AR enriches our world:
A comparison between creating annotations in AR and 2D

Jurjen Braam

Bachelor Information Sciences, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
j.e.braam@student.vu.nl

Supervisor: Victor de Boer

Abstract. Augmented reality (AR) being used for displaying annotations on objects has often been researched in various research areas. However, the specific task of creating annotations using augmented reality has not been described in the literature. Hence ,in this paper, I describe a research which is focused on discovering the suitability of creating annotations in AR in general and/or more specifically, annotating paintings in AR. This approach is compared to two traditional methods: Normal annotation (Traditional real-life annotation) and 2D annotation. In order to compare these methods, I describe the implementation of a prototype iOS app called AR Annotator. This app enables users to create annotations in three modes: Normal, 2D and AR. AR Annotator was used in an experiment which was performed by 6 participants on a location of the Vrije Universiteit Amsterdam. The generated annotations and accompanying questionnaire results are analyzed and presented. The results indicate that participants found the experience in AR mode more comparable to Normal mode than in 2D mode. The annotations show the same results for the distribution of the relative positions of the annotations. However, other metrics indicate the opposite. Namely, annotations in 2D mode are more comparable to annotations made in Normal mode than in AR mode. From the questionnaire it follows that factors like the limited size and instability of AR objects affect the usability and efficiency of annotating in AR. Therefore, further research is suggested with larger screen sizes and optimized AR experiences. Furthermore, research using larger sample sizes is suggested to provide statistical proof for the question if AR is suitable for the task of creating annotations.

Keywords: Mobile, Augmented reality, Task optimization, Annotation

1   Introduction

Due to recent developments in mobile computing, modern mobile devices are able to run complex augmented reality (AR) experiences. Augmented reality refers to the incorporation of computer-generated virtual objects into a real three-dimensional scene. In more detail, the three-dimensional scene is generated from the data stream of the

cameras and depth sensing sensors of the device [1]. A complex task that would not have been possible on a mobile device a decade ago. Both Apple and Google have released software development kits that enable developers to create applications with AR for mobile devices like smartphones and tablets. This creates possibilities for the use of AR in more different areas. However, it is still unclear what type of tasks are suitable to be assisted by AR. Moreover, the optimal characteristics of these applications incorporating AR are not fully discovered.

That AR applications have the potential to greatly enhance the user experience and performance is already shown in studies like [2] [3]. One task that has been the main focus of multiple AR studies is that of displaying annotations on top of real objects using AR [4] [5] [6]. The prototypes that came forth from these studies enabled participants to display annotations on real-life objects using AR, to fulfill their respective tasks. While the task of displaying annotations in AR has been researched, there has not been any research on the applicability of AR in the process of creating these annotations. Therefore, this research tries to find out if AR is suitable for this task.

The prospect of users fulfilling tasks in AR became closer to reality with the transition from expensive one-purpose AR devices to AR technology on mobile devices. As a result of this transition, millions of potential users are not only able to experience annotations in AR, even more so are potentially able to create annotations in AR.

The task of annotating can be described as the process of writing notes concerning the characteristics or contents of a piece of work. These annotations enable retrieval techniques of search engines. One of the most obvious types of works that can be subject to annotation is paintings. By describing the contents and characteristics of paintings the collection of works can be made manageable. However, the traditional process of annotating a painting is time-consuming since the annotator needs to be in the presence of every physical painting that is being annotated.

An alternative approach would be to annotate virtual objects representing the paintings. An example of this approach would be the Accurator project in which the goal was to support the annotation process of external annotators in order to enrich the collection of the Rijksmuseum Amsterdam [7]. In this project, participants were asked to annotate virtual paintings in the form of 2D images using a web-based tool. However, one could argue that paintings are originally made to be viewed in a real-life experience, in contrast to being watched as virtual representations in form of 2D images on a screen. I would argue that the experience of viewing a painting consists of looking at the painting from various angles and distances to find details that are overlooked at first glance. As a consequence, the process of annotating is also subject to the method of viewing the painting.

Following this logic, one could argue that the process of annotating could benefit from the support of AR to experience the painting as a virtual object and mimicking the experience of looking at a real painting. Another consideration is that the number of potential users has increased rapidly by this transition to AR on mobile devices. Therefore, this research is focused on the possibility for AR to assist the process of annotating paintings. The research question is thus stated as follows:

- Is AR a suitable alternative for the task of annotating paintings in comparison to traditional methods in real-life and 2D?

Following the assumption that paintings are originally intended to be viewed in real-life I state my hypothesis as follows: *The experience of annotating paintings in AR more closely resembles an experience in real-life than 2D. Therefore, the experience will create more similar annotations in comparison to 2D.*

In order to find an answer to the research question and to test the hypothesis, I will first describe the related work that has been done on AR and annotating. In this literature section, multiple prototypes of AR applications are described, and an annotation classification framework is discussed that will be used in the data analysis. In Section 3, I will outline the research method performed to gather the data. This includes a description of the process of developing the prototype iOS app AR Annotator and description of the experimental setup. In Section 4, I will present the results gathered from the annotations and the questionnaire. Following in Section 5, I will discuss how the results should be interpreted and the limitations of the research. Finally, in Section 6 I will present the conclusion.

2   Related work

2.1   Augmented Reality assisting tasks

Augmented Reality has been implemented for a wide variety of applications and in the future developers will find more use cases for AR. One of these use cases is the use of AR to provide more information on certain media. An example is the research in [3] where the researchers compared an app with a 2D map mode and paper map with an AR mode in a location-based game. In this game, the participants that used the AR-based program had the possibility to use the camera of the smartphone to gather information about a specific location on the paper map by pointing at the map. The most noticeable result was that the AR solution improved the collaboration of the participants. This shows one of the possibilities how AR can assist users in certain tasks.

Another example of where AR can assist in certain tasks is in training. In [2] it was shown that for learning assembly instructions had significantly lower completion time for the assemble part provided in AR in comparison to conventional 2D drawings for instructions. Though, both examples of AR do not incorporate any interaction with the AR objects, but rather use AR overlays to provide information. Whereas, the participants in this research will interact with the AR objects by walking around them and writing annotations about them.

2.2   Classification of user image descriptions

The number of images that are stored is increasing rapidly due to advances in storage techniques, the quality of cameras in smartphones and the increasing availability and speed of the world wide web. Therefore, more than ever there is the need for effective image retrieval techniques. Artificial intelligence image recognition has a lot of potential to assist in image retrieval, however, not all images can be analyzed by these algorithms. Where current AI systems can identify specific objects successfully, abstract concepts are much harder to identify by algorithms. Therefore, there is still the need user-generated descriptions. One type of user-generated descriptions are annotations which are the focus of this research. To compare the different types of annotations some sort of classification is needed. The research in [8] provides a classification framework for image descriptions by users. This framework is constructed from classification schemes and methods found in literature such as the Panofsky/Shatford model. I will use this framework for categorizing the annotations made in the experiment. This framework is not specifically developed for categorizing annotations. However, since user image descriptions and annotations have the same purpose I will use the framework to categorize the annotations. The framework depicts three viewpoints for categorizing user descriptions: nonvisual, perceptual and conceptual descriptions. The category of nonvisual descriptions entails descriptive information about the image. This is often called metadata. In the context of annotations, nonvisual annotations would describe objective information that cannot be directly derived from the painting. An example of such annotations would be the name of the painting or the creation date. The second category is the perceptual level which entails description that is directly derived from the visual characteristics of the painting. Examples of such characteristics are composition, color, shape and texture. The third category is the conceptual level. In this category, the annotations describe the semantic content of the image. This level consists of annotations that describe general, specific and abstract concepts. This description describes the information about the scene or an object in the scene. This framework will later be used to classify the annotations that are made by the participants. I will continue with a description of the research method used in this research.

3   Research method

To test if augmented reality is a suitable option for the process of annotating paintings I conducted an experiment in which participants used three different ways to annotate three different paintings. While I was looking for a suitable software solution for annotation in augmented reality I found that no off-the-shelf applications exist. Therefore, I chose to develop my own prototype iOS app. The reason that I chose for an iOS app is because my personal devices are iOS devices. Therefore, the development could be done on a MacBook Pro and the prototype could run on an iPhone X. The first part describes the development of the prototype iOS app AR Annotator. Secondly, I describe the experimental setup. Finally, I will describe how the data was analyzed.

3.1   Prototype: AR Annotator

The following requirements were stated for the prototype app AR Annotator:

- The application should handle three different modes of annotation: normal, 2D and AR.
- In normal mode, users should have the ability to write annotations.
- In 2D mode, users should be able to view the painting as a 2D image with control for zoom and movement and be able to write annotations for the painting.
- In AR mode users should be able to start an AR experience and display a virtual easel with a painting. Also, the user needs to be able to write annotations for the painting.
- The application should save annotations in a location suitable for retrieval grouped by the ID of the user.

Paintings

Fig. 1. Painting 1 (left), Painting 2 (center), Painting 3 (right)

The paintings used in the experiment were downloaded from Google. I chose pictures with a resolution of at least 1200 x 1100 pixels so that there is no visual difference between the printed version, the 2D image and 3D object of the painting. The paintings were also chosen to have sufficient elements and details for annotation. This relatively high number of elements ensures that the participants are not restricted to a limited amount of possible annotations. Otherwise, the research method will not be able to compare the methods successfully.

Google Firebase Realtime Database

From the previously stated requirements, it follows that the prototype should be able to store annotation data per participant. This data consists of the annotated painting, the annotation text and for 2D mode the specific X and Y coordinate where the annotation was pinned. This data needed to be accessible for analyzing at the end of the annotation

process. Therefore, the Realtime Database of Google's Firebase app development platform[1] was chosen as the service to store the data created by AR Annotation app. Besides providing a safe storage location for annotations, the Firebase Authentication handler ensures that the data is transferred securely. The Realtime Database allows for an easy integration into the app by providing simple functions to save, update and delete data. One of the biggest advantages of the Realtime Database is offline synchronization. In case of a network disconnection, the annotation data is stored locally until the network connection is restored and the changes are synchronized with the cloud version of the database. The database is only accessible with an admin account.

Blender
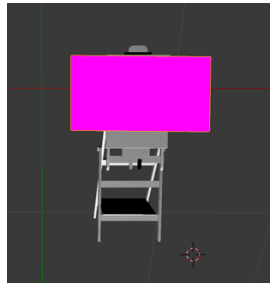
Fig. 2. Easel object with 2D image placeholder in Blender

Blender[2] is a free and open source 3D creation suite that enables the creation of 3D objects. Blender was used to create a 3D object consisting of two parts. The first part is an easel to hold the painting, which was downloaded from cgtrader.com. The second part is a 2D placeholder plane where the selected painting image could be plotted. The 3D object then was exported in DAE format to be imported into Xcode.

Apple ARKit

For iOS apps to be able to display augmented reality Apple's ARKit[3] software development kit is needed. This SDK enables apps to create augmented reality experiences, by providing the main components of AR experiences. ARKit uses the iOS device camera and motion features to create scenes in which virtual objects can be placed. As a starting point, I imported the ARKit demo provided by Apple. This demo provides a view that starts an AR session and provides the controls to add virtual objects. The ARKit demo app uses the camera data to find a horizontal plane and shows a yellow square on top of the camera feed when it has found one. I have replaced the default
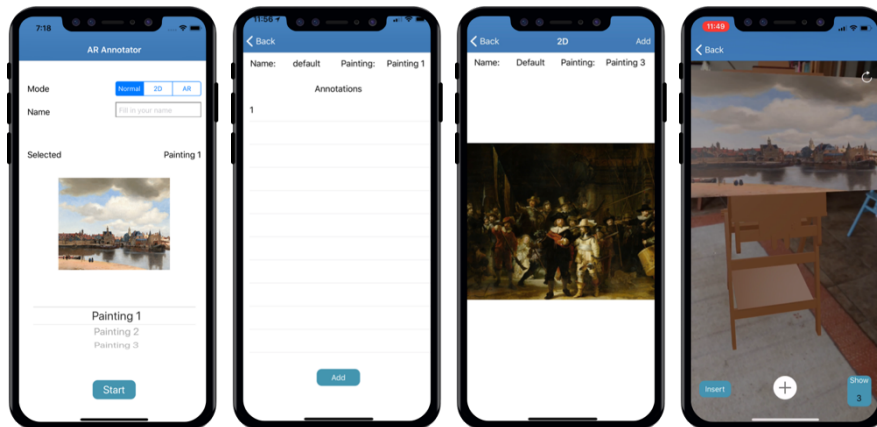
---

[1] Google Firebase (https://firebase.google.com)
[2] Blender (https://www.blender.org)
[3] Apple ARKit (https://developer.apple.com/documentation/arkit)

virtual objects with 3 easels 3D objects. The easel objects where converted from DAE to Scene object to be able to plot the painting image on the 2D plane.

Screens

Fig. 3. iOS Prototype start screen and 3 modes: Normal (left), 2D (middle) and AR (right)

The start screen of the prototype provides users the mode selector, the name field for holding the participant ID and the scroll selector for painting selection. The image view previews the selected painting. When the user is ready to begin annotating the Start button on the bottom starts the annotation.

Normal mode

In normal mode, the user is presented with a tutorial overlay that describes the process of adding annotations. When the user dismisses the tutorial, a list view is shown containing links to previously made annotations by that user. The annotations in this mode are created while looking at a real painting. When the user sees something to annotate then the user taps on the Add button which shows a new view with a text field in which the annotation can be written. When the user taps the save button the annotation text, the painting name and the participantID are saved in the database. Firebase then synchronizes the local changes in the database to the cloud version of the database.

2D mode

In the second mode, the user is presented with a tutorial overlay which describes the process of scrolling and zooming in the image view. Also, the tutorial explains how to add annotations. When the user dismisses the tutorial, the image view containing the painting is shown. The user can zoom using two fingers and move the painting with one finger. The experience is comparable to the experience of viewing an image in the

default image view on modern smartphones. Where it differs though, is that when the user taps a specific location on the painting for at least one second, an icon ⓘ is put on that location. When the user taps that button, the annotation view is shown where the user can fill in the annotation text. When the user taps the save button the annotation text, the painting name and the participantID are saved in the database. Firebase then synchronizes the local changes in the database to the cloud version of the database.

AR mode

Also, in the third mode, the user is first shown a tutorial overlay describing the annotation process. In this mode, the user is provided an image view which shows the live image of the device's camera. The user is instructed to point the device on the ground so that the ARKit session is able to scan a horizontal plane which is in most cases the floor. When this plane is found, a yellow square is displayed on this plane to indicate to the user that the session is ready to plot the easel object. When the user taps the Insert button, a virtual easel containing the selected painting is plotted on the horizontal plane. The user then is able to walk around the easel while pointing the device in the direction where the easel was inserted, to experience different viewing angles of the painting. By walking towards the easel, the painting can be scanned in more detail. Whereas walking away from the easel gives an overview of the full content of the painting. When the user taps the plus button, an overlay with a text field is displayed. In this text field, the user is able to write the annotation text. When the user taps the save button the annotation text, the painting name and the participantID are saved in the database. Firebase then synchronizes the local changes in the database to the cloud version of the database. Finally, the annotation overlay dismisses and the view continues the AR experience to enable further annotation.

3.2 Experiment

The experiment was conducted in a closed off office space at the Vrije Universiteit Amsterdam. In the office space, three different printed representations of paintings were arranged at eye level using the wall and a stand. The office also contained two desks and a couple of chairs. The experiments took place between 11:00 am to 1:00 pm The participants consisted of four students and two employees studying and working at the Vrije Universiteit Amsterdam. Three participants were male and the other three participants were female. The age of the participants was between 19 and 30 years. All participants study and/or work in the field of information technology. However, none of them are specialized in the field of augmented reality. The experiment was performed one participant at a time. The participants were asked to read the Experiment Instructions, which can be found in appendix A, that described what was expected from them during the experiment. I also asked the participants to sign the Experiment Instructions stating that they allowed me to store their annotations in Google Firebase and use them for this research. During the experiment, the participants annotated three different

paintings in three modes. The order of these modes was pre-determined. I chose this approach to counter the learning effect of annotation using the three modes after each other, which could potentially change the outcome. Also, the painting selected for each mode was pre-determined. This was chosen to counter the possibility that annotations made in a certain mode were influenced by the specifics of one of the paintings. Table 1 depicts this combination and order. The number under the respective mode columns indicates in what order the participant the modes performed. The numbers under the respective mode-painting columns indicate which painting the participant annotated in that mode.

Table 1. Combination of mode/painting and order of modes

Partici-pant ID	Normal painting	2D painting	AR painting	Nor-mal	2D	AR
1	1	2	3	1	2	3
2	2	1	3	1	3	2
3	3	2	1	2	1	3
4	1	3	2	2	3	1
5	2	1	3	3	1	2
6	3	1	2	3	2	1

The participants were asked to create 5 annotations per mode. No restrictions on the number of words were given. The participants were allowed to write in English or in Dutch. Furthermore, they were allowed to perform the task seated or standing. The process of annotating in the three different modes took between 10 and 15 minutes. When the participant was finished I asked them to fill in a questionnaire concerning the annotation process and their personal experience with AR. The questionnaire was created in Google Forms[4] and filled in using the laptop that I provided. The questions can be found in the appendix B. After finishing the questionnaire, I thanked the participants for their time. There was no reward given.
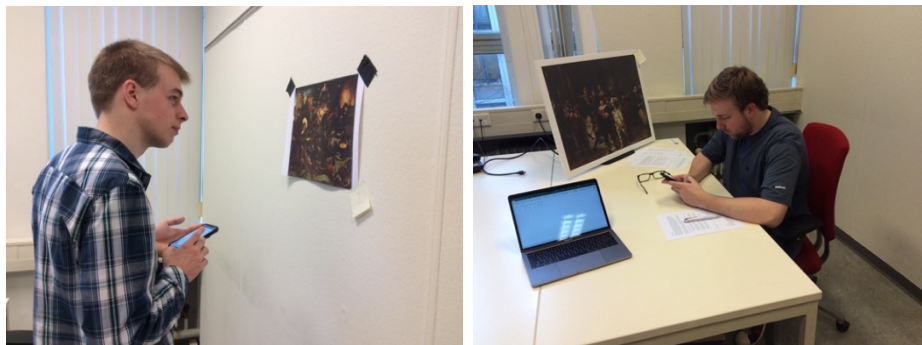
Fig. 4. Participant performing normal mode (left) and participant performing 2D mode (right)

---

[4] Google Forms (https://www.google.nl/intl/en/forms/about/)

3.3 Analyze Data

The annotation data was generated by the participants using the app AR Annotator. The app saved the data first in a local copy of the Firebase Realtime Database and then synchronized the changes to the cloud version. Firebase provides a JSON export of the Realtime Database which I used to download all annotations. From this JSON file I created a CSV file, to be able to import it into a Microsoft Excel spreadsheet. This gave me the opportunity to analyze the annotations using the functions Excel provides. Using these functions, I calculated the average word count per annotation. For the average word length, I had to use a Python script, because Excel does not provide functions for that. This Excel functions and Python code can be found in appendix C.

To get a better understanding of the annotations I calculated two more parameters for the annotations. The first parameter that was calculated was the image description class using the framework of [8]. This parameter is added to get a better understanding what categories of annotations were made per mode. The second parameter is the relative position of the element which is being annotated. Of course, this could only be done for the annotations describing perceptual characteristics or conceptual elements. To determine the relative position of the annotations four categories were defined: Top left, Top right, Bottom left and Bottom right. However, when assigning the annotations to their respective category I found that some annotations covered more than one category. That is, some annotations described elements or objects that spanned the top, bottom or complete painting. Therefore, three addition categories were added: Top, Bottom and All.

The responses to the questionnaire were gathered from Google Forms. The answers to the questions were analyzed to find out how the participants experienced the different modes, as well as their prior experience with AR and their view on using AR in the future.

4 Results

4.1 Annotations

Table 2 presents the results of the categorization of the annotations using the image description framework. The number of annotations per mode is organized in the classes Non-visual, Perceptual and Conceptual. 91% of the annotations described conceptual elements, 8% described perceptual aspects and 1% described non-visual information.

Table 2. Annotations per class per mode

Class	Normal	2D	AR
Non-visual	0	0	1
Perceptual	4	2	1
Conceptual	28	29	27

The results of the annotation metrics, average number of annotations and average word length are presented in Table 3. The metrics are grouped by mode.

Table 3. Annotation metrics per mode

Metric / Mode	Normal	2D	AR
Average number of words	3,84	3,81	3,59
Average word length	5,84	5,79	5,78

Finally, Table 4 displays the relative distribution of the annotations per mode. The values of the table represent the proportion of annotations in that position for the respective mode.

Table 4. Annotation positions Normal mode

Mode	Top left	Top right	Bottom left	Bottom right	Top	Bottom	All
Normal	3,13%	3,13%	28,13%	15,63%	3,13%	21,88%	25,00%
2D	9,68%	9,68%	32,26%	25,81%	0,00%	9,68%	12,90%
AR	3,45%	13,79%	17,24%	24,14%	3,45%	13,79%	24,14%

4.2 Questionnaire

The responses of the questionnaire gave the following results. Five of the six participants indicated that the AR mode was most enjoyable to use, where one participant found the 2D mode most enjoyable. Three participants indicated that 2D mode was least enjoyable, and two participants indicated that Normal mode was least enjoyable. One participant did not fill in the least enjoyable mode. When asked to order the modes with respect to efficiency, five of six participants indicated 2D mode to be the most efficient. One participant responded that Normal mode was most efficient. Four participants indicated that AR mode was the least efficient and one person indicated Normal mode was the least efficient. One participant did not fill in the least efficient mode. When asked what the participants enjoyed about annotating in the Normal mode they indicated that the overview of the painting was preserved and that there was no blocking text field covering the painting. On the other hand, the participants indicated that they did not enjoy that there was no zoom control and that the physical painting needed to be in the room. For 2D mode, the participants indicated that the ability to zoom in on a specific location to find details and the ability to create annotations without a physical painting were good aspects. However, they indicated that the bad aspects of 2D mode are that the annotation feels static and the focus is on one single area at a time. When

asked about the AR mode the participants indicate that it provides an immersive experience with zoom control, while providing the full overview of the painting. Furthermore, the mode does not require the presence of the physical painting. On the negative side they indicated that the text field blocks a large part of the view, the easel was not always stable and the disability to lock on a specific part of the painting.

When asked about prior experiences with AR five participants indicated to have one to three experiences and one participant did not have any prior experiences. All participants indicated not to use AR on a daily basis. On a Likert scale from 1 to 5, five participants indicated that they found the AR experience level 4 convincing. Whereas one participant found the AR experience level 5 convincing. Figure 5 shows the responses to the questions on the comparability of the experience of annotating in 2D and AR mode and the likeliness of choosing AR over 2D for future tasks.
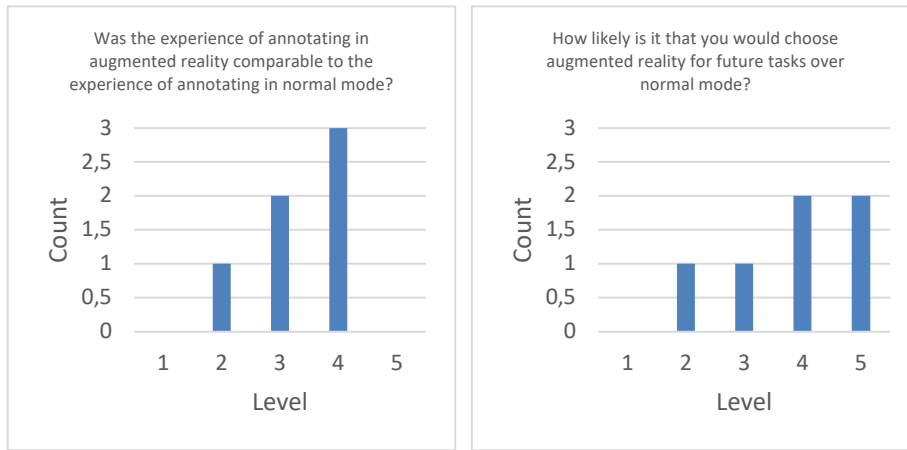
Fig. 5 Graphs denoting responses on Likert scale for comparability AR and 2D modes (left) and likeliness of choosing AR over 2D for future tasks (right)

5 Discussion

Because the sample size of six participants is too small the results do not have any statistical legitimacy to specify conclusions about the whole population of potential annotators. However, the results do give insights into the aspects of the experience of using augmented reality for annotation. When analyzing the first derived parameter of the annotations, the image description classes, we find comparable results as found in the research in [8]. Namely, that the majority of annotations described conceptual elements, about 10% described perceptual aspects and only 1% describes non-visual information. The distribution of classes of the annotations per mode does not indicate a significant difference between the three modes.

The metrics about the annotations describe that the annotations created in Normal mode consist on average of more words when compared to 2D and AR. The annotations created in 2D consisted on average of more words than the annotations created

in AR. This indicates that the participants described the annotations in more detail by using more words in Normal mode than in 2D or AR mode. The same can be concluded about the average word length. The average word length is higher for the annotations made in Normal mode when compared to 2D and AR. The fact that the annotations made in Normal mode have more and longer words could indicate that users are more focused on the contents and characteristics of the paintings as opposed to the technology to view paintings on a mobile device. One aspect that would support this view would be that the screen size of the mobile device limits the participant in their ability to have a full overview of the painting.

This could also be an explanation for the distribution of the annotations relative positions for each mode. The results indicate that the annotation created in Normal mode describe elements and characteristics spanning larger parts of the painting in comparison to the annotations created in 2D or AR. Although the results indicate that the positions of the annotations in Normal and AR mode are more comparable than the annotations made in 2D. The participants annotated elements that were in a specific location as opposed to elements that spanned larger parts of the painting. This could be due to the limited view of the painting when zoomed in.

From the responses on the questionnaire, I found that the participants enjoyed using AR mode the most although indicating that this mode was least efficient. The main reasons indicated where the instability of the painting and the blocking of view by the text field.

These results show that one of the limitations of this research was the use of a small screen (5.8 inches), to display virtual painting, which affects the annotations being made. Further research could be done with devices containing larger screens like tablets. An example would be Apple iPads which have displays of sizes up to 12.9 inches. This research could compare the results found in this research to the annotations made on these types of devices where the text fields do not fully block the view of the painting. Furthermore, research could be done on improvements of augmented reality technologies to bridge the gap between experiencing real and virtual objects.

Other limitations of this research were the use of printed painting representations, the use of a generic image description framework, the limited relative size of the virtual painting in AR mode and most importantly, the limited sample size. In further research larger sample sizes could be used to find statistical proof of the suitability of AR for creating annotations. This sample size could be chosen to consist of a group of people representative for the population of a crowdsourcing initiative for annotating objects. Alternatively, it could be chosen to use a sample with only domain experts for the objects that are annotated.

6   Conclusion

The task of creating annotations using augmented reality was not studied before. This research set out to find an answer to the question if current augmented reality on mobile

devices is suitable for creating such annotations. The results gathered from the experiment using the prototype app AR Annotator do not enable me to provide closing proof nor rejection of my hypothesis that annotations made using AR mode are more comparable to Normal mode than the annotations made in 2D mode. Whereas the majority of participants indicated that experience of annotating in AR mode indeed was comparable to annotating in Normal mode and the distribution of the relative positions of the annotations of AR mode closers resembled Normal mode than 2D mode. However, the derived parameters and metrics of the annotations indicate the opposite. Namely, that annotations made in 2D mode are more comparable to annotations made in Normal mode than in AR mode. Furthermore, the sample size of the experiment does not allow for a statically solid answer if augmented reality is suitable for the task of creating annotations on mobile devices. Further research has to be done with larger sample size to provide a statistical answer. Also, the usability aspects of applications using AR has to considered for further research. Although, when considering advances in augmented reality technology in the future and the massive adoption of mobile devices the question if augmented reality is a suitable alternative for all sorts of tasks will become increasingly relevant. I hope that this research inspires further investigation on this relatively new subject.

Acknowledgements

References

1. Pan, Z., Cheok, A. D., Yang, H., Zhu, J., & Shi, J. (2006). Virtual reality and mixed reality for virtual learning environments. Computers & Graphics, 30(1), 20-28.
2. Boud, A. C., Haniff, D. J., Baber, C., & Steiner, S. J. (1999). Virtual reality and augmented reality as a training tool for assembly tasks. In Information Visualization, 1999. Proceedings. 1999 IEEE International Conference on (pp. 32-36). IEEE.
3. Morrison, A., Oulasvirta, A., Peltonen, P., Lemmela, S., Jacucci, G., Reitmayr, G., ... & Juustila, A. (2009, April). Like bees around the hive: a comparative study of a mobile augmented reality map. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1889-1898). ACM.
4. Rose, E., Breen, D., Ahlers, K. H., Crampton, C., Tuceryan, M., Whitaker, R., & Greer, D. (1995). Annotating real-world objects using augmented reality. In Computer graphics (pp. 357-370).
5. Wither, J., DiVerdi, S., & Höllerer, T. (2009). Annotation in outdoor augmented reality. Computers & Graphics, 33(6), 679-689.
6. Rekimoto, J. (1998, July). Matrix: A realtime object identification and registration method for augmented reality. In Computer Human Interaction, 1998. Proceedings. 3rd Asia Pacific (pp. 63-68). IEEE.
7. Oosterman, J., Dijkshoorn, C., Leyssen, M. H. R., Traub, M. C., & Nottamkandath, A. (2012). Accurator: nichesourcing specific knowledge for cultural hertitage institutions.
8. Hollink, L., Schreiber, A. T., Wielinga, B. J., & Worring, M. (2004). Classification of user image descriptions. International Journal of Human-Computer Studies, 61(5), 601-626

Appendix A

Instructions experiment

Dear participant,

This experiment is part of a bachelor project with the goal of determining the current state of mobile Augmented Reality. With the increase of computing and graphical performance of mobile devices, new use cases emerge for augmented reality. The goal of this experiment is to test if current mobile augmented reality applications offer a comparable experience as a real-life experience. And by doing so could be a better alternative for certain tasks in comparison with 2D based applications as mainly used on mobile devices.

For this experiment, the task will be annotating a painting. This means that you are asked to describe the contents and/or characteristics of the selected painting. The task of annotating will consist of three parts, each part with a different painting.

Note the order of the parts will vary.

1. Normal mode: The first part consists of annotating a real-life painting (or picture containing the painting by adding annotations in the input field.

2. 2D mode: The second part consists of annotating a 2D image version of a painting displayed on the phone. And adding annotations to the image by long pressing the location for the annotation.

3. AR mode: In the third part, you will experience a painting using augmented reality. The application will display a virtual easel and the painting. You are asked to provide annotations for the painting.

4. Questionnaire: Finally, you will be asked to fill in a questionnaire about you experience annotating paintings.

ParticipantID: _____   Signature: _____

Disclaimer

For this experiment, the provided annotations will be stored on a private Database hosted by Google. All the data given will be anonymized and will not be used for any other purpose then the research. During the experiment pictures can be made. These pictures are only used for presentation purposes and to visualize the experiment.

Appendix B

Questionnaire

1. Order your enjoyment of the modes in descending order.
2. Order the efficiency of annotating in the modes in descending order.
3. What did you find good about the experience in Normal mode?
4. What did you find bad about the experience in Normal mode?
5. What did you find good about the experience in 2D mode?
6. What did you find bad about the experience in 2D mode?
7. What did you find good about the experience in AR mode?
8. What did you find bad about the experience in AR mode?
9. How much experience do you have with augmented reality?
10. Do you use augmented reality on a daily basis?
11. Did you find the augmented reality experience convincing?
12. Was the experience of annotating in augmented reality comparable to the experience of annotating in normal mode?
13. How likely is it that you choose augmented reality for future tasks over normal mode?

Appendix C

Program code and links to program code

Github link: https://github.com/JurjenBraam/AR-Annotator

Excel formula

```
=LENGTE(D2)- LENGTE(SUBSTITUEREN(D2;" ";"")) + 1
```

Python code

```python
import json
import csv

with open('file2.csv', 'rb') as csvfile:
  data = csv.reader(csvfile, delimiter = ';')

  for row in data:
    words = row[3].split()
    average = float(sum(map(len, words)))/len(words)
    print "%.2f" % average
```

Appendix D

Annotations

Paint ing	Par- tici- pant ID	Mode	Annotation	Position	Num- ber of words	Aver- age word length	Class
1	1	Normal	Churches	Bottom	1	8,00	Con- ceptual
1	1	Normal	City gate	Bottom	2	4,00	Con- ceptual
1	1	Normal	Clouds	Bottom left	1	6,00	Con- ceptual
1	1	Normal	People walk- ing	Bottom right	3	6,50	Con- ceptual
1	1	Normal	River	Bottom right	1	5,00	Con- ceptual
1	1	Normal	Some ships	Bottom	2	4,50	Con- ceptual
1	1	Normal	View on a dutch city doen across the water	Bottom right	9	3,78	Con- ceptual
1	2	2D	Een paar bo- ten aan de kade, rustig aan de kade is waarschijnlijk geen haven	Bottom	14	4,43	Con- ceptual
1	2	2D	Het lijkt op een stad met een rivier	Bottom right	8	3,62	Con- ceptual

1	2	2D	Kooplieden aan het water	Bottom left	4	5,25	Con-ceptual
1	2	2D	Mooi kasteel-tje, of toe-gangspoort	Bottom	4	7,50	Con-ceptual
1	2	2D	Opvallende witte kerkto-ren valt erg op tussen de grauwe stads-muren	Bottom	10	5,70	Con-ceptual
1	3	AR	Brug	Bottom	2	4,00	Con-ceptual
1	3	AR	Kerktoren	Bottom	1	9,00	Con-ceptual
1	3	AR	Klein kasteel-tje met twee torens	Middle	5	5,60	Con-ceptual
1	3	AR	Twee pra-tende vrou-wen aan de kade	Bottom left	6	4,67	Con-ceptual
1	3	AR	Zeilbootje	Bottom	1	10,00	Con-ceptual
1	4	Normal	Large city in background	Bottom right	4	5,25	Con-ceptual
1	4	Normal	Large ship in harbour	Bottom	4	4,50	Con-ceptual
1	4	Normal	People painter small	Bottom left	6	6,33	Con-ceptual

1			in fore-groundbby river				
1	4	Normal	Possible church in center of painting	Bottom	6	5,33	Conceptual
1	4	Normal	River running past city	Bottom	4	5,00	Conceptual
1	5	2D	Boat	Top right	1	4,00	Conceptual
1	5	2D	Church	Top	1	6,00	Conceptual
1	5	2D	Defense wall	Bottom left	2	5,50	Conceptual
1	5	2D	Moving bridge	Bottom	2	6,00	Conceptual
1	5	2D	Shopping basket	Bottom	2	7,00	Conceptual
1	6	2D	City life in medieval times	Bottom	5	4,60	Conceptual
1	6	2D	Dutch Landscape	Bottom	2	7,00	Conceptual
1	6	2D	Golden age	Bottom	2	4,50	Conceptual
1	6	2D	Large amount of detail	Bottom left	4	4,75	Perceptual
1	6	2D	Large amount of light	Bottom right	4	4,50	Perceptual

1	6	2D	Peasants‚Äô and Work-ers‚Äô life	Bottom	4	9,00	Con-ceptual
2	1	2D	Fire	Top right	1	4,00	Con-ceptual
2	1	2D	Funnel	Bottom	1	6,00	Con-ceptual
2	1	2D	Jesus	Top left	1	5,00	Con-ceptual
2	1	2D	Mouth	Top right	1	5,00	Con-ceptual
2	1	2D	Tent that is Also a hat	Top right	6	3,00	Con-ceptual
2	2	Normal	Druk en chaostisch schilderij	Bottom left	4	6,50	Per-ceptual
2	2	Normal	Een heilige links die in gevecht lijkt met demonen	None	9	4,67	Con-ceptual
2	2	Normal	Een man die gevild wordt linksonderin	Bottom left	6	5,33	Con-ceptual
2	2	Normal	Een man rechtsboven die opgehan-gen is	Middle	6	5,33	Con-ceptual
2	2	Normal	Links van het midden een bankier(?) die geld leent aan iemand	Middle	11	4,64	Con-ceptual

2	3	2D	Dobbelstenen	Top right	1	12,00	Conceptual
2	3	2D	Monstertje	Bottom right	1	10,00	Conceptual
2	3	2D	Overleg twee mannen aan een tafel	Middle	6	4,67	Conceptual
2	3	2D	Soldaat	Top right	1	7,00	Conceptual
2	3	2D	Soldaat die mensen vermoord	Top right	4	6,00	Conceptual
2	4	AR	Burning building	Bottom right	2	7,50	Conceptual
2	4	AR	Gigantic funnel	Top left	2	7,00	Conceptual
2	4	AR	Lot of naked people	Top left	4	4,00	Conceptual
2	4	AR	Man hanging from some sort of fruit possibly	Bottom	8	4,62	Conceptual
2	4	AR	Weird man creature with open mouth	Bottom left	6	4,83	Conceptual
2	5	AR	Aerobics	Middle	1	8,00	Conceptual
2	5	AR	Bugs	None	1	4,00	Conceptual
2	5	AR	Burning	Middle	1	7,00	Conceptual

2	5	AR	Casino	Top left	1	6,00	Conceptual
2	5	AR	Hungry	None	1	6,00	Conceptual
2	6	AR	Agony	None	1	5,00	Conceptual
2	6	AR	Large amount of detail	Top left	4	4,75	Perceptual
2	6	AR	Physical pain	Bottom left	2	6,00	Conceptual
2	6	AR	Religious motives	Top right	2	8,00	Conceptual
2	6	AR	Torture motive	Bottom left	2	6,50	Conceptual
3	1	AR	Beard	Bottom	1	5,00	Conceptual
3	1	AR	Feather	Bottom right	1	7,00	Conceptual
3	1	AR	Helmet	Bottom	1	6,00	Conceptual
3	1	AR	Spear	Top	1	5,00	Conceptual
3	2	AR	De mannen in het midden zijn in gesprek, de linker vertelt	Bottom	11	4,36	Conceptual
3	2	AR	Een trommelaar rechts op het schilderij	Bottom right	6	5,67	Conceptual

3	2	AR	Er ligt een dier op de grond, lijkt op een hond	Bottom right	11	3,36	Con-ceptual
3	2	AR	Het is de Nachtwacht van Rem-brand	All	6	4,67	Non-visual
3	2	AR	Links loopt een jongen weg met iets in zijn han-den, geen idee wat	Bottom left	13	4,08	Con-ceptual
3	3	Normal	Een hond	Bottom right	2	3,50	Con-ceptual
3	3	Normal	Engeltje	Bottom left	1	8,00	Con-ceptual
3	3	Normal	Man met trommel	Bottom right	3	4,33	Con-ceptual
3	3	Normal	Rode ge-kleede man met wapen	Bottom left	5	4,60	Con-ceptual
3	3	Normal	Twee leiders op de voor-grond	Bottom	5	4,80	Con-ceptual
3	4	2D	Duke or other important person talking	Bottom left	6	5,50	Con-ceptual
3	4	2D	Ghost dog maybe not very clear what this is	Bottom right	9	3,89	Con-ceptual

3	4	2D	Girl bathing in white light	Bottom left	5	4,60	Con-ceptual
3	4	2D	Green and yellow banner	Top left	4	5,00	Con-ceptual
3	4	2D	Musical in-strument	Bottom right	2	8,50	Con-ceptual
3	5	Normal	Discussions	All	1	11,00	Con-ceptual
3	5	Normal	Feathers	Bottom left	1	8,00	Con-ceptual
3	5	Normal	Girl	Bottom left	1	4,00	Con-ceptual
3	5	Normal	Pointing fin-gers	Bottom right	2	7,50	Con-ceptual
3	5	Normal	Richness	All	1	8,00	Con-ceptual
3	6	Normal	Aristocratical and military clothing	All	4	8,25	Con-ceptual
3	6	Normal	Dark colours and atmos-phere	All	5	6,00	Per-ceptual
3	6	Normal	Large detail of facial ex-pressions	All	5	6,00	Per-ceptual
3	6	Normal	Lively with lots of move-ment painted	All	6	5,17	Per-ceptual
3	6	Normal	Military pa-rade	All	2	7,00	Con-ceptual