

Augmented Reality-based Indoor Navigation: A Comparative Analysis of Handheld Devices vs. Google Glass

Umair Rehman and Shi Cao

***Abstract*— Navigation systems have been widely used in outdoor environments, but indoor navigation systems are still in early development stages. In this paper, we introduced an augmented reality-based indoor navigation application to assist people navigate in indoor environments. The application can be implemented on electronic devices such as a smartphone or a head-mounted device. In particular, we examined Google Glass as a wearable head-mounted device in comparison to handheld navigation aids including a smartphone and a paper map. We conducted both a technical assessment study and a human factors study. The technical assessment established the feasibility and reliability of the system. The human factors study evaluated human-machine system performance measures including perceived accuracy, navigation time, subjective comfort, subjective workload, and route memory retention. The results showed that the wearable device was perceived to be more accurate, but other performance and workload results indicated that the wearable device was not significantly different from the handheld smartphone. We also found that both digital navigation aids were better than the paper map in terms of shorter navigation time and lower workload, but digital navigation aids resulted in worse route retention. These results could provide empirical evidence supporting future designs of indoor navigation systems. Implications and future research were also discussed.**

***Index Terms*—Head-Mounted Display, Augmented Reality, Indoor Localization and Navigation, Markerless Tracking, 3D Environment Scanning**

I. Introduction

A. Overview

NAVIGATION is an area that has demonstrated successful human-machine system integration. Modern navigation systems use electronic devices to determine user's location, find appropriate routes, and in some cases also autonomously supervise vehicles to the destination. Currently, most navigation systems use satellite signals from Global Positioning System (GPS), which works in outdoor environments but has difficulty indoors due to reduced signal strength. Alternative technologies such as Wi-Fi-based and image-based methods have been proposed for indoor navigation; however, a definite solution for the industry has not been established. As the prevalence of smart mobile devices and location-aware applications [1], [2], indoor navigation systems become highly valuable for both personal use and applications in many industries [3] such as retail, entertainment, healthcare, and manufacturing [2].

On the machine side of indoor navigation systems, the most important goal is to achieve accurate localization. Compared with outdoor cases, indoor navigation faces a lot of technical challenges such as Non-Line-of-Sight (NLoS) conditions, high attenuation and signal scattering, greater concentration of physical impediments, transitory environment changes, and higher demand for accuracy. To address these challenges, different technologies have been introduced with various levels of accuracy, cost, and scalability. In order to find a suitable navigation technology for a particular application, designers need to align the performance parameters to the requirements of the users [4].

On the human side of indoor navigation systems, few studies have examined the human factors and usability issues. Part of the reason is that the technology itself is still being developed. In contrast, human factors regarding outdoor navigation devices and interfaces have been investigated in many previous studies. However, since the technologies (such as sensors) used in indoor navigation devices are very different and currently less reliable than outdoor navigation devices [5], findings pertaining to outdoor navigation cannot be directly applied to indoor environments. As a result, there is a strong need to test and evaluate the human factors of indoor navigation technologies and devices [6].

The focus of the current study is on Head-Mounted Display (HMD) and augmented reality (AR) interfaces. Wearable devices such as HMDs have been extensively investigated in research laboratories, and they now have a rapidly growing global market [7]. HMDs can be worn on the head as a spectacle or as a part of a helmet. They essentially contain a display optic unit in front of one (monocular HMD) or both eyes (binocular HMD) [8]. Some HMDs only show computer-generated virtual scenarios, whereas other HMDs can superimpose images on real-world views or camera feed. Systems combining HMDs and head movement tracking technologies could be highly valuable for navigation applications [3][4], because such technologies can directly show the route in front of the user's eyes and allow hands to perform other activities. Previous studies using HMDs [9][10][11][12] were often conducted in controlled laboratory

This work was supported in part by NSERC Discovery Grant RGPIN-2015-04134.

environments [13] due to the large size of the devices and their wired connections. Recently, however, companies such as Google and Microsoft have released their prototype versions of HMDs [7], which allow researchers to conduct more practical studies in natural environments. In an HMD, sensor data are utilized to automatically track head orientation and position, whereas with a handheld device, users need to hold the device with particular orientation and position for proper navigation view. Therefore a handheld device entails more cognitive and physical demands. We therefore believe that there is a strong need to conduct comparative studies on HMDs and hand-held devices in order to investigate the systems from cognitive ergonomics and human performance standpoints, identifying best practices of interface designs for indoor navigation applications; because most previous studies related to indoor navigation have focused on analyzing or improving localization techniques rather than human factors issues such as workload, comfort, and memory retention [14].

An imperative aspect of an indoor navigation system is the user interface design. With the traditional interface used in most electronic navigation systems, users had to mentally match the directions shown in the display to directions in the real world. With AR, this mental effort is reduced, because an AR interface can directly superimpose directions on a real-world view, therefore making the directions easier to perceive [15][11][12]. Many AR-based applications have been developed for a wide range of work domains including healthcare, defense, intelligence, and transportation [11]. AR interfaces for indoor navigation have been implemented on handheld devices and evaluated in previous studies [14][16]. These studies found that AR could support accurate localization and improved user experience [17]; however, for handheld devices, users need to hold the devices in an appropriate manner (specific orientation and position) for the applications to work properly [16]. This requirement may influence usability, navigation accuracy, and user satisfaction.

B. Research Questions

The overall research focus of the current study and our previous work [18][19] was on the design, development, and evaluation of an advanced and intuitive indoor navigation system. We concentrated our efforts towards developing a workable prototype, which could be used to investigate complexities confronting both the human and machine sides of indoor navigation research. The motivation for this research was to analyze whether it was possible to build an AR-based indoor navigation solution that could be implemented on both wearable devices (HMDs) and traditional hand-held cell phones. We were also interested in figuring out whether it was possible to achieve the above AR solution using methods that did not require physical infrastructure installation during pre-deployment stage (e.g. Bluetooth beacons, Wi-Fi routers, and fiducial markers). These initial motivations led us to shortlist and then further investigate the following research questions:

1) Can the AR-based indoor navigation solution pass technical assessments to ensure that it is workable and does not cause much glitches and fluctuations during usual walking scenarios?
2) Will the implementation on a wearable device result in better performance, lower workload, and better route retention than the hand-held implementation and paper maps in an indoor navigation task?

C. Contributions

The technical solution developed in the current study was a novel design of indoor navigation systems that utilized advanced feature tracking and augmented reality approaches towards navigation. The system used a pre-scanned 3D map to track environmental features. These features contained directional information so that instructions could be superimposed on the live visual feed at appropriate places. During navigation, directional information was presented to the user via both the visual channel (arrow and icons) and the auditory channel (speeches).

After developing the technical solution, we comprehensively tested the application in two studies, a technical assessment study and a human factors experiment. The technical assessment focused on the efficiency and feasibility of the technology in normal and fast walking scenarios. A real office environment was used to test the feature tracking technology.

The same prototype was then deployed on both a handheld device (Samsung Galaxy S4) and a wearable device (Google Glass). The human factors experiment focused on perceived accuracy, comfort, subjective workload, efficiency (traversal time), and route retention error. Specifically, by analyzing the data from the user study, we examined the AR indoor navigation prototype implemented on a wearable device vs. a handheld device, with a paper map as a baseline in comparison. The test of route retention was important because it reflected the extent to which users overly relied on the navigational aids. It could also reflect the performance of how users would act if the assistance devices were removed. It is necessary to consider such situations, especially for users in extreme environment such as firefighting and combating. Previous studies have identified some negative effects of too much navigational aid on route retention [20]. Therefore, route retention error was included in the current study.

II. Background Of Technology

Technologies used for indoor positioning can be generally categorized into two groups, wireless transmission methods and computer vision methods. Wireless transmission methods use technologies such as Ultra-wide Band (UWB), Wireless Local Area Networks (WLAN), and Radio Frequency Identification (RFID) to localize a device. These technologies often require physical infrastructures, such as Wi-Fi routers and Bluetooth beacons, to be deployed and installed in the indoor environment [4]. Most of these solutions are not very accurate and contain substantial localization errors, though these errors could be reduced by incorporating inertial sensor based positioning approaches and probabilistic techniques such as particle filtering [21]. Some technology solutions such as Bluetooth and infrared methods also have high latency during the detection phase [22]. Although these technologies are popular localization solutions, they have difficulties in estimating the user’s orientation, and therefore are not ideal for AR applications [23]. In contrast, computer vision techniques are more suitable for AR-based applications, and previous studies have found computer vision technologies to be more accurate in comparison to Wi-Fi based fingerprinting [22].

Many techniques have been developed to provide localization and navigation using computer vision. SLAM (Simultaneous Localization and Mapping) is one popular technique that stemmed out of the robotics community for autonomous vehicles [24]. The SLAM mapping process attempts to obtain spatial data (e.g., Received Signal Strength and 3D Point Clouds) of the environment in order to build a global reference map while simultaneously tracking the position of the subject [25]. There are many different SLAM algorithms that pertain to different technologies such as Wi-Fi, Bluetooth, feature tracking, and image recognition [24][25]. All these data types may be utilized for SLAM. However, the focus of the current study is on navigation situations such as in hospitals and office buildings where environment mapping can be done in advance. As a result, we did not use SLAM methods. Instead, the 3D maps were built offline before the navigation tasks.

A commonly studied vision-based indoor positioning approach involves image recognition of the real environment through live camera feed. These images are referenced against a pre-collected sequential database of orthographic images of the same environment. The pre-collected images are annotated with their locations, and the inertial sensors of the device can help deliver orientation [26]. This technique can therefore be used to deliver successful AR-based directional instructions as well as user localization. An issue with this technique, however, is that it requires extensive computational power because a large database of images is being utilized, which may cause delays during navigation [13].

Another computer vision based approach, widely studied before [27][28][15][10][13][11], uses physical markers for optical tracking. Physical markers such as ID markers, barcodes, and QR (Quick Response) codes use fiducial tracking [29] for detection. These markers are easily recognizable due to their unique geometric shape and/or high contrast. Other physical markers such as picture markers need to have enough unique visual contents to be distinctly recognizable. Physical markers often need to be positioned strategically to cover the entire indoor environment. In some cases, distinct features within the environment such as furniture and signs could also be used as picture markers. An issue with most physical markers is that they have to be physically placed in the environment so that they are all visible during navigation. For vision-based localization methods in general, there is a risk that the visual scenes might be changed, which could impair navigation performance [30].

Recent studies have also examined 3D markerless tracking approaches as an advanced form of optical tracking [30]. 3D maps are created by scanning the area of interest. Once adequate visual information of trackables (i.e., 3D point clouds at different camera angles) is collected, they could be used for AR information overlay. This approach is not very computationally exhaustive for mobile devices and also has some degree of resilience against changes in the environment. Identifying distinct point cloud patterns in an indoor area is easier than identifying a specific picture marker. A picture marker is difficult to see clearly from farther away. In contrast, point cloud patterns can cover a large area and are easier to detect from relatively farther distances. Directional information can then be overlaid on the trackables using AR technologies, which can produce a very accurate navigational experience. Therefore in the current study, we utilized 3D point cloud tracking technology on a wearable head-mounted display with an augmented reality interface to assist users in indoor navigation.

III. Proposed System

A. System Design

The major function of the system is to assist people navigate in indoor environments using environment tracking technology and augmented reality instructions (both visual and auditory). The system design is developed to achieve optimal performance for a mobile device or a head-mounted display. The head-mounted display used is Google Glass. It is suitable for the augmented reality application in this study because it has sensors (gyroscope, accelerometer, and magnetometer) that can facilitate the identification of device orientation. Algorithms based on sensor readings can help maintain the required position for the visual overlay to be displayed properly. This delivers a very rich experience where the virtual contents can be seamlessly integrated with the real environment. Developing applications on Google Glass is straightforward as Glass Development Kit (GDK) is an add-on to the Android SDK; thus the Android platform is used. The development of 3D point cloud localization requires a pre-deployment stage, where the indoor environment has to be 3D scanned. We developed our indoor navigation application using Metaio SDK [31] that provides a multilayered environment to build AR applications on Android platform.

B. System Overview

The pre-deployment data were collected and configured in Metaio SDK. The scanned environment that consists of visual features (3D point clouds) is stored as trackables. In a database, these trackables are associated with their corresponding locations and navigation related information, which can be superimposed on visual feed during the navigation aid process. The camera and inertial sensors of the device are used to track the 3D point clouds and device orientation. Based on the trackables identified from the camera feed, the current location and orientation of the user are determined. Then the route is calculated. The potential routes in this study, supplemented with directional instructions in a chronological order, are pre-stored in the application. The routes covered a floor of a mid-size office building. We kept the routes within a manageable size because the wearable device (Google Glass) has limited battery resources. The application presents AR-based navigation instructions including both visual and auditory cues, leading the user to the destination. As the user moves, location and navigation aids are updated in real time. Using gravity measurement from inertial sensors for pose estimation, the application positions the visual instructions at suitable screen locations, preventing any incongruity that could create confusion between augmented and real world environments.

C. 3D Environment Scanning

The location chosen for the experiment was the Games Institute at University of Waterloo. Nine different areas on each route were scanned using Metaio Toolbox [32] to develop the environment map. Crucial objects were shortlisted for potential tracking. We did not intend to scan the entire environment because that would have created a lot of data to process, which would have been highly strenuous on the battery of Google Glass. We established that the minimum area to be scanned would be 2 m in length so that trackables from far away could also be easily detected during the navigation aid process. This design choice would ensure that no discrepancy occurs when AR-based positional information is overlaid. Although all distinguishable surfaces within the environment were taken into consideration, highly textured surfaces were preferred in order to maximize the number of visual features (3D point clouds) within a scanned area. Environmental objects such as tables, chairs, bulletin boards, and signs were scanned from different angles. We also established that the minimum number of features to be scanned within an area would be 1500 so that the environment map could get adequately populated with trackables. Areas where a potential turn was expected were more comprehensively scanned for higher accuracy. All areas, once scanned with trackables, were gravity-aligned using the inertial sensors of the device. The process concluded once sufficient features on a route had been scanned.

The number of points within the 3D point clouds that were scanned at each location mainly affects how easy it is to identify the current location seen by the camera. In the extreme case, if there are too few points, the algorithm will not be able to distinguish between similar locations; therefore the system will fail to provide any aid. If there are enough points but they are scattered around, the user will need to scan around the location in order to see enough points for location recognition. When there is a large number of points at the location, recognizing it will take a shorter time because it does not require the user to scan around the scene. Finally, after the points reaching a certain number, further adding more would not help because location recognition has reached its minimal time duration. Since the focus of the current study is not on recognition algorithms, we did not test the optimal number of points at each location. In general, we expect that reducing the number to 500 or below will significantly decrease performance. Adding more points to the current level will not increase performance. Regarding battery energy consumption, the difference between processing more or fewer points is very minimal; the major energy consumption comes from the camera and the display.

D. Information Overlay and Tracking

After the routes were fully scanned, the images were exported to Metaio SDK for AR information overlay. The 3D scans of all areas were placed in a sequential order to develop a movie-like timeline progressing from the start to the end of each route. The next step is to add directional instructions on the trackables (e.g., shown in Figure 1). Three forms of assistive information were overlaid on the scanned areas. Visual arrows were the first information added. The arrows were superimposed as augmented information on the camera feed, which was then shown to the user via the display devices (for both smartphone and Glass cases). In the Glass condition, it was not implemented as a see-through display. We used giant, glossy, and green-colored arrows in order to achieve high visual salience on small displays such as mobile phones and Google Glass. Three forms of auditory instructions—"turn right, go straight, and turn left"—were also added to the scenario on appropriate places. Finally, text-based visual instructions (same contents as the auditory instructions) was also superimposed on the trackables, providing additional assistance. Other forms of augmentation, such as haptics that could better support people with either hearing or vision impairments, could also be considered in the future; however, the current study was geared towards the normal population. The trackables were properly translated, rotated, and scaled to ensure that AR information was correctly positioned.

Fig. 1. Information overlaid to the scanned 3D point clouds of different areas within the test environment [18]. The point clouds were only displayed in the development stage for testing but not shown to the users in the human factors experiment.

The design decisions were made following general guidelines and previous designs in this research field [14] [33]. Based on these studies we concluded that the major elements for an AR interface in this application should have the following characteristics.

1) Elements should be easy to discern.
2) Voice augmentation should be added to complement visual instructions.
3) All major areas should have adequate information to prevent navigation errors.
4) Virtual content should be meaningful, simple, commonly used, and context aware.
5) The most suitable tracking method should be utilized.

Our application used elements which were easily discernible; turn by turn voice augmentation was also added; navigation instructions were comprehensively distributed on the route; the virtual content such as arrows and audio instructions were meaningful, simple, commonly used, and context aware; and we utilized 3D point cloud tracking as that seemed to be the most appropriate option for indoor navigation scenarios.

When the application was tested on the testing site using both Google Glass (HMD) and Samsung Galaxy S4 (handheld), the interface updated navigational cues in real time as the user moved through the areas (Figure 2). The trackables were quickly detected, and the application processing was swift.

(a) (b)

Fig. 2. Screenshots of the application interfaces implemented on (a) Samsung Galaxy S4 and (b) Google Glass. Visual information aids (arrows and words) were superimposed onto the camera feed, which was then shown to the user via the display devices of the smartphone and Glass.

IV. Technical Assessment Results and Discussion

Technical performance assessment was conducted to evaluate the technology in terms of its feasibility and efficiency. We carefully measured the time needed for successful feature detection, processing of those features, and the subsequent display of auditory and visual instructions. Since we needed to quantify very short durations of time, a separate software program was developed to record important time stamps. Feasibility was determined by analyzing the application's ability to detect the percentage of features in a walking-speed controlled scenario as well as analyzing the walking speed threshold. The technical assessment was conducted on nine evenly distributed areas of a route. The height of users and the height where they held the phone camera were not considered as independent variables in this study. Participants generally held the phone around the shoulder or neck level. Participants' variation in heights also represented the same fact from the general user population.

A critical factor determining localization accuracy is how many features (rather than pixels) can be recognized in each camera view [34] [35]. Ideally, a considerable number of features should be tracked in a minimal amount of time so that AR information could be accurately overlaid without any noticeable delay. However, there are concerns with specific usage scenarios. For example, if a user is walking very fast and expected to take a turn, but the system still needs more time to identify sufficient features, a delay in information delivery could happen, which could negatively affect overall performance and user experience. In some possible but rare situations, if a user passes a target location way too swiftly, there will not be enough time for the camera to adequately capture the trackables, preventing the system from working properly. We used the percentage of recognized features as the measure because it allows results to be compared across different locations and camera views. System time responses were also measured.

In the current study, as the first step towards testing AR-based indoor navigation systems, we extensively scanned the testing area with visual features in nine areas that were uniformly distributed along the route. All technical experimentation was done in these nine areas where each area was roughly equal to 2 m in length. For experimental purposes, feature detection and AR overlay processes would only initiate after the user was physically present in the area. A total of four different assessments were conducted on the testing route. The assessments were conducted first on Google Glass, which is the focus device of this study, and then on a smartphone.

In the first assessment, we wanted to figure out the minimum percentage of features that are needed to initiate AR overlay processing for the application. In this assessment, the user started from a fast walking pace and gradually reduced the speed until there was enough time to collect the minimum number of features. The first assessment was repeated four times, and the results from different repetitions were very similar. We programmed a separate internal script that could record the number of tracked features. The results showed that on average, the minimum feature percentage needed was approximately 45%, with some variation across different areas. Regarding the corresponding actual number of features, that was on average about one feature in each 2.3 degree horizontal by 2.3 degree vertical visual field of view. Not all directional information was successfully overlaid on the trackables but adequate information was conveyed to the user, leading the user to the destination successfully. Overall, the speed threshold (i.e., the fastest pace that the user can walk without causing system localization failures) was found to be around 6.4 km/h to 7.6 km/h. Previous studies found that the general walking speed is around 3.4 km/h to 5.5 km/h [36], which is below the threshold speed. As a result, we could expect our system to be feasible for practical use at normal walking speed.

In the second assessment, we wanted to test the feasibility of the application in a fast walking scenario. For this assessment, our test user maintained an average walking speed of 6.4 km/h, which is much faster than the normal walking speed (about 30% more). We conducted four trials with this speed on the route and found out that the user was spending on average 0.7 s per area. Therefore, we wanted to test the percentage of features the application could successfully detect in 0.7 s. The results indicated that on average 50.6% of features were successfully detected, allowing navigation aids to be displayed correctly and promptly without any major issue. The results validated the application's effectiveness at a faster walking pace.

The third assessment was conducted to figure out the average speed and maximum time the application would require to work ideally. The ideal condition is when 95% of the features are detected at a particular position because 95% of features could seamlessly communicate all navigational instructions as well as process future instructions. This assessment was repeated six times and the maximum time for the system to identify 95% of features was mostly under 1 s at all areas while walking at an average speed around 4.3 km/h and nothing going below 3.8 km/h. The average speed of 4.3 km/h was within the general walking speed range, so it validated that this application could operate ideally with maximum efficiency at a slower walking pace. In particular, this result showed that the user travelled 1.2 m on average before the system detected 95% of the features.

Analyzing the time needed for each type of AR display was also crucial to determine the efficiency of the technology. As previously introduced, the two types of navigational assistance include visual direction (arrows and texts) and auditory direction (speech). We developed a testing program that could estimate the time for processing each kind of navigational assistance. This assessment was repeated five times and overall, the average time for Google Glass to produce audio augmentation was 0.18 s, and for visual direction arrows and texts, it was 0.14 s. The average distance travelled was less than 0.5 m during this time period.

After examining the application on Google Glass, we also wanted to examine the same application's performance on a handheld smartphone/cell phone. A Samsung Galaxy S4 cell phone running the Android operation system was used in the test. Below we listed the specifications of the two devices (Table 1). The comparative performance results were listed in Table 2, which shows similar results from both devices.

TABLE I
HARDWARE SPECIFICATIONS OF THE DEVICES

Specifications	Google Glass	Samsung Galaxy S4
Form-Factor	Monocular	Slate
Weight	50g	130 g
CPU	OMAP 4430 SoC, dual-core	Soc Exynos 5 Octa 5410, 1.6 GHz quad-core Cortex-A15
Operating System	KitKat for Glass	Android 4.2.2 "Jelly Bean"
Storage	16 GB flash memory total (12 GB of usable memory)	32 GB (8 GB used by the system) and 64 GB microSDXC
Memory	2 GB RAM	2 GB LPDDR3 RAM
Power	570 mAh Internal lithium-ion battery	2600 mAh External lithium-ion battery
Display	Prism projector, 640x360 pixels, covering 13° × 7.3° of the visual field	Super AMOLED, 1920x1080 pixels
Sound	Bone conduction transducer	Qualcomm DAC
Camera	5 MP Camera, f/2.48 aperture, focal length of 2.8mm, FoV (75.7˚ x 58.3˚) with 2528 x 1856 pixel resolution. During video recording, image gets encoded to 1280 x 720 pixels at 30fps (720p)	13 MP Camera, f/2.2 aperture, focal length of 4.2mm, FoV (69˚ x 49.6˚) with 1920 x 1080 pixels at 30fps (1080p HD)

In summary, the technical assessment showed that the navigation application implemented on both Google Glass and the Android cell phone was feasible and efficient in detecting, processing, and displaying AR-based navigational information. The application could operate well at normal walking speed and work satisfactorily at a fast walking pace. Regarding the time response and delay, it took about 140 ms to display the visual aid information and about 200 ms to play the auditory aid information. Since there is a lack of studies in this specific area that can provide a benchmark or user acceptance level of delay or lag, we consulted studies in the related human-computer interaction and virtual reality fields. It has been estimated that users' tolerance for key-press response delay is around 150 ms [37]. In the virtual reality setting, auditory delay around 240 ms has been shown to be tolerable without significant impact [38]. As a result, the delays in the current application seem to be tolerable. During the tests, the system responded promptly without any apparent delay that would affect navigation.

TABLE II
COMPARATIVE ANALYSIS OF TECHNICAL PERFORMANCE ASSESSMENTS CONDUCTED ON GOOGLE GLASS AND AN ANDROID CELL PHONE USING THE SAME AR-BASED NAVIGATION TECHNOLOGY

	Google Glass	Samsung Galaxy S4
Minimum percentage of features needed to initiate AR overlay processing	45.0% on average for the nine locations (respectively 34%, 37%, 34%, 49%, 53%, 35%, 46%, 67%, 50%) with the speed between 6.4 km/h to 7.6 km/h	42.7% on average for the nine locations (respectively 31%, 32%, 22%, 36%, 47%, 33%, 46%, 72%, 66%)with the speed between 6.0 km/h to 7.9 km/h
Percentage of features detected at a fast walking pace	50.6% on average for the nine locations (respectively 57%, 45%, 44%, 53%, 40%, 42%, 50%, 47%, 77%)with an average speed of 6.4 km/h and minimum speed of 5.5 km/h	44.3% on average for the nine locations (respectively 36%, 38%, 40%, 24%, 55%, 63%, 48%, 41%, 54%)with an average speed of 6.5 km/h and minimum speed of 5.3 km/h
Time taken to detect 95% of features	95% of features detected under 1 s for all nine areas (respectively 0.81 s, 0.93 s, 0.92 s, 0.84 s, 1.07 s, 0.96 s, 0.88 s, 0.89 s, 0.74 s) with an average speed of around 4.3 km/h	95% of features detected under 1 s for all nine areas (respectively 0.76 s, 0.74 s, 1.01 s, 0.85 s, 0.92 s, 0.99 s, 0.93 s, 0.9 s, 0.82 s) with an average speed of around 3.9 km/h
Time needed to generate each type of navigational information	0.18 s on average for all nine locations (respectively 0.17 s, 0.25 s, 0.2 s, 0.19 s, 0.13 s, 0.14 s, 0.2 s, 0.19 s, 0.16 s) to generate audio augmentation; 0.14 s on average for all nine locations (respectively 0.12 s, 0.11 s, 0.15 s, 0.19 s, 0.17 s, 0.15 s, 0.12 s, 0.1 s, 0.15 s) to generate visual direction arrows and texts	0.22 s on average for all nine locations (respectively 0.2 s, 0.27 s, 0.18 s, 0.23 s, 0.15 s, 0.24 s, 0.23 s, 0.26 s, 0.24 s) to generate audio augmentation; 0.13 s on average for all nine locations (respectively 0.09 s, 0.16 s, 0.13 s, 0.11 s, 0.1 s, 0.12 s, 0.1 s, 0.2 s, 0.17 s) to generate visual direction arrows and texts
Framerate	About 12-18 fps	About 14-23 fps

V. HUMAN FACTORS STUDY

The overall goal of the human factors study was to test and evaluate the human performance and workload of using the AR-based indoor navigation system, by comparing the results across the three types of navigational aids including AR navigation implemented on Google Glass, AR navigation implemented on a smartphone, and a traditional paper map. The paper map was included as a baseline condition. The digital navigation devices (Google Glass and cell phone) use an egocentric perspective whereas the paper map uses an exocentric perspective [39]. Participants were recruited to navigate an indoor environment using the three aids in a within-subject design. The human factors measures included traversal time, perceived accuracy, subjective workload, and route retention error.

To navigate successfully, people rely on spatial knowledge and cognitive abilities that can build and use such knowledge. Human spatial knowledge in topographic contexts includes three levels – landmark knowledge, route knowledge, and configurational knowledge [40]. As people navigate, they tend to build spatial knowledge about the area into cognitive maps that represent the real world area [41]. When more cognitive resources and attention efforts are used to process spatial information and build the cognitive maps, the results often leave a stronger and keeper trace in memory.

Digital navigation aids (Glass and cell phone conditions in the current study) provide turn by turn guidance and use an egocentric perspective, which is similar to the perspective of mental route knowledge represented as a sequence of egocentric visual images of landmarks with directions [42]. Users' cognitive maps formed while using digital navigation aids are often limited because of the ease to use the same egocentric perspective and the lower level of cognitive processing involved in passively following directions. In contrast, using a paper map involves much more cognitive processing and efforts. It requires spatial information to be mentally converted from the exocentric to the egocentric perspective. This helps the user develop comprehensive spatial cognitive maps [43]. While navigating with a exocentric map, users often need more cognitive processes such as mental rotation and zooming to establish correspondence between the map and the real world view [44]. This is why navigation with the exocentric perspective is often more difficult and time consuming than egocentric navigation [44][45]. However, active and deeper mental processing helps the learning and retention of cognitive maps [46].

Based on the theories and previous research findings, we expected that digital navigation aids would require less mental workload and time and would be perceived as more accurate when compared against the paper map; however when using the paper map, participants would retain more spatial knowledge and hence would have less route retention error. Due to the natural characteristics of HMDs, we expected that Google Glass would be better at conveying AR directional information than the handheld cell phone.

A. *Method*

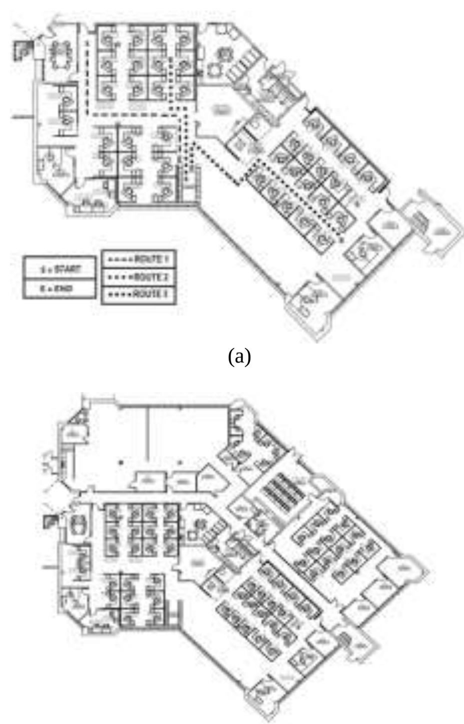
1) *Participants*

Thirty nine adults (24 males and 15 females), all of whom were students from University of Waterloo, participated in this study. None of them had any previous experience with mobile navigational aids in indoor environments; however, all were well aware of mobile navigational aids and had experienced them in outdoor environments. The majority of the participants stated that they were confident in navigating in indoor environments with or without navigation aids. All had normal or corrected-to-normal visual and auditory acuities. The participants had various levels of familiarity with the testing environment. Some of them were very familiar with the environment, whereas others had never been there before. This individual difference should not affect the results because a within-subject design was used.

2) *Tasks and Materials*

Three different routes (Figure 3a) were formulated and optimized for the experiment to ensure that navigational instructions were added at the most appropriate places. Once the user interface was properly designed, it was deployed on both the handheld device (Samsung Galaxy S4-Android Cell Phone) and a wearable device (Google Glass). The third navigational aid was a paper map, which was a CAD (computer-aided design) version of the entire floor plan.

The tasks required the participants to navigate through the test location and find specific books located on different shelves using different types of aids. Such tasks are typical representations of indoor navigation. When the participants approached the shelf using AR based digital aids, the audio channel informed the participant the target shelf number, and the visual channel pointed an arrow at that shelf alongside the text showing the shelf number. While using the paper map the user read the shelf number from the paper and visually searched for it. In the map retention test after the completion of the experiment (completing all three routes), participants were given a similar but not identical version of the floor plan to re-draw the routes (Figure 3b) as they remembered.



(a)

(b)

Fig. 3. (a) Three different routes used in the experiment. In the paper map condition, this map was given without the start points and the routes. Only the end points were shown. (b) The version of map that was used in the map retention test. No start point, end point, or any route was shown.

3) *Experimental Design and Measures*

The experiment used a within-subject design. The independent variable was the type of navigation aids, including three conditions − paper map, cell phone (handheld), and Google Glass (wearable). The order of experiencing the three navigational aids was balanced across subjects using a Latin square design. In addition, each navigational aid was equally tested on the three routes. The dependent variables included subjective workload ratings using NASA-TLX (raw overall score), perceived accuracy, contextual retention error, and efficiency (i.e., traversal time/task completion time). Each dependent variable was individually measured for the three navigational aid conditions. With the hand-held cell phone, the application would automatically re-orient the display in landscape or portrait based on user preference. Majority of the participants used it in portrait. The Glass view was landscape.

In order to measure unprepared route memory retention performance, the participants were asked to re-draw all the three trajectories only after completing all the three routes. Since the order of experiencing the three aids were balanced, the carryover effects should be controlled. Distance errors resulting from participants' map drawing were used to quantify the route retention error. The three target routes (Figure 3) had the shortest distance to their destinations, and therefore any extra distance drawn by the participants meant error. We compared the target routes on the map with the routes drawn by the participants, by superimposing both of them on a single map. The additional distance drawn by the participants was recorded as map retention distance error. In order to measure efficiency performance, we recorded the time taken by each subject to complete a single route (traversal time) for each device and calculated the average value for each aid condition. In addition, perceived accuracy was obtained through a questionnaire (5-point Likert scale) conducted after the experiment. Perceived accuracy here refers to how accurate the users perceived the navigational aids to be. It is not about the accuracy of 3D feature tracking algorithms used in this study. We used 3D feature tracking as an established method. Regarding the measurement and verification of 3D feature tracking accuracy, previous studies have documented the technical details, for example, benchmarking with corresponding ground truth poses or benchmarking with device data including inertial sensor data (e.g., gravity, acceleration, and rotation rate), camera properties (e.g., shutter time, gain, and focus), and time stamps [47]–[51]. We did not cover the details here due to limited space in this paper. The questionnaire in the current study also included other subjective evaluation questions for wearability comfort, usability control comfort, display comfort ratings, and subjective workload (raw NASA-TLX, without the weighting procedure).

4) Procedure

First, the participants read the information letter that described the details of the experiment, and then they filled the consent form and the pre-experiment questionnaire. Short practice for about 5 minutes was provided for them to get familiar with the devices. Most participants had not used Google Glass before, so we gave them adequate time to practice with the navigational technology until they felt fully confident to initiate the formal experiment. In each of the three trials, each participant was instructed to navigate using one of the three aids (wearable, handheld phone, and paper map) from the start location to the end location, taking the shortest route. Each end location was a locker at the test location. They were instructed to arrive at the destination as quickly as possible with a reasonable and safe walking speed in the same way for all three navigation conditions. Although different individuals may have different baseline walking speed, it should not affect our results because we used a repeated measures design. The experimenter shadowed and timed the participants. Once the participants completed testing the three aids, they were asked to fill the post-experiment questionnaires. Finally, they were given a blank floor map (Figure 3b) and were requested to draw the three routes as they remembered during the experiment. The participants drew all the three maps at the end after they had finished navigating all the routes and spent a few minutes filling the post experiment questionnaire.

B. Results

Initially, repeated measures MANOVA (multivariate analysis of variance) was conducted using SPSS (Version 22) to determine the effect of navigational aid type on the dependent variables, which included traversal time (task completion time), perceived accuracy, NASA-TLX (workload score), map retention distance error, and subjective evaluation scores (wearability comfort, display comfort, and usability control comfort).

Preliminary assumption checking revealed that there was no univariate or multivariate outlier, as assessed by boxplot and Mahalanobis distance, respectively; there were linear relationships, as assessed by scatterplot; no multicollinearity was present as assessed by Pearson correlation. The data was not normally distributed, as assessed by Shapiro-Wilk's and Kolmogorov-Smirnov's test ($p < 0.001$). The assumption for homogeneity of variance/covariances, as assessed by Box's test of equality of covariance ($p < 0.001$), was also not met. However, MANOVA is robust to violations of multivariate normality and violations of homogeneity of variance/covariance, if groups are of nearly equal size [52]–[54]. Since our groups were indeed of an equal size, we continued with the analysis. The MANOVA result showed that the effect on the dependent variables combined was significant, $F(12, 220) = 9.735$, $p < 0.001$; Pillai's Trace = 0.694; partial $\eta^2 = 0.347$.

Then we followed it up with repeated measures ANOVA (analysis of variance) using SPSS (Version 22); pairwise comparisons were conducted (with Bonferroni correction) to compare the three types of aids. One-way repeated measures ANOVA is also considered to be very robust against the violation of normality; Greenhouse-Geisser correction was consulted when the sphericity assumption was violated [55]–[57]. The effect of aid type on perceived accuracy was significant, $F(2, 76) = 29.622$, $p < 0.001$, $\eta^2 = 0.438$ as shown in Figure 4a. The wearable aid (4.46) was perceived to be more accurate than both cell phone (3.67) and paper map (3.00) conditions (p values < 0.001); difference of perceived accuracy found between the cell phone and paper map conditions was also significant ($p = 0.011$).

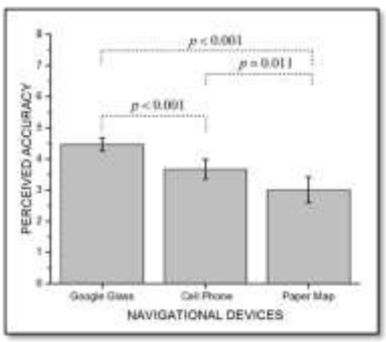
The effect of aid type on map retention distance error was also significant, $F(2, 76) = 11.056$, $p < 0.001$, $\eta^2 = 0.225$. No significant difference was found between the wearable (1.67 m) and cell phone (1.54 m) conditions ($p = 1.000$), but both conditions had significantly larger retention error than the paper map (0.63 m) condition (p values ≤ 0.001) as shown in Figure 4b.

Similarly, the effect of aid type on NASA-TLX overall workload score was significant, $F(2, 76) = 40.239$, $p < 0.001$,

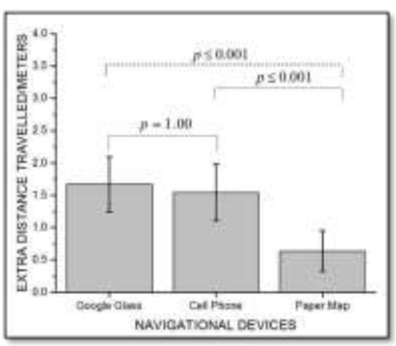
$\eta^2 = 0.514$. No significant difference was found between the wearable (21.52) and cell phone (28.53) conditions (p = 0.059), but both of them had significantly smaller overall workload than the paper map (52.39) condition (p values < 0.001), shown in Figure 4c.

The effect of aid type on traversal time (task completion time) was significant, $F(1.371, 52.116) = 10.515$, $p = 0.001$, $\eta^2 = 0.217$, using the Greenhouse-Geisser correction $\hat{\varepsilon} = 0.686$, because Mauchly's Test showed that the sphericity assumption was violated, $p < 0.001$. No significant difference was found between the wearable (111.26 s) and cell phone (118.03 s) conditions (p = 1.000), but both of them had significantly shorter completion time than the paper map (219.21 s) condition (p values ≤ 0.008) as shown in Figure 4d.

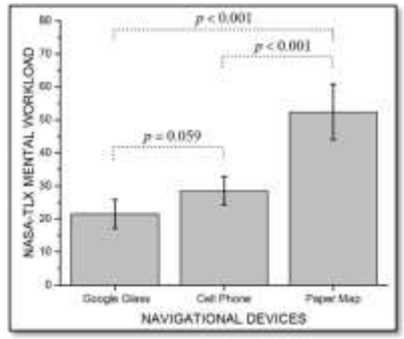
No significant effect was found on the wearability comfort ($p = 0.162$, $\eta^2 = 0.047$) between the wearable (3.46), cell phone (4.05), and paper map condition (3.64). Similarly no significant effect was found on usability control comfort ($p = 0.224$, $\eta^2 = 0.078$) between the wearable (3.97), cell phone (3.74), and paper map condition (3.58). Also no significant effect was found on display comfort ratings ($p = 0.221$, $\eta^2 = 0.039$) between the wearable (3.36), cell phone (3.79), and paper map condition (3.69).



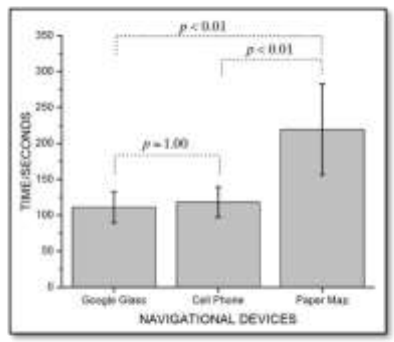
(a)



(b)



(c)



(d)

Fig. 4. Effects of navigation aid type on (a) perceived accuracy, (b) route retention error, (c) NASA-TLX overall workload rating, and (d) efficiency. Error bars represent 95% confidence interval.

C. Human Factors Study Discussion

In this human factors experiment, the wearable device (Google Glass) was perceived to have the best accuracy. A potential explanation for this would be that the camera of the wearable device was located at a higher position than the handheld cell phone; the high position gave it a wider view for feature tracking, and it was also a more natural viewing angle. The camera of the cell phone was usually held at the mid-body level that is different from the normal viewing angle, and therefore it may be perceived as unnatural and less accurate. Also the HMD on the wearable device made the AR experience more intuitive. The virtual representation of directional instructions on the camera feed was directly concentrated on the pupil of the eye, and the camera also adjusted naturally with head movement. This feature enhanced the navigational experience of the wearable device as its interface became more focused and adaptive.

A disadvantage of the cell phone condition is that it has to be held in an upright position, which makes users' arm tired. The way users held the mobile phone while navigation is not an ergonomic posture to maintain while walking. In contrast, HMD (such as Glass) does not have this issue. The results from the current study, however, did not show this disadvantage of the cell phone, probably because the route and test time were not long enough. Future studies need to test and compare the devices in longer routes with longer test duration to investigate this issue.

The traversal time was not significantly different between the wearable and the cell phone conditions. The traditional paper map, however, was a very slow medium for directional assistance. It took participants almost twice as much time as the two electronic device conditions. An explanation is that when using the paper map, users have to mentally understand and rotate the map and then translate it to the contextual environment. This is same as our expectation based on previous study findings.

No significant difference was found on subjective comfort ratings (wearability comfort, usability control comfort, and display comfort) across the three aid types. This is possibly because each individual device had certain drawbacks that influenced the participants' experience. The cellphone had to be kept at a certain position and orientation in front of the head for the augmented information to match the real-world perspective. Glass has a display resolution smaller than the smartphone, and the display contrast may be low due to background glare. For the paper map condition, the floor plan was not easily explicable because the paper map had excessive information that made discerning the area of interest challenging.

The NASA-TLX results showed that navigation using the paper map caused the highest workload. The participants had to analyze where they were on the map with respect to the environment and also identify their target location; then they had to constantly analyze the surrounding for potential clues. All this yielded a heavy toll on the time taken to complete the experiment and raised participant dissatisfaction. The workload values in the wearable and cell phone conditions were lower since neither was a cognitively strenuous exercise.

Another key aspect we wanted to evaluate was route retention in case the user had to navigate the same routes without the navigational aids. We concluded that the wearable device and the cell phone performed poorly in this test as the retention errors were larger than the paper map condition. In the map retention test, we used a paper map similar (but not identical) to the one used in the navigation condition (Figure 3). Alternatively, a blank piece of paper could be used. The advantage of using a blank paper is that it would not provide any reminder of the paper map used in the navigation test. However, the disadvantage of using a blank paper is that it would be very difficult to quantify map retention error without the necessary spatial and distance references (e.g., walls and corridors). As a result, we chose to use a similar paper map in the retention tests with design considerations to minimize its potential disadvantages. The navigation activity using the paper map was for a relatively short period of time (several minutes). There was a time delay from using the map as a navigation aid to the map retention test (at least 10 minutes). The participants were asked to complete other survey and workload questionnaires before finally asked to complete the memory retention test, minimizing any trace of the navigation map in the working memory. Participants were not told that there would be a map retention test until after all the navigation tests, so they should not have strong motivation to memorize the map. The navigation map did not contain start points or the shortest route information. Moreover, previous studies that administered a similar sketching question, on a blank paper, also reported results indicating that users of digital navigation devices had poorer understanding of the routes as compared to those who used paper maps [58]. Nevertheless, it is a potential limitation that the retention test paper map looks similar to the navigation aid paper map. An improvement in future studies could be adding the use of a blank paper as the first step of retention test, followed by the second step using a map with necessary spatial information. Combining the two methods may give a more comprehensive evaluation of map retention. Since the routes used in the current study were relatively short and simple, all the participants were able to reach all the three destinations, and nobody was lost during the task. There were very few cases where participants made a wrong turn, so navigation error was not regarded as a dependent variable. In such error cases, it was often only a couple of steps away from the correct route. When the digital aids were used, they could provide cues for participants to turn back and return to the correct path. When the paper map was used, we found that participants would stop and look around, and finally they can correct themselves back on track. However, this stop would increase the total task completion time, so this time variable was used as the dependent variable. Although it was not strictly measured, we observed that the digital aids could help people recover faster from such small errors when they happened.

When using digital devices for navigation, participants get used to simply following the navigational instructions and are not involved in actively processing the surrounding environmental information. In contrast, when using a paper map, the participants have to analyze the environment alongside the map in order to navigate successfully. Automated navigation aids, resulting in worse map retention performance, could become a problem when they become dysfunctional, especially for users in critical situations like rescue workers or fire fighters. Regarding the paper map, it requires deeper understanding and mental processing of the environment. These processes increase navigation time and workload but at the same time they equip the user with adequate cues that help make future navigation a lot easier. A potential solution could be to develop adaptive automation aid systems that could balance the need for navigation aid and the need for map memorization and retention. Future studies are needed to identify better design solutions.

VI. Conclusion and Future Work

With respect to our research questions, the results showed that, first, the developed solution passed the technical assessments and worked well when tested during usual walking scenarios. Second, the human factors study showed that the HMD aid was perceived to be more accurate, with similar performance and workload results to the handheld smartphone, but both had worse route retention when compared to the paper map.

In the first technical assessment, the results showed that on average, the minimum average feature percentage needed to conduct appropriate navigation on the route was approximately 45%. In the second assessment, walking on the route at a faster speed than the general walking speed, we found that 50.6% of features were successfully detected on average, therefore detecting more features than the minimum needed. Both the first and second assessments found that the general walking speed to be lower than the threshold speed that was maintained during experimentation, therefore indicating that our developed system was feasible for practical use at moderately fast walking speeds. The third assessment was conducted to figure out the average speed and maximum time the application would entail to work ideally (detect 95% of features). The maximum time for the system to identify 95% of features was under 1 s at all areas with an average speed of around 4.34 km/h, which validated the fact that this application could operate ideally with maximum efficiency at normal walking speeds. In the last assessment, we measured the average time it took Google Glass to produce audio augmentation and visual direction information, which was 0.18 s and 0.14 s respectively. This result confirmed that the application was highly efficient and able to quickly process and display the directional information.

In the human factors experiment, the wearable device (Google Glass) was perceived to have the best accuracy. The traversal time was not significantly different between the wearable and the cell phone conditions; however, the paper map condition was comparatively time consuming. No significant difference was found on subjective comfort ratings (wearability comfort, usability control comfort, and display comfort) across the three aids. The NASA-TLX results showed that navigation using the paper map caused the highest workload. We concluded that the wearable device and the cell phone performed poorly in the memory retention test as their errors were much larger than the paper map condition. The wearable device was perceived to be more accurate, but objective performance and subjective workload results indicated that the wearable device condition was not significantly different from the handheld cell phone condition. This result might be explained by the fact that the current experiment was conducted in a simple indoor environment and used relatively shorter routes. We also faced technical difficulties as the Google Glass had limited battery life, and 3D scanning during the pre-deployment stages was time consuming and complicated, which hampered our ability to conduct large scale tests. Based on the current results, we concluded that augmented reality indoor navigation implemented on the wearable device was neither worse nor better than the cell phone implementation. However, we still expect that the wearable implementation would be preferred if the task was performed for longer duration in a more complex environment. The current study, however, would form the basis for future research that could aim to use technologically superior wearable devices with better battery life and higher computational powers.

In future studies, an alternative route retention test could be used as a way to avoid the need of using the paper map again. This route retention test could require participants to re-walk the routes without any assistive aids, and their time and route errors are recorded. It will be interesting to see which route retention test is better.

It would also be a unique idea (thanks to an anonymous reviewer) to examine an improved assistive design that adds a small version of the area map in a corner of the AR or handheld display [59]. When the size of the map is properly selected, it might potentially improve map retention results.

We would also be interested in examining the time duration taken for completing the route retention exercise and the effects of different navigational aids on this retention task time. The time measure would be especially meaningful in military and firefighting situations in which quick reaction is very important.

References

[1] P. Coelho, A. Aguiar, and J. C. Lopes, "OLBS: Offline Location Based Services," in *Next Generation Mobile Applications, Services and Technologies (NGMAST), 2011 5th International Conference on*, 2011, pp. 70–75.

[2] Tony Costa, Sarah Rotman Epps, Thomas Husson, Julie A. Ask, Peter Sheldon, Carlton A. Doty, Colin Campbell, and Andia Vokshi, "Indoor Venues Are The Next Frontier For Location-Based Services," Jan. 2013.

[3] G.-M. Jeong, W.-S. Choi, G. Y. Han, D.-K. Suh, and J.-Y. Yeo, "Paradigm Shift and the State of the Art of LBS in the Advent of Smartphone," in *Security-Enriched Urban Computing and Smart Grid*, R.-S. Chang, T. Kim, and S.-L. Peng, Eds. Springer Berlin Heidelberg, 2011, pp. 251–258.

[4] Mautz, Rainer, "Indoor Positioning Technologies." ETH Zurich, Department of Civil, Environmental and Geomatic Engineering, Institute of Geodesy and Photogrammetry, 2012.

[5] K. Pahlavan, X. Li, and J.-P. Makela, "Indoor geolocation science and technology," *Commun. Mag. IEEE*, vol. 40, no. 2, pp. 112–118, 2002.

[6] M. Brown and J. Pinchin, "Exploring Human Factors in Indoor Navigation," in *The European Navigation Conference*, 2013.

[7] H. Moustafa, H. Kenn, K. Sayrafian, W. Scanlon, and Y. Zhang, "Mobile wearable communications [Guest Editorial]," *Wirel. Commun. IEEE*, vol. 22, no. 1, pp. 10–11, 2015.

[8] S. Emmitt and K. Ruikar, *Collaborative Design Management*. Routledge, 2013.

[9] S. L. Joseph, X. Zhang, I. Dryanovski, J. Xiao, C. Yi, and Y. Tian, "Semantic Indoor Navigation with a Blind-User Oriented Augmented Reality," 2013, pp. 3585–3591.

[10] M. Kalkusch, T. Lidy, M. Knapp, G. Reitmayr, H. Kaufmann, and D. Schmalstieg, "Structured visual markers for indoor pathfinding," in *Augmented Reality Toolkit, The First IEEE International Workshop*, 2002, p. 8–pp.

[11] J. Kim and H. Jun, "Vision-based location positioning using augmented reality for indoor navigation," *Consum. Electron. IEEE Trans. On*, vol. 54, no. 3, pp. 954–962, 2008.

[12] J. B. Kim, "A personal identity annotation overlay system using a wearable computer for augmented reality," *Consum. Electron. IEEE Trans. On*, vol. 49, no. 4, pp. 1457–1467, 2003.

[13] S. Kasprzak, A. Komninos, and P. Barrie, "Feature-based indoor navigation using augmented reality," in *Intelligent Environments (IE), 2013 9th International Conference on, ACM*, 2013, pp. 100–107.

[14] A. Mulloni, H. Seichter, and D. Schmalstieg, "Handheld augmented reality indoor navigation with activity-based instructions," in *Proceedings of the 13th international conference on human computer interaction with mobile devices and services*, 2011, pp. 211–220.

[15] L. C. Huey, P. Sebastian, and M. Drieberg, "Augmented reality based indoor positioning navigation tool," in *Open Systems (ICOS), 2011 IEEE Conference on*, 2011, pp. 256–260.

[16] A. Möller, M. Kranz, S. Diewald, L. Roalter, R. Huitl, T. Stockinger, M. Koelle, and P. A. Lindemann, "Experimental evaluation of user interfaces for visual indoor navigation," 2014, pp. 3607–3616.

[17] N. Bhanage and A. Zhong, "Improving User Experiences in Indoor Navigation Using Augmented Reality," EECS Department, University of California, Berkley, Thesis UCB/EECS-2014-73, May 2014.

[18] U. Rehman and S. Cao, "Augmented Reality-Based Indoor Navigation Using Google Glass as a Wearable Head-Mounted Display," in *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on*, 2015, pp. 1452–1457.

[19] U. Rehman and S. Cao, "Experimental Evaluation of Indoor Navigation Devices," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2015, vol. 59, pp. 1783–1787.

[20] J. Holmquist, "Navigation aids in route training: Increase navigation speed, decrease route retention?," University of Central Florida, 2005.

[21] Plamen Levchev, Michael N Krishnan, Chaoran Yu, Joseph Menke, and Avideh Zakhor, "Simultaneous Fingerprinting and Mapping for Multimodal Image and WiFi Indoor Positioning," presented at the Indoor Positioning and Indoor Navigation, 2014.

[22] J. Z. Liang, N. Corso, E. Turner, and A. Zakhor, "Image based localization in indoor environments," in *Computing for Geospatial*

Research and Application (COM. Geo), 2013 Fourth International Conference on, 2013, pp. 70–75.
[23] JZ Liang, E Turner, A Zakhor, and N Corso, "Image-Based Positioning of Mobile Devices in Indoor Environments," in *Multimodal Location Estimation of Videos and Images*, Springer International Publishing, 2015, pp. 85–89.
[24] G. Gerstweiler, E. Vonach, and H. Kaufmann, "HyMoTrack: A Mobile AR Navigation System for Complex Indoor Environments," *Sensors*, vol. 16, no. 1, p. 17, Dec. 2015.
[25] T. Bailey and H. Durrant-Whyte, "Simultaneous localization and mapping (SLAM): Part II," *IEEE Robot. Autom. Mag.*, vol. 13, no. 3, pp. 108–117, 2006.
[26] M. A. Lakhani, "Indoor Navigation based on Fiducial Markers of Opportunity," 2013.
[27] S. S. Chawathe, "Marker-based localizing for indoor navigation," in *Intelligent Transportation Systems Conference, 2007. ITSC 2007. IEEE*, 2007, pp. 885–890.
[28] B. A. Delail, L. Weruaga, and M. J. Zemerly, "CAViAR: Context Aware Visual Indoor Augmented Reality for a University Campus," 2012, pp. 286–290.
[29] D. Amin and S. Govilkar, "Comparative Study of Augmented Reality Sdk's," *Int. J. Comput. Sci. Appl.*, vol. 5, no. 1, pp. 11–26, Feb. 2015.
[30] C. Koch, M. Neges, M. König, and M. Abramovici, "Natural markers for augmented reality-based indoor navigation and facility maintenance," *Autom. Constr.*, vol. 48, pp. 18–30, Dec. 2014.
[31] "Metaio SDK," May-2013. [Online]. Available: http://www.metaio.com/sdk/.
[32] "Metaio Toolbox," 28-Aug-2012. [Online]. Available: http://dev.metaio.com/storage-folder/toolbox/toolbox/. [Accessed: 14-Apr-2015].
[33] M. Billinghurst, R. Grasset, and J. Looser, "Designing augmented reality interfaces," *ACM Siggraph Comput. Graph.*, vol. 39, no. 1, pp. 17–22, 2005.
[34] R. B. Rusu, Z. C. Marton, N. Blodow, M. Dolha, and M. Beetz, "Towards 3D Point cloud based object maps for household environments," *Robot. Auton. Syst.*, vol. 56, no. 11, pp. 927–941, Nov. 2008.
[35] A. Irschara, C. Zach, J.-M. Frahm, and H. Bischof, "From structure-from-motion point clouds to fast location recognition," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, 2009, pp. 2599–2606.
[36] K. Fitzpatrick, M. Brewer, and S. Turner, "Another Look at Pedestrian Walking Speed," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1982, pp. 21–29, Jan. 2006.
[37] L. Bass and B. E. John, "Supporting the CANCEL command through software architecture," 2002.
[38] D. S. Brungart, B. D. Simpson, R. L. McKinley, A. J. Kordik, R. C. Dallman, and D. A. Ovenshire, "The interaction between head-tracker latency, source duration, and response time in the localization of virtual sound sources," 2004.
[39] C. D. Wickens, C.-C. Liang, T. Prevett, and O. Olmos, "Egocentric and exocentric displays for terminal area navigation," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 1994, vol. 38, pp. 16–20.
[40] M. Raubal and M. J. Egenhofer, "Comparing the complexity of wayfinding tasks in built environments," *Environ. Plan. B Plan. Des.*, vol. 25, no. 6, pp. 895–913, 1998.
[41] B. Kuipers, "The' map in the head' metaphor," *Environ. Behav.*, vol. 14, no. 2, pp. 202–220, 1982.
[42] S. Gillner and H. A. Mallot, "Navigation and acquisition of spatial knowledge in a virtual maze," *J. Cogn. Neurosci.*, vol. 10, no. 4, pp. 445–463, 1998.
[43] F. Filimon, "Are All Spatial Reference Frames Egocentric? Reinterpreting Evidence for Allocentric, Object-Centered, or World-Centered Reference Frames," *Front. Hum. Neurosci.*, vol. 9, Dec. 2015.
[44] K. Harwood and C. D. Wickens, "Frames of reference for helicopter electronic maps: The relevance of spatial cognition and componential analysis," *Int. J. Aviat. Psychol.*, vol. 1, no. 1, pp. 5–23, 1991.
[45] W.-C. Lee and B.-W. Cheng, "Effects of using a portable navigation system and paper map in real driving," *Accid. Anal. Prev.*, vol. 40, no. 1, pp. 303–308, 2008.
[46] J. Z. Bakdash, S. A. Linkenauger, and D. Proffitt, "Comparing decision-making and control for learning a virtual environment: Backseat drivers learn where they are going," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2008, vol. 52, pp. 2117–2121.
[47] S. Lieberknecht, S. Benhimane, P. Meier, and N. Navab, "A dataset and evaluation methodology for template-based tracking algorithms," in *Mixed and Augmented Reality, 2009. ISMAR 2009. 8th IEEE International Symposium on*, 2009, pp. 145–151.
[48] K. Pentenrieder, P. Meier, G. Klinker, and others, "Analysis of tracking accuracy for single-camera square-marker-based tracking," in *Proc. Dritter Workshop Virtuelle und Erweiterte Realitt der GIFachgruppe VR/AR, Koblenz, Germany*, 2006.
[49] D. Kurz, S. Lieberknecht, S. Benhimane, and others, "Benchmarking inertial sensor-aided localization and tracking methods," in *The second international workshop on AR/MR registration, tracking and benchmarking*, 2011.
[50] D. Kurz and S. Ben Himane, "Inertial sensor-aligned visual feature descriptors," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, 2011, pp. 161–166.
[51] D. Kurz, T. Olszamowski, and S. Benhimane, "Representative feature descriptor sets for robust handheld camera localization," in *Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on*, 2012, pp. 65–70.
[52] H. Finch, "Comparison of the performance of nonparametric and parametric MANOVA test statistics when assumptions are violated," *Methodology*, vol. 1, no. 1, pp. 27–38, 2005.
[53] "One-way MANOVA in SPSS Statistics | Laerd Statistics Premium." [Online]. Available: https://statistics.laerd.com/premium/spss/owm/one-way-manova-in-spss.php. [Accessed: 05-Apr-2016].
[54] N. L. Leech, K. C. Barrett, and G. A. Morgan, *IBM SPSS for intermediate statistics: Use and interpretation*. Routledge, 2015.
[55] G. Norman, "Likert scales, levels of measurement and the 'laws' of statistics," *Adv. Health Sci. Educ.*, vol. 15, no. 5, pp. 625–632, 2010.
[56] D. Howell, *Statistical methods for psychology*. Cengage Learning, 2012.
[57] "One-way repeated measures ANOVA using SPSS Statistics | Laerd Statistics Premium." [Online]. Available: https://statistics.laerd.com/premium/spss/owrma/one-way-repeated-measures-anova-in-spss.php. [Accessed: 05-Apr-2016].
[58] T. Ishikawa, H. Fujiwara, O. Imai, and A. Okabe, "Wayfinding with a GPS-based mobile navigation system: A comparison with maps and direct experience," *J. Environ. Psychol.*, vol. 28, no. 1, pp. 74–82, Mar. 2008.
[59] T. Ni, D. A. Bowman, and J. Chen, "Increased display size and resolution improve task performance in information-rich virtual environments," in Proceedings of Graphics Interface 2006, 2006, pp. 139–146.

Umair Rehman is a MASc candidate within the Department of Systems Design Engineering at University of Waterloo. He received his BS in Electrical and Computer Engineering from WPI, Worcester, MA in 2014. He is an HCI (Human-Computer Interaction) researcher who has special interest towards augmented reality, 3D interfaces and wearable computing.

Dr. Shi Cao is an Assistant Professor within the Department of Systems Design Engineering at University of Waterloo. He received his PhD in Industrial and Operations Engineering from University of Michigan, Ann Arbor in 2013. His research areas include human performance and workload modelling, transportation human factors, and human-computer interaction.