Heuristic Guidelines for Wearable Augmented Reality Applications

Nathan Gale
UOIT
2000 Simcoe St. N.
Oshawa, ON, L1H 7K4 Canada
Nathan.Gale@uoit.ca

Pejman Mirza-Babaei
UOIT
2000 Simcoe St. N.
Oshawa, ON, L1H 7K4 Canada
Pejman.m@acm.org

Isabel Pedersen
UOIT
2000 Simcoe St. N.
Oshawa, ON, L1H 7K4 Canada
Isabel.Pedersen@uoit.ca

CHI PLAY 2015, October 03-07, 2015, London, United Kingdom
ACM 978-1-4503-3466-2/15/10.
http://dx.doi.org/10.1145/2793107.2810309

Abstract
In recent years, new wearable platforms and peripherals requiring unique modes of interaction have been emerging in record numbers. Watches, necklaces, glasses, even pants [1] are beginning to incorporate technology. Wearable devices also offer many new ways for users to interact; therefore, more research is needed to evaluate these novel methods of interactions. Our research focuses on proposing a heuristic list to guide design and evaluation of wearable devices in playful applications. In this work-in-progress (WiP) article we show a prototype of a unique, playful, interaction application, *META Museum*, and propose the development of a heuristic list to evaluate it. This WiP reports on the early evaluation results based on user's inputs and discusses some common mistakes that users make and in what way this relates to how the application is designed.

Author Keywords
Augmented Reality; Heuristic Evaluation; Wearable Technology.

ACM Classification Keywords
H.5.2. Information interfaces and presentation: User Interfaces – Evaluation/ Methodology.

Introduction

The development of heuristic evaluations for wearable augmented technology, to date, has been relatively limited. Some heuristic lists that exist for augmented reality are too focused on a single utilization and are therefore not usable by general applications, such as Wearables User Experience Heuristics for the Enterprise [5]. A heuristic that focuses on Augmented applications is needed, as using existing, more general, heuristics to evaluate their usability could lead to weak inter-rater reliability, as is shown White et al [9]. One of the more comprehensive and broad heuristics for augmented reality, Developing Usability Evaluation Heuristics for Augmented Reality Applications [2], is, likewise, not currently validated. By reviewing current heuristic lists, we are working towards proposing a set of heuristics specific to evaluation for playful wearable-based applications. This WiP reports on our initial heuristic lists and showcases an application that has been developed as an evaluation case in order to expand, test, and validate our heuristic list. The ultimate goal is to create a validated set of heuristics that can reliably be used to test wearable augmented reality applications.

Background

Augmented Reality (AR) has taken a backseat to Virtual Reality (VR) in the public eye in terms of gaming applications. VR headsets such as the *Oculus Rift* have gained massive popularity while AR headsets like *Google Glass* faded from the limelight. Although AR has many uses, much of the development has been focused on more business and social aspects while VR has focused on the gaming scene, which has caused much more excitement and interest. Due the fact that it is fully immersive, VR is seen as an easier way to create a full gaming experience without having to overcome the constraints that come with having to take into account the real world [8]. With upcoming devices like the *Microsoft HoloLens*, AR will potentially grab everyone's attention again.

Meta Museum

The application being developed is designed to augment museum spaces (see figure 1). A primary goal is to include as many features that the chosen AR headset allows. This makes it suitable for testing and improving a heuristic list. The application includes QR code object generation, object recognition, finger and gesture recognition, and audio.

In its current state, the application has the ability to recognize QR codes on objects and generate pop up information on them. Users can manipulate this information with natural gestures such as swiping, pressing, pinching, and grabbing. In addition to the interactions with on-screen objects there is also the use of audio and visual cues to allow off-screen interactions. With these cues users will realize that the augmentation is not only on the objects they are looking at, but all around them.

META Museum adds a playful experience to museum spaces by adding interaction to a space commonly meant just for viewing. Users are able to grab and play around with virtual representations of artifacts and interact with simulated versions of historical figures as well as scroll through additional information about the museum objects.

The application is being developed in accordance with the heuristic list we are developing and will be used as

a case to improve our working list of heuristics. The application has been evaluated in accordance with our generated list and other relevant preexisting lists (see Current Heuristic Progress). The knowledge gained from these evaluations will result in improvements to both the META museum application and the heuristic list itself.

Figure 1. A participant trying the META Museum experience with the use of the META Spaceglasses developer kit.

The overall development of the heuristic list and prototype adheres to the following four stages:

- Stage 1: Review relevant heuristic evaluations and gather heuristics important to wearable augmented reality.

- Stage 2: Rapid test/development with a prototype using both user testing and professional heuristic experts. This will include onsite testing at the Royal Ontario Museum (ROM).

- Stage 3: Refine list and prototype with knowledge gained from the tests. This is where new items will be added or removed from the list and will influence design of the application.

- Stage 4: Repeat until complete.

Current Heuristic Progress

For the development and validation of the heuristic, we will follow the design proposed in Rusu et al [7] and is currently in the explicative stage.

The development of the heuristic began by looking at multiple heuristics that could be related to use in playful AR applications. The *Nielsen Heuristic* [4] was the initial starting point, but upon research into other heuristic lists we found the most useful and relatable to AR was the *Playability Heuristics for Mobile Games* proposed by Korhonen and Koivisto [3]. The reason for choosing the Korhonen and Koivisto heuristic as a starting point was both because it incorporates the Nielsen heuristic and because it extends to include an awareness of the physical space around the user. The mobile games heuristic is also a highly cited paper leading credence to its evaluation. Similar to mobile devices, AR devices "do not dictate where and when [they] are played" [3].

By examining and evaluating the META Museum application with the use of this previous heuristic, certain items on the list were found to be irrelevant for an AR experience. The major heuristics that were investigated, alongside the Nielsen and Mobile Games heuristics, were the SoPlay Heuristics [6], Wearable User Experience Heuristic [5] and Usability Evaluation Heuristics for Augmented Reality [2]. Investigation into

the Wearable User Experience Heuristic helped to inform choices made to acknowledge user safety. The heuristics that were discarded related mainly to game progression and story. Though it is possible for playful AR experiences to contain these aspects, they are not requirements and thus take away from the broadness that we are trying to achieve.

Although this current iteration of the heuristic list refers to games and players it could be applied to playful wearable AR applications. See table 1 for our initial proposed list of heuristics. Below we provide some details/explanation on some of the violations we found to be more applicable to wearable AR. Through information gathered from testing with users during conferences as well as observations by those with expertise in heuristic evaluation, we have discovered that these violations occurred most often (see Table 1):

- GU4: Visibility of indicators – GU4 becomes more important when the device is wearable, as when a user is wearing the device they may not be able to tell if it is functioning properly and must therefore be displayed appropriate **feedback**.

- GU7: Consistent interactions with control keys– GU7 holds significance because of the lack of any formal standard controls for AR at the moment. For development of the application we have tried to keep **interactions** close to how a person would naturally interact with similar objects in the real world. We have incorporated natural gestural movements like waving, grabbing, and touching to make the application instinctively usable.

Wearable Augmented Reality Heuristics	
ID	**Definition**
GU1	Audio-visual representation supports the game
GU2	Screen layout is efficient and visually pleasing
GU3	Device UI and game UI are used for their own purposes
GU4	Indicators are visible
GU5	The player understands the terminology
GU6	Navigation is consistent, logical and minimalist
GU7	Control keys are consistent and follow standard conventions
GU8	Game controls are convenient and flexible
GU9	The game gives feedback on the player's actions
GU10	The player cannot make irreversible errors
GU11	The player does not have to memorize things unnecessarily
GU12	The game contains help
MO1	The game and play sessions can be started quickly
MO2	The game accommodates with the surroundings
MO3	The interruptions are handled reasonably
GP3	The players are rewarded and rewards are meaningful
GP4	The player is in control
GP6	The first time experience is encouraging
GP8	There are no boring or repetitive tasks
GP9	The players can express themselves
GP10	The game supports different playing styles
GP11	The game does not stagnate
GP12	The game is consistent
WT1	Play length should match comfort
WT2	Allow as many modalities as is sensible
WT3	Put user safety first

Table 1. Our initial Wearable AR heuristics list, generated from the heuristics for mobile games and our own summary suggestions to create a base list. The list consists of game usability (GU), mobility (MO), gameplay (GP) and our own addition wearable technology (WT).

- MO2: AAccommodation with surroundings - As it relates to AR, MO2 can include how the generated objects interact with the world, but more importantly how a user is able to interact with gestural motions in a given space. To **accommodate** users in our application we have included both swiping as a more natural movement when given enough space and buttons to press when there is not enough room to swipe.

- GP10: Support for diversity in playing styles – GP10 covers the fact that users may try to interact with similar things differently, or may not be able to interact in certain ways, and must therefore be allowed to interact differently. This adds **diversity** to the application so it can reach a wider audience. For our application we have put thought into the users who may not be able to use their arms. For this we have the ability to do a majority of the actions simply with staring.

In addition to the important heuristics that were adapted from exiting lists there are three that we explicitly propose for wearable technology:

- WT1: Comfort in play length – WT1 is important so that the device does not become uncomfortable during the users play period. Testing should be done on a per device basis, and applications for these devices should be designed around the level of **comfort**. Our application is designed for a short experience with a specific exhibit or piece of art. Players are able to enjoy the experience in multiple separate interactions with breaks from the device, or if comfortable to them experience it all together.

- WT2: Enablement of many modalities - User's should be able to interact with the augmented environment as closely as possible to the real world. WT2 means including **multiple interaction methods** for single interactions. Although it may be impractical to replicate every possible interaction it is important that common, natural interactions should be possible.

- WT3: Safety - With the possibility of AR being used in public places it is important that the device does not hinder or annoy users during tasks that require their attention. For example, should a user trip and fall, the screen should become blank so as to not obscure vision and worsen the fall. This is to make sure the **safety** of a user is never in peril because of the device or application.

This list of heuristics serves as a starting point that will be iterated upon as the accompanying application is too iterated upon.

Planned Study Design – next stage

As previously stated, up to this point we have developed a heuristic list and application to help improve it. We have done user testing at multiple conferences as well as internal testing.

The planned design of the study is to invite two separate user groups with similar user experience

levels to evaluate the application, with one group using our developed heuristic list and the other group using the Nielsen heuristic, which is a good benchmark for any design of user interface. This is in accordance with the heuristic development proposed in Rusu et al [7]. This study will be run multiple times throughout the life of the project, comparing different list and gathering together a final heuristic list.

Users of both groups will record which heuristics are being violated and this data will be collected and compared. Results should show that more violations are identified by the group using the wearable tech list when compared to the preexisting list. The study can be run comparing any existing heuristic list to our list to uncover any discrepancies and common problems that may have been overlooked.

Conclusion

The contributions by this paper include the development and validation of a heuristic for wearable augmented reality user interfaces as well as the development of an augmented reality application designed to create an immersive museum experience

Acknowledgements

We thank the engineers at META for allowing us to get an early version of the developer's kit and the participants at the We Are Wearables show whose observed interactions and informal suggestions helped develop a starting point for the heuristic list and application improvements. We thank the Royal Ontario Museum. We acknowledge the funding of NSERC and the Canadian Foundation for Innovation Fund (CFI).

References

[1] Dorrier, J. Forget Google Glass – It's All About Google Pants. (May 31, 2015). http://singularityhub.com/2015/05/31/forget-google-glass-its-all-about-google-smart-clothes/

[2] Kalahati, J. Developing Usability Evaluation Heuristics for Augmented Reality Applications. (2015).

[3] Korhonen, H., & Koivisto, E.M.I. Playability Heuristics for Mobile Games. Mobile HCI '06 (2006).

[4] Nielsen, J., & Molich R. Heuristic Evaluation of User Interfaces, Proc. ACM CHI '90 Conf. (Seattle, WA, 1-5 April 1990), 249-256.

[5] O'Broin U. Wearable User Experience Heuristics for the Enterprise 1.0. http://www.slideshare.net/uvox/user-experience-heuristics-for-wearables-in-the-enterprise

[6] Paavilainen, J. SoPlay Heuristics for Design and Evaluation of Social Network Games. CHI 2013 (2013).

[7] Rusu C., Roncagliolo S., Rusu V., & Collazos, C. A Methodology to Establish Usability Heuristics. ACHI 2011 (2011).

[8] Schick, S. Why developers should be making a play for AR and VR mobile games ASAP. (April 13, 2015). http://www.fiercedeveloper.com/story/why-developers-should-be-making-play-ar-and-vr-mobile-games-asap/2015-04-13

[9] White, G. R., Mirza-Babaei, P., Mcallister, G., and Good, J. "Weak Inter-rater Reliability in Heuristic Evaluation of Video Games." In *CHI '11 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '11). ACM, New York, NY, USA, 1441-1446. DOI=10.1145/1979742.1979788