A taxonomy of handheld augmented reality applications

Tor Gjøsæter
Department of Information Science and Media Studies
University of Bergen
Bergen, Norway
Tor.Gjosater@infomedia.uib.no

Abstract— **This paper provides a straightforward layer based taxonomy for categorizing research and development in handheld augmented reality (HAR) applications. The taxonomy provides a framework for articulating and differentiation of research and development specific to HAR applications within a visual model. A literature review is used to place examples of previous research into the model. In addition we show two use cases where we illustrate how we use the taxonomy to position our own current research into the model. We conclude that the model enables a more focused view on topics specific for HAR within the field of AR.**

Index Terms—**Handheld augmented reality, model, layers, taxonomy**

I. Introduction

Augmented reality (AR) as a research field is currently flourishing and has grown extensively over the last ten years [1]. With the advent of ever more powerful mobile devices, interest into handheld augmented reality (HAR) is growing. This paper aims to provide a general taxonomy of handheld mobile augmented reality applications to frame ongoing-, previous and future research into this specific domain.

The premise of this paper came from an explicit need to frame our own research into HAR in a consistent manner. A model to communicate where our research into HAR had its focus points was needed to pinpoint where our research interest and contribution to field was.

We present a taxonomy in this paper we use as scaffolding when we demonstrate the relation and contribution our user studies and application development has to the handheld augmented reality field. In some cases HAR studies does not fit into, or envelops too broadly in the proposed "…five main research areas in augmented reality

- (1) Tracking techniques
- (2) Interaction techniques
- (3) Calibration and registration
- (4) AR applications
- (5) Display techniques

…" [1] found by Zhou et al. The field may benefit from a taxonomy that takes into account these categories, but adds specific categorization aimed at and directly related to a generic HAR application.

II. Method

A common approach to understanding the augmented reality research field is to review literature published under ISMAR. In a "Survey of user-based experimentation in augmented reality" [2] Swan et al gave an overview of how usability experiments where performed at ISMAR as well as in other scientific communities. The scope of that study was to survey how user-based experimentation manifests itself in the AR community. Most of the studies Swan investigates are informal in nature; however they provide a reference list for exploring interaction in AR and are studies concerned with interaction specifically. This paper will use examples from relevant literature in the field of both AR and HAR that illustrate the layers in the proposed model (Fig. 1). Some groundbreaking research that benefits the field of HAR does not necessarily emerge from a handheld device, however the studies that emerge from the general AR research sited in this paper translate directly from the general AR paradigm to the HAR paradigm. Illustrations, figures and citations will be used to clarify where different AR and HAR research and fit into the taxonomy.

III. Scope

The scope of this taxonomy model is to encompass applications and interaction that can be said to be HAR and "...have the following three characteristics: 1) Combines real and virtual 2) Interactive in real time 3) and Registered in 3-D" [3]. It is important to note that the model proposed in this paper does not attempt to frame other subcategories of mobile AR like wearable AR but rather focus specifically on HAR. HAR applications can be recognized by that they employ a registration device like a camera or other sensors in a handheld form factor to augment the reality on a display mounted on the same device.

We are aware that this model may be at some level applicable to a general taxonomy of AR, however we believe that the niche of HAR is a long-term prospect and is similar but different from AR. The similarities come from the main research areas stated by Zhou et al [1], while the differences come from the form factor, and technological limitations in, and context of use for handheld augmented reality. Some argue [4] that the pinnacle of AR can never be achieved through HAR as its immersiveness may be severely limited. This may be true if your goal is immersive AR, however the utility of handheld AR has not yet met its potential as a niche within the field, and total immersiveness may not be a goal in itself. We believe that HAR will not cease to exist in its current arrangement, however by the ongoing improvement and understanding of the different properties of HAR may lead to better arrangements and future applications.

978-1-4673-4369-8/12/$31.00 ©2012 IEEE

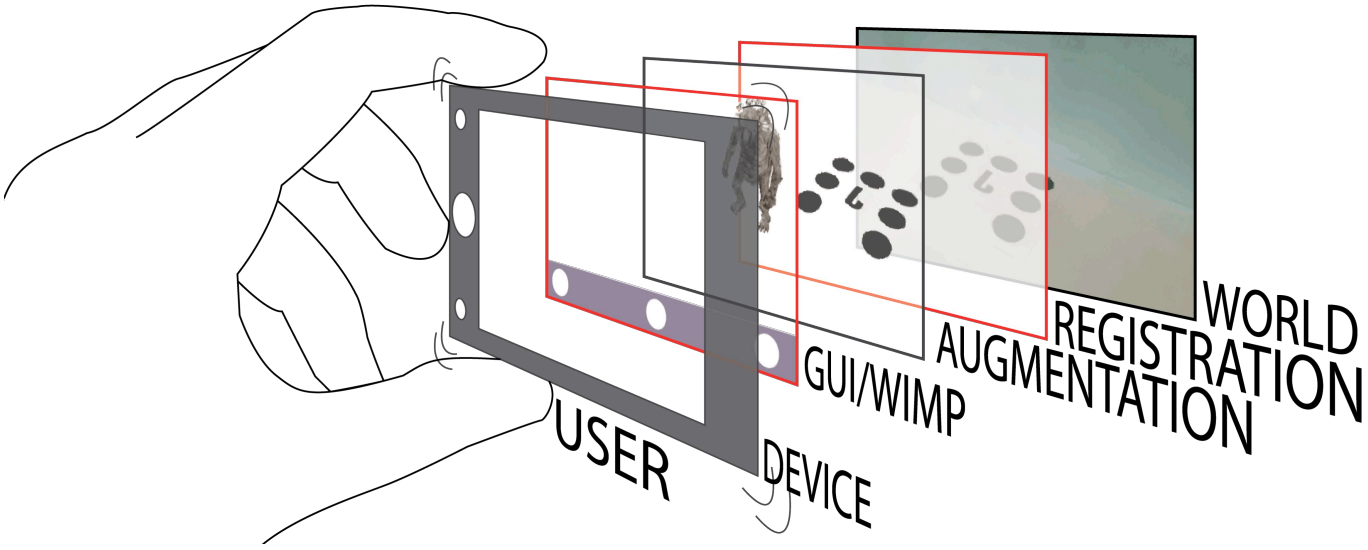

Fig. 1. A visual representation of a taxonomy of handheld augmented reality

IV. Taxonomy of HAR

Figure 1 illustrates our taxonomy and how we separate a HAR application into different layers. Its outer layer is the World layer, following from right to left Registration, Augmentation, GUI/WIMP, Device and finally User. The purpose of the model is to connect the visual and tangible properties of a HAR application with a visual coherent model grounded in published material surrounding the AR and HAR field. The taxonomy proposes a model for describing all the aspects of a HAR system from the user, through the design and technical aspects to the world where the system will operate.

The following sections will describe the layers and their rationale for being in the model by referencing studies that directly fits in the taxonomy. It is difficult to predict future applications of HAR, however we believe that studies cited in this paper illustrates the current focus of research. The current research enables us to speculate about fruitful future research into the topics native to the layer in the taxonomy.

A. World

The world layer is categorized by research and design that consists of trying to improve or understand how objects that exists in real environment space affect the mobile AR application. This may constitute, but is not limited to research into design and affordance of markers like improving the usability of markers or optimal placement of markers in ads or on posters. Schmalsteig et al [5] reports on being mindful of the physical affordance markers have on the user experience and how it for instance affects gameplay on handheld devices. There will always be something in the world to track (otherwise we are not augmenting the world and moves toward a VR state), and this object whatever it may be will exploit some affordance, however it is difficult to predict what. Some recent research into the combination of worldly robots in combination with AR have great promise [6] and opens up new ways of looking at properties of worldly objects in combination with AR.

In the case of Fig. 1 the "world" layer shows a simple marker on a white A4 sheet of paper. There is no additional graphics surrounding the marker or in the view.

B. Registration

This layer is responsible for illustrating the registration of world that is an important part of a HAR application. Within the field of registration we encompass research that deal with the hard science of registration. This encompasses user input by the common manipulating of marker, natural feature or any other tracked surface or object. Within technical computer vision Wagner et al [7] stands out with a self-contained handheld AR system developed as early as in 2003. In addition we mention other approaches to user input like registration and tracking of fingertips and gestures [8] and for instance letting the user sketch on pieces of paper for the application to interpret [9]. Other interesting contributions may be Lang et al [10] with their use of accelerometer and gyroscope to achieve inertial tracking on a mobile device. Other types of input for registration will emerge as the new technology to sample the world in real time is discovered.

In Fig. 1 the registration layers is illustrated by using a marker subtracted from the background, this is familiar reference in the AR research community.

C. Augmentation

Whereas the registration layer deals with interpreting the world, this layer is concerned with the representation of the input received from the registration layer. How content and augmentations is presented and perceived [11] is an ongoing effort within the field of HAR. The augmentation is crucial as it gives the application its purpose. In some cases different approaches to rendering can improve the usability of the application by adjusting the quality of the displayed image [12] or trying to achieve greater photorealism [13]. This layer borrows extensively from the current state of the art in real time visualization of 3D graphics. By virtue of existing on a handheld device, HAR applications inherit the current technological constraints as well as the mobile user in action. How we can create meaningful augmentations to mobile users

TABLE I. CLASSIFICATION OF RESEARCH RELEVANT TO HAR

Study	Topic	U	D	GUI	A	R	W
Kruijff et al [17]	Proposed development of device for handheld AR. A tech note on how to create an ergonomic device (D).		◆				
Veas et al [16]	Evaluation and the design of Vespr. This paper includes a user study (U) focusing on the user experience of the device (D).	◆	◆				
Nishina et al [13]	More realistic representation of augmentations. In this study the author illustrates a setup where you can create more photorealistic augmentations (A).				◆		
Klein et al [12]	Camera adjusted representations of augmentations. This study enhances the augmentations by aligning their look and feel closer to how the camera samples. We argue that this research use the knowledge of (D) to improve (A).		◆		◆		
Hürst et al [8]	User study of gestures in mobile AR. This study shed light on the users (U) satisfaction with registered (R) gestures in relation to the world (W) and the presentation (A) of the tracked gestures.	◆			◆	◆	◆
Lang et al [10]	Technical paper describing registration (R) using inertial tracking on mobile devices.					◆	
Wagner et al [7]	First handheld AR, focus on registration (R) and representations of augmentations (A).				◆	◆	
Hagbi et al [9]	Registration (R) of user input via sketching.					◆	
Wagner et al [11]	How users (U) perceive and relate to humanoid augmentations (A).	◆			◆		

on limited hardware in different environmental, social and cultural contexts is interesting topics for future research.

A 3D model of troll figure projected in 3D represents the augmentation layer in Fig. 1.

D. GUI/WIMP

This layer is concerned with the understanding of how to translate knowledge from the GUI and WIMP paradigms to an AR interface. Buttons and direct onscreen input methods are still needed to interact with most AR applications. Wagner et al [7] application running on the IPAQ in 2003 show buttons named “show map”, “options” and “exit” put on top of the screen similar to a menu on the desktop paradigm. These overlays are common and can be found in several of the studies cited in this paper. In Schall et al [14] paper they show a figure where different tools can be select from a pop up overlay. However, we find vanishingly little rationale for the context information some studies [15] provide when showing the GUI overlays in their applications.

While the field of AR may want to move towards computer vision (CV) based interfaces employing novel gesture tracking in relation to the augmentations [8], variations of the WIMP (windows, icons, menus and pointer) paradigm are still ubiquitous on handheld devices and used as tools for navigation in program structures. A better understanding of how we can use this paradigm to support interaction with the augmentations and features of a HAR application is important. Research into this may be seen as stale or even to a great degree “solved” and is perhaps less interesting than developing groundbreaking features of HAR, yet we find it important to not neglect the obvious need in some applications to return to traditional paradigms to enable a better user experience.

Some simple buttons and a frame represent this layer in the model in Fig. 1.

E. Device

Multi-touch smartphone enabled devices - with their benefits and limitations - is the current platform for most of the research into HAR. This form and standardized electronics comes with inherent limitations. Research into different form factors and novel use of electronics, camera [12] and so forth is crucial to further the field. All innovative technology embedded in handheld devices with a display is relevant to this layer. Ranging from the development of electronics like processors, technology for range finding and novel sampling of the world, new inputs like buttons or sticks and output like vibration. In addition we find that ergonomics has always been an issue in AR in general, from the clumsy head mounted displays to the lacking display qualities of mobile devices. Research from Veas et al [16] and Kruijff [17] gives an example of focused research on ergonomics and input methods of the Vesp´r device more than the AR application itself.

The illustration shows the device with a display, including buttons and vibration in Fig. 1.

F. User

This layer represents the focus on users and user studies within the realm of mobile AR. With HAR being an emerging field, Swan et al [2] shows that formal and deep user studies on AR applications is an integral part of the research contribution to the AR field in general. While innovations in software and technology increase the potential in HAR applications we need to understand how this impacts the use and user experience. More user studies just looking at the overall qualities of an application and user studies focusing on distinct parts of the application is needed to further our understanding of HAR.

A hand symbolizes the users input and interaction with the HMAR device and interface in Fig. 1.

V. CLASSIFICATION

In table 1 we have categorized a selection of the studies cited in this paper by categorizing what areas of HAR the studies shed light on. The selection was done to illustrate the extensive range of HAR research and how the model allows addressing of these important topics within the field.

The abbreviations U, D, GUI, A, R and W reflect the layers in the model User, Device, GUI/WIMP, Augmentation, Registration and World, and the bullet points signify the focus area(s) of the study.

Some studies are technical and provide algorithms or technical descriptions of how to utilize an algorithm, whereas other are focused towards understanding the users perception of augmentation or how they interpret the affordances of tools

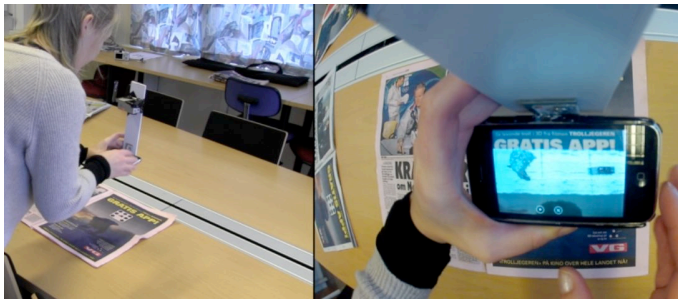

Fig. 2: View from a recorded think aloud session focusing on the users perception on the GUI (small icons on the mobile device)

provided. While the studies commonly focus on one major topic they commonly touch upon several aspects of HAR when exploring their main topic. The primary focus may be to create or improve new registration possibilities, but it is difficult to determine the users perception of the innovation without user studies.

None of the studies target GUI, however in our literature review we find several references [18] to applications creating GUI's for their applications. However, there lacks any detailed rationale behind the design choices and its inception often mentioned in a side note and to a small degree discussed in the studies. We find little deliberation in the literature in regards to how the classical GUI overlay affects the user experience or if it is even understandable. We do not believe a HAR application that employs a classical GUI to support interaction is inferior to one that has; on the contrary we believe it may be necessary in some instances. However, we found no studies focusing mainly on this aspect of HAR. We would like more light to be shed on how one successfully fuse the post-wimp interfaces in HAR with the traditional GUI paradigm.

VI. USE CASES

In this section we will apply our model taxonomy to two use cases of ongoing research to illustrate how we frame our research. The two studies where completed using design research as a framework [19]. The illustrations below consist of data from gathered from the evaluation step in the design research process. The cases use video based research in combination with think aloud to allow us to observe both the user, and the application in action by synchronizing the video streams. This method allowed us to look at the direct interaction with the artifact use the transcripts of the think aloud to infer how the users perceived the application.

A. ARad

Figure 2 shows video captured from a usability study where we focused in particular on the GUI/WIMP of the HAR application and how the user interpreted this GUI to play a small game placed on a marker in a newspaper. The application we studied is a commercially available application for iOS we developed called ARad, available for free on the AppStore.

By applying our model to our research we can articulate our contribution to HAR more clearly. We formulate our research effort by using our taxonomy. "How does the quality of our GUI affect the overall user experience of HAR?" By framing our research in this manner, we point directly to the *GUI/WIMP* layer and its relation to the *User* layer.

Fig. 3: Researching the users perception of augmentations in relation to a worldly artifact.

This may seem inconsequential, but as argued previously in some cases we cannot escape the GUI paradigm and we think it is important to understand how this layer affects the user experience.

In addition we can apply our model from a development perspective. Since this application platform also should serve as a mobile commercial platform for delivering AR content to a newspaper in Norway, particularly for advertisement. We believe this model can be helpful in communicating how AR technology works technically on a handheld platform, and how the integration of the different parts of a HAR system creates the user experience. In one instance we found it difficult to communicate the inner workings of AR, particularly the registration of markers to a client unfamiliar with AR technology. We had to go through several sessions to clearly communicate how registration works, and how good registration enables a good user experience. This stemmed from a newspapers desire to print markers as small as possible to save space for other content or advertisements. Too small markers and not enough whitespace around the markers led to poor tracking in our case (Fig. 2). In this particular project the designers of the printed ad, the 3D content creators, developers and customers ordering the AR ad never physically met during the execution of the project, and had a different understanding of how HAR works and should work. By having a model that separates the inner parts of the system into explicit explainable parts - all the way from the *World* through to the end *User* - can be helpful in similar situations where users and clients are unfamiliar with development of HAR solutions.

B. ARTar

Figure 3 shows an still image from a user study done with five users in a masters project [20]. In this particular project the GUI/WIMP was not needed, and hence not implemented. The device running the application was a crude prototype of makeshift components; web camera, lightweight touchscreen monitor and cables connected to a laptop to do the image processing. The problems with the device itself is discussed in the thesis, however, the device is not the focus area of the research.

This research was focused on the art-artifact, a carefully designed tree being augmented in the world. The main focus in this study was on the artifact in the *World* in relation to the *User* and the *Augmentation*. By using the taxonomy we direct our research towards understanding the interplay between the world layer, where the tree and markers lives, the augmentation layer and how we represent the augmentations in direct relation to the tree-artifact in the world layer and how this in its entirety is perceived by the user.

VII. Discussion

Though there exists a few models to describe placement of research and design in the field of HAR they may in some cases not be suited to place research precise in relation to topics relevant for HAR. In this section we will discuss previous work aimed towards creating tools for categorizing and modeling research and development in AR and HAR.

Tools for framing development of AR in general exists, for instance Dubois et al [21]. However, the model proposed in our paper does not attempt to give design advice or design guidelines, rather it resembles a taxonomy and provides conceptual knowledge explicitly about HAR. This may be useful in in particularly design science research aimed at HAR as this model can help in "…identifying essences in the research territory and their relationships" [22]. While it does not give any design advice, as this can be found in user studies, it does give an overview of the components of a HAR system and what must be considered when designing for HAR. For someone that is a novice in the field of HAR, our model proposes a straightforward and illustrative description of the significant parts of a HAR system.

Zhou et al [1] five categories may give the entire field of AR suitable categories to a certain extent to place their research. HAR research benefits from general AR research to a great degree and some of the research in the categories proposed by Zhou et al can be transferred directly to the field of HAR. While these categories can readily be used to categorize general AR research, it does not necessarily capture work aimed distinctly towards HAR. We find the proposed categories by Zhou et al to be either too narrow, by not encompassing device ergonomics for instance, or to wide by categorizing any user study in the "AR Applications" category. We maintain that a taxonomy directly aimed at HAR enables us to discuss the nature of HAR and it characteristics in more specific terms related to its inherent qualities without having to place our research in categories aimed towards the broad field of AR.

Milgrams continuum [23] is quite often used to place entire AR applications on a line between a real environment and virtual environment. This may give a notion of where a study fits within the field of mixed reality, but the focus may be lost in a discussion about a level of "ARness". A World layer in an AR application may be suitable to put very close to the real environment in Milgrams continuum and the augmentation layer can very simplified be put anywhere on the continuum depending on how much of the screen real estate is being used to render it. However, this rhetorical placement we just proposed does little to help the understanding of the interplay between these radically different spheres of topics within HAR.

If we take the two case studies presented in this study and try to place them on Milgrams continuum, they will fall more or less on the same point. However, these two applications differ greatly in how they approach AR research and development. By applying the taxonomy proposed in this paper it allows us to address the distinct qualities of HAR and clearly illustrate where our research effort contributes to the field.

It allows us determine how different features of a HAR system work together to create a fun or useful handheld augmented reality user experience, and we believe it is essential to be aware of these different aspects when evaluating and developing HAR applications.

We find this interplay between the different layers interesting and it enables us to ask questions like how does the augmentation affect the user experience? Does the ergonomics of the device enable better handling and aiming at trackables? Does the metadata provided in the GUI enable the user to understand CV tracked gestures they just performed by interpreting vibration from the device? We find these questions important, and we believe the taxonomy proposed in this paper empower the field to direct their attention and question directly to topics within without HAR without situating research on a line between real and virtual or within categories general to the entire field of AR.

VIII. Conclusion

We conclude that the taxonomy proposed in this paper is a useful addition to already existing models and categories aimed at specifying research and development within the field of AR and HAR. The taxonomy proposed in this paper identify the distinct parts that a HAR system is composed of, and the layer-based model proposed may help researcher to better frame their research in HAR by pinpointing their contribution directly to important topics within the field.

The existing models for framing AR research [1], [23], [21] does not necessarily carry over to HAR research specifically, nor do they address the interplay between the user and the underlying technology. The taxonomy presented in this paper allows a clear demonstration of where the research is focused and enables communication between researchers, designers and developers working with HAR technology. Both hard technical and softer user-focused research may benefit from the model since it allows HAR research to be classified within a taxonomy that takes into account features of HAR separate from other topics within AR.

IX. Future work

As the field matures, and mobile devices expand their possibilities it may very well be needed to revise this taxonomy or expand it accordingly.

References

[1] F. Zhou, H. B.-L. Duh, and M. Billinghurst, "Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR," presented at the ISMAR '08: Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, 2008, pp. 193–202.

[2] J. Swan and J. Gabbard, "Survey of user-based experimentation in augmented reality," presented at the Proceedings of 1st International Conference on Virtual Reality, 2005.

[3] R. Azuma, "A survey of augmented reality," *Presence-Teleoperators and Virtual Environments*, vol. 6, pp. 355–385, 1997.

[4] L. J. Rosenblum, S. K. Feiner, S. J. Julier, and J. E. Swan, "The Development of Mobile Augmented Reality," in *Expanding the Frontiers of Visual Analytics and Visualization*, no. 24, London: Springer London, 2012, pp. 431–448.

[5] D. Schmalstieg and D. Wagner, "Experiences with Handheld Augmented Reality," presented at the ISMAR '07: Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, 2007.

[6] C. K. A. W. Woo, "ARMate: An Interactive AR Character Responding to Real Objects," in *Edutainment Technologies. Educational Games and Virtual Reality / Augmented Reality Applications*, vol. 6872,

Springer Berlin Heidelberg, 2011, pp. 12–19.
[7] D. Wagner and D. Schmalstieg, “First steps towards handheld augmented reality,” presented at the ISWC '03 Proceedings of the 7th IEEE International Symposium on Wearable Computers, 2003, p. 127.
[8] W. Hürst and C. Wezel, “Gesture-based interaction via finger tracking for mobile augmented reality,” *Multimedia Tools and Applications*, pp. 1–26, Jan. 2012.
[9] N. Hagbi, R. Grasset, O. Bergig, M. Billinghurst, and J. El-Sana, “In-Place Sketching for content authoring in Augmented Reality games,” presented at the Virtual Reality Conference (VR), 2010 IEEE, 2010, pp. 91–94.
[10] P. Lang, A. Kusej, A. Pinz, and G. Brasseur, “Inertial tracking for mobile augmented reality,” presented at the Inertial tracking for mobile augmented reality, 2002, vol. 2, pp. 1583–1587.
[11] D. Wagner, M. Billinghurst, and D. Schmalstieg, “How real should virtual characters be?,” presented at the Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology, Hollywood, California, 2006, ACM. p. 57.
[12] G. Klein and D. Murray, “Compositing for small cameras,” presented at the ISMAR '08: Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, 2008, IEEE Computer Society. pp. 57–60.
[13] Y. Nishina, B. Okumura, M. Kanbara, and N. Yokoya, “Photometric registration by adaptive high dynamic range image generation for augmented reality,” presented at the ISMAR '08: Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, 2008, IEEE Computer Society. pp. 53–56.
[14] G. Schall, E. Mendez, E. Kruijff, E. Veas, S. Junghanns, B. Reitinger, and D. Schmalstieg, “Handheld Augmented Reality for underground infrastructure visualization,” *Personal and Ubiquitous Computing*, vol. 13, no. 4, pp. 281–291, Dec. 2009.
[15] Y. Tokusho and S. Feiner, “Prototyping an outdoor mobile augmented reality street view application,” *Int’l Symp Mixed and Augmented Reality (ISMAR09)*, 2009.
[16] E. Veas and E. Kruijff, “Vesp’R: design and evaluation of a handheld AR device,” presented at the ISMAR '08: Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, 2008, pp. 43–52.
[17] E. Kruijff and E. Veas, “Vesp'R - Transforming Handheld Augmented Reality,” presented at the ISMAR '07: Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, 2007, pp. 1–2.
[18] J. Gimeno, F. Pardo, P. Morillo, and M. Fernández, “A Mobile Augmented Reality System to Enjoy the Sagrada Familia,” presented at the Proceedings of Eurographics, 2010.
[19] A. R. Hevner, S. T. March, J. Park, and S. Ram, “Design science in information systems research,” *Mis Quarterly*, vol. 28, no. 1, pp. 75–105, 2004.
[20] T. Nordbø, “Bergen Open Research Archive: ARTree. Designing & evaluating handheld augmented reality art,” 2011.
[21] E. Dubois, P. Gray, and L. Nigay, “ASUR++: a design notation for mobile mixed systems,” in *Human Computer Interaction with Mobile Devices*, vol. 2411, Springer Berlin Heidelberg, 2002, pp. 22–24.
[22] J. Iivari, “A paradigmatic analysis of information systems as a design science,” *Scandinavian Journal of Information Systems*, vol. 19, no. 2, 2007.
[23] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino, “Augmented reality: A class of displays on the reality-virtuality continuum,” *Proceedings of Telemanipulator and Telepresence Technologies*, vol. 2351, no. 34, pp. 282–292, 1994.