 2010). Schmitt categorized experience as sense, feel, think, act and relate experience. On the other hand, Marc Hassenzahl described pattern of experience as different experiences that happen in different situation which have the same essence. For example, both solving a puzzle in a game and coming up with a neat algorithm when making a computer program have the same essence of "solving riddle" experience. By categorizing experience and figuring out experience's pattern, one can design a product that give similar experiences to many people.

The fourth property of experience is dynamic. The way people perceive their experience changes overtime. How people think their experience of an activity when they are doing the activity can be different from when they are done with the activity. An experience can be averaged by looking at the worst part of the moment and at the end part of the moment, this is called Peak-End-Rule (Hassenzahl, 2010). The Peak-End-Rule is caused by human's own memory. A person is likely to remember something better when there's something outstanding (peak). The Recency Effect (Neath, 1993) explains why a person remembers better the last thing that happened to him/her.

The last property of experience would be 'worthwhile'. Designing experience should focus on giving valuable feeling to the user. Worthwhile or valuable does not necessarily mean positive. Positive experience is indeed worthwhile because it is what a person wants to have. However, negative experiences that gave us valuable insight can also be worthwhile.

2.4.2 UX Model

Hassenzahl proposed that UX can be seen from two different dimensions: pragmatic and hedonic (Hassenzhal, 2003). Pragmatic aspect of a product means how the product can help the user to achieve their do-goals; to make a phone call, to watch a movie, or to work on their thesis. On the other hand, hedonic aspect of a product means how the user can achieve their be-goals; to feel closer to someone, to be entertained, to be an educated person. As an example, the user experience of a wristwatch seen from pragmatic dimension can focus on how good the wristwatch can tell the current time to the person who is wearing it; how readable it is, how accurate it is, or does it stay accurate over the time. From hedonic dimension, a wristwatch might give classy feeling due to its brand or perhaps invoke memories if it was given by someone close to the person wearing the wristwatch.

Furthermore, people see pragmatic and hedonic as separate aspect of a product (Hassenzhal, 2001). Two cars, one being 10-years old car and the other one being current year model, have similar pragmatic aspect; to transport a person to another place. However, one might be feeling prouder when they are driving a newer model car compared to the old car.

2.4.3 Measuring UX and User Experience Questionnaire

When designing a product, measuring that the product has good enough user experience is important. Good user experience will allow people to keep using the product (increase loyalty) and people would be willingly telling other people about how good the product is. Furthermore, measuring user experience of a product might also reveal improvements that can be done to the product. Hence, it is crucial to be able to measure user experience of a product.

There are many ways to measure the level of user experience of a product. Measuring user experience can be done during the requirement gathering phase, ideation phase, concepting phase, prototyping phase, and even during the time the product has launched. There are also numerous methods to perform user experience evaluations. One blog website has listed 26 UX evaluation methods (Wilson, 2014) and a study by Vermeneen et al. listed 96 UX evaluation methods (Vermeeren, et al., 2010). The list that Vermeneen et al. have gathered can be seen at http://uxems.shorturl.com even though the website now only list 86 UX evaluation methods. The techniques that are incorporated in different methods also vary including questionnaire, self-report, think-a-loud, psycho-physiological measure, and heuristic.

In this thesis work, the author used User Experience Questionnaire (UEQ) (Laugwitz, Held, & Schrepp, 2008) to evaluate the work. UEQ was designed to have quick, comprehensive UX impression, and be simple to use. UEQ consists of 26 questions that cover six aspects of a product; attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty. Attractiveness is about the impression of a product, whether it is appealing to user or not. Perspicuity is the extend of how easy a person can become familiar with a product and understand how to use the product. Efficiency is about how well a product can fulfil a person’s need without requiring unnecessary effort.

Dependability is about how secure and in control a person feels when using the product. Stimulation is about how exciting and motivating a product can be for a person. Novelty tells about how innovative a product is in a person's mind.

To allow the questions to be answered with less central tendency bias, UEQ uses seven stage scale for each question (See Figure 2.)

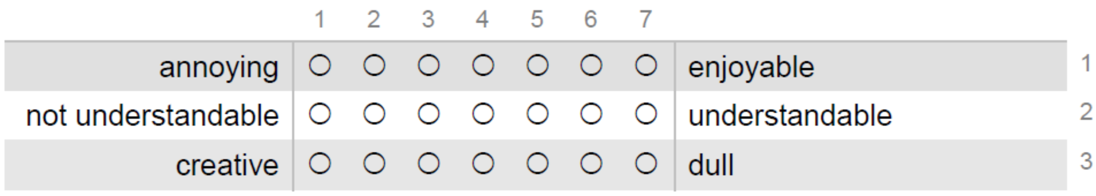


Figure 2. Some of UEQ questions.

Five of the six aspects of UEQ cover both pragmatic quality (perspicuity, dependability, efficiency) and hedonic quality (novelty, stimulation) of a product. The sixth aspect, attractiveness, is a valence dimension. There are 26 questions in UEQ. Attractiveness aspect is covered by 6 questions and the other aspects are measured with 4 questions (Schrepp, 2015) (See Figure 3).

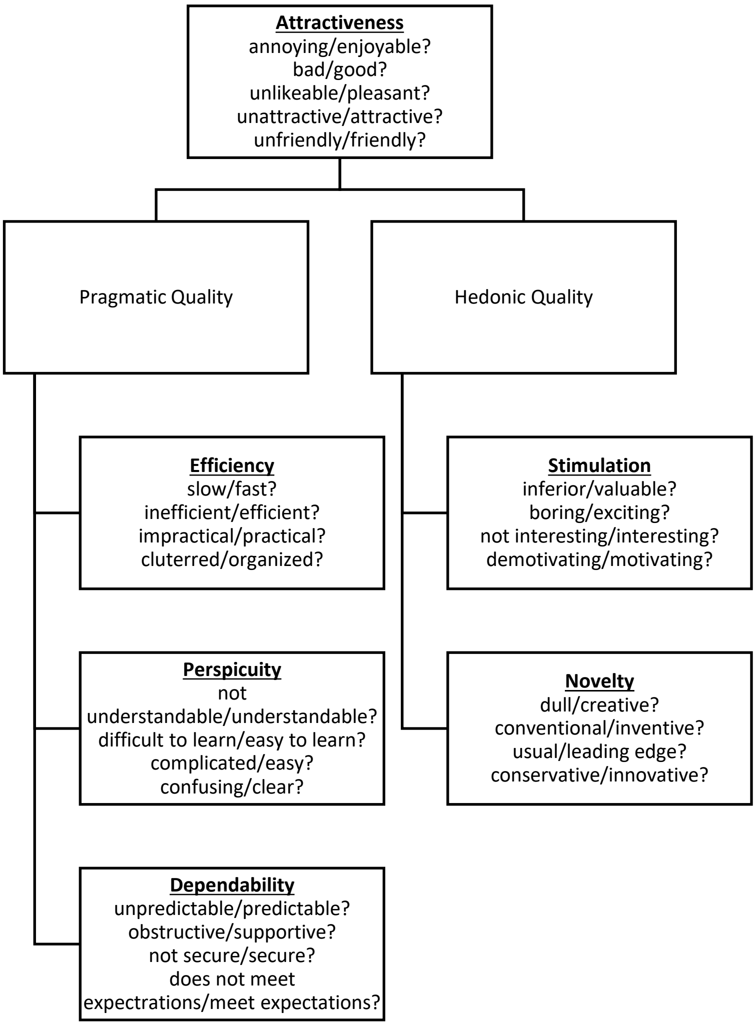


Figure 3. Questions and categorization in UEQ (Schrepp, 2015).

The authors of UEQ provide files that contain the UEQ questions and spreadsheet files to record and to interpret data from user studies (http://www.ueq-online.org/). The spreadsheet files can calculate the mean, variance, standard deviations of the user study data, produce graphs, and other interpretations. In this thesis work, the author used the provided files to record and interpret the user study data.

3 Design and Implementation

The aim of this thesis work is to see if video can be utilized as navigation aid inside indoor environments. The author would also like to see the differences between video and map, if there is any. To find out, two navigation videos and two maps for two different routes in a building were made. Next, these navigation videos and maps were put into use in a user study.

This chapter describes the creation process of the navigation videos, starting with figuring out how the findings from literature can be incorporated into visual representation. Then, the process of how the videos were recorded is described. Lastly, the editing process to put navigation elements into the videos is explained.

3.1 Initial Design of Navigation Video

The design of navigation video aims to allow people to understand the route they need to take easily in a brief time. Furthermore, the design should also be used for any kind of routes including those which have numbers of turns, walking up and down the stairs, and going through doors.

There were several things that were designed in the initial design phase. First, the location where the user study would be conducted was selected. It was planned that the user study will be done in the main building of University of Tampere. The building was chosen as it has complex layout and contains many rooms that are not easy to reach. A university building was also considered a place where many people would need help to navigate, for example, new students, university guests, or other visitors. Lastly, the main building of University of Tampere has a projector with a computer set up for a public display. This helped the process of setting up a user study.

Second, the routes and rooms used in the video recording were selected. To have a balanced data, multiple navigation videos were planned to be made each with a different route. Two rooms were chosen to be used for the videos. Both rooms require routes that have similar features. The route to reach the first room contains 5 turns, 1 door, and involves climbing up 1 staircase. The second room's route has 5 turns, 2 doors, and involves climbing up 2 staircases. Both routes start from the same location.

Third, the navigation elements that would be put on top of the videos were chosen. From the previous findings, several navigation elements were selected: You-are-here indicator, first person perspective, path line, direction arrow, and landmark highlighting. You-are-here is depicted by including the view of the environment which user would see when they are standing in front of the screen where the video would play (Figure 4).

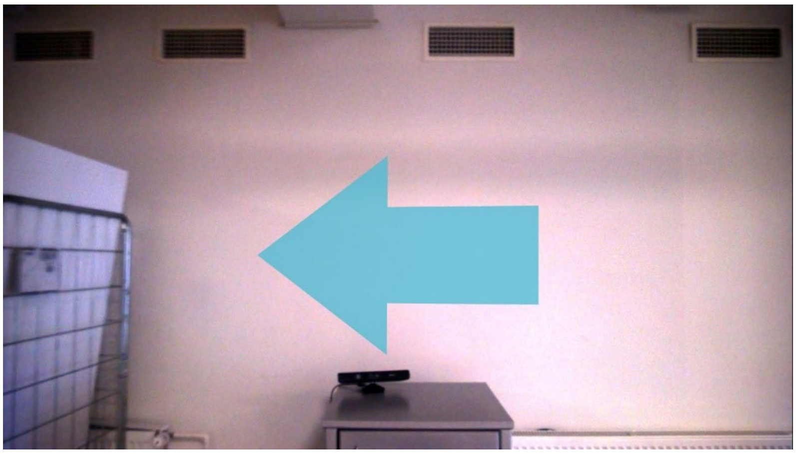


Figure 4. Depicting you-are-here by including the environment which user would see when they watch the video.

As can be seen in Figure 4, the videos would be recorded near the eye level of a person at approximate of 120 cm from the ground.

To indicate the route that the user needs to take to reach their destination, lines are superimposed flat on the ground level (Figure 5). The lines are translucent to make them less obstructive to the real environment while still visible at the same time. The color of the line is bright and vivid to help the user distinct the line from the real environment.

Figure 5. Lines to indicate route

To indicate turns that the user need to make, arrows are superimposed to the video with the same color as the lines, translucent and bright color (Figure 6). The placement of the arrows is floating in the air and in the middle of the video to act like a barrier or a wall to the user. In the initial design, turns were not uniformed in 90 degrees.

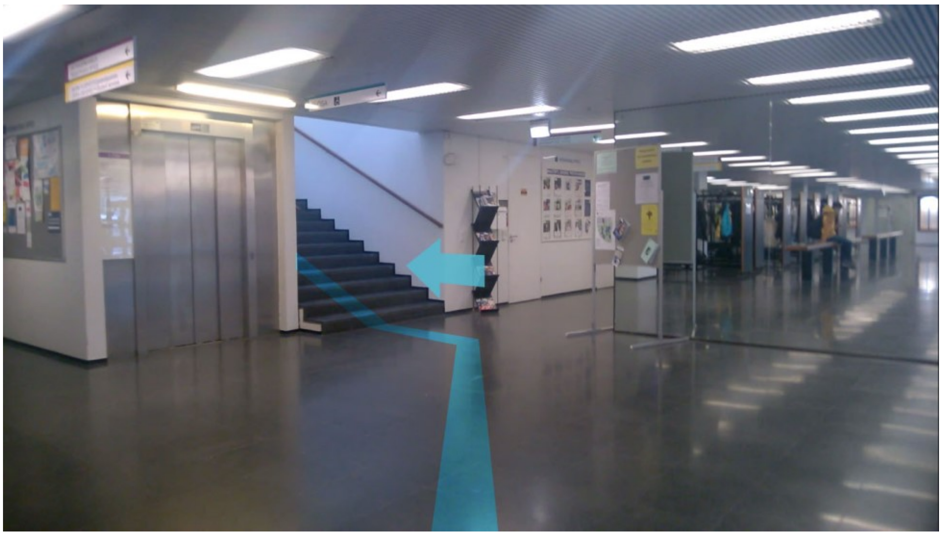


Figure 6. Arrows to indicate turns.

Landmark highlighting was done by using both frame-based highlight and soft highlight (Figure 7). In the initial design, prominent landmarks such as main entrance were highlighted, even though they are not directly related with the route that the user needs to take.

Figure 7. Landmark highlighting (the orange rectangle)

3.2 The First Revision of Navigation Video

After the initial design had been created, a discussion was held with the author's supervisor. The discussion resulted in several points for further development. First, the two routes that were chosen were accepted. Second, the way lines and arrows are presented was accepted as well. Third, the number of landmarks that were highlighted in the initial design was reduced. Landmarks that are not directly related with user's route might distract instead of helping the user, hence their highlights were removed. (See Figure 8)

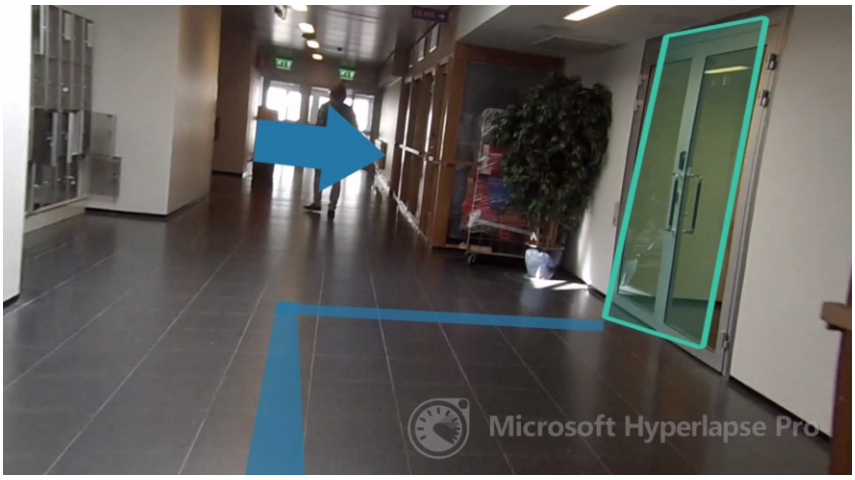


Figure 8. Only highlight landmarks that are directly related to route.

3.3 Video Recording Process

Videos were recorded using a GoPro camera with resolution of 1920x1080 pixels at 29 fps. The GoPro camera was chosen as it has a wide field of view which allows further editing at later stage. The camera was held by hand at the height around 120 cm above the ground. The videographer recorded the video by walking at normal pace from the starting point to the destination room. The recording of every turn was done in a similar manner, which was approximately 90 degree turn. This uniform turning manner was chosen to allow people to realize that they need to make a turn. Furthermore, when making a turn, the videographer stayed in the same position so that the turn is close to 90 degrees. The length of the first video is 2 minutes and 17 seconds at 29 fps. For the second video, the length is 2 minutes and 26 seconds at 29 fps.

The recording was done at the time when there were not that many people in the building. However, there are still people doing their own activities (sitting, walking around, and standing still) recorded in the video. At one time, there are maximum of three persons in a single frame. In the first video, in total there are 8 persons visible in the video and a maximum of 3 persons at a time. In the second video, there are 10 persons visible (including the videographer's reflection on the doors) with a maximum of 3 persons at a time. The recording was also done around noon that both natural light and lamps were the source of light which illuminated the interior of the building.

3.4 Video Editing Process

The editing process aimed to achieve three goals: reduce the shakiness in the video, reduce the length of the video by increasing its playback speed, and put navigation elements into the video. Because the videos were recorded without any stabilization aid, there was shakiness in the video. A software called Microsoft Hyperlapse (Microsoft, n.d.) was used in an attempt to stabilize the video. Stabilizing the video has a downside which is a reduced field of view of the video (Figure 9). Both horizontal field of view and vertical field of view were reduced to about 69% of the original value. However, the wide field of view of GoPro camera compensate enough this downside.

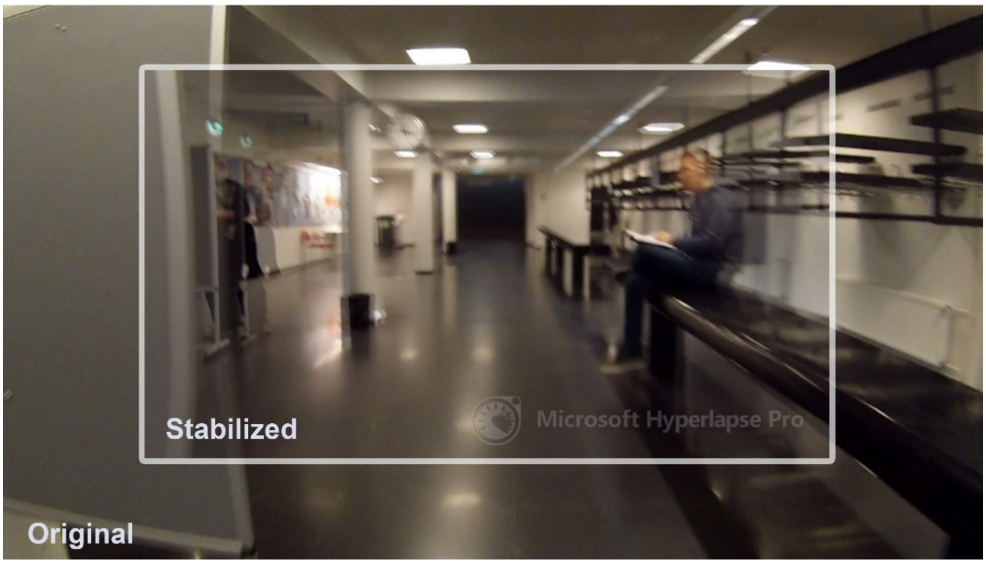


Figure 9. The stabilized image superimposed on the original image that shows the field of view difference.

To save user's time when consuming the navigation video, the length of the videos was shortened by increasing the playback speed of the video four times faster than the original speed. Increasing the video playback speed was done also using Microsoft Hyperlapse software.

After shortening and stabilizing the videos, navigation elements such as lines, arrows, door highlights, and room numbers were put into videos using Adobe Photoshop. Adobe Photoshop was used as it was the editing tool that the author is familiar with.

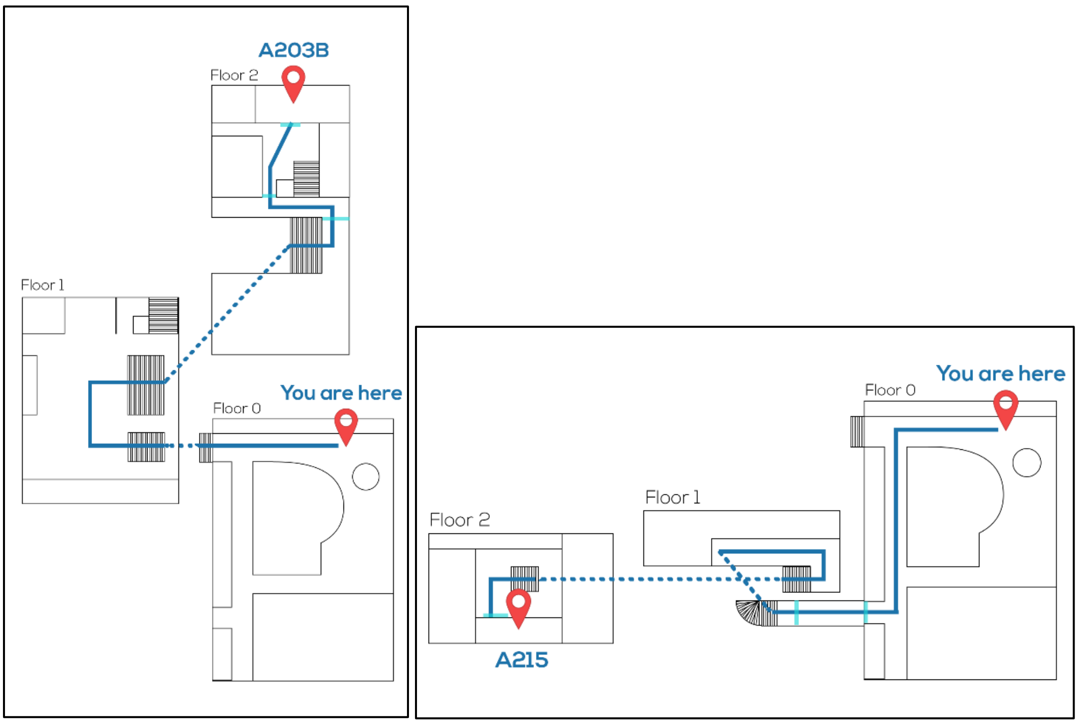
3.5 The Second Revision of Navigation Video

Before using the video in a user study, another discussion was held with the author's supervisor to find anything that can be improved. There were three main points extracted from the discussion. First, the shakiness was not fully eliminated but it was acceptable. Second, the speed of the video was considered good. Third, pause was added before entering closed doors. This was intended to allow the user to better orient themselves.

The videos produced for the second revision were the ones used in the user study reported in Chapter 4. After the editing process, the length of the first video was reduced to 40 seconds at 30 fps (original video was 2 minutes and 17 seconds) and the second video was 45 seconds at 30 fps (original video was 2 minutes and 26 seconds). The resolution of both videos was 1920x1080 pixels. Despite the same resolution, the field of view of the final videos was narrower compared to the original videos due to stabilization attempt. The final versions of the navigation videos can be found in the appendix section of this thesis.

3.6 Map Design

As it will be described in Chapter 4, maps were used as another way to present navigation instruction in the user study of this thesis. At the time of planning the user study, the author could not find a map of the main building of University of Tampere. For this reason, the author created two maps that have the same route information as the navigation videos (Figure 10). The maps also contain similar navigation elements such as you-are-here indicator, lines, arrows, and door highlights.



Map 1 **Map2**

Figure 10. Two maps produced for the user study

4 User Study

This chapter explains how the user study was planned and conducted. There are several points regarding the user study that will be explained in this chapter. First, the goal of the user study and the general plan is described. Second, the questionnaire that was used and how it was presented is explained. Third, the hardware setup that was prepared for the user study is described. Fourth, the participants of the user study are described. Fifth, the tasks that the participants needed to do are defined. Lastly, the type of data and how it was collected from the user study is explained.

4.1 User Study Overview

The aims of the user study were to find if there are differences in the way people navigate in a building and their experience in receiving navigation instruction from the navigations videos and from maps. Furthermore, the user study was performed to understand whether the produced navigation videos from this thesis work are enough to help people to perform navigation. The user study was conducted by asking participants to navigate to specific rooms in main building of University of Tampere using both a navigation video and a map as source of information. After going to a specific room, they were asked to fill in a questionnaire regarding their experience in using the navigation instruction they just used. To conclude the user study and to get further information from the participant, an interview was conducted at the end of the user study.

The independent variable in this user study is the media where navigation information was presented; map and video. The dependent variables are the time that participant use to consume the navigation instruction, the time that they take to reach the destination rooms, the number of wrong turns the participants make during their navigation, number of attempts to consume navigation instruction and perform navigation, and the UEQ score for each navigation instruction they received.

There were ten people who participated the user study. To counterbalance the user study, five of them used map to reach room 1 and video to reach room 2. While the other five participants used video to reach room 1 and map to reach room 2 (Table 1). The order of destination was always Room 1 on the first navigation session and Room 2 on the second navigation session.

Participant	Media	
	Destination: Room 1	Destination: Room 2
1	Map	Video
2	Video	Map
3	Map	Video
4	Video	Map
5	Map	Video
6	Video	Map
7	Map	Video
8	Video	Map
9	Map	Video
10	Video	Map

Table 1. Counter balanced user study order

4.2 User Study Setup

The navigation video and the map were loaded into a desktop computer with specification of Intel Core i7 processor, 16GB of RAM, with Windows 7 SP 1 OS. Both navigation instructions were displayed on a projector with resolution of 1920 x 1200 pixels and view size of approximately 200 cm in width and 120 cm in height. Windows media player was used as the video player and Windows' own picture viewer software was used to present the map.

Figure 11. User study setup.

4.3 Participants

A call for participants was posted to a social media website and this approach gave 3 participants. Besides that, the author also asked directly author's acquaintances and recruited another 7 participants. Prior to accepting a participant, all participant candidates were asked a question of "Are you familiar with the main building of UTA?". Only those who answered negative were invited to participate the user study. In total, 10 participants took part in this user study. There were 4 males and 6 females. A small appreciation token was promised to be given to each participant at the end of their user study session.

4.4 Tasks

Each participant was asked to go to two different rooms; room 1 and room 2, using two different methods of navigation instruction; map and video. Counter balancing was applied by asking half of the participants to read a map to go to room 1 and to watch a video to go to room 2. The other half of the participants were asked to watch a video to get to room 1 and to read a map to get to room 2.

Every participant was given a time limit of 15 minutes to consume navigation instruction and walk to the designated destination. Within the time limit, the participant could read or watch the instruction and attempt to walk to the destination as many times as needed.

During the navigation session, the moderator recorded how long a participant read a map or watched a video until they were ready to go to the room. Also, the moderator recorded the time the participant needed to get to the room as well as the number of wrong turns the participant did.

The navigation session had two possible results: successful and unsuccessful. Successful is when a participant reached the destination room within the time limit. Unsuccessful is when the participant could not reach the destination room within the time limit. After the participant managed to reach the room or after the time limit was passed, a questionnaire that also includes UEQ was given to the participant.

4.5 Procedure

For each user study sessions, there were one moderator and one participant. Each user study sessions lasted for about 45 minutes. The general flow of a user study session was to sign a consent form, fill in a background questionnaire, perform the first navigation session, perform the second navigation session, and lastly to have an interview.

In general, the user study was divided into four parts: background questionnaire part, navigation part 1, navigation part 2, and interview part. In the beginning of a user study session, the moderator explained to the participant the purpose of the user study briefly. Additionally, the general information of what a user study session would consist was also explained to the participants. After giving the introduction, the moderator asked the participant to sign a consent form printed on paper which allowed the moderator to record the user study session. After the consent form was signed, a background questionnaire printed on paper was handed to the participant to be filled in. Background questionnaire was given to record general information about the participant. After the background questionnaire was filled in, the user study session continued to the navigation parts.

In a navigation part, firstly the participant was asked to consume a navigation instruction presented either as a map or a video. The timer to measure the participant's performance was started right after the navigation instruction was presented on the screen. If the navigation instruction was presented as a video, the video kept playing in loop automatically during the navigation part. After the participant thought that he or she was ready, the participant could start walking to find the room that was mentioned in the navigation instruction. The moderator also went together with the participant when the participant was looking for the room. This was done to allow the moderator to observe the participant's behavior and to determine when the navigation session can be ended. A time limit of 15 minutes in total was given to consume a navigation instruction and to find the room. After the participant found the room or after the time limit had passed, the moderator ended the navigation session.

After each navigation session, the moderator asked the participant to go back to the starting point and then fill in the UEQ. The UEQ was given as an online form that the participant can fill in using a laptop that was prepared by the moderator. Having the UEQ as an online form eased data collection process at the later stage of the study. After filling in the UEQ, the moderator started the second navigation part. Only after both navigation parts were performed, an interview was held to get more insight from the participant. At the end of the user study session, the moderator thanked the participant and gave the promised appreciation token to the participant.

4.6 Data Collection

Several types of data were collected during the user study. In the beginning of a session, the participant was given a background questionnaire in a paper form to see their familiarity with using map and video as navigation aids.

Several performance data were recorded during the navigation parts of the user study. First, the time that each participant needed to consume navigation instructions was recorded using a stopwatch. Second, the time that was needed for a participant to reach the destination room was also recorded using a stopwatch. Third, the number of wrong turns were recorded manually on a paper by the moderator. A wrong turn was recorded when a participant walked to different direction than what the navigation instruction had instructed. Fourth, the number of attempts that a participant did in a navigation session was also recorded manually on paper by the moderator. An attempt was recorded when a participant consumed a navigation instruction and then walks the route. If the participant could not find the destination and walked back to the starting point to read the navigation instruction again, it was counted as another attempt. These performance data were also written in an online form at the end of each navigation sessions.

Besides recording the time that participants take to consume navigation instructions and to reach destination rooms, the user study also measured the user experience of the media in which navigation information was presented. To do this, UEQ was used. The UEQ was given using prepared online form that is based on the English version of the UEQ form (http://www.ueq-online.org/ueq-download/). The UEQ was given after a participant had consumed a navigation instruction and attempted to reach the destination room.

At the end of the user study, a short interview was conducted to get more information from the participant. The findings from the interview were recorded on paper. The interviewer asked below questions both for map and video:

- How did you find the map/video instruction?
- What about the information that was in the map/video?
- Did you find any difficulty in understanding the map/video?
- What can be improved in the map/video to help you navigate better?

5 Results and Analysis

This chapter presents data that was gathered from the user study. As a summary, there are significant differences between maps and navigation videos in the time taken by the participants to reach their destinations, the number of wrong turns the participants did during navigation, as well as the result of UEQ. Furthermore, the last part of this chapter describes the interview results.

5.1 Background Questionnaire

Regarding how often the participants use map to navigate, five of them answered few times a week, four of them said few times a month, and one of them replied less than a few times a month. The participants' skills to read map were rather uniform as nine of them evaluated their own map reading skill as good. Only one of the participants evaluated his map reading skill as excellent. However, when the participants were asked how often they use video to navigate, only two of them replied more rarely than a few times a month. The other eight participants never used video to navigate. Furthermore, background questionnaire also revealed that the age range of participants was between 20-36 years old. (See Table 2)

Participant	Age	Gender	Q1: How often do you use map to navigate?	Q2: How do you evaluate your map reading skills?	Q3: How often do you use video to navigate?
1	20	F	2	2	4
2	36	F	2	2	5
3	29	F	3	2	5
4	27	M	2	2	5
5	24	F	3	2	5
6	30	F	3	2	4
7	23	M	2	2	5
8	27	F	3	2	5
9	26	M	2	2	5
10	32	M	4	1	5

Legend

Q1, Q3:	**Q2:**
1 - Daily or nearly daily	1 - Excellent, I can understand map easily
2 - Few times a week	2 - Good, I can read map but it takes time
3 - Few times a month	3 - I can't read map
4 - More rarely than few times a month	
5 - Never	

Table 2. Background questionnaire result

5.2 Performance Data

This chapter describes the performance data result and analyses it by calculating the count, average, standard deviation (STDev), and standard error of the mean (S.E.M) (the variability of standard deviation between samples from the same population) of the data. Furthermore, T-Test and Wilcoxon paired signed rank test analysis were applied to the performance data to see if there was any significant difference between using map and using video during navigation. The software to perform those tests was Microsoft Excel. The p-value (significance value) used in the data analysis is 0.05. However, since there were 5 statistical tests performed to the same data set, Bonferroni correction was used with adjusted p-value of (0.05 / 5) 0.01.

5.2.1 Number of attempts

The average number of attempts when participants tried to reach Room 1 after reading the map was 1.6 times with a standard deviation of 0.55 and S.E.M of 0.24. Video allowed participants to reach Room 1 with the average number of attempts of 1 with 0 standard deviation and 0 S.E.M. To reach Room 2, the map allowed the participants to go there with the average number of 1.2 attempts, standard deviation of 0.45 and S.E.M of 0.2. The average of number of attempts to reach Room 2 after the watching video was 1 with 0 standard deviation and 0 S.E.M. (See Table 3). The graphical view of the data can be seen in Figure 12.

Participant	Room 1		Room 2	
	Map	Video	Map	Video
1	2			1
2		1	1	
3	1			1
4		1	1	
5	1			1
6		1	1	
7	2			1
8		1	1	
9	2			1
10		1	2	
Average	1.6	1	1.2	1
STDev	0.547723	0	0.447214	0
Count	5	5	5	5
S.E.M	0.244949	0	0.2	0

Table 3. Number of attempts.

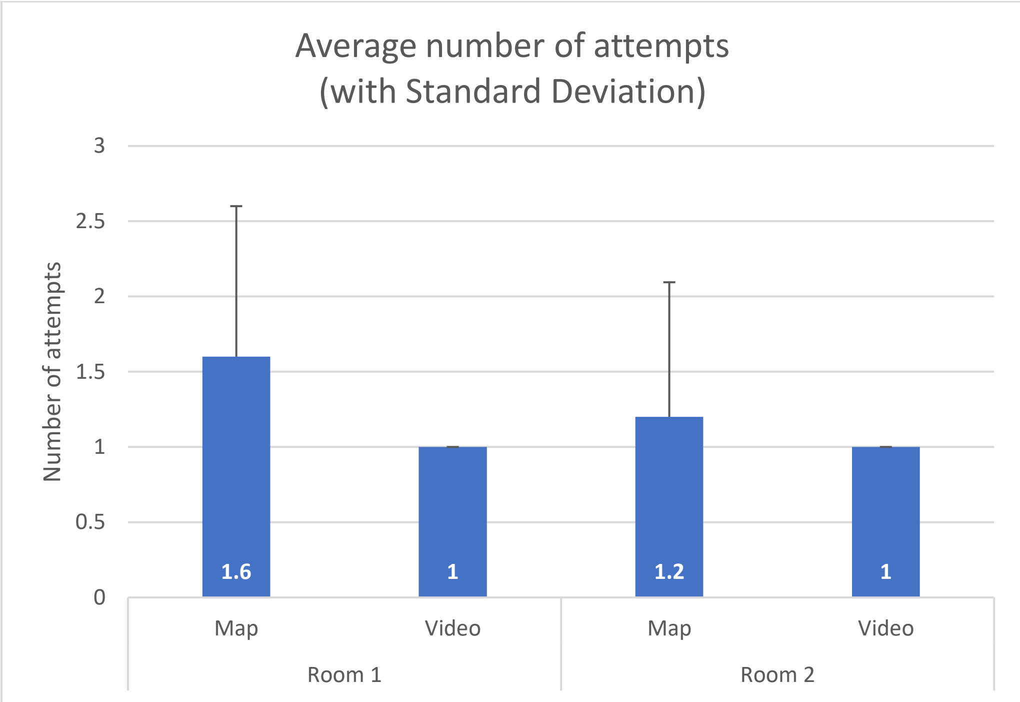


Figure 12. Average number of attempts (with Standard Deviation)

To see if there was any statistically significant difference on the number of attempts during navigation sessions, the data is directly compared within participant between after consuming a map and after consuming a video. (See Table 4)

Number of attempts

Participant	Map	Video
P1	2	1
P2	1	1
P3	1	1
P4	1	1
P5	1	1
P6	1	1
P7	2	1
P8	1	1
P9	2	1
P10	2	1
Mean	1.4	1
STDev	0.516398	0

Table 4. Direct comparison of number of attempts after consuming map and after consuming video.

A Wilcoxon signed rank test was conducted using Bonferroni adjusted p-value of 0.01 (0.05/5) to compare the number of attempts during navigation test sessions using a map and a video. There was not a significant difference in the number of attempts after consuming a map and after consuming a navigation video, $W=3$, $p<0.01$.

5.2.2 Number of wrong turns

The average number of wrong turns when participants tried to reach Room 1 after reading a map was 2 times with the standard deviation of 1. Video allowed participants to reach Room 1 with the average number of wrong turn of 0. To reach Room 2, a map allowed participants to go there with the average number of 1.4 wrong turns and the standard deviation of 0.89. The average number of wrong turns to reach Room 2 after watching the video was 0 (See Table 5). The graphical representation of this data can be seen in Figure 13.

Participant	Room 1		Room 2	
	Map 1	Video 1	Map 2	Video 2
1	1			0
2		0	2	
3	3			0
4		0	0	
5	1			0
6		0	2	
7	2			0
8		0	1	
9	3			0
10		0	2	
Average	2	0	1.4	0
STDev	1	0	0.894427	0
Count	5	5	5	5
S.E.M	0.447214	0	0.4	0

Table 5. Number of wrong turns.

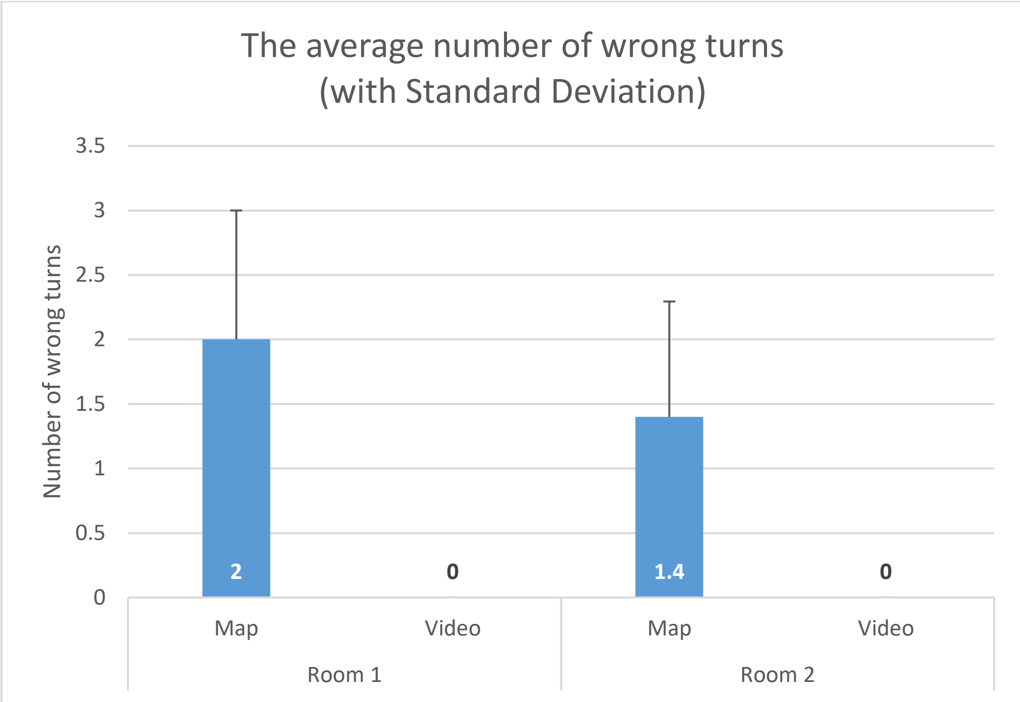


Figure 13. The average number of wrong turns (with Standard Deviation)

To see if there was any statistically significant difference on number of wrong turns, the data was directly compared within participant between number of wrong turns after consuming map and after consuming navigation video.

Number of wrong turns

Participant	Map	Video
P1	1	0
P2	2	0
P3	3	0
P4	0	0
P5	1	0
P6	2	0
P7	2	0
P8	1	0
P9	3	0
P10	2	0
Mean	1.7	0
STDev	0.948683	0

Table 6. Direct comparison of number of wrong turns after consuming maps and video.

A Wilcoxon signed rank test was conducted using Bonferroni adjusted p-value of 0.01 (0.05/5) to compare the number of wrong turns made by the participants during navigation after consuming a map and after consuming a navigation video. There was a significant difference in the number of wrong turns after consuming map and after consuming navigation video, W=3, $p<0.01$.

5.2.3 Time to consume navigation information

The average time to consume a map to reach Room 1 was 78.4 seconds with a standard deviation of 42.05 seconds, while the video that has information to reach Room 1 was consumed in 95.2 seconds in average with a standard deviation of 84.76 seconds. For Room 2, the map was consumed in 70 seconds in average with a standard deviation of 24.8 seconds. The mean of time to consume the video of Room 2 was consumed in 95 seconds with a standard deviation of 38.79 seconds. For the length of the videos, the video that shows how to reach Room 1 is 40 seconds long and the video for Room 2 is 45 seconds long (See Table 7). The graphical view of the data can be seen in Figure 14.

Participant	Room 1		Room 2		
	Map	Video (40 secs)	Map	Video (45 secs)	Which one takes shorter time?
1	81			48	Video
2		79	62		Map
3	56			87	Map
4		245	110		Map
5	22			90	Map
6		57	66		Video
7	101			94	Video
8		44	42		Map
9	132			156	Map
10		51	70		Video
Average	78.4	95.2	70	95	
STDev	42.05116	84.75966	24.81935	38.79433	
Count	5	5	5	5	
S.E.M	18.80585	37.90567	11.09955	17.34935	

Table 7. Time to consume navigation instruction

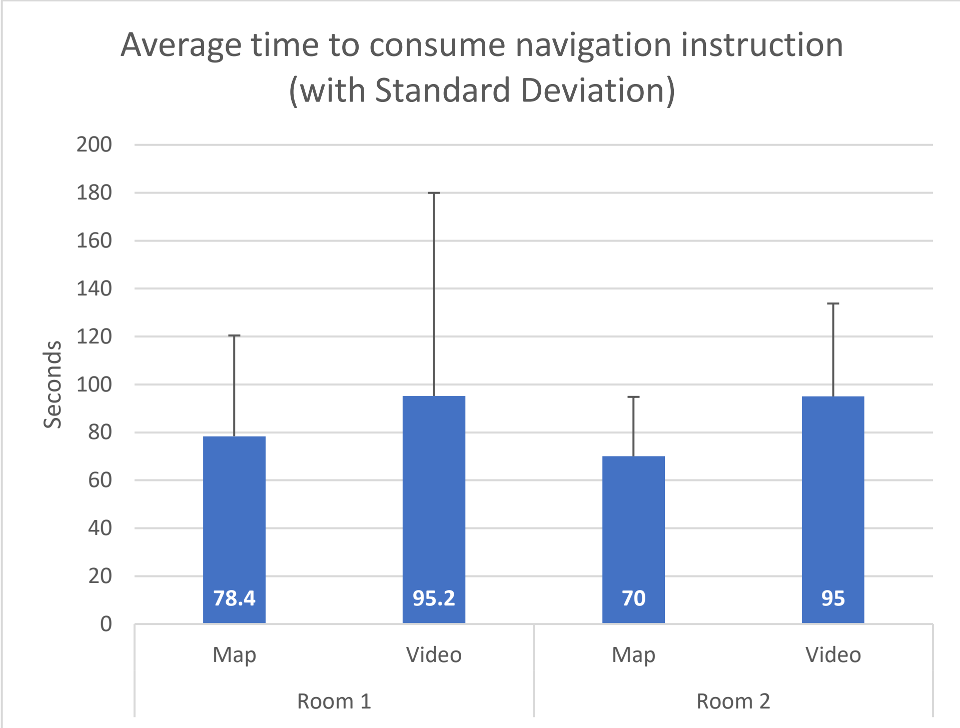


Figure 14. Average time to consume navigation instruction (with standard deviation)

To see if there was any statistically significant difference on time to consume navigation instruction, the data is directly compared within participant between time to consume map and to consume video. (See Table 8)

Time to consume navigation information (in seconds)

Participant	Map	Video
P1	81	48
P2	62	79
P3	56	87
P4	110	245
P5	22	90
P6	66	57
P7	101	94
P8	42	44
P9	132	156
P10	70	51
Mean	74.2	95.1
STDev	32.85253	62.14401

Table 8. Direct comparison of time to consume map and navigation video

A paired-samples t-test was conducted using Bonferroni adjusted p-value of 0.01 (0.05/5) to compare the time to consume navigation information from map and from video. There was not a significant difference in the time used to consume map (M = 74.2, SD = 32.85) and to consume navigation video (M = 95.1, SD = 62.14); t (9) = -1.3401, p = 0.213.

5.2.4 Time to reach destination room

The average time to walk to Room 1 after reading a map was 272 seconds with a standard deviation of 134.78 seconds. Video allowed participants to reach Room 1 in average of 86.2 seconds with a standard deviation of 9.98 seconds. The mean time to walk to Room 2 using a map was 318.8 seconds with a standard deviation of 204.8 seconds. Video allowed participants to reach Room 2 in average of 114.4 seconds with a standard deviation of 44.04 seconds (See Table 9). The graphical view of the data can be seen in Figure 15.

Participant	Room 1		Room 2		
	Map	Video	Map	Video	Which one takes shorter time?
1	292			89	Video
2		86	564		Video
3	197			91	Video
4		84	108		Video
5	94			88	Video
6		87	478		Video
7	326			113	Video
8		101	124		Video
9	451			191	Video
10		73	320		Video
Average	272	86.2	318.8	114.4	
STDev	134.7832	9.984989	204.8492	44.04316	
Count	5	5	5	5	
S.E.M	60.27686	4.465423	91.61135	19.6967	

Table 9. Time to reach destination rooms.

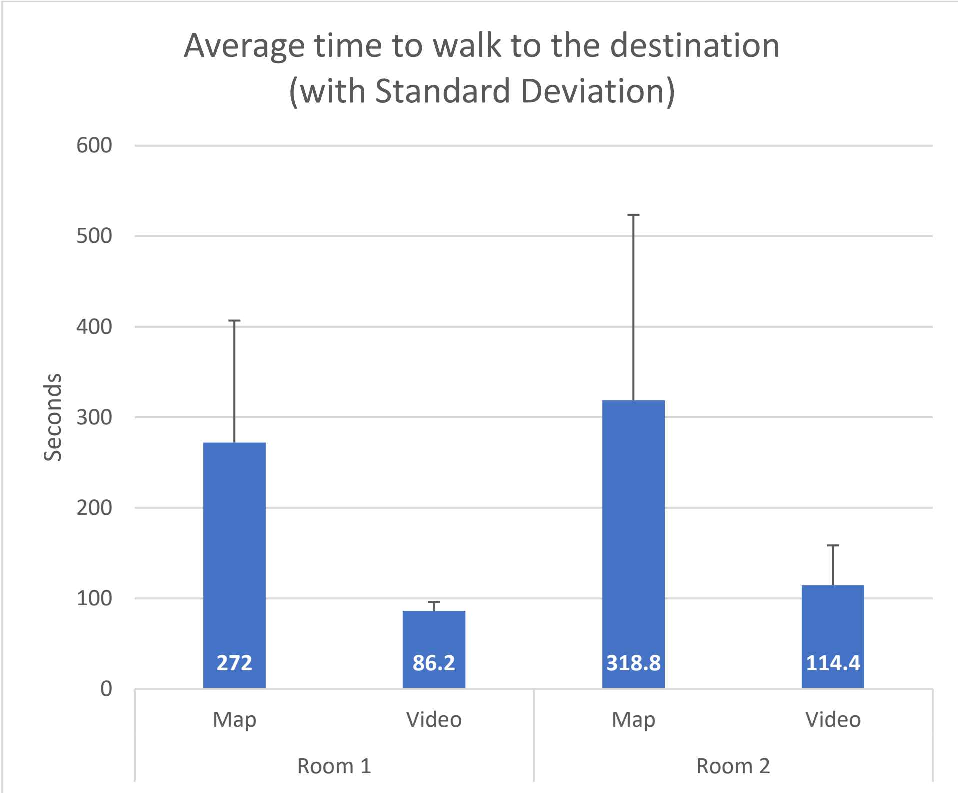


Figure 15. Average time to walk to the destination (with standard deviation)

To see if there was any statistically significant difference on time to reach destinations, the data is directly compared within participant between time to reach destination after consuming map and after consuming navigation video. (See Table 10)

Time to reach destination (in seconds)

Participant	Map	Video
P1	292	89
P2	564	86
P3	197	91
P4	108	84
P5	94	88
P6	478	87
P7	326	113
P8	124	101
P9	451	191
P10	320	73
Mean	295.4	100.3
STDev	165.3261	33.57595

Table 10. Direct comparison of time to reach destination after consuming map and after navigation video.

A paired-samples t-test was conducted using Bonferroni adjusted p-value of 0.01 (0.05/5) to compare the time to reach destinations after consuming map and after consuming navigation video. There was a significant difference in the time used to reach destination after consuming map (M = 295.4, SD = 165.32) and after consuming navigation video (M = 100.3, SD = 33.58); t (9) = 3.87, p = 0.0037.

5.2.5 Total navigation time

The average total navigation time to reach Room 1 using a map was 350.4 seconds with a standard deviation of 176.71 seconds. Video allowed participants to reach Room 1 in average of 181.4 seconds with a standard deviation of 83.78 seconds. The average total navigation time to reach Room 2 using a map was 388.8 seconds with a standard deviation of 199.48 seconds. Video allowed participants to reach Room 2 in average of 209.4 seconds with standard deviation of 80.86 seconds (see Table 11). The graphical view of the data can be seen in Figure 16.

Total navigation time

Participant	Room 1		Room 2		
	Map	Video	Map	Video	Which one takes shorter time?
1	373			137	Video
2		165	626		Video
3	253			178	Video
4		329	218		Map
5	116			178	Map
6		144	544		Video
7	427			207	Video
8		145	166		Video
9	583			347	Video
10		124	390		Video
Average	350.4	181.4	388.8	209.4	
STDev	176.7138931	83.77529469	199.477317	80.85975513	
Count	5	5	5	5	
S.E.M	79.02885549	37.46545075	89.20896816	36.16158182	

Table 11. Total navigation time to reach destinations room.

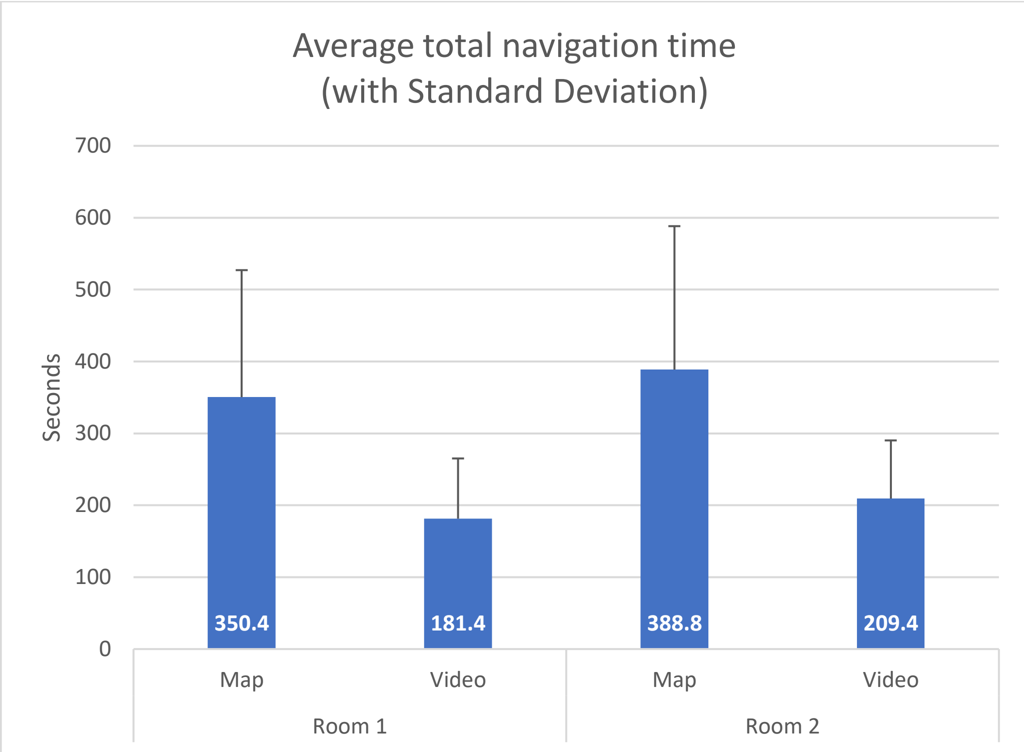


Figure 16. Average total navigation time (with Standard Deviation)

To see if there was any statistically significant difference on total navigation time to reach destinations, the data is directly compared within participant between total navigation time to reach destination after consuming map and after consuming navigation video. (See Table 12)

Total navigation time (in seconds)

Participant	Map	Video
P1	373	137
P2	626	165
P3	253	178
P4	218	329
P5	116	178
P6	544	144
P7	427	207
P8	166	145
P9	583	347
P10	390	124
Mean	369.6	195.4
STDev	178.8116327	79.01223534

Table 12. Direct comparison of total navigation time to reach destination after consuming map and after navigation video.

A paired-samples t-test was conducted using Bonferroni adjusted p-value of 0.01 (0.05/5) to compare the total navigation time to reach destinations after consuming map and after consuming navigation video. There was not a significant difference in the total navigation time used to reach destination after consuming map ($M = 369.6$, $SD = 178.81$) and after consuming navigation video ($M = 195.4$, $SD = 79.01$); $t(9) = 2.91$, $p = 0.017$.

5.3 User Experience Questionnaire (UEQ)

Spreadsheet files that were provided by the UEQ website were used to calculate the result of UEQ. The data from the online questionnaire was downloaded and manually copied into the UEQ spreadsheet files.

The UEQ result from using a map to navigate shows that out of scale from -3 to 3, Attractiveness mean value is 0.303, Perspicuity mean value is -0.025, Efficiency mean value is 0.425, Dependability mean value is 0.425, Stimulation mean value is -0.025, and Novelty mean value is -1.625. (See Table 13)

Confidence intervals (p=0.05) per scale						
Scale	Mean	Std. Dev.	N	Confidence	Confidence interval	
Attractiveness	0.303	1.552	10	0.962	-0.658	1.265
Perspicuity	-0.025	1.656	10	1.026	-1.051	1.001
Efficiency	0.425	1.496	10	0.927	-0.502	1.352
Dependability	0.425	1.323	10	0.820	-0.395	1.245
Stimulation	-0.025	1.534	10	0.951	-0.976	0.926
Novelty	-1.625	1.101	10	0.682	-2.307	-0.943

Table 13. UEQ result from using map

Meanwhile, the UEQ result from using video to navigate shows Attractiveness mean value is 1.7, Perspicuity mean value is 1.9, Efficiency mean value is 1.608, Dependability mean value is 1.542, Stimulation mean value is 1.433, and Novelty mean value is 1.525. (See Table 14)

Confidence intervals (p=0.05) per scale						
Scale	Mean	Std. Dev.	N	Confidence	Confidence interval	
Attractiveness	1.700	0.955	10	0.592	1.108	2.292
Perspicuity	1.900	0.966	10	0.599	1.301	2.499
Efficiency	1.608	0.843	10	0.522	1.086	2.131
Dependability	1.542	0.964	10	0.598	0.944	2.139
Stimulation	1.433	0.838	10	0.519	0.914	1.953
Novelty	1.525	0.901	10	0.558	0.967	2.083

Table 14. UEQ result from using video

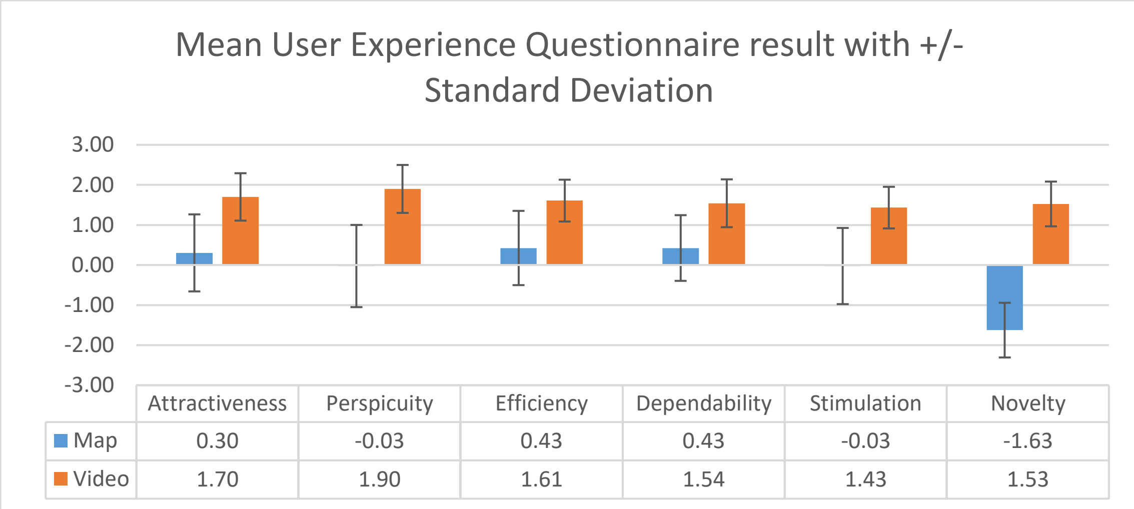


Figure 17. Comparison graph between UEQ result from map and video

T-Test revealed statistically significant result in all aspect of tested user experience between using map and video for navigation with alpha level of 0.05.

Alpha level:	0.05	
Attractiveness	0.0285	**Significant Difference**
Perspicuity	0.0065	**Significant Difference**
Efficiency	0.0466	**Significant Difference**
Dependability	0.0461	**Significant Difference**
Stimulation	0.0195	**Significant Difference**
Novelty	0.0000	**Significant Difference**

Table 15. T-Test result between using map and video.

5.4 Interview

Comments and feedback from participants were written down on paper during the interview sessions and any other time during the user study. After all participants gave their comments, similar comments and improvements ideas that are kept reappearing during the interviews were grouped together. Questions that were asked in chapter 4 were also used to identify comments that had similar themes.

5.4.1 Feedback of Video

The videos were thought to be friendly by three participants; P1, P2, and P7. P1 even said that the video creator might care about other people since he/she went through the hassle of recording the videos. P4, P6, P9 thought that the navigation videos were 'cool' and one of them compared the videos with Google Street View. However, P10 did not like the idea of using video to help navigation and mentioned that the idea is 'too new' for P10.

The pace of the video was considered appropriate by P1, P6, and P8. On contrary, P2, P4, and P10 thought that the playback speed was too fast. Pauses were considered good by P4 and P7, but P1 felt that the pauses should be shorter.

The videos were also considered easy to understand by P5, P6, P8, and P9. One of the participants thought that the video might be aimed for 'dummies' as they were very easy to understand. However, P2 had difficulty in understanding where he/she needed to go. P2 and P4 also could not recognize that the beginning of the video is the same location where he/she was standing while watching the video.

Regarding the navigation video that uses real environment, there were comments coming from P3, P6, P8, P9 that say being able to see real things in the video is good. It makes the video "feel more alive". Additionally, P6 said that having the video in first person view also gave the feeling that it was the participant himself/herself who walks through the route. Showing the real environment might also invoke memories as indicated by P8 who remembered some places that he/she previously saw even though he/she was not familiar with the building itself. Another benefit of showing the real environment mentioned by P1 and P2 was that they could choose and remember landmarks on their own and used the chosen landmarks to understand when to make turns. However, two participants mentioned that being able to see real environment in a real situation adds distractions to the video such as people sitting or walking around. The real environment video also raised P2's expectation that everything on the video should match the real condition which was not fulfilled by the placement of the text of the room number in the video and in real condition.

Regarding the augmenting elements such as arrows, lines, and highlights; P4 and P5 mentioned that the arrows were clear, helpful, and understandable. On the other side, P1 did not even remember there were arrows in the video. Highlights on landmarks were considered good by P6. Additionally, P2 and P3 asked for more highlights on other landmarks to help them understand the route better. The text to indicate the destination room seems to need improvements as P7 and P8 did not see the text of the destination room and P2 also asked for more textual information.

One feedback concerning the way the video was recorded came from P7 who thought that the turning should be more natural, instead of 90° turns. Another feedback from P4 was about the quality of the video itself that was still too shaky.

The participants also gave their ideas on how to improve the navigation video. One idea was mentioned by P1 and P2 to include also a map in the video. This might help people to relate navigation video to the building itself. Another idea by P4, P7, and P9 is to add text information when people need to make turns such as "Turn Left". Regarding landmark highlighting, P5, P7, and P8 mentioned that there should be more highlighted landmarks specially when the route is a long one. To enhance the augmented notations, P7 suggested the use of a more contrasting color. Showing the arrows earlier to allow people to anticipate making a turn might also be better according to P7. To allow the text to be more noticeable, P7 suggested to make it blink, or P6 suggested to use different color than the lines and arrows. Another novel idea came from P3 who suggested that instead of showing the entire journey, the video could jump to only important parts of the journey, such as the turns. To make watching the navigation videos more enjoyable, P6 suggested the addition of auditory cues saying "Turn left!" and perhaps even music. One idea that tries to utilize the latest technology came from P9 who mentioned that the video could be a 360^{o} video.

Other questions also raised by some participants include; "What if the video was recorded in different season?" (P2) and "Will there be different video for those who have disabilities?" (P3).

5.4.2 Feedback on Map

P1 mentioned that the design of the map was nice, easy to remember, and easy to understand. P10 understood the given lines are to indicate the route he/she needed to take. P10 also commented that he/she felt more secure from using map and map was easier than video for him/her to focus on understanding the given information.

On the other hand, many participants had difficulty understanding the given map. P1, P4, P6, P7, P8, and P9 found it hard to understand the route lines. The dotted lines that indicate moving to a different floor were not understood by some participants either. The circling lines (Figure 18) were confusing for P4. There were also other comments regarding unclear information. Unclear you-are-here symbol was mentioned by P4 and P10. Some comments about how the map should have the right scale was mentioned by P8 and P10. P3 mentioned that the map was confusing.

P1, P5, and P10 thought that the map had enough information. However, P2, P3, P4, P6, P7, and P9 mentioned that the provided map still needed more information. P2 and P6 mentioned that having a legend to explain the notation would improve the map. P4 and P7 commented that including landmarks in the map would make it more helpful. Difficulty in orientation due to lack of information was also mentioned by P6 and P9.

Another observed behavior was that P6 and P9 decided to take a photo of the map using their mobile phones. Finally, P6 somehow deviated from the provided route and decided to explore the building by using solely the room number as a guidance.

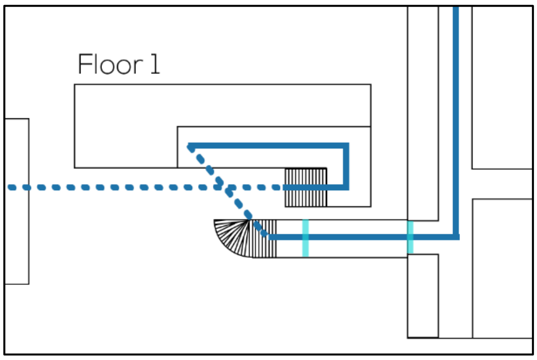


Figure 18. Confusing circling line in the map to reach Room 2

6 Discussion

This chapter compares the performance of the participants between using a video and a map as navigation instruction. Furthermore, the results from UEQ and interviews are also analyzed. Lastly, the reflection of how the thesis work was carried out is given.

6.1 Navigation Performance

The number of attempts to perform navigation was higher when a map was used as navigation instruction even though it was not statistically significant. The number of wrong turns was also higher when using a map compared to a video. This suggests that the given maps were not clear enough for the participants even though the navigation elements that were used were similar in videos (lines, door highlights, room number, you-are-here).

The average time to consume maps was shorter than to consume videos but not for all participants. Having the navigation instruction in video format requires people to watch the entire video at least once to get the complete information. Furthermore, when the person did not understand the video the first time, he/she would watch the entire video again which doubles the consumed time. On the other hand, having the navigation instruction in map format allowed some participants to understand the complete information in a very brief time (22 seconds was the fastest time in the user study). However, for some other participants, the map was not easy to understand and required more time than watching the video before they could start walking the route.

Regarding the time to walk to destinations, all participants reached their destinations faster after receiving navigation instruction in video format compared to the maps. One reason could be that there were no wrong turns done by the participants after watching a video navigation to understand the route.

When we consider the total navigation time that was required to reach a destination, most of the participants could reach their destination faster after watching navigation videos compared to reading maps. However, the average total navigation time difference was not statistically significant. This might be due to the fact that having instruction in video format requires more time for the participants to receive the whole information before starting to walk to the destination.

As a summary, a navigation video allowed the participants to perform navigation in shorter time compared to a map, although not statistically significant. A navigation video also allowed the participants to understand the information better indicated by fewer wrong turns during navigation. However, consuming navigation information from a video might take longer compared to a map.

6.2 Navigation Video Design

In the design phase, several navigation elements were chosen to be used in navigation videos: You-are-here, first person perspective, path line, arrow, and landmark highlighting. The user study showed that those elements are indeed considered to be helpful when applied to a navigation aid video. However, further improvements are still needed on the current design of applying navigation elements to the navigation videos created in this thesis.

In the navigation videos, the You-Are-Here element was only represented by making the video start from the same location as the person watching it. This does not seem to be enough as two participants who mentioned that they did not realize the navigation video started from the same location where they were watching the video. More design improvements are needed to have a good and understandable You-Are-Here element on a navigation video.

Regarding the perspective of the navigation video, the interview results showed that egocentric perspective is indeed a good way to present navigation guidance in video format. Being able to see the real environment in the video the same way as when the participant walked the route also make the video "feel more alive" and might make people remembers things related to the route that they saw previously. Further improvement can be done in the way the video is recorded. Besides having an egocentric perspective, the movement of the camera might also need to be natural as one participant suggested the turns should not be 90°.

Presenting a path with lines and arrows was also considered good by the participants. The comments about the video as "easy to understand" might have come from this. The participants also suggested that the lines and arrows could have more contrasting colours to make them more visible.

Landmark highlighting was also considered useful by the participants. However, the interview revealed that the current design that only highlights doors was not enough. Landmarks which indicate a place where people need to make turns can be one example of landmarks that need to be highlighted.

Participants also gave several comments that may improve navigation videos. Including more related landmarks might be more helpful for people to understand the route better. Another participant mentioned that adding a mini map to the video might help people to orient themselves during the video playback. One participant also commented that a navigation video may jump from one important section to another important section (such as sections where people need to make turns) to shorten the video when the route is too long.

6.3 User Experience of Navigation Video

From the result of UEQ, the Attractiveness value of a navigation video is higher than that of a map. This is supported by the interviews where some participants mentioned that the video looked ‘cool’ and friendly.

The navigation videos were easier to understand compared to the maps indicated by a higher level of Perspicuity score for navigation video in UEQ. The interview results support this finding with four of the participants commenting that the navigation videos were easy to understand. The right pace and the provided pauses might have also affected how easy the videos were to understand. On the other hand, the low level of Perspicuity of the maps also reflected in the interviews with many comments stating how difficult it was to understand the provided maps.

The Efficiency level of a navigation video was higher than that of a map. This is supported by the performance data showing that using video to understand a route allowed participants to perform navigation without making wrong turns.

The level of Dependability was higher for navigation videos than for maps. This result might be related to the observation that videos allowed participants to perform navigation with fewer errors. However, one participant thought that if the video was displayed in handheld device, it might take too much attention of the person and make navigation less safe.

The Stimulation level was higher in video. The properties of the videos that use real environment was exciting and interesting for some participants. One participant thought that the way the video was recorded in first person view made him/her feel like it was him/her walking in the video. Such thought might also stimulate a person’s mind.

The Novelty level of video was much higher than map. This might be because map has been used for navigation for much longer so it doesn’t feel new anymore. Some participants considered that it was good to use a new method to receive navigation information. However, for some others, using video for navigation is ‘too new’ and they hesitate to use it.

Looking from the point of view of UX model, most participants indicated that navigation videos give good user experience both on the pragmatic and hedonic aspect. On the pragmatic aspect, performance data shows that navigation videos allowed the participants to navigate faster with fewer errors compared to maps. The UEQ further confirms that navigation videos have good pragmatic aspect showed from the high score of Perspicuity, Efficiency and Dependability scores. On the hedonic aspect, the higher number of Stimulation and Novelty scores from UEQ indicate that navigation videos possess good hedonic aspect. The comments from interview about navigation videos such as ‘cool’, ‘friendly’, and comparison with Google Street View may also indicate that navigation videos possess good hedonic aspect.

6.4 Navigation Video Production

Recording video using GoPro was simple and easy. The small size and the light weight of GoPro also eased the recording activity. Stabilizing and making the playback speed faster was also easily done using Microsoft Hyperlapse. Microsoft Hyperlapse provided enough playback speed options. Furthermore, the stabilization improvement done by Microsoft Hyperlapse was acceptable for this thesis work. However, the stabilization sacrificed a lot of the field of view of the video has been explained in Chapter 3. To add navigation elements into video, Adobe Photoshop was not the best tool to use as it was difficult to apply the navigation elements. Adding lines, arrows, and highlights required almost frame-by-frame editing and it was a tedious task. A software that is intended to apply visual effects to video should be used instead.

A better video quality is an obvious improvement that needs to be considered when producing navigation videos. A better video quality includes a wider field of view, higher definition of image quality, more stable movement, and more accurate superimposed navigation elements.

6.5 Comparison with Related Studies

This thesis work explored the approach of helping people to be more prepared before they walk on the route they need to take. This approach is similar to the use of 3-D digital environment for virtual walking (Schaik, Mayouf, & Aranyi, 2015). Although the user study in this thesis did not test nor directly compare the 3-D virtual walking to the navigation video, there are similarities and differences that can be deducted.

Both solutions, the navigation video and the 3-D virtual walking, allow a person to experience the route that they need to take to reach their destination. Furthermore, both studies presented the virtual walking in egocentric perspective. This is a benefit of both solutions compared to maps.

The navigation video and the 3-D virtual walking allow people to use landmarks better to help them in their navigation. However, the 3-D virtual walking inform landmarks differently than the navigation video. The 3-D virtual walking study required artificial landmarks to be installed in specific sections of the building as well as shown in the 3-D environment. On the other hand, navigation video does not require such artificial landmarks installation but utilizes the existing landmarks in the building. Furthermore, since the navigation video uses real footage of the building, it allows people to see and choose other landmarks that they find helpful for their journey.

Both studies also used animation instead of user-controlled playback to present the navigation information to a person. Such method might have a role in a better navigation performance as people can just watch the animation without having to think too much or make decisions when they are understanding the route.

The effort to produce 3-D virtual walking is also different than the navigation video. In the beginning, creating the 3-D digital representation of a building and making navigation videos which includes recording and editing videos for different rooms in a building may take similar amount of time and effort. However, if the layout of the building change, updating the content may be done faster for the 3-D environment while navigation video will require new recordings and editing.

6.6 Methodological Reflections

Using public space to conduct a user study showed that timing is important. The time to conduct a study should not be during rush hour when the environment is too crowded. However, the condition also should not be too empty so that the user study also has sufficient correspondence to real conditions. Measuring time using stopwatch in the user study was difficult when there was only one person conducting the user study. However, using a smartphone as a stopwatch was helpful and allowed the recorded time to be sent via email and organized easily.

UEQ was helpful to put experience that users had into understandable numbers which can be used in comparisons. The provided spreadsheet files also fasten the process of analyzing the UEQ results.

Conducting interview was beneficial to the user study as it gave more information in what the participants thought about the tested products. The result of the interview also confirmed and explained the result UEQ.

7 Recommendations for Navigation Video Production

This chapter explains the recommendation to produce navigation videos based on the findings in this thesis work. The recommendations include details on the way navigation elements can be applied as well as the properties of the video itself.

7.1 Navigational Information

7.1.1 You-are-Here

You-Are-Here information can be given in the beginning of a navigation video to allow a person to orient himself/herself before the route information is provided. One recommendation is to give the You-Are-Here information by starting the video from the same location facing the same direction as the person who is watching the video. To make it clearer that the video starts from the same location as the person, a short pause or a slower playback speed can be introduced in the specific section of the video that shows the You-Are-Here information.

7.1.2 Lines and Arrows

After the person understand where he/she is standing in the video, the next information that can be given is the direction the person needs to turn to and the path the person needs to take.

To indicate the path that a person need to walk on, lines can be augmented to the ground in the video. The line should follow the perspective of the video, wider on the closer side to the person and thinner on the further side. The width of the line should also be wide enough that the further side of the line can still be visible. It is suggested that the line is always visible on the middle of the video. (See Figure 5)

To indicate turns that a person needs to make, arrows can be augmented floating on the front of the view perpendicular to the path. Having an arrow on the front of the view would make the arrow clearly visible to a person and prevent them from missing the turn. (See Figure 6)

The arrows and the lines need to be contrast enough from the rest of the video while at the same time look blended with the real environment recorded in the video. One way to achieve that requirement is by giving a vivid and translucent color for arrows and lines.

7.1.3 Landmarks highlighting

Choosing landmarks that are directly related with the route is necessary. In this thesis work, the highlighted landmarks are the doors that a person need to enter along their route. However, the interview showed that more landmarks needs to be highlighted. Landmarks that indicate when to make a turn and that can assure the person that he/she is still going in the right direction are important and need to be highlighted.

The number of landmarks that may be highlighted shall be limited to small number. One landmark in a frame should be enough. This would limit the distractions in the video. Furthermore, the interview also revealed that people can choose other landmarks themselves even though the landmark is not highlighted in the video.

To highlight landmarks, this thesis work used both lines around the landmark (frame-based highlight) and a translucent color on top of the landmark (soft highlight) (See Figure 8). Using both methods to highlight landmarks was enough to get people's attention while at the same time allow the landmarks to be visible. Even thought there was no direct feedback from participants regarding the landmark highlighting method, it might have contributed to the interview results which stated that the videos were easy to understand.

7.2 Properties of Video

7.2.1 Egocentric Perspective and Height

The perspective of the navigation video may be recorded in egocentric perspective (first person view) to require minimal effort of converting navigation information to match with the real experience of walking a route. As for the height of the view, there has not been a deep research for it in this thesis work. However, with the height of the video being around 120 cm from the ground, there was no complaint about it of being too low or too high.

7.2.2 Wide Field of View

It is suggested that the video has a wide field of view. A wide enough field of view would allow people to get more navigational information such as landmarks and other cues to help them to orient themselves. This thesis work did not measure the exact value of field of view of the produced navigation video. However, it was wide enough to allow participants to understand the route.

Another benefit of recording a video with wide field of view is that it can also act as a buffer when further editing is required. In this thesis work, the stabilization of the video was done using a software and it required some part of the video to be cut-out. The wide field of view of the original video allow the produced video to still have a wide enough field of view.

7.2.3 Steadiness and Movement

It is important to have stable and smooth movement in the video. It is suggested to have the horizon of the view stays in the middle of the video the whole time during moving straight and making turns. When walking up the stairs, it is suggested that the end of the stairs stays in the middle of the video.

To reduce shakiness, it is suggested to already record the video as stable as possible. This can be done by using a camera stabilizer when recording the video. Another way to produce a stable video is by using computer software. However, this might come with a down-side of losing some parts of the video which require to be cut-out and resulted in reduced field of view of the video.

Regarding movement in the video, it is suggested to have straight movement as much as possible. Unless the route requires a person to make a turn, then the video should always move straight forward along the route and point to the front. When making a turn, it is recommended to have a 90^{o} turning movement. This would make the turning movement look different enough than the straight movement. There was also a feedback from the interview that said the turning movement should be more natural instead of 90^{o}. This would still need further investigation.

7.2.4 Playback Speed and Length

To save people's time, the playback speed of a navigation video may be increased. At the same time, it is also important not to make it too fast so that the video still be understandable. For this thesis work, the video playback speed was increased 4x faster than its original speed. There was mixed feedback about this being an appropriate speed or being too fast. Based on this finding, it is recommended to set 4x as the maximum playback speed of a navigation video.

The total length of a navigation video should not be too long. This thesis work produced 40 seconds and 45 seconds videos. While there was no feedback whether the videos were too long or not, it is recommended to keep the total length of a navigation video to be shorter than 1 minute. If the route is too long to be put into a 1 minute video, one idea would be to split the route into different videos and then spread the video into separate locations along the route.

7.3 Video Production

7.3.1 Video Recording

It is recommended to record navigation video outside rush hour or even when the building is closed to minimize the number of people visible in the video. The interview revealed that having people walking around in the navigation video is distracting and might create difficulties for a person to understand the information properly.

Using a camera with the ability to record video in high resolution, at least 1920x1080 pixel, should be highly considered. This will ensure the result video is clear enough for people to see even if the video is put on a public display.

7.3.2 Video Editing

Using appropriate software to apply navigation elements and perform other necessary video editing is important to reduce workload to produce navigation video. In this thesis work, video editing was done using an image processor software. It was not a very efficient way to perform video editing and the results were not in the best possible quality. A more appropriate software would have the ability to apply digital visual effects and to design motion graphics into a video. In real life deployment, it is highly suggested to have a dedicated software which purpose is to create navigation video. Such software will allow a more effective navigation video production.

8 Conclusions

This thesis work explored the possibility of video to be used as an alternative way to present navigation information and to become a guidance for people to navigate inside a building. Two navigation videos were produced and then tested in a user study participated by 10 people. The user study measured and analyzed the navigation performance as well as the user experience of using video as navigation guidance compared to maps. The analysis of performance data showed that there was significant difference with faster time to reach destinations and fewer number of wrong turns after watching video compared to after reading map. However, there was no significant difference in number of attempts of consuming and performing navigation between using video and map. Neither was there significant difference in the time to consume navigation video compared to understanding map nor difference in the total navigation time. Furthermore, most of the participants showed high acceptance to this rather novel way of providing navigation instruction. Most participants mentioned that the experience of using video to receive navigation information was good.

This thesis work also came up with a set of recommendations to produce good navigation videos. These recommendations are explained in detail in Chapter 7. In summary, the recommendations are:

- A navigation video should start from the same place as the person who watches the video.
- A navigation video should indicate the path and turns that a person needs to follow. This can be done by showing lines to indicate the path and arrows to indicate the turns.
- Important landmarks that can help people to identify the route they need to take, to tell the place to make a turn, and to assure people that they are going in the right route should be highlighted.
- A navigation video should be recorded in first person view.
- A navigation video should be recorded at the height around 120 cm from the ground.
- The field of view of the navigation video should be wide enough. This will allow people to see and choose many different cues that could help their navigation.
- A navigation video should be as stable as possible, maintain the horizon of the view in the middle of the video, and when climbing stairs, have the stairs in the middle of the video.
- The playback speed of a navigation video may be increased up to 4x of its original speed.
- The length of a navigation video is suggested to be 1 minute at maximum.

Besides the benefits of video over map that are explained above, there are also several limitations that come with using video as navigation guidance. The production of navigation videos seems to be more laborious than creating a map. The high-quality properties of video are important for navigation videos which mean a better recording tools (camera, stabilizer) and better video editor software are required which may not come in cheap. Additionally, a longer time might be needed to receive complete navigation information from a video as it is required to watch the whole video at least once to understand the information and this could take longer than reading a map.

8.1 Future work

One of the aims of this thesis work was to explore if video can be used as an alternative to traditional maps to provide navigation information inside buildings. While this thesis work showed that video is a feasible medium to present navigation information, further considerations are still necessary to allow videos to replace maps in helping people navigating in a building. One consideration is about how to produce the video in bigger quantity as efficient as possible. A building might have tens of rooms and locations that might need to be included in an information system. Recording and editing videos for each room would take a lot of effort hence a more efficient method to record and edit navigation videos for them is necessary. Route selection should also be considered whether the route in the video should be the shortest one or perhaps the easiest one to understand. Another consideration is to find out the system that can manage the videos and allow people to find a proper video to help them reach their destination.

The way a navigation video is presented in a public place is also another task that needs to be discussed. How big the screen that play the video should be and whether the video itself should be in landscape or portrait format. Furthermore, it might also be good to find out if allowing people to control the video playback would be beneficial or not. Overall, if all above considerations can be answered, video seems to have a great potential to be used as indoor navigation guidance.

9 References

Apple Inc. (n.d.). *About High Definition Video Formats*. Retrieved 2017, from Final Cut Pro 7 Professional Formats and Workflows: https://documentation.apple.com/en/finalcutpro/professionalformatsandworkflows/index.html#chapter=1%26section=2%26tasks=true

Apple Inc. (n.d.). *What Is Frame Rate?* Retrieved 2017, from Final Cut Pro 7 Professional Formats and Workflows: https://documentation.apple.com/en/finalcutpro/usermanual/index.html#chapter=D%26section=1%26tasks=true

Barberis, C., Bottino, A., Malnati, G., & Montuschi, P. (2014). Experiencing Indoor Navigation on Mobile Devices. *IT Professional*, 50-57.

Brush, A., Hammil, K., & Levi, S. (2010). User experiences with activity-based navigation on mobile devices. *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services (MobileHCI 2010)*.

Carver, C. S., & Scheier, M. F. (1989). *On the self-regulation of behavior.* New York: Cambridge University Press.

Darken, R. P., & Sibert, J. L. (1996). Wayfinding strategies and behaviors in large virtual worlds. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 142-149. Retrieved from http://doi.acm.org/10.1145/238386.238459.

Darken, R., & Peterson, B. (2001). Spatial Orientation, Wayfinding, and Representation. *Handbook of Virtual Environment Technology*.

Goodman, J., Gray, P., Khammampad, K., & Brewster, S. (2004). Using Landmarks to Support Older People in Navigation. *Proceedings of Mobile HCI 2004*, 38-48.

Hassenzahl, M. (2010). *Experience Design Technology for All the Right Reasons.* Morgan & Claypool. doi:10.2200/S00261ED1V01Y201003HCI008

Hassenzahl, M., & Tractinsky, N. (2006). User experience – a research agenda. *Behaviour & Information Technology*, 91-97.

Hassenzhal, M. (2001). The effect of perceived hedonic quality on product appealingness. *International Journal of Human-Computer Interaction*, 479–497. doi:10.1207/S15327590IJHC1304_07

Hassenzhal, M. (2003). The thing and I: understanding the relationship between user and product. In *Funology* (pp. 31-42). Kluwer Academic Publishers.

Hile, H., Grzeszczuk, R., Liu, A., Vedantham, R., Košecka, J., & Borriello, G. (2009). Landmark-Based Pedestrian Navigation with Enhanced Spatial Reasoning. *Pervasive Computing: 7th International Conference, Pervasive 2009, Nara, Japan, May 11-14, 2009. Proceedings*, 59-76. doi:10.1007/978-3-642-01516-8_6

Kallioniemi, P., Hakulinen, J., Keskinen, T., Turunen, M., Heimonen, T., Pihkala-Posti, L., . . . Raisamo, R. (2013). Evaluating Landmark Attraction Model in Collaborative Wayfinding in Virtual Learning Environments. *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia*, 33:1--33:10. doi:10.1145/2541831.2541849

Klippel, A., Freksa, C., & Winter, S. (2006). You-are-here maps in emergencies – The danger of getting lost. *Journal of Spatial Science*, 117–131.

Knemeyer, D., & Svoboda, E. (n.d.). *User Experience - UX*. Retrieved 11 2016, from Interaction Design Foundation: https://www.interaction-design.org/literature/book/the-glossary-of-human-computer-interaction/user-experience-ux

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and Evaluation of a User Experience Questionnaire .

Liben, L., & Downs, R. (1993). Understanding person-space-map relations: Cartographic and developmental perspective. *Developmental Psychology, 29*, 739-52.

Media College.com. (n.d.). *Video Resolution*. Retrieved 2017, from Media College.com: http://www.mediacollege.com/video/resolution/

Merriam-Webster. (n.d.). *Field Of View | Definition of Field Of View*. Retrieved 2017, from Merriam-Webster: https://www.merriam-webster.com/dictionary/field%20of%20view

Microsoft. (n.d.). *Microsoft Hyperlapse Pro*. Retrieved from Microsoft Research: https://www.microsoft.com/en-us/research/product/computational-photography-applications/microsoft-hyperlapse-pro/

Möller, A., Kranz, M., Diewald, S., Roalter, L., Huitl, R., Stockinger, T., . . . Lindemann, P. A. (2014). Experimental Evaluation of User Interfaces for Visual Indoor Navigation. *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, 3607-3616. doi:10.1145/2556288.2557003

Mulloni, Alessandro, Seichter, Hartmut, Schmalstieg, & Dieter. (2011). Handheld Augmented Reality Indoor Navigation with Activity-based Instructions. *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, 211-220. doi:10.1145/2037373.2037406

Münzer, S., & Stahl, C. (2011). Learning Routes from Visualizations for Indoor Wayfinding: Presentation Modes and Individual Differences. *Spatial Cognition & Computation*, 281-312. doi:10.1080/13875868.2011.571326

Neath, I. (1993). Distinctiveness and serial position effects in recognition. *Memory & Cognition*, 689-698.

Norman, D., & Nielsen, J. (n.d.). *The Definition of User Experience (UX)*. Retrieved 12 2016, from Nielsen Norman Group: https://www.nngroup.com/articles/definition-user-experience/

Oxford University Press. (n.d.). *map - definition of map in English*. Retrieved from Oxford Dictionaries: https://en.oxforddictionaries.com/definition/map

Oxford University Press. (n.d.). *navigation - definition of navigation in English*. Retrieved March 3, 2016, from Oxford Dictionaries: http://www.oxforddictionaries.com/definition/english/navigation

Oxford University Press. (n.d.). *path - definition of path in English*. Retrieved from Oxford Dictionaries: https://en.oxforddictionaries.com/definition/path

Oxford University Press. (n.d.). *video - definition of video in English*. Retrieved from Oxford Dictionaries: https://en.oxforddictionaries.com/definition/video

Pedro, B. (2013, March 19). *ISO 9241-210 | API UX*. Retrieved 12 2016, from API UX: http://apiux.com/tag/iso-9241-210/

Rooke, C., Tzortzopoulos, P., Koskela, L. J., & Rooke, J. (2009). Wayfinding: Embedding Knowledge In Hospital Environments. *Proceedings HaCIRIC 2009:Improving Healthcare Infrastructures*.

Ruggiero, G., Iachini, T., & Ruotolo, F. (2009). Spatial Memory: The Role Of Egocentric And Allocentric Frames Of Reference.

Schaik, P. v., Mayouf, M., & Aranyi, G. (2015). 3-D route-planning support for navigation in a complex indoor environment. *Behaviour & Information Technology*, 713-724. doi:10.1080/0144929X.2015.1004649

Schmitt, B. (2000). *Experiential Marketing: How to Get Customers to Sense, Feel, Think, Act, Relate.* Free Press.

Schrepp, M. (2015, September 21). User Experience Questionnaire Handbook.

Shepard, R., & Hurwitz, S. (1984). Upward direction, mental rotation, and discrimination of left and right turns in maps. *Cognition*, 161-93.

Siegel, A. W., & White, S. H. (1975). The Development of Spatial Representations of Large-Scale Environments. *Advances in Child Development and Behavior, 10*.

Striegnitz, K., & Majda, F. (2009). Landmarks in Navigation Instructions for a Virtual Environment. *In: Proceedings Of The First NLG Challenge On Generating Instructions In Virtual Environments*.

Tolman, E. C. (1948). Cognitive Maps in Rats and Men. *Psychological Review*, 189-208.

Vermeeren, A. P., Law, E. L.-C., Roto, V., Obrist, M., Hoonhout, J., & Väänänen-Vainio-Mattila, K. (2010). User Experience Evaluation Methods: Current State and Development Needs . *Proceedings: NordiCHI 2010*, 521-530.

Walters, J. (2002, March). *Timeline*. Retrieved 11 2016, from Earlycinema.com: http://www.earlycinema.com/timeline/index.html

Warren, D., & Scott, T. (1993). Map alignment in traveling multisegment routes. *Environment and Behavior*, 643-66.

Wilson, C. (2014, February 28). *Method 26 of 100: The Critical Incident Technique (CIT)*. Retrieved from Designing the User Experience at Autodesk: http://dux.typepad.com/dux/2014/02/method-26-of-100-the-critical-incident-technique-cit.html

Appendix

Appendix A - Background questionnaire

The questionnaire that was used to record background information of the participants.

Master Thesis Work | University of Tampere

BACKGROUND QUESTIONNAIRE

A STUDY OF THE DIFFERENCES BETWEEN MAP AND VIDEO TO PROVIDE NAVIGATION INSTRUCTION

BACKGROUND INFORMATION:

Age : ________________

Gender : [ ] Male [ ] Female

NAVIGATION

How often do you use map to navigate?

[ ] Daily or nearly daily
[ ] Few times a week
[ ] Few times a month
[ ] More rarely than few times a month
[ ] Never

How do you evaluate your map reading skills?

[ ] Excellent, I can understand map easily
[ ] Good, I can read map but it takes time
[ ] I can't read map

How often do you use video to navigate?

[ ] Daily or nearly daily
[ ] Few times a week
[ ] Few times a month
[ ] More rarely than few times a month
[ ] Never

Appendix B - Navigation session questionnaire

The questionnaire to record the result of a navigation session. The questionnaire records both navigation performance data as well as user experience data based on UEQ.

1 You are the participant number...

To be filled by the researcher :)

2 And this is session number... *

3 Which tool did you just use to find the place? *

○ Map ○ Video

4 Did you find the place? *

○ Yes ○ No

5 How many times you went to a wrong direction?

Wrong direction means...
- Going to the left or straight when you're supposed to go right.
- Going into a door or place that is not included in the navigation route

6 How long did it take to understand the instruction?

In minutes and seconds, please :)

7 How long did it take for you to reach the destination?

In minutes and seconds, please :)

The next questions are about what you think regarding the method you just used to help you navigate.

Please decide spontaneously :)
And remember: there is no wrong or right answer!

Was it... (annoying/enjoyable) ?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

Annoying — Enjoyable

Was it... (not understandable/understandable)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

not understandable — understandable

Was it... (creative/dull)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

creative — dull

Was it... (easy to learn/difficult to learn)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

easy to learn — difficult to learn

Was it... (valuable/inferior)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

valuable — inferior

Was it... (boring/exciting)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

boring — exciting

Was it... (not interesting/interesting)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

not interesting — interesting

Was it... (unpredictable/predictable)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

unpredictable — predictable

Was it... (fast/slow)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

fast — slow

Was it... (inventive/conventional)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

inventive conventional

Was it... (obstructive/supportive)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

obstructive supportive

Was it... (good/bad)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

good bad

Was it... (complicated/easy)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

complicated easy

Was it... (unlikable/pleasing)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

unlikable pleasing

Was it... (usual/leading edge)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

usual leading edge

Was it... (unpleasant/pleasant)?

1	2	3	4	5	6	7
○	○	○	○	○	○	○

unpleasant pleasant

Was it... (secure/not secure)?

1	2	3	4	5	6	7
○	○	○	○	○	○	○

secure not secure

Was it... (motivating/demotivating)?

1	2	3	4	5	6	7
○	○	○	○	○	○	○

motivating demotivating

Did it... (meet expectations/not meet expectations)?

1	2	3	4	5	6	7
○	○	○	○	○	○	○

meets expectations does not meet expectations

Was it... (inefficient/efficient)?

1	2	3	4	5	6	7
○	○	○	○	○	○	○

inefficient efficient

Was it... (clear/confusing)?

1	2	3	4	5	6	7
○	○	○	○	○	○	○

clear confusing

Was it... (impractical/practical)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

impractical practical

Was it... (organized/cluttered)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

organized cluttered

Was it... (attractive/unattractive)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

attractive unattractive

Was it... (friendly/unfriendly)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

friendly unfriendly

Was it... (conservative/innovative)?

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7

conservative innovative

34 Almost there! Any other comments or feedback?

Thank you for your participation!
Here, have a bar of happiness~ :)"

Submit

Never submit passwords! - Report abuse

Appendix C - Links to navigation videos

Links to the videos that were produced and tested for this thesis work.

Navigation video 1:
https://www.youtube.com/watch?v=VwUh6H7YObI

Navigation video 2:
https://www.youtube.com/watch?v=2d6LDLVIo1U