eyeWalkie : Indoor Navigation System for Impaired Students in University Buildings

Annemarie J. Collijn
Faculteit van Exacte Wetenschappen, Vrije Universiteit.
Amsterdam, the Netherlands
email a.j.collijn@student.vu.nl

ABSTRACT

Navigating in university building is at first complicated and confusing for every student. For visually and physically impaired student this is a daily challenge. In this paper we will present the system *eyeWalkie*, an indoor navigating system specifically designed for this specific group of users. *eyeWalkie* will not only provide directions, but also avoid obstacles which might not seem like obstacles to other users; e.g. large open spaces, stairs and crowds.

Keywords

Ambient Intelligence, Mobile, indoor navigation, mobile devices, system design, disabled students

1 INTRODUCTION

"In the near future our homes will have a distributed network of intelligent devices that provides us with information, communication, and entertainment." (Philips)

For this project we have looked into the domain of Ambient Intelligence and Mobile systems. There are several sorts of ambient intelligence which are already at work in different environments. These environments differ from educational to healthcare facilities and many more.

There are several research papers to be found with examples of Ambient Intelligence (AmI) in different situation. For example (Ntoa, Antona, Margetis, & Stephanides, 2013) show an AmI system which monitors a students learning abilities and progress during several years. Another example is an AmI system which is discussed by (Kartakis, Sakkalis, Tourlakis, Zacharioudakis, & Stephanidis, 2012), where a hospital room is transformed into an ambient intelligent room. Techniques are used to allow the patient to control the environment and interact with present facilities, while vital signs are monitored by a clinically oriented interface.

We will focus primarily on systems aiding people with disabilities, namely visually or otherwise physically impaired. With technological possibilities developing as they are, designers become more and more able to design systems that can aid people with visual impairments or other disabilities. In this paper we will propose a system for indoor navigation, using a handheld, or mobile, device. This system, called *eyeWalkie*, is able to help users navigate indoors without running into obstacles, and finding the shortest and/or easiest and safest route to different resources, e.g. offices, classrooms, elevators, etc.

The basis for our design will mainly be the work of (Burigat & Chittaro, 2012), in which they describe the navigational needs of disabled students in a university setting. We found their work a good basis for a design of our own for a very specific target group. We chose this group because we ourselves have come across some difficulties navigating on a new campus. We assume that having a physical disability, like blindness or motoric difficulties, would cause more difficulties in navigating on already complicated grounds.

Another paper which has greatly influenced our design decisions, is from (Fallah, Apostolopoulos, Bekris, & Folmer, 2012). In their paper the authors discuss the most commonly used techniques and methods for navigation systems, stating both negative and positive points for each technique or method discussed.

In the first part of this paper we will discuss the topics we chose as our main interest, i.e. Ambient Intelligence and Mobile, as well as Indoor Navigation. These first two topics were chosen from a list of ten topics, the latter is our proposed system's main focus. A literary review will be part of this, and will at the same time review the state of the art. Next a description of our interaction and system design for *eyeWalkie* will follow. In the final part we will evaluate our proposed system.

2 RELATED WORK

2.1 Ambient intelligence

As mentioned in the introduction, Kartakis et al. propose a system which allows patient to interact with their environment, in their hospital room, while a different interface in the same system allows hospital staff to monitor vital signs of the patient. Their system has an interface for the patient, as shown in Figure 1.

Several functionalities are present, for example the schedule for this patient, a button to call the nurse, and some environmental controls. This is a good example of ambient intelligence when looking at the following definition.

"The objective of AmI is to broaden the interaction between human beings and digital information technology through the usage of ubiquitous computing devices." (Alcañiz & Rey, 2005) give the above definition in their

Figure 1. Patient UI (Figure 2 from Kartakis et al.)

paper, along with an explanation of the main focus in developing AmI systems. According to the authors, one should focus on user-friendly and low-cost solutions with a high level of network security. This will help make sure it is profitable for companies to develop such a system, looking at Human-Computer Interaction will help make sure a system is used by the intended user group, and network security will help create a more stable system, which adds to both points.

One of the main requirements for AmI systems is being unobtrusive, personalized, adaptive and anticipatory. (Emiliani & Stephanidis, 2005) explain how these requirements should be implemented. Being unobtrusive means a system has to be imbedded in the environment, and not intruding on a user's consciousness until needed. A system should be personalized, i.e. tailored to the user's needs. It should be able to adapt in response to a user's actions, and finally able to anticipate a user's desires and environments with as little interference as possible.

Emiliani et al. discuss how AmI systems appear to have a wide range of benefits, but also several challenges that will have to be overcome to create a successful and usable system. They focus on the distribution of interaction over devices and modalities, the balance between automation and adaptation and direct control, the identification of contextual dependencies among services, health and safety issues, privacy and security, and social interaction in ambient intelligence environments.

Burigat et al.'s research is based on interviews with our intended target group, and is done after they discovered there were a lot of research and system designs for this group, but no literature on their specific needs. Their sample consisted of 7 students with different degrees of motor disability, 3 with visual disability, 1 with dyslexia, and 2 with multiple disabilities (motor and visual), totalling at a sample of 13 students.

The most important conclusion from their research state the relevance of navigational barriers. A system should be able to navigate around these barriers. Barriers can be stairs, but also empty spaces. They also stress the point of a system being mobile, for it to work and fit needs better. Being mobile gives the system a more adaptable character. It allows users to change direction and destination, when for example encountering a change in the school schedule.

The need for guidance towards resources is another point at which developers should look. It is often difficult for disabled people to find resources such as tables, sockets, etc. available to them; a system should be able to provide directions.

2.2 Mobile systems

A constraint for a navigational system for (visually) impaired people is the use of mobile devices. Even though most people nowadays have a mobile phone, not everyone has a smart phone. And also, a touch screen phone might not be usable for some people in the target group. According to Burigat et al., the need to get feedback through the pushbuttons of the phone is necessary for them to work the phone properly. However, as Tommy Edison, also known as the Blind Film Critic, shows in a video on YouTube (Edison, 2012), it is possible for blind people to use touch screen, when using the voice-over software already available in most smartphones. The video shows an iPhone, but other phones have the same option from the Accessibility menu which can be found in the Settings menu in most mobile phones.

(Bouwer, Visser, Nack, & Terwijn, 2013) performed four studies in the domain of navigation aids in complex indoor environments. This research was done in the context of the SmartINSIDE project (Visser, 2010-2012). The SmartINSIDE project aimed at developing a complete integrated solution for indoor navigation services within large sites. According to Visser *"indoor navigation will be cheaper, more efficient, and accessible to any user with a standard mobile phone."* Bouwer et al. imply in their research that people at a fair are likely to travel in groups of 2 to 5 people. Of course, when new to University, one probably does not know who will be in their class and who is looking for the same room, but it is feasible that in this environment one would seek people to travel with.

2.3 Indoor navigation

(Guerrero, Vasquez, & Ochoa, 2012) present a prototype of an indoor navigation system for visually impaired people. They have focused mainly on usability of the application, as well as the suitability for deployment of the system in other settings. In order to be able to avoid obstacles like chairs, tables, stairs, cabinets, etc, they propose to use an infrared sensor which is positioned on the user's cane. The cane also supports the navigation system, as is it built on the cane. While an interesting development, since their system is focused on visually impaired people, as opposed to our system, which is also focused on otherwise physically impaired people, we will not use the same methods with add-ons for the used mobile devices.

A system which provides for the same target group we focus on is Drishti (Helal, Moore, & Ramachandran, 2004). The goal of Drishti is to provide the blind user with an "*augmented pedestrian experience*". This augmented experience includes enough inputs to provide a comfortable walk from one location to another. Although the proposed system has several good features, which include avoiding obstacles and a smooth adjustment between outdoors and indoors navigation, there is one big issue with this system. The prototype weighs approximately 8 lbs. As mentioned before, we do not want any add-ons, for this can hinder the user.

(Stein, Gottshall, Donofrio, & Kling, 2011) introduce us to iDocent; an indoor navigational smartphone application for the blind which uses already existing Wi-Fi access points to locate an individual through triangulation and route the user to a desired location. This design is later further developed by (Guotana, Kim, Partlo, & Thiagarajasubramanian, 2011). This application can not only help a user navigate through a university building on request; it also detects the smartphone of the user and starts automatically.

Several techniques with which to develop an indoor navigational system are reviewed and proposed in literature. For example (Evennou & Marx, 2006) discuss how inertial navigation systems can be combined with Wi-Fi triangulation. The Wi-Fi signal will aid in achieving higher accuracy in localization through dead-reckoning using inertial navigation systems.

Fallah et al. summarize different approaches for localization, dividing these approaches in four distinct groups;

1. *Dead-reckoning*
   This technique estimates the users location based on a previously estimated or known position. Movement of the user is measured using odometry readings from sensors such as accelerometers and gyroscopes.
   Dead-reckoning has a problem with accuracy because of the estimates. If an estimation result is incorrect, this leads to an accumulation of errors in localization.
   Some of the problems with accuracy can be overcome by combining this technique with other localization techniques.
2. *Direct-sensing*
   This technique localizes the user through the sensing of tags or identifiers which are installed in the environment. There are two approaches to this: (1) the tag itself stores the location and environment information; or (2) this information is stored in a separate database and retrieved using the tags' identifier, which is unique.
   There are different tags available: Radio Frequency Identifier Description (RFID), Infrared (IR), Ultrasound identifiers (USID), Bluetooth beacons and Barcodes.
   The main drawback is that the user has to carry the sensor with them. It has to be placed on them or their equipment.
3. *Triangulation*
   This technique localizes the user by triangulating the tags installed in known locations, using at least three known positions. Lateration uses the distance between a user and at least three known points, while angulation uses angular measurements from these three points to localize the user. For outdoor navigation, this technique is used (GPS). In situations where GPS is not available, as it is not indoors, there are alternatives. For example Cell-tower positioning, which uses the known locations of cell towers, using the provided signal strength; or wireless local area network (WLAN), which uses triangulation using the location of wireless base stations using the provided signal strength.
   Both non-GPS techniques have a lower precision than GPS, which is a large drawback.
4. *Pattern-recognition*
   This technique uses data from one or more sensors carried by the user, to localize the user. It compares the data from these sensors with prior collected data, which has been coupled with an environment map.
   There are several different sensing techniques: Computer vision based, which requires the user to carry a camera, either embedded in a hand-held device or not. This camera captures images of the environment, which it matches against prior collected images of know locations, and then can determine the user's position and orientation. This method can be impeding on mobility, because of the need to carry supporting computing equipment, which is often required.
   Signal distribution or fingerprinting, compares the unique signal data from a source sensed at a particular location with a map of prerecorded data. It basically does the same as computer vision, but with a different method. A so called training phase is needed, where the received signal strength is acquired and stored to create a map. While navigating, the received signal strength is measured and compared with this data to find the closest match. One could use the signal strength of WLAN access points, which would need a relatively small number of base stations to localize the user.

Besides the different techniques for localization, Fallah et al. also describe ways for path planning, representation of the system and interaction issues, which are useful for our own project.

An example of direct sensing as a localization technique is explained by (Ozdenizci, Ok, Coskun, & Aydin, 2011).

They use Near Field Communication (NFC). According to them, their system, called NFC Inertial, increases the usability of indoor navigation systems. By using NFC tags and an NFC enabled mobile phone, one can navigate easily indoors. NFC tags are also used in McDonald's restaurants, but in a different way. There they are used to be able to play using the tabletops as a racing venue (McDonald's - Happy Table / Table game - case study (2013), 2013).

(Liu, Chen, Pei, Guinness, & Kuusniemi, 2012) present a system, HIPE, which is a smartphone indoor positioning engine. HIPE uses measurements from smart phone sensors combined with wireless signals to determine the user's position. The authors claim HIPE can be used in other systems to localize a user.

A way to create more accurate localization is shown by (Feng, et al., 2012). In their system, designed for visually impaired people, they use a pre-built map in fingerprinting methods to localize a Wi-Fi device. Their testing results show that the proposed tracking system leads to significant improvement over other positioning and tracking systems. Experimental results suggest that the proposed system can be used to guide visually impaired subjects to their desired destinations.

(Navarro, Peuker, & Quan, 2010) show how localization through fingerprinting is done for a quite different purpose. Their research goal is to proof a concept for a playground child tracking system. With the developments of WLAN, positioning techniques based on Wi-Fi are becoming available. As the two main options for Wi-Fi localization, they name triangulation and fingerprinting. For their purpose, triangulation showed the least promise, whereas fingerprinting had promising results.

People are increasing demands for wireless location services. (Feng & Liu, 2012) also discuss two positioning techniques based on WLAN: triangulation and fingerprinting. They claim using Wi-Fi allows the implementation of WLAN without wires. This reduces costs significantly. Wi-Fi signal strength is not so much interfered by environmental issues. It can be used in complex environments.

According to Feng & Liu, mobile geographic information systems (Mobile GIS) has become over the past few years one of the mainstream research directions with regards to indoor navigational systems. This is why they evaluate both fingerprinting and triangulation with Mobile GIS and speech recognition. Their results show that accuracy of positioning using Wi-Fi signal is higher than using GPS. Wi-Fi can also be used in many environments, not only outdoor, but indoor as well. They conclude that fingerprinting as a localization technique is more accurate than triangulation. Therefore, fingerprinting can be used in complex (indoor) environments. We will implement this technique in our system.

3 INTERACTION DESIGN

As mentioned in the introduction, we have focused mainly on the research of Burigat et al. to find the requirements of our system. A mobile system which leads a user not only to certain rooms, but also to available resources, should preferably be available for different types of smart phones.

3.1 Meet Mike and Susan

In this section two scenarios will be presented that demonstrate the potential use of the aimed for system.

In the following paragraphs, you will meet Mike and Susan, who will be present throughout different parts of our design.

Scenario 1- Mike

Mike is a 23 year old student of the University of Amsterdam. He is starting University this year, and has yet to learn his way around campus. Besides being new to the campus, he has the problem that he is blind.

To help him navigating campus, he has found the app *eye-Walkie*, which uses voice guidance to navigate to the correct room. He can tell the app where he wants to go, and the app shows him a route to his destination. By asking for more description, he can select a different route, or accept this route, if he wishes. When coming across an obstacle, he can tell the app to present a different route. The storyboard[1] in Figure 2 shows how Mike uses the application.

Figure 2. Storyboard

---

[1] A larger image of Figure 2 can be found in Appendix A – 1. Storyboard

Scenario 2 - Susan
Susan is 20 year old student, and like Mike, new to the University. She is not able to walk, and therefore uses a wheelchair to get around. She uses *eyeWalkie* which shows her the easiest route to the desired classroom. This route will direct her to the nearest elevator, if applicable. The app uses voice and images to give directions. When encountering an elevator out of order, she can tell the app to show her another route.

3.2 Use Cases

The following Use Case Diagram (Figure 3. Use Case Diagram) shows the functionalities available to the user.

The user of the system has several functionalities he or she can use. Before actually using the system, it is important to *set* the *preferences* for use. Of course it is possible to *alter preferences* at a later stage. There are different options available, depending on being visually or otherwise impaired. For example, for visually impaired people, the directions have a smaller interval, i.e. the directions will be repeated more often. However, for someone in a wheelchair, this might not be necessary; they could use the map images as additional feedback.

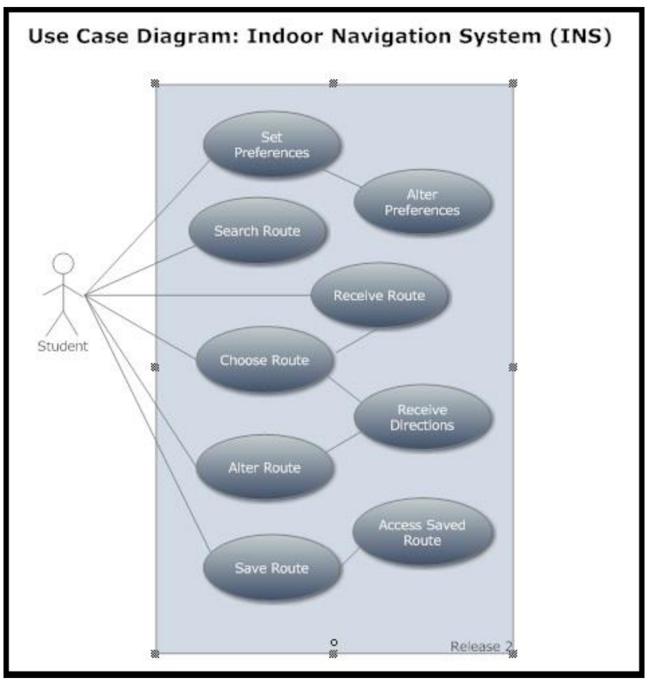


Figure 3. Use Case Diagram

The user is able to *search* a *route* to his or her destination by using voice commands. Simply telling the system to "please show me the route to room G3.03", will make the system search for the requested location, and deliver the best available route.

After searching the best available route, the user will *receive* the best *route*, based on the set preferences. This route will be the shortest path to the destination, considering eventual necessary objects like elevator and obstacles such as stairs. After this, the user will be able to *choose* this *route*, or choose an alternative route. This could be, for example, the user deciding he or she wants to use the stairs instead of the elevator. After choosing, the user will *receive directions* to the desired location.

User Action	System Response
User turns on application	System localizes user System greets user System gives description of present location System requests destination
User states destination	Systems searches available route(s) System checks user preferences System selects best route System gives best route and description System requests decision on route
User chooses given route or alternative route	Systems starts navigation
User follows directions	System gives directions
User reaches destination	System ends navigation System tells user destination is reached System requests route to be saved
User responds positive or negative	System does or doesnot save route System requests further navigational needs
User responds with new destination	System Searches available route
User responds no further needs	Systems says goodbye

Figure 4. Concrete Use Cases

In the occasion of an obstacle blocking the route, for example an out-of-order elevator, the user is able to *alter* the *route* by using a voice command.

After reaching his or her destination, the user will be able to *save* the used *route*. When using the system again, the user can *access* this *saved route*.

(Figure 4. Concrete Use Cases) shows the system's responses to the user's action.

As shown in Figure 4, the system's first action is to *localize* the *user*. Based on the literature study, we decided the best method to do this is to use fingerprinting through Wi-Fi hotspots already available in the building. After localizing the user, the system *greets* the *user* with a voice greeting, and *gives* a *description of the present location*. This greeting is done through speech. Next the system *asks* the user for the desired *destination*.

After receiving this information, the system *searches* it's maps for the *available route(s)*. Also a *check* is done for the *users* set *preferences*. The best *route* is *selected* based on these preferences. System relates this route to the user by *giving* a *description* of the desired location and available *route*.

Here the user has a choice, he or she can either accept this route and start navigating; ask for a more detailed description of the route; or ask for another route. For this concrete use case I have chosen to describe the decision of starting navigation with the given route. After choosing to start navigation, the system *starts navigating*. This involves giving directions to user about where to go, when to take a turn, etc.

When reaching the users destination, the systems *ends navigating*. The system then *tells the user* that destination has been reached, and *asks* if the destination should be *saved*. Also the system *asks* if a *new route* should be *searched* for. If a new route is needed, the system *searches* for a new *route*. If the user states no new route is needed, the system *says goodbye*.

3.3 Content diagram

In order to show the workflow of the interaction with system, we have added the following content diagram (Figure 5. Content Diagram)[2]. As seen in this diagram, the system starts after a prompt of the user. This prompt is the launching of the app.

There are several points in the interaction where the system needs input from the user. The system will always ask for this input, in preferences it will be able to set whether this should be via text or speech. For a blind person like Mike from scenario 1, it is vital the system uses speech, for he cannot see directions given in text or images. However, for Susan from scenario 2, speech is not necessarily needed. It might be convenient in some situations, for example if her hands were busy navigating the wheelchair, but it is not vital.

[2] A larger image of Figure 5 can be found in Appendix A – 2. Content Diagram

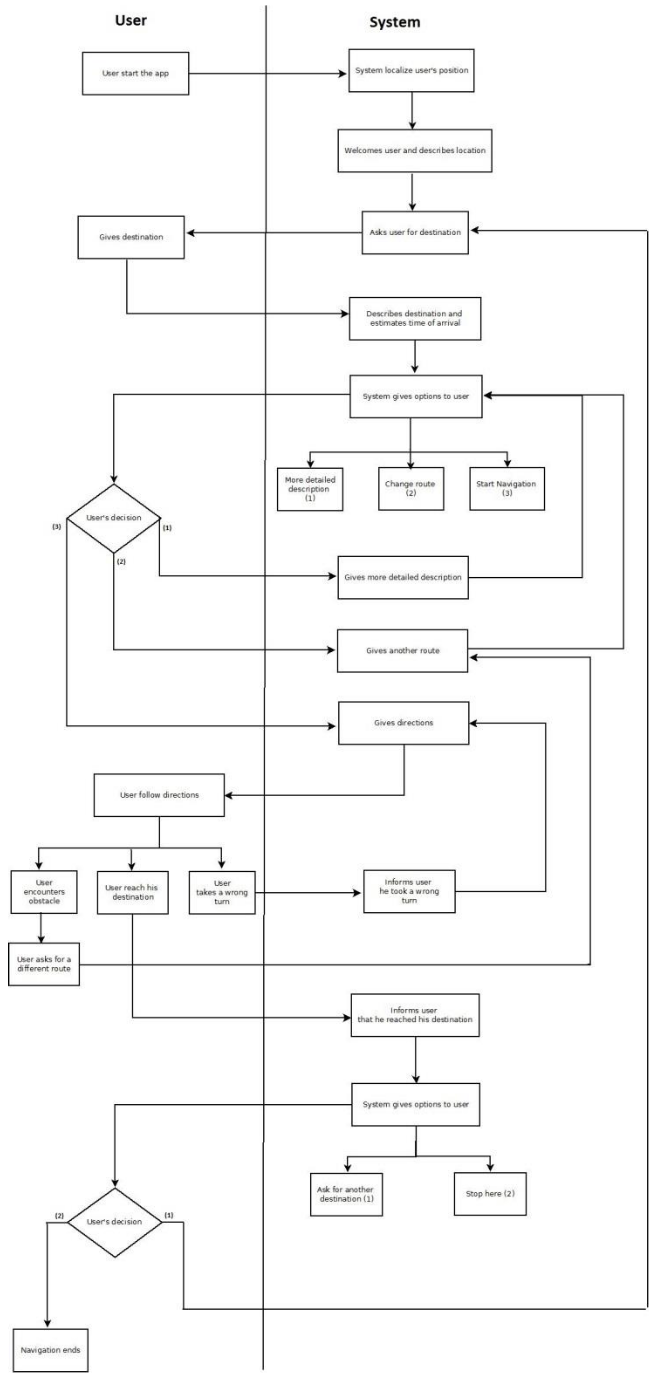


Figure 5. Workflow diagram

Also a few decision points are part of using of the system. Here also, speech is involved. Each decision prompts a different action from the system. One decision might lead the system to an earlier step, prolonging the time it takes to effectively use the system. This could be a negative point in terms of user experience. However, we feel these decision points are needed to make the overall user experience better, in effect they help the user to better reach their goal. The feeling of having influence on the system could also help better the user experience.

To provide positive feedback, the system tells the user when reaching the desired destination. The user ends the system with a prompt.

3.4 Lo-Fi Prototype

In order to show the user's interaction with the system, we have also created images of the screens the user will encounter while using the application. These are shown in Figure 6 and 7. Figure 6 shows the interaction for a visually impaired user whereas Figure 7 shows the interaction for a physically impaired, but seeing user[3].

The first screen (Figure 6 and 7. Screen 1) shows the menu screen of the mobile phone. When clicking the navigation app for the first time, the user will be directed to the welcoming screen (Figure 6 and 7. Screen 2). In this screen, the user is prompted to define his or her preference for voice guidance by speaking into the microphone of the phone. After doing this, the preferences screen opens (Figure 6 and 7. Screen 3). This screen is always available during use, but for first time use it is necessary for the user to provide these settings. The system uses these preferences to find the "best route" for the user.

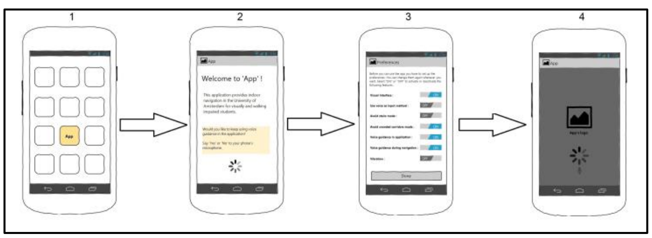


Figure 6. Interfaces for visually impaired user

The following preferences are present for the user:

- Visual interface (on/off)
- Use voice as input method (on/off)
- Avoid stairs mode (on/off)
- Avoid crowded corridors mode (on/off)
- Voice guidance in application (on/off)
- Voice guidance during navigation (on/off)
- Vibration (on/off)

If voice guidance in application is turned on, the user will get asked each preference. With his answers to these preference questions, the preferences are set. After setting one preference, the next question is asked. After the preferences are set and saved (by clicking the "Done" button or saying "Save Preferences", depending on whether or not the user opted for voice guidance in application) the user is shown to the next screen (Figure 6 and 7. Screen 4). Preferences have to be set only on first-time use, in each additional use the visually impaired students will immediately go from Figure 6. Screen 2 to Figure 6. Screen 4. Each physically impaired user will immediately go from Figure 7. Screen 2 to Figure 7. Screen 4, unless he or she has set the preferences to turn visual interface off.

For visually impaired users, the only screen available from this point is the right-most screen in Figure 6 (Figure 6. Screen 4). For the physically impaired but seeing user, further screens are available. For both users the functionalities are equal, the difference is in the operating mode. While visually impaired users receive voice guidance to guide them through the application, the seeing user will see the options available. For easier description the screens for the seeing user are explained in more detail.

If the option for voice guidance in the application is turned off, the user will only see Figure 7. Screen 4. If the user chose voice guidance on, the user will hear the system say: "Welcome! You are at …" (Figure 6 and 7. Screen 4). The available options here are Choose Destination and Saved Destinations, which will open a screen with all saved destinations. Also, the preferences screen is available from here (Figure 7. Screen 5). When choosing Saved Destinations, the user will go to the next screen (Figure 7. Screen 6). Here an overview of the previously saved destinations is shown for easy access and navigating to these locations. When choosing Choose Destination, the user will go to the next screen (Figure 7. Screen 7).

This screen shows where the user can enter his or her destination when voice guidance in application is turned off. This could be applicable for physically impaired people, not visually impaired. The user provides the system with the building he or she wants to go to, the floor, and finally the room number. This information can usually be derived from the notation of the room in the class schedule. For example, Room G3.02 is in building G, floor 3 and room 02.

Since this notation may be difficult to understand on first use, it is advised to use the app with voice guidance in application turned on. The speech recognition software in the system will automatically derive this information from the spoken "room G302".

After selecting the right destination, the user has to either click OK or Cancel. Clicking OK will lead to the next screen (Figure 7. Screen 8); while clicking Cancel clears all fields and returns the user to screen 4.

With voice guidance on, screen 4 (Figure 7. Screen 4) waits for the user to speak the destination as described above. After not recording any spoken language for a period of 3 seconds, the system will assume the destination is named, and will move on to screen 8 (Figure 7. Screen 8). If no input at all is collected, or the input cannot be understood, the system will repeat the question. In screen 8, the user will hear the system say: "Your destination is …" (description of destination) and "What do you want to do?".

[3] The final sequences of the screens for each interface, i.e. for visually impaired or for physically impaired user, are also visualized in Appendix A – 3. Interfaces.

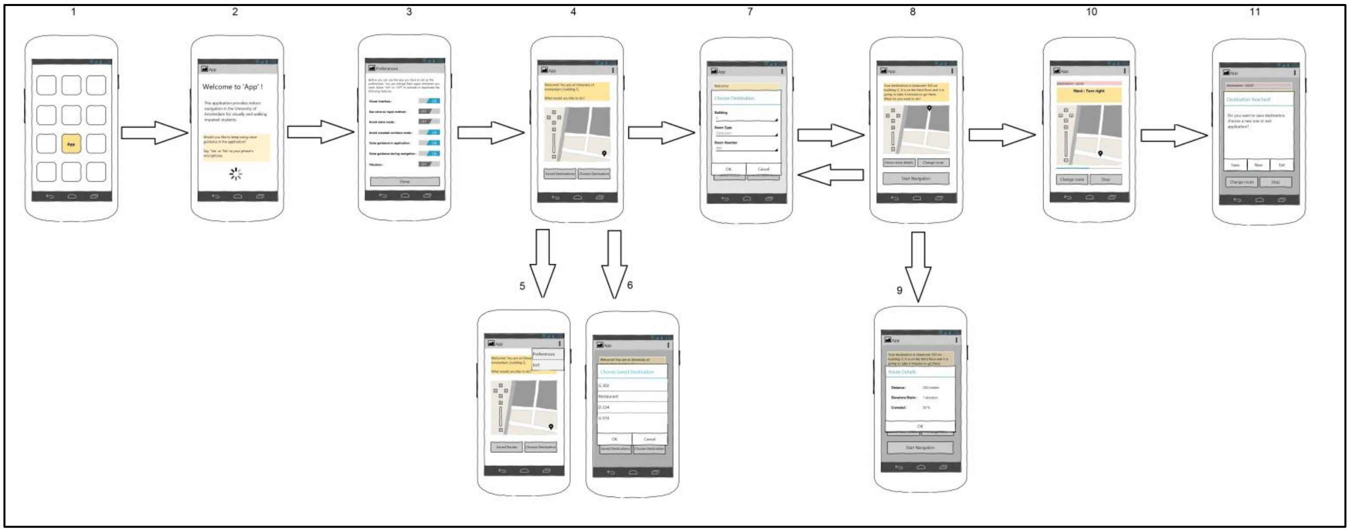


Figure 7. Interface for seeing user

Screen 8 (Figure 7. Screen 8) gives the user three options:

1. More detailed information
2. Change route
3. Start navigation

Option 1 leads to a screen with more detailed information about the route and destination (Figure 7. Screen 9). Option 2 leads back to screen 8, but with an alternative route to the destination available. Option 3 leads to screen 10 (Figure 7. Screen 10 and Figure 6. Screen 4). From this point, the actual navigation starts. The user will receive directions via audio or visual output, or a combination of both, depending on the set preferences.

After reaching the destination, the final screen is shown (Figure 7. Screen 11), which gives the user the option of saving the destination, choosing a new destination or exiting the application. When choosing 'Save', the destination will be saved in the list of Saved destination and is available later on from the Saves destinations (Figure 7. Screen 6). When choosing 'New', the user is guided back to the screen where a destination is chosen (Figure 7. Screen 4). When choosing 'Exit', the application will end and the user returns to his or her mobile phone menu (Figure 6 and 7. Screen 1).

4 SYSTEM DESCRIPTION

The system design we propose for *eyeWalkie* uses systems from different sources. It will need a speech recognition system to make voice commands and audio feedback possible. Maps will have to be available in the software. Also, our system needs to employ an algorithm to calculate the best route for our intended user, based on the information in the provided maps. In calculating this best route, certain specific needs need to be taken into account, for example the avoidance of large open space, the avoidance of crowded areas and the need for an elevator. A text-to-speech system is needed to provide the user with the actual (spoken) directions. To be able to support all used systems, and save user information, a server will be needed for system information storage (i.e. maps, additional information about points on the map, speech recognition, text-to-speech), while in the app there will have to be storage available for locally saved data (i.e. users preferences, saved destinations).

In Figure 8 the different processes necessary in the system are shown. Each process needs a different system to work. The image is derived from an image created by Ozdenizci et al. Their image showed in very clear way what the different stages of the system where, and could therefore be easily adapted. The following paragraphs will describe the used systems in each process in more detail.

The first step in the system is the localization of the user. In order to track the user and his or her progress in getting to the destination, localization is done using fingerprinting as shown by Navarro et al (2010). We propose to combine this method with dead reckoning as described by (Kothari, Kannan, Glasgow, & Bernardine Dias, 2012).

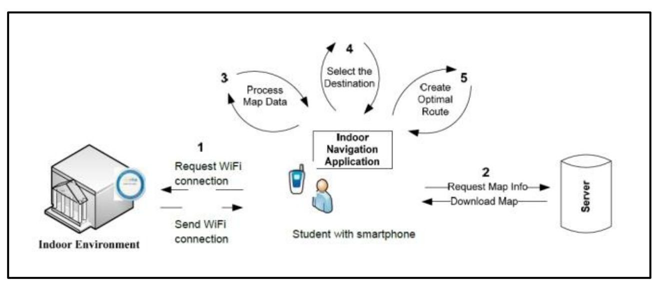


Figure 8. System overview

4.1 Localization

Fingerprinting is a method of localization using pattern recognition as discussed in our related work section. As explained by Feng et al. fingerprinting is more accurate then triangulation. For our intended user, accuracy is a high priority requirement. Figure 9 shows how the fingerprinting process is done.

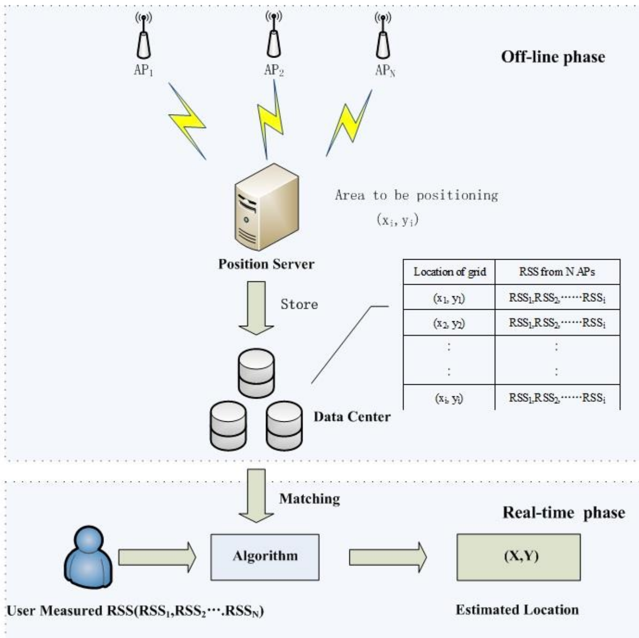


Figure 9. Fingerprinting (Figure 3 from Feng et al)

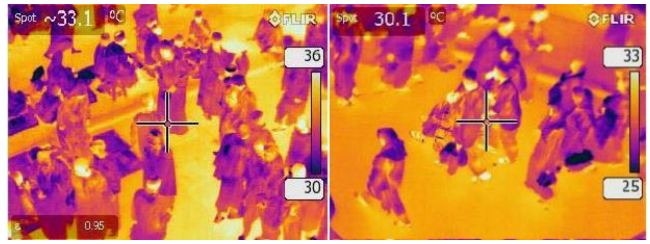
First there is an off-line phase, in which a location fingerprint database is constructed. Each entry in the database is a mapping between a position and a location fingerprint (e.g. doorway, wall). This off-line phase is the training phase for the system. Each map will first have to be divided into grids. Each cell in this grid contains information about the location which creates decision points (fingerprints). These decision points are then taken into consideration while planning the path to the desired destination. This is the real-time phase.

With the map and position of resources such as elevators known, a path can be calculated, using Dijkstra's algorithm for ShortestPath. Dijkstra (Dijkstra, 1959) has created this algorithm in 1959, but it is still widely used for navigational purposes. We will also use this algorithm in our system.

A query has to be done for the right map. Building maps, inertial navigation systems using accelerometers and gyroscopes already present in the mobile phone and Wi-Fi spots are used to get the location of the user and the appropriate map. The map is then downloaded.

4.2 Maps

The following step is to process the map data. This is where the layered maps will be. We propose to use thermal cameras to flag crowded areas (Abuarafah, Khozium, & AbdRabou, 2013). Thermography represents the electromagnetic radiation of an object. The amount of radiation emitted by an object increases with temperature. Therefore thermography allows us to see variation in temperature. When viewed through a thermal imaging camera warm objects stand out well against cooler backgrounds. Humans and other warm-blooded animals become easily visible against the environment day or night.



10. Thermal camera image

The system monitors and estimates the density of a crowd in real time using infrared thermal video sequences. This technique consists of three steps: (1) Analyzes video sequences captured by a thermal camera (for example FLIR Camera); (2) Calculates the occupied area of land in the captures screens; and (3) Calculates the crowd density. The camera output will be continuously analyzed in real-time to always have an accurate image of the crowded areas.

4.3 Path Planning

After the user selects the destination (Figure 8. Step 4) the collected data from the thermal cameras can be layered on the fingerprint map, and used as a parameter in the Dijkstra algorithm (Figure 8. Step 5). Also, the location of different resources can be used as parameters in the algorithm. This is done by adding weight to different nodes such as elevators, walls, doors, and other objects in the location. This weight can be considered in the algorithm. We add these different weights to the parameters because the shortest path is not always the best path.

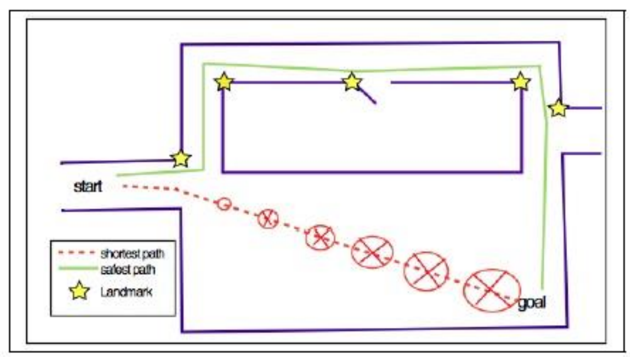


Figure 11. Shortest Path

Take for example the situation in Figure 11 (Fallah, Apostolopoulos, Bekris, & Folmer, 2012). For a blind person, it feels safer to walk next to walls, they give a sense of location and direction, both very important points for our intended user. Therefore, walking into an open space is not the safest path to a location for a blind person, even though it might be the shortest.

The algorithm will make a decision about the route each time a decision point is reached. In Figure 12, this is shown as yellow stars (landmarks). These landmarks can be corners, elevators, doors, extremely crowded areas, large open spaces, etc. This decision is then compared to a database with possible decisions, and output to the user by either text or audio. With the algorithm, the shortest and best path is calculated, where each decision point generate directions based on the decision made. This needs to be converted from text to audio, which is output to the user, if audio feedback is turned on. We will expand on this part of the system further along this chapter.

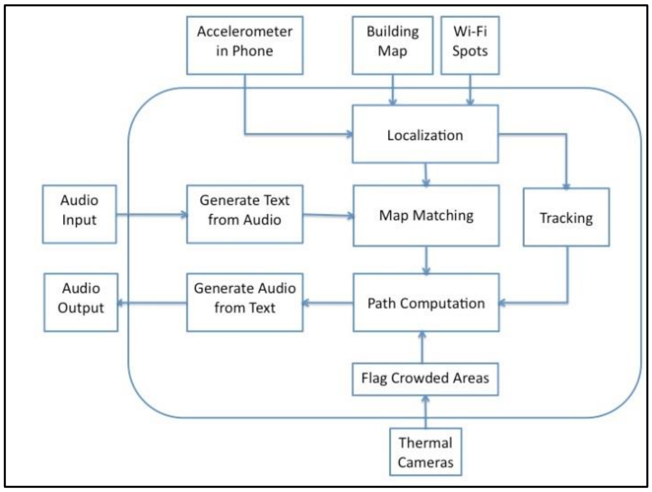


Figure 12. System input and output

Our system can be used with audio input as well as text input, therefore the box Audio input in Figure 12 can be changed for Text input. In that case, the step to generate Text from audio can be skipped.

4.4 Representation

For the audio input, there has to be voice recognition. We choose to use the open source voice engine Julius for this. Julius is a complete voice engine. The system Julius works as follows (Figure 13. Overview of Julius) (Lee & Kawahara, 2009).

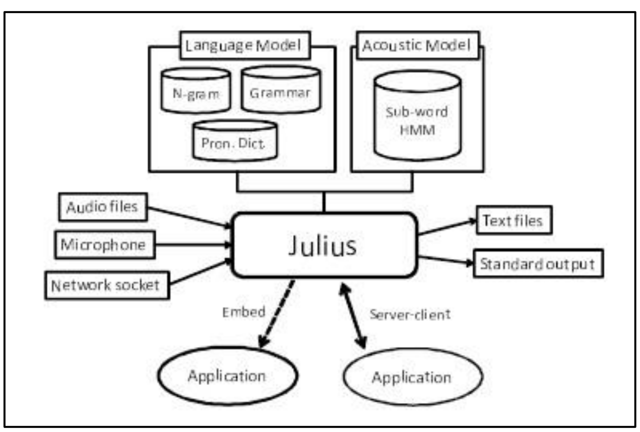


Figure 13. Overview of Julius (Figure 1 from Lee et al.)

From the voice input, text has to be generated. This comes from an audio file, which is provided from the microphone in the mobile phone as the user speaks his commands or query. Julius converts this audio input to text files using a language model and an acoustic model. The acoustic model we propose to use is VoxForge[4]. This is also open source software. To create a language model, we can use the SRI Language Modeling Toolkit (SRILM)[5]. This open source software is free to use in any non-profit purpose.

In order to have little troubles translating the user's commands to text, we will use a controlled vocabulary for the available commands. This might lead to a shorter learning curve for the speech recognition system. For example, for the system to start navigating, the user simply states "Start Navigation".

For audio output, a different type of system is needed. Here we will use a Text-to-Speech system (TTS-system). There are many TTS-systems available. A TTS-system basically transfers normal language text files to audio files by concatenating prerecorded speech from a database.

A database with prerecorded speech is needed for this step. These recordings will come from a controlled vocabulary. The TTS-system will turn the text from the directions to audio for the user (Figure 14. Text-to-Speech). The system we propose to use for this process is eSpeak[6].

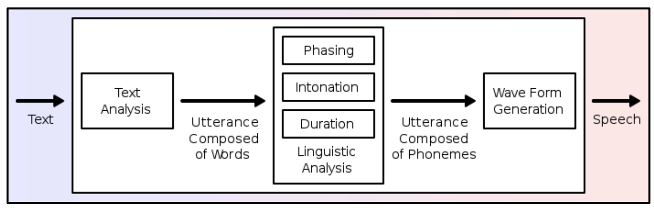


Figure 14. Text-to-Speech

For storing all necessary information to use the system, a server is needed. This server contains all the map and fingerprint information but also all speech-related information such as command list, controlled vocabulary and spoken directions. The application will need local storage as well, to store the user's preferences and destinations.

5 DISCUSSION AND CONCLUSION

We have had many different ideas of how the system's interaction should work, what system components to use, and what would be better design decisions. For example the

[4] VoxForge: open source acoustic model. Additional information can be found on http://www.voxforge.org/nl.

[5] SIRLM: open source language model. Additional information about SRILM can be found on http://www.speech.sri.com/projects/srilm/

[6] eSpeak: open source Text-to-Speech system, available in multiple languages. Additional information about eSpeak can be found on http://espeak.sourceforge.net/

choice of having visuals on screen for a blind person. Of course, the blind person cannot see this, and it should not be present "in case of" getting lost and needing help from a seeing person. However, it is a very good possibility that people will use the app while in a group. This is shown by the research done by Bouwer et al (2013). Also, the design of iDocent by Stein et al. and later Guotana et al. has interfaces for other users then the originally intended user. The question for us was if the intended user alone should dictate the entire look of the application or not, especially if this intended user is visually impaired and cannot see the interface. In our case there are two different interfaces since there will be seeing and non-seeing users.

Another issue is that we feel that there has not been done enough research on the specific requirements for our intended target group. While discussing the needs assessment done by Burigat et al (2012), the question arose if these needs are very specific for impaired students, or for any impaired person. We believe more research is needed to make this and specific requirements clear, the research done by Burigat et al is perhaps too generic for our purpose and does not lead to a specific enough requirements list.

As stated by Emiliani et al. (2005), it is important to involve user representatives into shaping the design. Involving them will ensure the system meets the specific needs of (visually) impaired students. It is easier to state requirements and make sure the application has higher usability. Since we have not done any research on the user's needs besides looking at the research already done by Burigat et al., we believe the testing phase should be the place where we ask the potential users for their experiences with and expectations of the application.

The question of whether or not the users would become too dependent on the system, and would perhaps withdraw themselves from social interaction because they can use the application for guidance is also an important one. We believe this should not be the case. Of course there will be some dependency on the system for providing directions, but this does not have to mean withdrawing from social interaction. One of our main goals is to make the user feel more confident while navigation campus grounds by him- or herself. The application would ensure they can do this without additional help, but this does not mean all other social interaction is rejected simultaneously.

Something that is not yet present in our system, but would be a good addition in our opinion, is to automatically generate altered routes when encountering broken down elevators. Elevators send out information about their state to a server. The information from this server could be implemented in our system, thus avoiding broken down elevators, or automatically altering a route when an elevator breaks down while already navigating towards this elevator. There would have to be audio feedback explaining the sudden change of course with regard to the directions. At the moment the system is designed as such that the user has to ask for a new route when encountering obstacles.

Another addition could be software for sign language. This would not per se heighten the ambient intelligence of the system, but would add to the accessibility. This software is available, as KPN, a large Dutch telecom party, will implement this Total Conversation concept[7] in 2013[8]. By making the application usable via sign language, a whole new group of users could be reached; i.e. hard hearing and deaf users. Although one could question the actual benefits of this, since most deaf or hard hearing users can see, and could thus use the application as is.

Choosing the right localization technique for our purpose has proven to be very difficult. We have altered our ideas from using dead reckoning and triangulation, to the more accurate fingerprinting in combination with dead reckoning. We eventually based our decision on which technique to use on findings in the research of Fallah et al (2012) and more detailed explanations about fingerprinting from Kothari et al (2012).

Adding the thermal cameras to our system might seem like an ad hoc decision, since we have been focusing mainly on "cheaper" alternatives for our system. While thermal cameras can be very expensive, we believe that for our intended user group, they can provide essential information in order to create a 'best' route. These user specific needs, are very important in creating a system to make sure that it will actually be used. Since crowded areas in university buildings are mostly dependent on peak hours, for example between classes or lunch hour, the thermal cameras can be thus programmed that the use is optimized.

While conceptualizing our system, we have not paid much attention to constraints such as costs. We have tried to look for open source systems to use in our system, where possible, but further than that, we have not. Of course, when actually building a system, this can be an important and often deciding factor in designing. The used techniques and implemented systems in our proposed system are decided upon based on functionality, and whether or not it fit our needs. When adding cost as a factor, these decisions might need to be looked at again, and perhaps made anew.

All the objects we want to assign weight to for our path planning module, need to be saved in our system. For fingerprinting, large storage is needed. Users can save their visited destinations, this information needs to be stored. Spoken commands need to be stored for giving the route description and all text available in the system. All this

[7] http://www.itu.int/en/ITU-T/studygroups/com16/accessibility/Pages/conversation.aspx

[8] http://www.dovennieuws.nl/index.php/nieuws/binnenland/1001-kpn-brengt-total-conversation-dienst-voor-doven-in-h2 (Dutch article)

needed storage facility can be a constraint for our system. When collecting data from large databases, the system could get slower than intended. Also the more storage is needed, the larger the application gets. The solution we proposed for this is implementing a server on which the data is saved, and then collect it from there, instead of using only locally saved data. Since certain information is dependent on the specific user, this information needs to be stored locally; i.e. saved destinations and preferences.

For our final project presentation[9] we have built an interface mock-up, or demo, which showed the interfaces a seeing user as well as a non-seeing person would encounter while using *eyeWalkie*. While our feedback came only from people with normal to good vision, the feedback was useful. Remarks were made about the volume of the spoken guidance in the application, which was fairly low for a crowded and noisy area. It would perhaps take too much concentration to focus on the spoken guidance, thus taking away the user's ability to focus on other things. The in-phone microphone picks up these background noises, therefore the speech recognition software needs to consider these as well.

In response to the demo it was stated that going through the preferences screens every time the application is used is not very user-friendly. This remark could easily be countered by explaining that the preferences have to be set only one time.

5.1 Conclusion and Future Work

Our proposed system has the potential to be very useful for the intended user. The techniques and systems discussed in relation with the system are, in our opinion, the best choices for the system.

As shown above, our current approach leaves room for future work. Since we only conceptualized our system, this concept can be used by others to actually build the system. A prototype should be made to experiment with in for example the University of Amsterdam, using the intended user group as test subjects. A needs assessment after this testing phase should render any additional needs which can then be implemented into the final system design. Even in the showing of our interface mock-up, several remarks were made that need to be taken into consideration in future prototype(s). After the initial testing with a prototype in University of Amsterdam, the system could be extended to other universities.

More ambient intelligence could be implemented in the system; i.e. using information available from machines that are present in the environment, such as the state of elevators, coffee machines, Xerox machines etc. The possibilities here should be researched and if possible implemented in future versions of the system.

[9] Poster from the final presentation can be found in Appendix B

Finally, the addition of sign language software to broaden the accessibility of the system is worth investigating.

In short, future work could consist of the following points:

- Build and test prototype
- Add any eventual requirements after the test phase
- Add more ambience intelligent options
- Extend the accessibility
- Enrol in bigger or other environments

6 ACKNOWLEDGMENTS

My appreciation goes to Amna Alhitmi, Konstantinos Chronopoulos, Giorgos Lilikalis and Lily Martinez Ugaz, fellow students from Vrije Universiteit Amsterdam and Universiteit van Amsterdam, for working with me on this project, which was part of the course Intelligent Interactive Systems in the Master course Information Sciences. Finally to Frank Nack, professor for this course, for providing us with useful feedback throughout the project.

7 REFERENCES

McDonald's - Happy Table / Table game - case study (2013). (2013, 08 20). Recalled on 09 08, 2013, from Adland.tv: http://adland.tv/commercials/mcdonalds-happy-table-table-game-case-study-2013#sL8XuJQkW1BW0bD3.03

Abuarafah, A. G., Khozium, M. O., & AbdRabou, E. (2013). Real-time Crowd Monitoring using Infrared Thermal Video Sequences. *Journal of American Science*, 133-140.

Alcañiz, M., & Rey, B. (2005). *New Technologies for Ambient Intelligence.* Recalled on 09 12, 2013, from Neuro VR - Immersive Virtual Telepresence in Health Care: http://www.neurovr.org/emerging/book5/01_AMI_Alcaniz.pdf

Bouwer, A., Visser, A., Nack, F., & Terwijn, B. (2013). Location Awareness, Orientation and Navigation: Lessons Learned from the SmartInside Project. *Proceedings of the Third International Workshop on Location Awareness for Mixed and Dual Reality*, (pp. 1-4). Santa Monica, California, USA.

Burigat, S., & Chittaro, L. (2012). *Mobile Navigation and Information Services for Disabled Students in University Buildings: a Needs Assessment Investigation.* Udine, Italy: Workshop Accessibility MOBILE HCI 2012.

Dijkstra, E. (1959). A Note on Two Problems in Connexion with Graphs. *Numerische Mathematik*, no 1, p. 269-271.

Edison, T. (2012, 04 10). *How Blind People Use Twitter & You Tube on the iPhone 4S.* Recalled on 10 01, 2013, from YouTube: http://youtu.be/c0nvdiRdehw

Emiliani, P. L., & Stephanidis, C. (2005). Universal access to ambient intelligence environments: Opportunities and challenges for people with disabilities. *IBM Systems Journal*, volume 44, p. 605-619.

Evennou, F., & Marx, F. (2006). Advanced Integration of WiFi and Inertial Navigation Systems for Indoor Mobile

Positioning. *EURASIP Journal on Applied Signal Processing*, 1-11.

Fallah, N., Apostolopoulos, I., Bekris, K., & Folmer, E. (2012). Indoor Human Navigation Systems - s Survey. *Preprint submitted to Interacting with Computers*.

Feng, C., Valaee, S., Wain Sy Au, A., Reyes, S., Sorour, S., Markowitz, S. N., et al. (2012). Anonymous Indoor Navigation System on Handheld Mobile Devices for Visually Impaired. *Int J Wireless Inf Networks*, 19:352–367.

Feng, J., & Liu, Y. (2012). Wifi-based Indoor Navigation with Mobile GIS and Speech Recognition. *International Journal of Computer Science Issues*, vol. 9, Issue 6, No. 2, p. 256-263.

Guerrero, L. A., Vasquez, F., & Ochoa, S. (2012). An Indoor Navigation System for the Visually Impaired. *Sensors*, no 12, p. 8236-8258.

Guotana, K., Kim, V., Partlo, A., & Thiagarajasubramanian, S. (2011). *iDocent (Indoor Digital Orientation Communication and Enabling Navigation Technology) (Phase 2).* Michigan State University.

Helal, A. (., Moore, S. E., & Ramachandran, B. (2004). *Drishti: An Integrated Navigation System for Visually Impaired and Disabled.* Gainesville, Florida.

Kartakis, S., Sakkalis, V., Tourlakis, P., Zacharioudakis, G., & Stephanidis, C. (2012). Enhancing Health Care Delivery through. *Sensors*, no 12, p. 11435-11450.

Kothari, N., Kannan, B., Glasgow, E. D., & Bernardine Dias, M. (2012). Robust Indoor Localization on a Commercial Smart Phone. *Procedia Computer Science*, vol. 10, p. 1114-1120.

Lee, A., & Kawahara, T. (2009). Recent Development of Open-Source Speech Recognition Engine Julius. *Proceedings : APSIPA ASC 2009 : Asia-Pacific Signal and Information Processing Association* (pp. p. 130-137). Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP.

Liu, J., Chen, R., Pei, L., Guinness, R., & Kuusniemi, H. (2012). A Hybrid Smartphone Indoor Positioning Solution for Mobile LBS. *Sensors*, no 12.

Navarro, E., Peuker, B., & Quan, M. (2010). *Wi-Fi Localization Using RSSI Fingerprinting.*

Ntoa, S., Antona, M., Margetis, G., & Stephanides, C. (2013). UAHCI/HCII 2013, Part III. *Addressing Learning Disabilities in Ambient Intelligence Educational Environments* (pp. 231–240). Crete, Greece: HCI International 2013.

Ozdenizci, B., Ok, K., Coskun, V., & Aydin, M. N. (2011). Development of an Indoor Navigation System Using NFC Technology. *The Fourth International Conference on Information and Computing Science.*

Philips. (sd). *The technologies of Ambient intelligence*. Recalled in 09 2013, from Philips Research - Technologies: http://www.research.philips.com/technologies/projects/ami/breakthroughs.html

SRILM - The SRI Language Modeling Toolkit. (sd). Recalled on 10 04, 2013, from SRI International: http://www-speech.sri.com/projects/srilm/

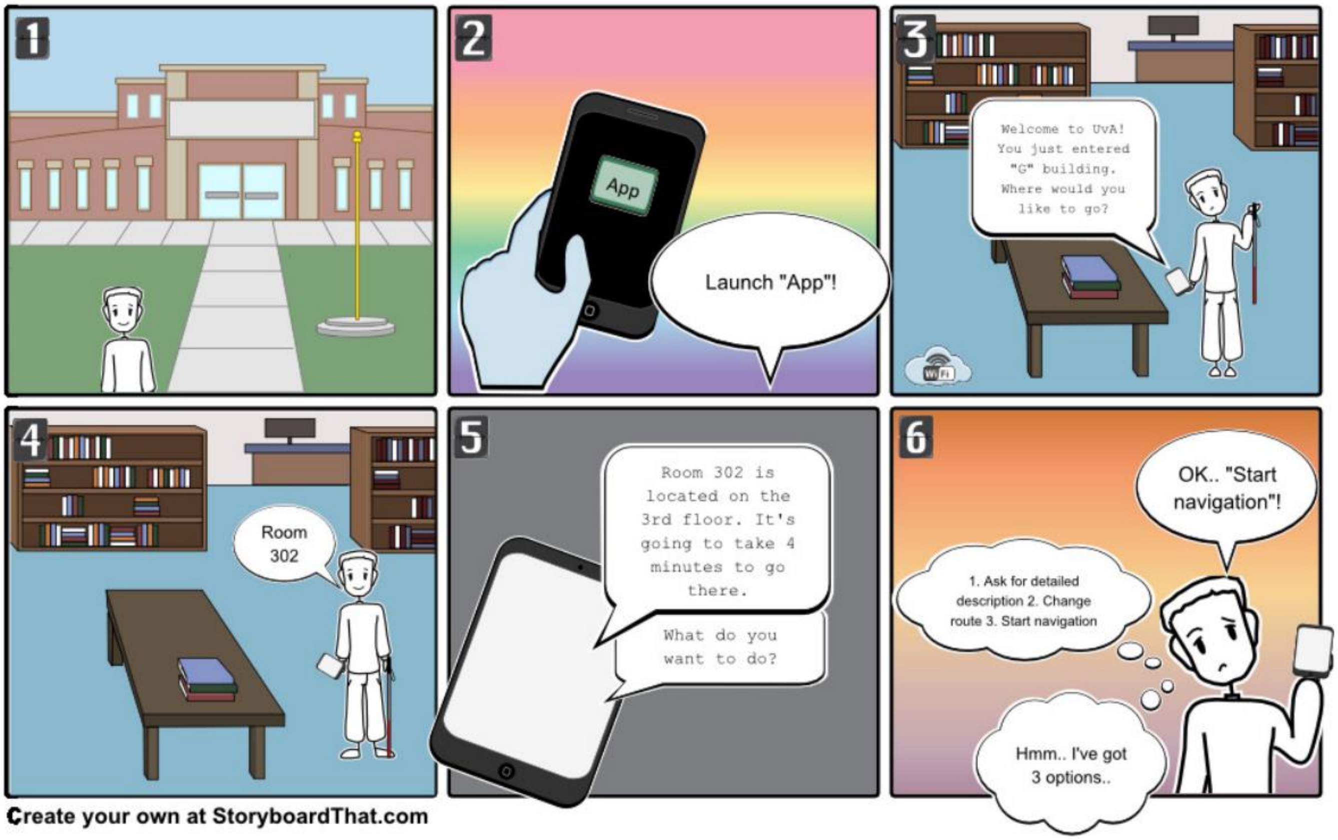
Stein, G., Gottshall, M., Donofrio, J., & Kling, A. (2011). *iDocent : Indoor Digital Orientation Communication and Enabling Navigational Technology.* Michigan State University.

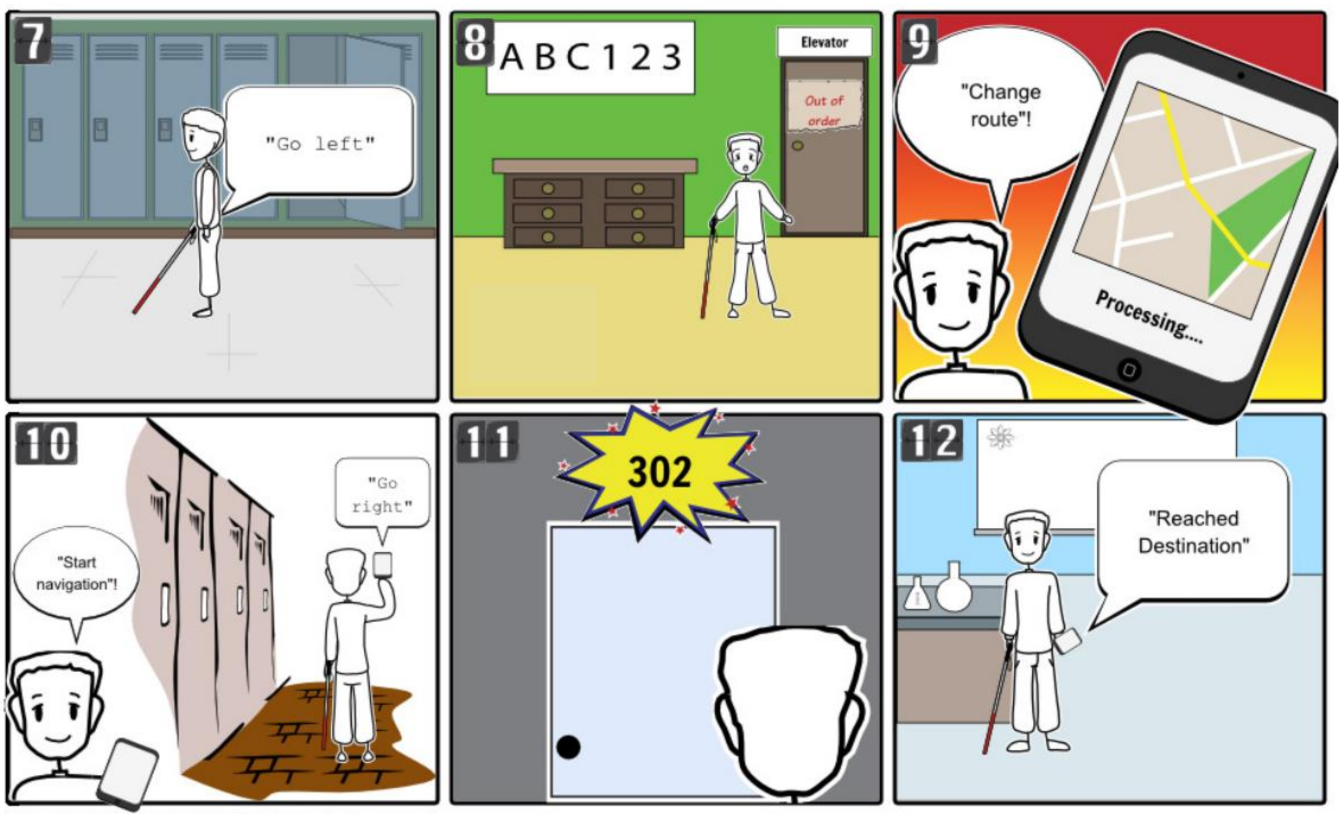
Visser, A. (2010-2012). *Smartphones with Innovative Navigation Service for a unique InDoor Experience.* Recalled on 09 26, 2013, from Faculty of Science - UvA: http://www.science.uva.nl/~arnoud/research/SmartInside/

VoxForge. (sd). Recalled on 09 30, 2013, from VoxForge: http://www.voxforge.org/nl

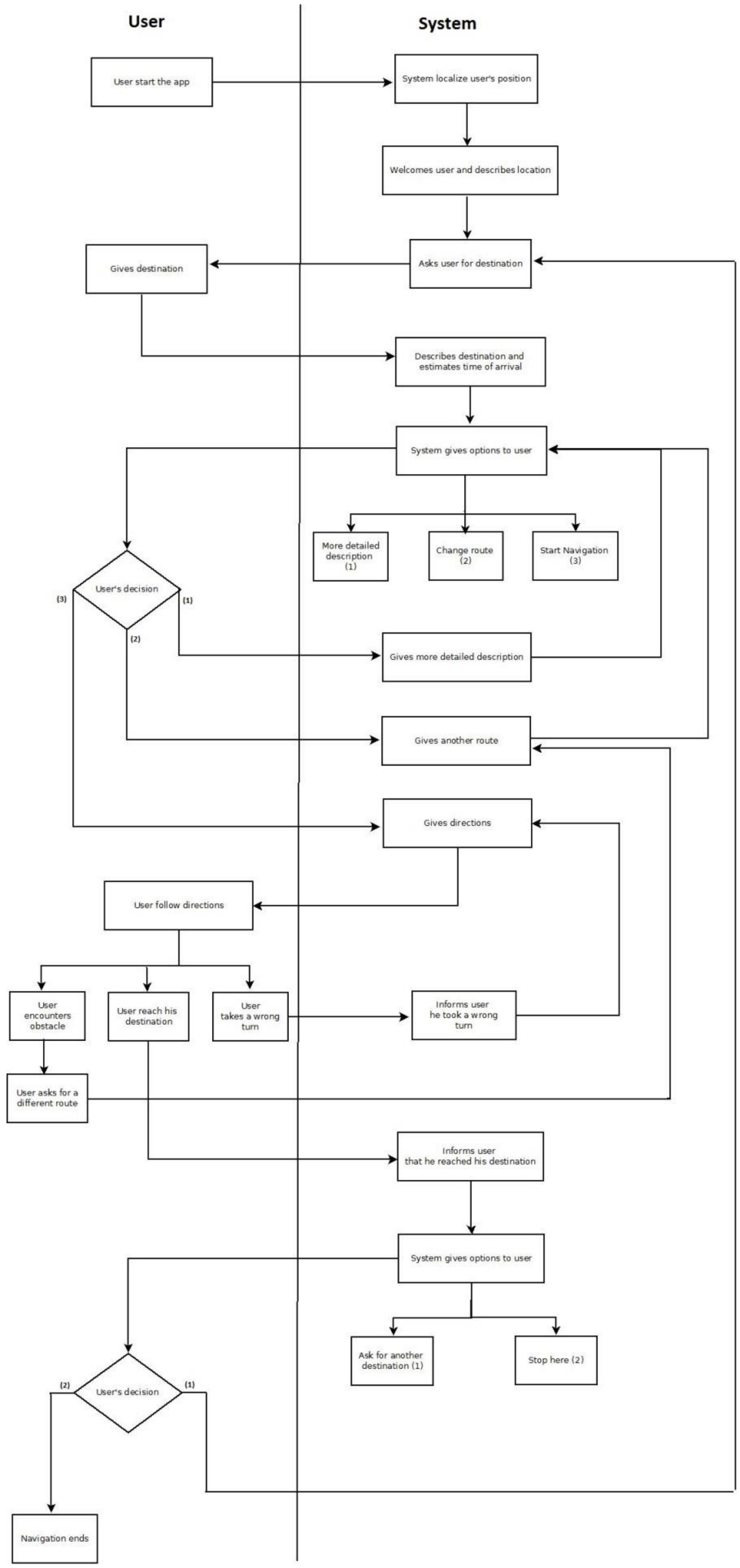
APPENDIX A – INTERACTION DESIGN

1. **Storyboard**

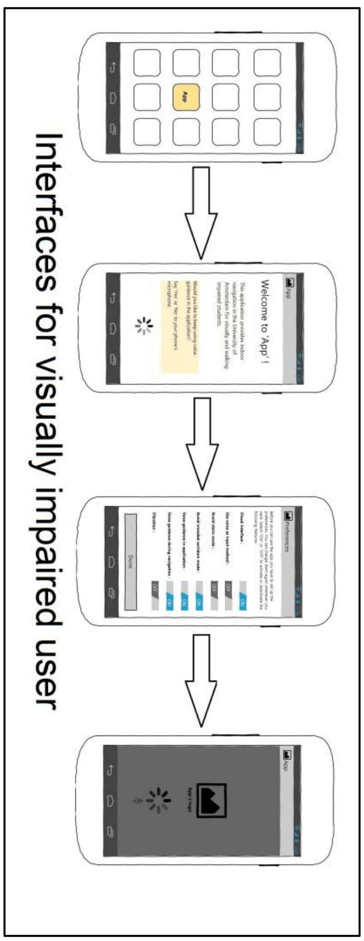


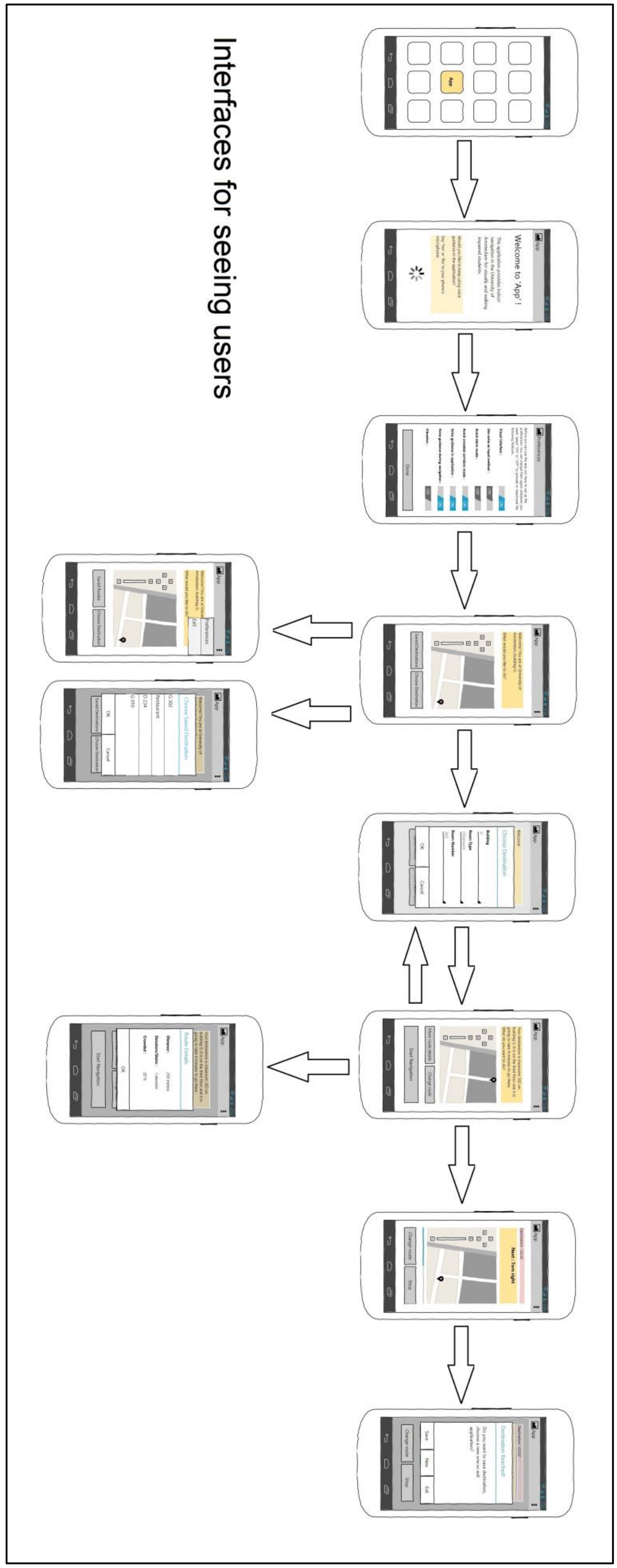


2. Content Diagram



3. Interfaces





APPENDIX B – POSTER FOR POSTERPRESENTATION

eyeWalkie Indoor Navigation System for Impaired Students in University Buildings

Amna al Hitmi, Annemarie Collijn, Georgios Lilikakis, Lily Martinez Ugaz, Kontantinos Chronopoulos

1. Introduction

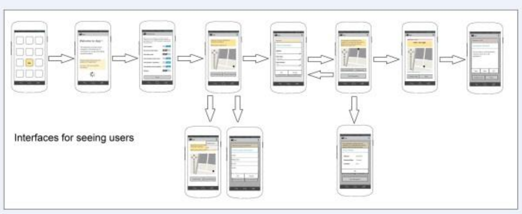
Navigating in university buildings for physically or visually impaired students is a daily challenge, therefore we designed eyeWalkie an indoor navigation system for these type of students.

2. Interaction design

EyeWalkie caters to the specific needs of the physically and visually impaired primarily inside university buildings. Given the fact that a large population of impaired individuals own smart phones, eyeWalkie will use smart phone as the medium of interaction with the users.

Two different interfaces will be available; one for visually impaired and one for non-visually impaired users.

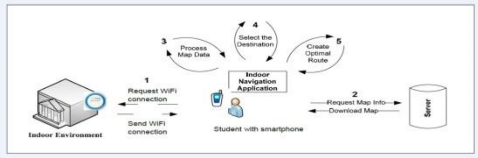
Interfaces for visually impaired user



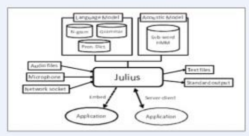
Interfaces for seeing users

3. System Design

The system will take audio input from and produce audio output for the visually impaired user. The data collected from sensors in the smart phone and different Wi-Fi spots will be used to locate the user (fingerprinting technique) and to track his movements. In addition to locating and tracking the user, ambient intelligence will be used to suggest safe and convenient routes to the user according to his needs and preferences.

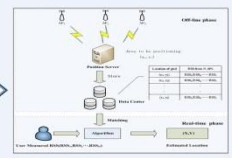


Audio to text



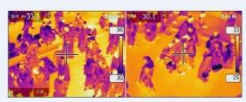
We choose to use the open source voice engine Julius to convert the audio to text and finally to take the destination string to the search route query.

Localization of the user



The localization of the user is done using fingerprinting.

Thermal cameras

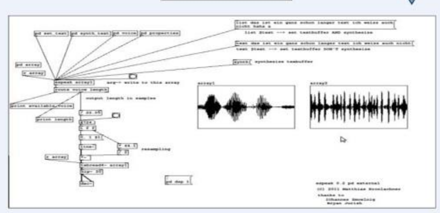


The system monitors and estimates the density of a crowd in real time using infrared thermal video sequences.

Path planning

For the calculation of the Shortest Path, Dijkstra's algorithm is used, taking as a parameter the collected data from the thermal cameras as well as the location of different resources.

Navigation of the user



This is the time that the actual navigation of the user starts.

4. Conclusion

Clear system design for the visually and walking impaired students who are attending the University of Amsterdam, but still in need of further development and testing.

Future work proposals for the system:

- Experimenting with a prototype in the UvA
- Needs assessment investigation after test phases
- Extend the system to other universities aside from the UvA
- Cheaper alternative in relation to the use of thermal cameras

VU VRIJE UNIVERSITEIT AMSTERDAM

UNIVERSITEIT VAN AMSTERDAM

Information Studies: Human Centered Multimedia | Intelligent Interactive Systems 2013/2014 | October 18th, 2013