Do you see it yet?: Guiding Users to Augmented Projections via a Feedback System

Jessie Lai
HCI-E MSc Final Project Report 2016
UCL Interaction Centre, University College London
Supervisor: Simon Julier

ABSTRACT

Usability is a seldom researched area in Augmented Reality (AR). Dunser et al found that only 10% of research on AR mentioned some type of user evaluation **[5]**. The present study builds on what we do know about usability in AR, and expands our knowledge in how people interact with image recognition-based AR systems. A simple AR application with interesting content based on buildings within the University College London (UCL) campus is created. An exploratory user study using the app found that the biggest usability challenge users face in using this type of app is the difficulty in finding virtual projections and the lack of feedback the system provides to users when they can't find the projections. Based on this finding, a feedback system that gives users directional instructions that guide them to optimal locations for stable projections was created. Results from a usability test comparing an AR app with and without the feedback system showed that the feedback system helped users find projections significantly faster, and users scored the app with the feedback system as significantly more usable. However, the feedback system failed to guide users to the optimal location to fix unstable projections. Reasons for the failure, implications of the results and suggestion for future research are discussed.

Author Keywords
Augmented reality; mobile augmented reality, feedback system

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

MSc Contribution Type
Interaction Techniques

MSC HCI-E FINAL PROJECT REPORT
Project report submitted in part fulfilment of the requirements for the degree of Master of Science (Human-Computer Interaction with Ergonomics) in the Faculty of Brain Sciences, University College London, 2016.

1. INTRODUCTION

Augmented reality (AR) generally refers to the combination of real and digital information in a way that appears to users as if they are in one cohesive environment [8]. AR falls within the continuum of mixed reality, where the physical and virtual world mix, and real and virtual objects interact and complement one another [14]. AR is usually understood as augmenting the physical world around us with computer-generated information or objects, as opposed to putting real-world information or objects into the virtual world [1].

There are a few ways to display AR, such as with head-mounted displays like Microsoft HoloLens[1] and Google Glass[2]. However, as these devices are not available to the average consumer, nor are they widely used due to their limited availability, this paper will focuses on mobile augmented reality (MAR) instead. The technology in our mobile devices such as smartphones and touchpads are continuously becoming more sophisticated and more advanced. With the technology necessary for AR, such as a camera, sensors, display screen, fast internet and GPS etc., already in most of these devices, plus them being relatively cheap and easily obtainable, they become the ideal way for a large portion of consumers to get their hands on AR technology [17]. In fact, Wither et al. expects that in the next few years, use of consumer-level AR applications will dramatically increase [27].

Despite its leap and bounds in technological advances and its increasing adoption in the consumer world, not a lot is known about the user experience in AR [5,22], especially how users interact with current MAR. A high number of research has gone into developing technology within AR, such as algorithms that identify and track physical objects in the real world so that the merging between real and virtual worlds can be more seamless (e.g. [1,25]), or the variety of ways to display augmented information and objects (e.g. [14,21]). As these researches' focus is on the technology itself, little attention is often paid to the technology's usability and convenience [5,13,15].

While AR systems are created and published frequently, some never mention testing or user evaluation (e.g. [9]),

[1] www.microsoft.com/microsoft-hololens/en-us
[2] www.google.com/glass/start/

and some are tested by the creators themselves [6,19,23] or by their colleagues [20]. Swan and Gabbard conducted a study compiling all user-based studies in the field of AR till 2005 (266 studies in total) [22], and found that only 14.3% of them had elements of user-centred design, and only 7.9% studies conducted experiments with laymen users. Building upon Swan and Gabbard's research, Dunser et al. conducted a survey on evaluation techniques used in AR studies published between 1992 and 2007 [5]. Their results showed that only 10% of the publications involved some type of user evaluation, and for those who did, most of them were looking into issues regarding perception and cognition, or users' performance on specific set tasks.

Since researchers are using people with a certain level of knowledge in AR to test the systems, it is no surprise that difficulty in handling the system is rarely mentioned. For example, in Reitmayr and Schmalsteig's paper where a collaborative AR tool for outdoor navigation and information browsing is made [20], they only mentioned the lack of GPS positioning accuracy due to the poor signal leading to jumps in the augmented projections and controls not being tailored for the task as possible usability issues. But they did not go further into describing any of these issues or how the participants tried to deal with it.

These types of testing are usually done to make sure the technology they created works on a computing level, and they do not test whether an average user will be able to understand and handle such a system. There is no knowing how vast the knowledge gap is between these researchers and their colleagues, and a layperson who may not have ever tried AR or know the slightest thing about it. It is very possible that there are nuanced things that the researchers may think of as a given in terms of using the technology, but are completely foreign to a layperson.

As seen from the brief review of studies done on AR, examination of the average user's experience with the technology is extremely rare. This is perhaps because a lot of researchers believe AR provides a very 'intuitive' way of browsing information that are location or object based [20], even though that belief may not be grounded in exhaustive research, and it may not be as intuitive to a layperson as they believe it is.

To our knowledge, it is very difficult to find a paper that has focused on the topic of whether people know how to use an image recognition-based AR system. Image recognition-based AR systems use physical things in the real world as part of its interface, using them as trigger points to elicit information or object to pop up in virtual space. This type of AR is in contrast to those that mostly rely on GPS and sensors within the mobile device to elicit the corresponding location-based information.

Users of consumer-level AR are not likely to be technical, and with AR being such a new and advanced technology, there might be big discrepancies between how people think they should use the technology, and how it should actually be used. This may be especially true with outdoor AR that doesn't rely on any fiduciary markers or QR codes (i.e. a markerless approach), since there is less affordance to suggest where the user should aim their camera at in order to trigger projections.

As previously mentioned, there are hardly any papers that have explored in detail how people interact with image recognition-based AR systems specifically and what challenges they face using it. However, this knowledge is very important as AR technology is still in its infancy and consumer adoption of it is only just beginning to pick up. Bad usability can very well obstruct the wider adoption of the technology. In fact, Olsson and Salo stated that frustration caused by poor usability is expected to be one of the main reasons why users stop using AR applications [17]. Therefore, the present study aims to examine laypeople's interaction with AR and their understanding of the technology.

In the next section, the rare few papers that focus on user experience in AR will be briefly reviewed, and based on what they have reported, it gives us some understanding of what might be the possible usability challenges that users face when using AR technology, and more specifically, image recognition-based ones.

Then an exploratory user study is conducted to confirm those challenges, as well as to find other user requirements not yet discovered. After analysing the results from the user study, a main usability issue is concluded and a solution is designed to overcome it. Lastly, the solution is tested and evaluated.

2. LITERATURE REVIEW

As mentioned previously, user experience research in AR is very rare. In this section, the few papers that did focus on it are reviewed and their mentions of any usability challenges or requirements will be highlighted.

Olsson et al. conducted a study aimed at discovering expected user experiences and main user requirements with MAR, in the context of a visit to the mall [16]. Their angle of user experience mostly focused on emotional and hedonic elements, basing their approach on Desmet and Hekket, and Hassenzahl and Ullrich [4,7], who emphasised the importance of positive emotions in user experience.

They gave participants, who are in a shopping mall, a plastic phone-sized container that represents an "omnipotent" futuristic MAR device and asked them questions such as what addition information they felt were needed in that particular environment they were in, creation and interaction with AR information, environmental elements that they think would affect their use of MAR, perceived pros and cons of using MAR etc.

As the study did not use an actual AR device and the participants did not actually experience MAR, the conclusions drawn about user requirements on interaction with the technology are very limited. In fact, the researchers admitted that interactions were rarely discussed by the participants, as they were not provided with concrete interactions with the technology. The only user requirements the study concluded in regards to users' interaction were (1) user's ability to access content easily and immediately, and be allowed to interact with it flexibly, as well as (2) cues about available augmented information has to be subtle and sensitive to the context. Due to the lack of realistic interaction with MAR, user requirements that are perhaps more relevant to the AR technology currently available to us were not mentioned.

Olsson and Salo conducted an online user survey on MAR apps available in the market in spring 2011, to find out more about users' adoption and acceptance of the technology [17]. These consumer AR apps include Layar[3], Wikitude[4], Junaio[5] and Google Goggles[6].

They examined topics such as how often the apps were used, the reason for adoption, and the acceptance of the technology. The user interface and interaction with application were mentioned as strengths by respondents, but more people mentioned technical and functional problems, such as inaccuracies in locating, failings in hardware and functionality, unreliability of software resulting in bugs and crashes. The user interface and interactions were also presented as a problem more often than not, with respondents mostly having problems with 'bad usability in general' or 'lack of UI feedback'. Although such issues were highlighted, the paper did not delve further into the details of the issues, such as what particular aspects of the application were causing trouble for participants and frustrating them.

Olsson and Salo published another paper based on the same online survey, but this time focused more on the satisfying and unsatisfying experiences of using the AR apps [18]. The paper detailed what activities people were using AR apps for, their context of use, their feelings towards the AR apps that they've used. They still did not mention much about the actual usability aspect of it. But they did note that unsatisfying experiences are mostly resulted from the app not performing adequately and it not meeting users' expectations. For example, respondents saying that Google Goggles did not recognise anything from a well-known scenery or an items they were trying to capture. Participants expressed frustration, strong irritation or disappointment when they were faced with functionality problems or other limitations. These issues include feeling like they had done things in vain, poor overly or mixing of virtual information and physical objects, the app not being usable or efficient enough in what it was advertised to do.

A little bit more light is shed on users' interaction with a MAR system in Walther-Frank and Malaka's study, where a mobile pedestrian navigation system prototype is made [24]. However, their prototype did not use a real-time video stream, as normal AR system have. Instead, a simplified AR experience is provided using visually augmented photographs. A set of panorama photos at waypoints along the route are put into the system. AR overlays were created and composited in a 3D software. The prototype receives and parses GPS sensor data, constantly checks for waypoint proximity, displays augmented overlays etc.

The aim of their experiment was to compare the usability of the AR interface and a map-based interface in a field evaluation. Researchers observed users' behaviour, recorded their satisfaction in a standard test, as well as interviewed them on their experience after the task. Results from their observation indicated that low signal quality is a cause for users' irritation in both conditions. More importantly, users in the map-based system were more able to judge signal quality because the position icon would move in an irregular way. In the image-based system on the other hand, users can only judge signal quality based on the absence of waypoint instructions. The paper noted that when no new instructions came up due to low signal quality, or when participants misinterpret instructions, they are often left stranded and unsure what to do.

In addition, because the image-based system only covered the waypoints with pre-recorded photographs, the system's functionality outside of them was emulated instead. So when participants wandered too far away from the pre-set route, a researcher would point them back to the closest waypoint. Perhaps because of this and the occasional absence of instructions due to poor signal, incidents of participants deviating from the route occurred more frequently and with more users in the image-based system than in the map-based system. In the image-based system, the researcher had to intervene 6 times, while all other instances of it (including the ones in the map-based system) were resolved by the participants themselves.

Lastly, Ko et al. conducted a very detailed usability study on MAR applications that are mainly based on GPS technology [11]. They compiled a list of existing usability principles from a few studies that were not strictly specific to AR and developed their own 22 heuristics for evaluating MAR. A lot of the heuristics focused on information display, and content and navigation hierarchy, but a number of them address usability problems mentioned in other studies mentioned in this section. These include 'error management', so when errors such as poor GPS signal, jumps in augmented projects, or bugs and crashes occur, users will be provided with ways to prevent or resolve them. Users experiencing those issues can also benefit from heuristics such as 'feedback', where they will be able to perceive the system's status and sequence of processing.

They then used those heuristics to conduct heuristic evaluation on three newly launched, popular AR apps from Korean app stores that provide users with various information based on their locational information. Six top common problems were concluded, and two representative problems were chosen to be improved upon in their prototype. The two problems were duplicated results were being shown on the screen and too little information is provided during search.

A usability test is conducted to validate results from their heuristic evaluation. They compared their prototype to two other consumer AR apps in two task scenarios. Results showed that users were more satisfied with the prototype than the two other apps, and it took users less time to

[3] www.layar.com

[4] www.wikitude.org

[5] www.junaio.com

[6] www.google.com/mobile/goggles

complete the tasks using the prototype than when they used other apps.

3. THE PRESENT STUDY

Judging from the usability issues raised in the studies mentioned above, users are often frustrated when the AR system is not performing optimally or not performing according to their expectations, either because there is poor GPS or internet signal, the system not seemingly able to register scenery or objects, poor overlay of augmented information etc. And this reflects a few usability heuristics Ko et al. created for MAR [11].

None of the studies mentioned above have a way of helping users identify and fix problems when the system fails, or when users aren't using the system correctly or are misinterpreting the system's instructions. Even in Ko's study, despite highlighting the importance of error management and feedback, the prototype they built focused on improving other aspects of MAR instead [11]. And as seen from Walther-Frank and Malaka's study, the way they had to help a user when they were truly lost was to let the researcher intervene and guide them back [24]. Consumers of regular AR apps available in the market are unlikely to get such assistance, therefore there is a need for a system that helps users identity and fix problems that they face while using MAR.

This is particularly important for AR because the technology is not very sophisticated and stable yet, and lots of errors can occur. Additionally, laypeople generally aren't very likely to understand the technology behind it, and since its interaction is so different from other technologies that people are more used to, it is possible that people will not know how to interact with it outside of the basic instructions given by an AR application.

Despite the studies mentioned and their contribution to our understanding of users' interaction with AR, a detailed user study on people's interaction with the technology is still needed, specifically with image recognition-based AR systems. Ko et al.'s usability heuristics for MAR can be used as a guide but cannot be solely relied upon in this study for a few reasons [11]. One is because their study focused on AR apps that utilise basic information about a user's location, while this study focuses on image recognition-based AR. The two different types of AR may suffer from different types of usability problems. For example, the two representative problems they chose to improve upon in their prototype are less likely to be an issue in image recognition-based AR. Because location-based AR like the ones mentioned in Ko et al.'s paper usually provide users with a large amount of information about anything that the system can pick up on in the user's vicinity [11]. Whereas with image recognition-based AR, the system is likely to only display information about the object right in front of the camera.

Secondly, as the usability testing done to validate their heuristics was conducted in a laboratory environment, their results might not be completely applicable to outdoor MAR experiences. Thirdly, their usability heuristics were adapted from usability research that weren't strictly about AR, and their heuristic evaluation on the three consumer AR apps were done by usability experts. It would be useful to conduct user studies on a group of laypeople and see if the usability problems that they face fit with what has been mentioned so far.

The user study in the present study should examine laypeople's understanding of the technology and reasons for their actions when interacting with AR, in order to know how to improve their user experience by building upon their understanding or the lack thereof. The study also serves to confirm whether there is a need for a system that identifies and helps users fix problems they face during interaction, a so-called 'error recovery system'.

After the user study, an intervention system was designed to solve the biggest usability problem identified from the user study. An AR application with the intervention system was then tested by users against an AR application without it to see whether it improved the AR application's usability.

4. EXTRACTING USER NEEDS

To gather insights into people's understanding and usage of the technology, an AR app based on the UCL campus was developed.

4.1 Methods

A pilot test was carried out with 5 participants to ensure the app worked as it was intended to, and to help refine the interview questions asked so that they were better able to extract user needs. It also highlighted the difference in thought process between people with and without a technical background or experience, which helped form the participant criteria for the experiment.

The interview questions centred around participants' understanding of the technology behind the AR app, such as how the system is able to give you location specific information or why do errors occur. Further questions probed their interaction with the device, as well as challenges they faced while using it and aspects of the application they would like to see improved.

9 participants (6 female, 3 male) were recruited from the UCL's psychology participant subject pool as well as through researcher's personal connections. The mean age of the sample was 24.7 (SD = 5.3), with an age range of 21-38. Since the objective is to assess lay people's understanding and needs of the technology, people with a technological background were excluded as participants (e.g. those who studied computer science, or have worked with image/video processing software before). Participants recruited through the subject pool were given a £5 department store voucher for their participation.

A simple AR application was developed using Unity and Vuforia, and was deployed on an iPad mini 3 running iOS9. Pictures of 5 buildings along an outdoor path within the UCL campus were taken, and interesting information about them were overlaid on top of the front of the building vertically in virtual space (the snippets of information are no longer than 3 sentences). So when participants aimed the iPad camera at the 5 buildings and the image captured matched the one previously fed into the programme, the corresponding information and related pictures would 'pop-up' in virtual space (as illustrated in Fig.1).

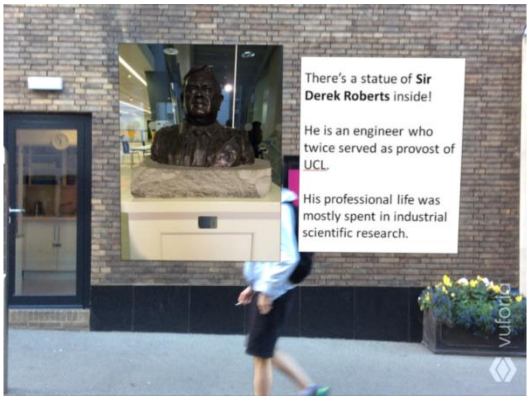

Fig. 1. An example of an AR projection in front of a building.

Each participant started with one practice trial at the beginning of the path. The researcher instructed them to lift the iPad in front of them so that the camera captures the front of the building, where a welcome message appears in virtual space. This showed participants how to use the app and what they should expect from it. After that, the researcher told them where the next target is after each stop, but otherwise no further instruction is given. The participants were told that they could move on to the next target as long as they had read all the information on the projections.

After the experiment, a semi-structured interview was conducted. Questions were mainly centred around people's understanding of the technology after having given a chance to use it, such as how the system is able to give you location specific information or why do errors occur. Further questions were about their interaction with the device, as well as challenges they faced while using it. They were also asked about their feeling towards their experience, as well as ways that they thought their experience could be improved. Participants were debriefed at the end. Each run took about 30 minutes.

For analysis, all interviews were recorded and transcribed. Thematic analysis [2] was conducted on the transcriptions so that main themes can be extracted.

4.2 Results

The interviews gave insights about people's understanding of the technology, their ways of approaching it, as well as captured what aspect of its usability they would most like to be improved.

4.2.1. Understanding of AR

System Understanding

Most people (6 participants) understood that the AR app is based on the camera recognising something in its frame. Although some understood this as the camera recognising a picture of the building's front that has previously been fed into the programme, most others thought that the camera recognizes specific features on the building, such as windows, wooden beams and signs next to building entrances. What the system actually recognises are features that are at a much lower level, such as spots, corners and cracks.

This understanding seems to come from previous experience with other similar technology that uses cameras to recognise specific things in its frame. *"It's like QR codes, when you scan them, information comes up on your screen"* (p=1). *"You know with Snapchat filters, it recognises specific things in your face. […] So for a building, they have like, windows, right? So the app is programmed to respond to that"* (p=7).

2 participants thought the app was utilising GPS. *"When you get to a point, it locates where you are, and the bubbles would come up. […] When you go to a different place and different things pop up, so it's about the location"* (P=3). It is of interesting note that all of these participants mentioned PokemonGo, a GPS-based game where gamers can choose to turn on the AR feature and see Pokemon augmented into their reality via mobile phone screens. *"When you're right next to the location, like in PokemonGo, you'd get access to that information about it"* (P=2).

1 participant considered both image recognition and GPS and could not come to a decision as to which one the app uses. 1 participant who eventually settled on it being image recognition-based also considered both during the trials. *"After I adjusted the angle to get a more stable projection and it worked better, I thought maybe it's image processing"* (p=6).

Error Understanding

The most prominent observable error in the AR system was the constant flickering of projections when the person was not standing at an optimal distance away from the building or the camera was capturing the front at a poor angle (as illustrated in Fig.2). The other was the projection flashing across the screen then disappearing again without returning even though the person maintained the same position.

4 participants mentioned that error occurred because people were walking in front of the camera. *"People walking pass buildings. Sometimes it works, sometimes it doesn't"* (p=6). Most people noted that projections flickered in time with pedestrians walking between their camera and the building, but did not mention that as an error.

Fig.2: The projection jumps to different positions on the screen as shown in the picture above. The jumping occurs so rapidly that it appears as if the projection is flickering on the screen.

4 participants suggested that they might not have been at the right distance away from the buildings. This belief is consistent with the idea that the system recognises something through the camera. *"I was too far away, so it wasn't properly able to recognise the thing, because there's too many other things in the shot"* (p=8). However, this point was also mentioned by both participants who thought the system was GPS based. *"I felt like maybe I wasn't having the door completely fit in the frame when I wanted to read the text, so it disappears"* (p=2). *"...when you get to some point, you need to fit the image in the screen, not too close or too far"* (p=3). This suggests a general confusion or misunderstanding about how GPS or image recognition technology works.

3 participants thought the angle also contributed to flickering. 3 thought it was due to the quality of the software. 1 participant thought it was an internet connection problem, and another thought it was because of the weather.

Most people (7 participants) contributed errors to multiple reasons, as shown in Table 1.

4.2.2 Interactions with AR

All participants moved the iPad around the front of the buildings (either moving to another physical spot or only moving their arms around) while trying to find the augmented objects or find a more stable projection. Different people had different reasons as to why they moved.

3 participants moved far away from the front of the building so that the camera can capture as much as it can, which they think will help the app work better. *"I started further away so I won't be missing anything"* (p=5). *"I try to get more into the picture. If the camera sees more, it'll find the thing that it responds to"* (p=7).

2 participants moved so that the camera captures what they thought were important in the frame. *"I moved to try to get the door completely in frame"* (p=2). 2 other participants did a mixture of both when they tried to move the iPad around to figure out what the application responded to, such as whether it worked based on the entire front of the building, or specific features on the building front, or GPS. *"At Institute of Making, I tried capturing the title, that didn't work so I go back and capture the front of the building, until I finally can make it work. I didn't know if it's some specific feature, or whole front of the building. I was trying to figure that out"* (p=9).

4 participants said they walked around looking for projections to appear. *"I'd wait for text to appear. If it didn't appear, I would change positon of the iPad, tried walking around the place to see if extra text would show up"* (p=1). 1 participant said they can't explain why they moved around. *"It's a little intuition"* (p=4).

2 participants mentioned tapping on the screen, and stopping once they realised it doesn't do anything. 2 participants said they saw projections flash across their screen then disappeared, so they tried recreating the moment where it happened. Both said their attempts didn't work.

	Participants								
Reason for error	1	2	3	4	5	6	7	8	9
People walking in front of camera				x		x	x	x	
Not the right distance away from buildings		x	x			x		x	
iPad was at the wrong angle		x			x				x
Poor software quality	x						x		x
Internet connection			x						
Weather				x					

Table.1: A table showing each participant's reason for why they thought error occurred while they used the system

4.2.3 Challenges and improvements wanted

Challenges

6 participants found the application frustrating to use. All of them mentioned the frustration of trying to find the projections and nothing appearing on the screen. *"If I did this by myself, like sightseeing or something, I'd just leave it and go away"* (p=3). *"It's tricky to find where the text would be on the picture. I don't know how far away I should be"* (p=5).

In relation to that, 2 participants mentioned getting a glimpse of the projection and being unable to find it again. *"Not getting anything is annoying but it's really annoying getting it for a very short amount of time and not being able to get it back. I thought I was doing what I was doing when it appeared but it never showed up again"* (p=7).

3 participants mentioned the projections flickering as a point of frustration. *"I tried aiming as best as I can but I still can't get a stabilised image. That was very frustrating. I didn't know what I was doing wrong"* (p=1).

Improvements wanted

All participants' answers revolve around wanting the system to give them instructions about what they should do. 7 of them said they want instructions to appear and help them with finding the projections. *"It should have an outline to guide you so you know which part you have to focus on. […] Tell me the correct thing to do, not just pointing your camera at the building. Tell you why you're not getting the pop up, and guide me to the correct position"* (p=7). *"It would be better if there's an indication of where the text would be and not have to guess. Like signpost to show where I can see it"* (p=5).

4 participants mentioned instructions that help them fix the flickering of the projections. *"I wish it would tell me what the problem was. Tell me to please stop moving or indicate I had to move the screen up or down"* (p=9). *"Tell people to move forward or backward to get things in frame, tell people you're not focusing on the right area. Errors can be caused by a number of issues so it's good to know which one it is"* (p=6).

4.3 Discussion

This exploratory study suggested that people are generally doing the correct actions when interacting with the system, e.g. moving around to different spots in front of the building or moving the iPad around to change the angle of the camera. Also it showed that they are doing this despite having different mental models for how they believe the AR system works.

However, even though they were all able to discover the correct ways to interact with the system, their actions did not always help them find the stable projection that they wanted. In fact, all of them expressed frustrations towards the apparent fruitlessness of their actions sometimes. The lack of feedback from the system made participants question whether what they're doing is making a difference or not, and leading to a feeling of helplessness.

It also appears that most participants want the most help with finding the projections, and less so with fixing the flickering. This makes sense as people need to find the projection before they can get to a flickering projection. In addition, flickering projections are usually caused by people standing at a less then optimal angle or distance away from the target. So if a system can help guide people to a good place for target acquisition, they are less likely to encounter flicking or other poor quality projections.

These findings reflect both the 'error management' and 'feedback' MAR usability heuristics listed by [Sang]. These two heuristics may go hand-in-hand in AR systems, as current AR systems are so opaque in their functioning, it's very difficult for users to know what might be an error and what might not be. Many times during trials, participants have expressed their doubt as to whether the app is working when no projections appear the first couple of times they tried. They might have persisted due to it being within an experimental context, but would not have done the same if it were, for example, a commercial application. In fact, some participants did give up on some targets that they couldn't find after numerous attempts.

5. FEEDBACK SYSTEM

As concluded from the exploratory study, the usability issue that is most important in making the AR experience more enjoyable to users is the implementation of some kind of feedback system that lets users know how they can achieve good quality AR projections.

Participants' inability to find good quality projections are mostly caused by them not standing at the optimal spot in front of the building, with other occasional errors caused by weather contrast that affects the system's ability to recognise the building. In addition, from the researcher's observations, the majority of people struggle to find good quality projections simply because their camera is aimed too close to the building.

Because of this, the most useful feedback to give users would be real time directional feedback that guides them to the optimal spot for each target. The directional feedback will not only serve to guide users to better spots, but also acts as a sort of system status updater that shows users that the application is in fact working and the reason why projections aren't showing up is because they need to find a better spot to aim from.

A Wizard of Oz (WoZ) system was developed for this purpose and tested to see if it helps improve the usability of the system, as well as help users find good quality projections faster. A WoZ system is one in which a human emulates the system's functionality and interacts with the user via a real or mock computer interface [10].

The WoZ method was chosen because the feedback provided by the fully computerised feedback system we created was not as robust and flexible as those given by a person. Also, there are key feedbacks, such as telling users they are standing too close to the building, that the computerised system is unlikely to be able to provide. This is because image recognition systems become quite unreliable when the camera captures a very close view of something (e.g. the camera simply sees the wall of a building with hardly any distinguishable features on it to tell the system what it is looking at).

To build a fully computerised feedback system that works as well as a WoZ system does would take far longer and might potentially require additional technology on top of image recognition, such as the ability for the device to sense how close it is to the target building through some kind of signalling from the building. Therefore, due to cost and the type of technology needed for the fully computerised system to work, the rapid-prototyping method of WoZ is more suitable [12,26]. This method however, suffers from the obvious downside of humans not being as accurate and consistent as a computer system is.

The hypotheses for this experiment are that:

1. Participants will be able to find projections quicker with the feedback system.

2. Participants will rate the application with feedback system as more usable.

3. Participants using the application with feedback system is more likely to fix the flickering error because of the feedback provided, compared to those without.

5.1 Methods

In the exploratory study, participants were told where the next target building is. However, in a real life situation, a user is likely to be using the application without any such assistance. So in order to make the test for the feedback system more ecologically valid, a paper map showing where the target buildings are is given to participants instead.

The map was adapted from UCL's official campus map, with the starting spot and the target buildings highlighted. Special care was taken to make sure the map didn't give hints to users as to where they should stand, so as to not influence participants' behaviour during the experiment.

To ensure the map was easy to understand and the transition between paper and iPad was smooth, so that its implementation does not affect participants' usability ratings for the feedback system, a pilot study with 6 participants with different iterations of the map was done.

For the final experiment, 27 participants (13 female, 14 male) were recruited through researcher's personal connections. The mean age of the sample was 26.3 (SD = 4.11), with an age range of 21-37.

This experiment used a between-subject design to avoid any learning effects that might affect users' performances. The independent variable was the presence of the feedback system, and the dependent variables are (1) the time it took for each participant to complete the task of finding the projections, similar to how [Sang] tested their MAR prototype; and (2) the rating they gave the system on the System Usability Scale [2]. The SUS was chosen to measure usability of the feedback system as it is a well-established and simple tool that has proven to be very reliable across different systems and products (Brooke, 2013). It has actually been used in other AR navigation research [24].

The control condition uses a Unity AR application that's mostly similar to the one used in the exploratory study. The only difference being that 5 building targets were cut down to 3: two of the most robust targets from the sample were kept and a new robust target along the same outdoor path was added. The other targets from the first study were discarded due to how unreliable they are.

A restriction on the projections' appearance on the screen was imposed on the AR application for both conditions, so that unless the participant was standing within 20 degrees of the centre of target building, projections would not appear even if the system was tracking the target (as seen in Fig.3). This was done to avoid a ceiling effect on the results, as all three buildings used were very robust targets, which means their projections should be quite easy to find.

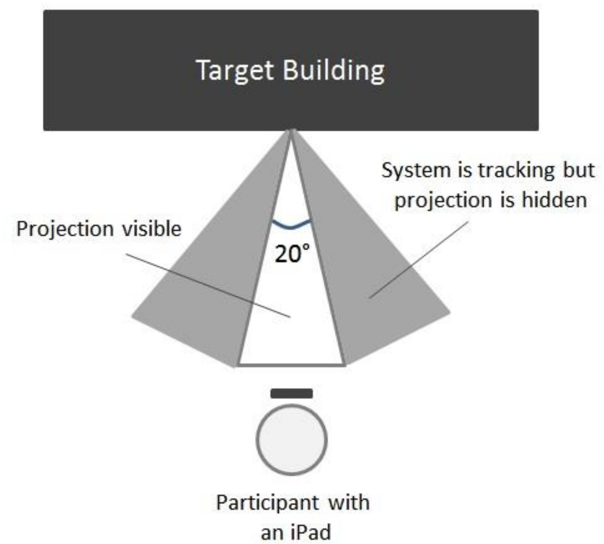

Fig.3: Diagram showing when projection is shown and when it's being hidden

The experimental condition had a WoZ feedback system that consisted of the same iPad used in the control condition and an iPhone 5S running on iOS9. A Vuforia AR application was installed on the iPad, it recognised targets and displayed projections the same way as the Unity app used in control condition did. In addition, it responded to commands from a websocket server. An Oz app is installed on the iPhone and can communicate to the Vuforia app on the iPad via the websocket server.

As many wayfinding AR applications developed in previous research use arrows to guide users where to go [3,19,24] a similar approach is adapted in the present feedback system. 2D green triangular arrows are used to tell users to move to the left or right. 3D forward and backward arrows were considered as a feedback (2D arrows in this situation would look too much like instructions to move the iPad up or down), however in the end, to eliminate any ambiguity as to what the 3D forward and backward arrows

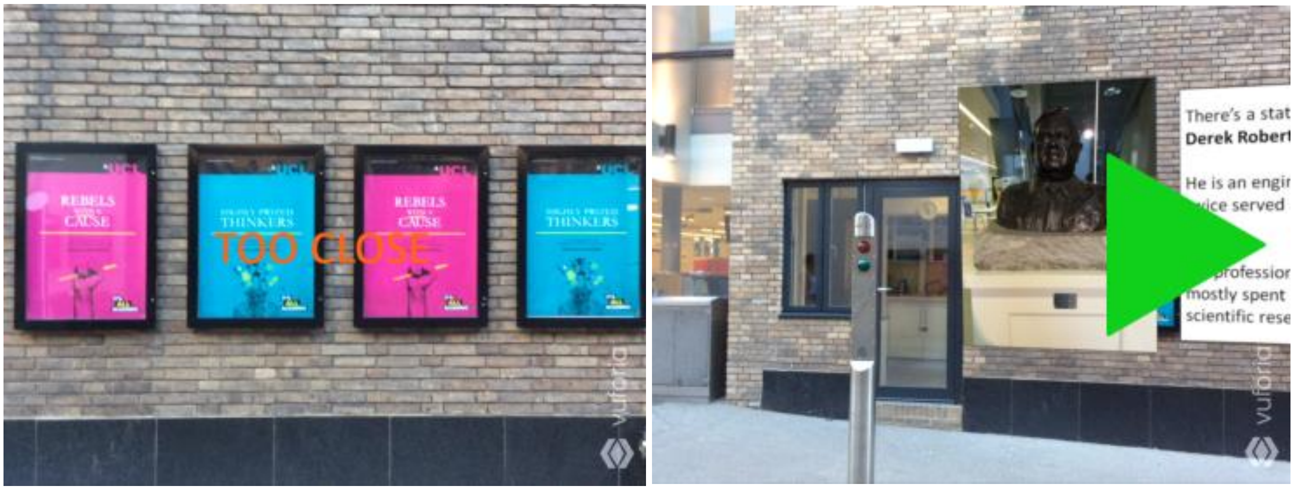

Fig.4: Feedback shown when a user is too close to a target, or aiming too far left of the target.

might be telling users to do, words were chosen as feedback instead. So the words “TOO CLOSE” or “TOO FAR AWAY” appear on the iPad screen when appropriate. An example of how the feedbacks looked is shown in Fig. 4.

The Vuforia and Oz applications were linked through a web server so that the iPhone can control what feedback appears on the iPad screen (as illustrated in Fig.5). A 4G connection was used on both devices. In the Oz app, there were four buttons that trigger the four corresponding feedback available. The response time on the iPad has at most a 1 second delay. Most of the time, feedback appeared on the iPad screen immediately after the button is pressed on the iPhone. When feedback was required, the researcher pressed the appropriate button repeatedly (so it would appear on the iPad screen as if the words or arrows were flashing) until participants did what the feedback was intended to make them do.

Both applications track the movements of the iPad through the device’s inbuilt sensors. All the following data were recorded from each trial: time since application started (in seconds), position of iPad (latitude, longitude, altitude), gyro attitude (the iPad’s orientation in 3D space), accelerations, angular velocities. Both applications also recorded when tracking is found or lost, and when the projections were being hidden or shown. Participants’ reaction to flickering projections when feedback was presented was observed and noted down by the researcher.

These sensor data were all tracked so that the time it took people to find the projections can be calculated, and people’s reactions were observed and noted down so that users’ behaviours could later be analysed together with whether or not the flickering error was fixed by their actions.

At the beginning of each experiment, the participant was handed the paper map with instructions at the top telling them to point their iPad camera at the buildings highlighted in order to find out information about them. Meanwhile, the researcher set up the connection between the iPad and the iPhone. The experiment started in front of the building that serves as a test trial (the same one from the exploratory study) that helped the participant familiarise what they have to do and what they can expect from the experiment.

For the control condition, the procedure happened exactly like it did in the exploratory study where the participants

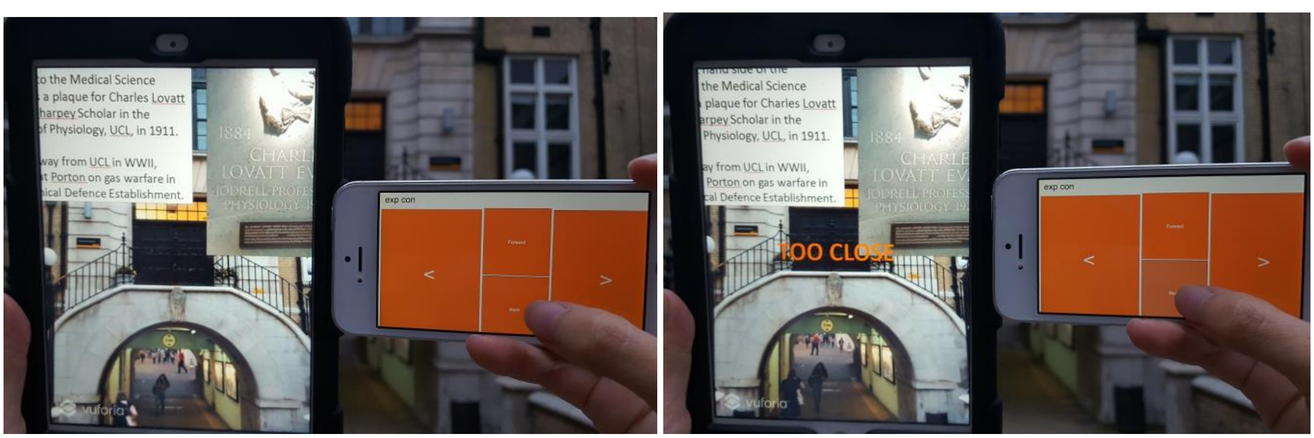

Fig.5: Pictures showing when the ‘back’ button on the iPhone is pressed, the feedback “TOO CLOSE” appears on the iPad screen.

went to the different target buildings with the researcher following silently and observing. The difference in the experimental condition was that the researcher stood behind the participant so they could see what the participant was seeing on the screen and give the appropriate feedback on their iPhone. The researcher sent feedback to the iPad screen until the projections occupy at least a third of the screen's real estate and is appearing at the centre of the screen. Participants could not see the researcher's iPhone screen throughout the experiment.

At the end of both conditions, participants were asked to use the System Usability Scale (SUS) to rate the feedback system only, discounting the paper map and the content of the projections. The SUS consisted of 10 questions that participants answer on a Likert scale. The positive and negative statements in the questionnaire alternate to prevent people from answering without thinking.

At the end of the SUS questionnaire, participants were given the option to leave some extra feedback. Each experiment took less than 15 minutes.

5.2 Results

The three main analyses for this experiment are the time taken to find projections, participants' rating of the system's usability, and whether the feedback helped participants fix flickering projections.

To calculate the time it took for a participant to find a projection, a 'time to target' value is needed. This value is calculated by looking at the time difference between when a participant lifted the iPad up in front of their face to look for the projection and the time when the projection is actually shown on the screen. Each participant has 3 'time to target' values, corresponding to the 3 targets in the experiment. 5 participants only have 2 because they weren't able to find one of the projections. 2 of these participants were from the experimental condition, 3 from control. 2 participants from the control condition were not able to find any projections at all and their data is excluded from this analysis. 1 participant's data was logged with significant bits missing and hence was also excluded. Each participant's average 'time to target' value is calculated.

As the data's distribution is positively skewed, the nonparametric test of Mann-Whitney U was conducted on the participants' average 'time to target'. The 'time to target' value in experimental condition (median = 3.13, IQR = 2.33) is significantly lower than the one in control condition (median = 6.54, IQR = 10.87), $U(22) = 17$, $Z = -3.16$, $p = .002$.

For the usability rating, the SUS gives each participant a perceived usability score that ranges from 0 to 100 based on their 10 answers. An independent-samples t-test shows that the SUS scores from the experimental condition ($M = 85.5$, $SD = 8.67$) is significantly higher than the one from the control condition ($M = 71.35$, $SD = 6.56$), $t(23) = -2.1$, $p < .05$.

From the researcher's observations, the feedback did not seem to have helped people fix the flickering projections. In general, the "TOO CLOSE" feedback is mostly used when no projections are appearing on the screen. It was displayed to all participants from the experimental condition and everyone followed the feedback and walked back immediately. Green arrows were much less used in this context, as participants typically were standing more or less around the centre of the target building's front anyway.

The green arrows and "TOO FAR AWAY" were mostly used when a flickering projection is in view and were used to signal to participants to move to a better location, in order to get a better track, and hence more stable projection. The green arrows were used much more often than "TOO FAR AWAY", due to the nature of the outdoor path having buildings on both sides, which means participants weren't likely to be that far away from the target buildings.

There are two types of responses to the appearance of green arrows when a flickering projection popped up on the screen. 6 participants followed the feedback with a delay, possibly to follow it once they've finished reading the information on the flickering projections. And when they do, they turn right or left, instead of walking right or left as the feedback was intended to make them do. The remaining 7 participants completely ignored the feedback. Therefore, none of the flickering projections were fixed, as no participants changed their location after seeing the feedback.

5.3 Discussion

The feedback system was created to help guide users to a better spot in front of target buildings and serve as a way of updating users more about the system's status. Our analysis showed that participants were guided to a better spot by the feedback so that projections would appear on their screens quicker. This supports our first hypothesis.

In addition, our analysis also showed that participants perceived the application with the feedback system as more usable than the one without. This presumably meant that the frustration and helplessness that participants from the exploratory study expressed were alleviated by at least a statistically significant extent by the feedback system.

In fact, a participant commented after the experiment was finished that this AR application was easier to use than a similar one they used while they were sightseeing in Beijing. They said with the app in Beijing, targets were sometimes easy to find, and sometimes very difficult. But the feedback in this app informed them that they were where they needed to be, and they were not seeing what they wanted to see just because they were not standing at the right place. Another participant volunteered a similar sentiment: "[…] *tells you it's too close so you can't capture the information, that helps as well, so I won't be confused when I'm close to the point and nothing shows*".

This supports the idea that the feedback system served as a way of informing the users about its system status, and the AR app doesn't feel as opaque as it previously was. This supports our second hypothesis that this new system with feedback is more usable.

The feedback system did guide participants to a better spot, however it was not able guide them to the best spot for a particular target building, as people were unresponsive to the feedback aimed at helping them fix flickering projections, or did actions that the feedback did not intend for them to do.

These responses can be broken down into a few points about what might need improvement in the system.

Firstly, there were those who ignored the feedback completely while reading information on the flickering projections. One way of interpreting that is that they understood the green arrows were there to help guide them to projections (most likely extrapolating from the "TOO CLOSE" feedback they saw previously), but they felt that they didn't need to follow them as they've deemed the flickering projection readable already. This interpretation was confirmed to be at least true for some participants, because when a participant was asked at the end of the experiment, why they ignored the green arrow flashing on the screen, they answered *"Because it was enough for me to read it"*. This response, compared to how participants tend to all follow feedback very quickly when no projections are visible on the screen, suggests that when no projections are visible, users will follow the feedback diligently; but when they are visible, some users will judge the quality of the flickering projection by their own standards before deciding whether to follow the feedback given to fix it.

There is a possible reason why so many participants chose to ignore the feedback. It may have been because of the 20 degrees visibility restrictions imposed on the system. As seen from Fig.3, if a participant is in the area outside of the 20 degrees zone in front of the target building, projections aren't shown even if the system has registered the building and is tracking off it. In fact, the system's tracking also gets worse the further from the centre a user stands and aims. So the quality of the projection inside the 20 degrees zone is actually better than outside of it. This means, the flickering seen by participants in this experiment might not have been as bad as the ones seen by those in the exploratory study. Hence, participants in this experiment might not have felt the need to fix the flickering, and therefore ignored the feedback that was supposed to aid them in doing so.

Secondly, for those who followed the feedback, they all followed them after a delay, likely because they were trying to read the texts on the flickering projection before tending to the feedback. It is less obvious as to what they thought the green arrow meant, but for the two participants who were asked about it after their experiment, they said that they thought it meant there were more projections on the left/right and the arrows were pointing them towards them. One of the two participants even said that after they turned left with the iPad and found no new projections, they thought perhaps the green arrow meant that she was standing at the right place.

Clearly, the arrows' meaning was too ambiguous, both in terms of what they were for and what they're asking users to do.

The belief that the arrows were pointing users to more projections might be based on the fact that they have previously followed feedback (e.g. "TOO CLOSE") that did lead them to new projections, therefore they thought feedback is only used for finding new projections. This suggests that perhaps feedback for finding projections and feedback for fixing flickering projections need to be distinguishable from each other. Doing so might clear up some of the ambiguity that is present in the current system.

The feedback for fixing flickering projections may need to be explicitly labelled as such, as users with little or no experience with AR may not know that the flickering is something that they are able to fix. Without specific wording or labelling, it may not be obvious to them that the feedback and the flickering are related at all.

Another ambiguity that needs to be cleared up is what the green arrows are telling users to do. It is clear that none of the participants took the arrows as meaning that they needed to move to the left/right, all of them turned their bodies to the left/right instead. Differently shaped arrows (such as one with a tail, e.g. →, instead of a simple triangle) or animations may lead to different responses from users, however, words may be the most straight forward and easy to understand feedback. For example, "TOO CLOSE" informs the users about what's wrong and implies the action they should take to correct it (e.g. walk back). From this experiment, it shows that feedback shown need to convey two things: why the feedback has shown up and what the user can do about it. Otherwise, it causes confusion.

6. GENERAL DISCUSSION

Because of the small amount that is currently known about average users' interaction with image recognition-based AR technology, this study set out to investigate their understanding of the technology, their ways of interacting with it and the reasons behind them.

Results from the exploratory user study showed that people generally are doing the correct actions to interact with the AR system, but those actions did not always help them find the stable projections that they wanted. Many participants expressed frustrations towards the fact that they seemed to be doing a lot of things in vain during their attempts to find stable projections. Hence, the biggest user need was found to be help with finding projections and not feeling stranded by the lack of response from the system when users can't find it themselves. A feedback system that provides users with real time directional feedback that guides them to the optimal spot for each target was created.

In the user testing for the feedback system, the user needs concluded from the exploratory user study were validated in that user experience was indeed improved with the presence of the feedback system. Participants were able to find projections faster with it, and reported higher usability ratings, which supports the first and second hypotheses for the user testing. The third hypothesis however, was not supported as the feedback system failed to guide users to the best spot for a particular target building.

This study's findings are mostly in line with what was already known about usability problems in AR applications, such as frustrations with their inability to make the app work and encountering errors that users don't know how to fix or prevent. Similar sentiments were also expressed in studies about AR usability previously mentioned [17,18,24]. It also validated two of the MAR usability heuristics proposed by Ko et al.: 'error management' and 'feedback'[11].

But in addition to that, the user study provided knowledge about how users understood AR as a technology. For example, many knew the technology was based on some form of image or feature recognition, while a couple thought it was GPS based. And they also had a variety of ideas as to why errors they encounter using the app occurred, most of them were actually correct.

Future research should create a similar AR application but have the app be GPS based instead. It would be worth investigating whether users are able to piece an understanding of the technology together with just brief usage of it, and see if their ways of interacting with it changes accordingly. That could further inform us on how well users are able to adapt to a foreign system, and what are the aspects of it that users cannot figure out by themselves and would need assistance from the technology.

Other than that, the usability test also provided some interesting lessons about designing a feedback system that helps users find targets when they fail to do so themselves. The results showed us that when no projections are visible, users will very likely follow the feedback given; but when projections are visible, some users will judge the quality of the flickering projection by their own standards before deciding whether to follow the feedback to fix it.

In addition, part of the reason why the feedback system failed to guide users to the best position for a given target was because of ambiguity in the feedback, both in terms of what they were for and what they're asking users to do. Our results point to the straightforwardness of feedback in words over graphical signs, however, future research should formally investigate the difference between words and graphical signs in conveying feedback.

At the same time, different ways of presenting feedback should be explored and have users' response to each of them investigated. For example, when brainstorming how our system can convey feedback to users, a heat map type of feedback system using 'percentage match' that the image recognition system provides was mentioned. In such a system, users can see how close they are to the optimal position in front of the target building. This incremental feedback can be represented in colour gradients, where if a user is far away from the optimal position, the screen will be tinted in red, and as the user gets closer, the colour gradually changes to green depending on the device's location. However, the data on 'percentage match' is not available to the public for the platform our system is built on. If this data is made available in the future, researchers looking to build a feedback system for their AR app should consider making use of it.

Other than different ways to present feedback that guides user to optimal positions, different types of system status or feedback can also be explored in the future. This is because errors or lack of response from an image recognition-based AR system is not always caused poor angle and distance. Factors such as weather/lighting contrasts, objects obstructing the target, target having changed its appearance due to construction work, poor internet signal if the system has to communicate with a server, they all can contribute to errors or unresponsiveness from the system. Even if some of these factors can't be reliably detected by the system, the system should still inform users of the possible reasons why it is not performing as proficiently as it should, such as by making a help menu available to users.

There is one limitation in this study that restricts its ecological validity and that is the robustness of the targets used in the experiment. The ensure that projections would show up consistently as long as the participants stood at the right spot, all the buildings used in the usability test are very robust targets. However, there are many buildings or sceneries in the wild that are very difficult for image recognition software to recognise. In fact, a lot of buildings along that outdoor path in UCL campus were poor targets and had to be discarded. It is possible that the results from the usability test may not be as positive if poor targets were used. Future studies should investigate whether a feedback system like this one is still capable of improving user experience when the AR system is faced with hard-to-recognise targets.

7. CONCLUSION

In this study, contributions were made to the knowledge about laypeople's understanding of image recognition-based AR technology, their interactions with it and reasons behind them. An exploratory user study showed that the biggest usability challenge users face is the difficulty in finding projections and the subsequent frustration of not being able to find anything and not knowing what to do. A feedback system that provides directional instructions that guide users to the optimal position to acquire a stable projection is created.

Usability testing comparing AR system with and without the feedback system showed that users found projections significantly faster with it, and thinks the system is significantly more usable with it. However, the feedback system was not able to guide users to optimal locations in order to fix unstable projections. Reasons for this may be due to some participants' higher tolerance to reading information from unstable projections as well as ambiguity in feedback signals used.

Overall, the present study showed that a feedback system that alerts users to errors and gives instructions for how to fix them significantly improves user experience.

ACKNOWLEDGMENTS

A huge thank you to my supervisor, Simon Julier, and my co-supervisor, Geraint Jones. I wouldn't have been able to do this amazing project without either of them. Also thank you to my friends and family who kept me sane and kept me company throughout these months. I also thank coffee and takeout.

REFERENCES

[1] Azuma, R. Tracking requirements for augmented reality. *Communications of the ACM 36*, 7 (1993), 50–51.

[2] Braun, V. and Clarke, V. Using thematic analysis in psychology. *Qualitative Research in Psychology 3*, 2 (2006), 77–101.

[3] Buchmann, V., Billinghurst, M., and Cockburn, A. Directional interfaces for wearable augmented reality. *Proceedings of the 9th ACM SIGCHI New Zealand Chapter's International Conference on Human-Computer Interaction: Design Centered HCI*, ACM (2008), 47–54.

[4] Desmet, P. and Hekkert, P. Framework of product experience. *International journal of design 1*, 1 (2007).

[5] Dünser, A., Grasset, R., and Billinghurst, M. A survey of evaluation techniques used in augmented reality studies. ACM Press (2008), 1–27.

[6] Gammeter, S., Gassmann, A., Bossard, L., Quack, T., and Van Gool, L. Server-side object recognition and client-side object tracking for mobile augmented reality. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*, IEEE (2010), 1–8.

[7] Hassenzahl, M. and Ullrich, D. To do or not to do: Differences in user experience and retrospective judgments depending on the presence or absence of instrumental goals. *Interacting with Computers 19*, 4 (2007), 429–437.

[8] Höllerer, T. and Feiner, S. Mobile augmented reality. *Telegeoinformatics: Location-Based Computing and Services. Taylor and Francis Books Ltd., London, UK 21*, (2004).

[9] Kähäri, M. and Murphy, D.J. Mara: Sensor based augmented reality system for mobile imaging device. *5th IEEE and ACM International Symposium on Mixed and Augmented Reality*, (2006).

[10] Klemmer, S.R., Sinha, A.K., Chen, J., Landay, J.A., Aboobaker, N., and Wang, A. Suede: a Wizard of Oz prototyping tool for speech user interfaces. *Proceedings of the 13th annual ACM symposium on User interface software and technology*, ACM (2000), 1–10.

[11] Ko, S.M., Chang, W.S., and Ji, Y.G. Usability Principles for Augmented Reality Applications in a Smartphone Environment. *International Journal of Human-Computer Interaction 29*, 8 (2013), 501–515.

[12] Landauer, T. Psychology as a mother of invention. *ACM SIGCHI Bulletin*, 18.4 (1987), 333–335.

[13] Livingston, M.A. Evaluating human factors in augmented reality systems. *IEEE Computer Graphics and Applications 25*, 6 (2005), 6–9.

[14] Milgram, P. and Kishino, F. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems 77*, 12 (1994), 1321–1329.

[15] Nilsson, S. and Johansson, B. User experience and acceptance of a mixed reality system in a naturalistic setting-a case study. *Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality*, IEEE Computer Society (2006), 247–248.

[16] Olsson, T., Lagerstam, E., Kärkkäinen, T., and Väänänen-Vainio-Mattila, K. Expected user experience of mobile augmented reality services: a user study in the context of shopping centres. *Personal and Ubiquitous Computing 17*, 2 (2013), 287–304.

[17] Olsson, T. and Salo, M. Online user survey on current mobile augmented reality applications. *Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on*, IEEE (2011), 75–84.

[18] Olsson, T. and Salo, M. Narratives of satisfying and unsatisfying experiences of current mobile augmented reality applications. ACM Press (2012), 2779.

[19] Reitmayr, G. and Schmalstieg, D. Location based applications for mobile augmented reality. *Proceedings of the Fourth Australasian user interface conference on User interfaces 2003-Volume 18*, Australian Computer Society, Inc. (2003), 65–73.

[20] Reitmayr, G. and Schmalstieg, D. *Collaborative augmented reality for outdoor navigation and information browsing*. na, 2004.

[21] Rohs, M., Schleicher, R., Schöning, J., Essl, G., Naumann, A., and Krüger, A. Impact of item density on the utility of visual context in magic lens interactions. *Personal and Ubiquitous Computing 13*, 8 (2009), 633–646.

[22] Swan, J.E. and Gabbard, J.L. Survey of user-based experimentation in augmented reality. *Proceedings of 1st International Conference on Virtual Reality*, (2005), 1–9.

[23] Takacs, G., Chandrasekhar, V., Gelfand, N., et al. Outdoors augmented reality on mobile phone using loxel-based visual feature organization. *Proceedings of the 1st ACM international conference on Multimedia information retrieval*, ACM (2008), 427–434.

[24] Walther-Franks, B. and Malaka, R. Evaluation of an augmented photograph-based pedestrian navigation system. *International Symposium on Smart Graphics*, Springer (2008), 94–105.

[25] Ware, C. and Balakrishnan, R. Reaching for objects in VR displays: lag and frame rate. *ACM Transactions on Computer-Human Interaction (TOCHI) 1*, 4 (1994), 331–356.

[26] Wilson, J. and Rosenberg, D. Rapid prototyping for user interface design. (1988), 859–875.

[27] Wither, J., DiVerdi, S., and Höllerer, T. Annotation in outdoor augmented reality. *Computers & Graphics 33*, 6 (2009), 679–689.