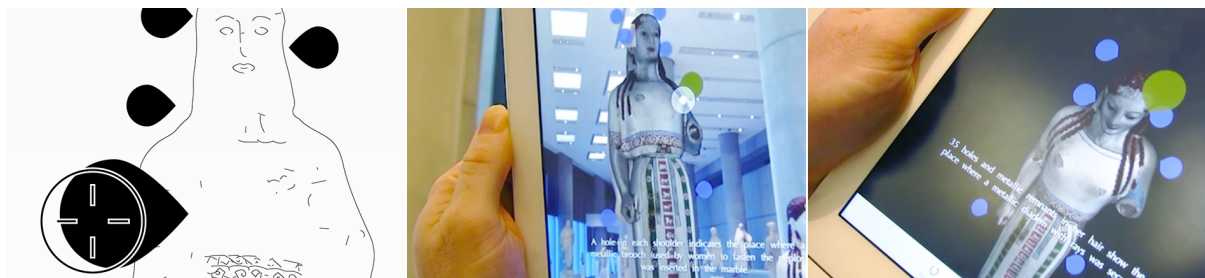


# Lean In or Lean Back? Aspects on Interactivity & Mediation in handheld Augmented Reality in the Museum

Jens Keil<sup>1</sup>, Timo Engelke<sup>1</sup>, Michael Schmitt<sup>1</sup>, Ulrich Bockholt<sup>1</sup>, Laia Pujol<sup>2</sup>

<sup>1</sup>TU Darmstadt & Fraunhofer IGD, Germany

<sup>2</sup>Universitat Pompeu Fabra, Spain



**Figure 1:** Left/Center: Pointing principle with a screen-centred and motion-controlled crosshair to trigger augmented targets. Right: Information persistence – AR turns into sole 3D presentations on a lowered device, preserving the information context.

## Abstract

With the idea to design augmented reality experiences that attract and inform, and which are also seamlessly incorporated into interactive museum narratives, this paper explores finding the appropriate balance between attraction, interactivity and information mediation from the user's point of view. Within the scope of research project CHES, we've implemented techniques that fuse interaction and mediation to enrich visits to cultural institutions by visual means and AR specific interactions. While it is AR's wow-effect that attracts, our findings show that users far too often struggle to cope with the system and interactivity instead of focusing on the information presentation in AR. We discuss our results of finding the right balance between interactive (lean-in) and non-interactive (lean-back) presentation and interaction techniques in AR

Categories and Subject Descriptors (according to ACM CCS): Information Interfaces and Presentation [H.5.1]: Artificial, augmented, and virtual realities—User Interfaces [H.5.2]: User Centered Design—

## 1. Motivation

The increasing capabilities of Augmented Reality (AR) technology in the area of tracking and rendering and the advent of mobile devices have raised audience expectations, advancing the use of mobile and wearable AR in cultural heritage (CH) settings. At the same time, the attention regarding the use of AR has shifted from attracting and entertaining audiences, to finding suitable ways of providing contextually relevant information: how do users perceive information in AR, and how do we support them in doing so?

While a multitude of studies and projects show that AR – although still immature – has the ability to engage while providing quantifiable learning outcomes (cf. [DCB\*08]), our own experiences raise questions concerning AR's impact on the objectives of the user experience. While it is AR's wow-effect that attracts, we argue that it is the added value that users willing to engage in AR in the long run, hence raising questions in terms of the design of appropriate interaction and mediation methods.

The paper critically questions the use AR and its interaction principles based on a tablet application developed for the museum context within CHESSE, a research project aiming to design and evaluate interactive storytelling for cultural sites (cf. [VKK\*12]). We present our evaluation results to answer the question in terms of finding the appropriate balance between interactive (lean-in) and rather non-interactive (lean-back) presentations in AR.

## 2. Background & Related Work

A multitude of AR applications in CH, most of which are enhanced museum guides, visually augment physical exhibits with background or interpretive information. The majority of these applications are object-oriented, based on the principle of the video-see-through or magic lens (cf. [MMT\*08], [FZB\*05], [SW07]). Due to their novelty and focus on research, these projects illustrate AR's general potential, leaving its integration in the pre-existing media landscape unaddressed; hence they constitute a stand-alone or isolated feature; either because of being a proof-of-concept, working only at single exhibits, or because they aim to replace existing digital mediators (e.g. [SW07]).

We argue that the need arises to develop AR applications for multi-media landscapes, where AR does not aim to replace existing mediators but instead merges in the communicational and spatial environment (cf. [MMM12]). Within this sense of integration, [LWD12] investigates AR's user experience. Although general design considerations for handheld AR are addressed, the work focuses on game design, missing recommendations for storytelling.

The few existing storytelling-driven projects use AR to convey the history of a place in the context of a guided tour, and are mostly implemented for the outdoors. In these cases, the mobile device is used to get AR views of a building to receive additional location-based information, or to listen to audio- and 3D-enhanced narrations ([WAS\*10], [PSAP\*05]).

## 3. Application & Conception Outline

### 3.1. Cultural Setting, Media integration & Content

We integrated and tested our application setup at Acropolis Museum's Archaic Gallery, where two major AR scenarios were integrated into the storytelling context of the entire project, reflecting the personalisation to different target groups (a detailed background can be found in [KPR\*13]).

Within the Archaic Gallery exhibits are arranged in thematic and chronological clusters. The content is composed through "activities" that comprise diverse multimedia assets (audio narration, images, 2D and 3D reconstructions, video, games, and AR). They are interweaved in alternative story plots, differing in terms of theme, graphic depictions information depth.

### 3.2. AR Activities

To gain insights into the balance of AR-based mediation and interaction within the storytelling context, we've implemented two AR activities for two different exhibits. Each is based on a different interaction principle that allows users to interact in different fashions at varying levels of engagement: while the first example is rather passive (lean back), where users must not do much more but pointing the tablet toward an exhibit, the second one is rather interactive (lean in), demanding users to be more active by either tap screen elements; physically move the device; or change their own position relative to the exhibit. The three interaction modalities differ in terms of their mental model and the interaction complexity on a usability level.

Scenario A is linked to a relief representing the mythological being Medusa. Users follow the linear explanations while seeing superimposed graphical effects. Adapted to a profile, in version one the device is as a protection shield against Medusa's powers, where AR superimposes glowing eyes that will virtually crack the screen. In the other, the ornament is reconstructed with illustrative 2D sketches and placed in its original position on top of a temple.

In Scenario B, a maiden statue is superimposed with its original colours and with narrative audio annotations which are spatially spread around it. To interact, user may either tap these annotations, target them by moving the device and a virtual pointer visible on the screen (cf. figure 1), or by solely moving physically around the statue, where the viewing angle triggers content (cf. section 3.3). Audio narrations, varying around 12-20 seconds, explain interesting aspects, such as missing elements and other history-related details.

Whenever the device is lowered, AR presentations turn into sole 3D ones, keeping the last AR view in 3D storing AR's last camera position in the 3D view. Users may use multitouch and touch to interact in 3D. We use the device orientation (from a vertical to a horizontal position) as a gesture to switch between these modes. With this, mediation may take place outside AR without users losing the context of presentation (keeping information persistent).

### 3.3. Setup and Implementation

The AR activities are incorporated inside a full storytelling experience, combined with other HTML-5 activities of different (media) types, stepwise delivering a narrative plot (cf. [VKK\*12]). Content and application logic are integrated through web technologies based on HTML templates. The web-app runs in a native AR framework for mobile devices [EBW\*13], which includes JavaScript bridge to control render and tracking engine from within the web-app.

The computer vision (CV) tracking is based on SLAM. In a post-processing step according to [WWK11], a 3D-feature map was generated and matched with the exhibits' 3D model

to connect the coordinate spaces of visualisation and tracking. On the device the 3D tracking runs on an average frame-rate of around 25 to 30 fps.

The pointing principle uses a simple target matching mechanism. For the angle-based paradigm superimposed annotations are matched in respect to a central position of exhibit and user (i.e. the virtual camera position). The annotations behave like points on an orbit that are triggered once users enter a defined area by changing their position relative to the exhibit. Instead of using the annotations' absolute position in the AR coordinate space, we created equidistant spaces for the interaction around the exhibit's vertical axis. With this we avoid that overlapping and uneven spread annotations cause jumps, which thus feels more seamless and more coherent to users while moving around.

## 4. Evaluation

### 4.1. Objectives

Visitors should be attracted, immersed and informed at the same time while using AR. Knowing that it might be disrupting for visitors to split attention between physical objects and mobile device [DPM07], [LH03], we've been interested in finding a proper balance between high interactive digital extensions (lean-in) and rather passive ones.

While we believe that AR-related interaction techniques could help to ease this tension, we question, whether users understand and anticipate these principles, perceiving them as helpful, or, if they are rather frustrated in terms of the user experience. We can argue that if users won't anticipate, they won't rate AR to be an added value and probably be not eager to use it continuously. Hence it is of interest to see, how each participant would start to use the system, how they would work with the paradigms, and if they would concentrate on the mediation after a little time.

### 4.2. Methodology & Procedure

We conducted a qualitative assessment by observing participants while using the system in the museum and debriefing them with structured interview afterwards. Thirty-six (36) selected participants (17 male, 19 female; the age ranged from 11 to 45 years) we grouped up. A group of eight (8) was exclusively focusing on interactions in AR. All were familiar with AR and have seen it in action; although three third hasn't personally used it before. The participants haven't been introduced to AR beforehand. Instead, an introduction screen was provided at the beginning of each activity.

Observation and debriefing was complemented by capturing a log file with application events and timestamps. With this, we recorded a) activity start/end, b) used AR interaction principles, c) tracking state, d) AR-3D mode toggle, e) audio playback duration, and f) general device movement.

## 5. Results & Discussion

Users had difficulties to control the experience, demanding us to interrupt and give more introduction. Interestingly, while participants found the startup-instructions not evident enough, observation shows that almost all skipped them too quickly. Anyway, they were eager to continue because they found the visual effects attractive enough.

As expected, holding the iPad high felt tiring after a while because of its weight (iPad Air). Surprisingly, comparing AR to the 3D "fallback", observation and log files reveal that users stayed longer in AR (in average 61% of the time) than in sole 3D (39%). Asking for length & duration, participants rated Scenario A with its 2-3 minutes duration over the free exploration at the statue.

Asking for the preferred interaction principles at the statue we got similar results to prior evaluations (cf. [KZE\*13]): all 8 participants tried to tap at first, but once explained, 7 preferred pointing over moving and tapping, stating that they a) were used to touch, b) didn't expect AR principles and found it tricky to understand the technique (especially for moving-around), c) confessed to be less eager to move much to interact, but however d) found it more comfortable not to tap while the iPad is raised up.

While observing users we found that too much moving around distracted them from the content – which was case, even when they felt more familiar after a while. When asked during the run, they confessed to be not focusing much but concentrating on handling the device. However, surprisingly, none of them found it disrupting nor uncomfortable. The same goes when comparing the flow of AR activities to others (e.g. interactive 2D images or video content), and when comparing it to the entire experience.

Participants left mixed impressions about the interactivity: AR effects (re-coloring, glowing eyes) clearly caught attention, but the more the app demanded to freely explore and interact, the less users were eager to do so. Almost all participants stopped following after 3-4 info chunks from a total six. One reason can be related to the visual design of the annotations, which, although a common visual AR tool, seemed not appropriate enough; more than a third found them too vague and imprecise and not doing much.

The passive Scenario A allowed to focus more on the content: 34 out of 36 recalled the core information ("mythical powers", visual restitution). Although mediation was rather shallow and highly visual, 7 from the AR interaction group stated that if of longer duration, they would have lost interest, since it felt limited in terms of interactivity.

In terms of the balance our results show that once there is too much to control inside the AR, people need more time to orientate before they concentrate on the exhibits and follow information & content mediation. Our participants admit to better recall the superimposed visuals rather than in-depth content and narrations given by the annotations. Although all

endorsed and favoured the inclusion of AR in the storytelling in general, a third argued that they would be interested in more content in it. But, comparing our observations with the debrief we see that the same participants weren't focused much, nor utilising AR to inform themselves.

## 6. Conclusion & Future Work

In this paper we have presented interaction and mediation techniques for handheld AR, that have been tested and implemented inside an interactive storytelling experience. With the idea to ease interaction between real and digital content through the use of AR, we critically explored the question about the appropriate balance between attraction, interactivity and information mediation from the user's point of view. While it is AR's wow-effect that attracts, our evaluation results show that when it comes to mediation, users far too often struggle to cope with the system and interactivity instead of focusing on the information and presentation in AR.

Having participant's rating of being rather a) more (lean in) or b) less interactive (lean back), we can say that we didn't get a clear result. People argued that they like to be interactive, but observation showed that in AR, they had the most difficulties to control the app and focus on the content. While it appears, as if AR stays far too much on the surface with regard to mediation, and complex with regard of interactions, compared to non-AR activities, we conclude to work with a stepwise process, that starts with superimposed visuals (lean-back) leading into more exploration and interactivity (lean-in), which would allow users to learn and become familiar while working with the system.

Be it rather passive or (inter-)active, on the one hand users shouldn't stand still for too long without doing much, but on the other, spatial movement of the user shouldn't be the main modality, as our results show that users inclined to become distracted. Hence, for the time being we'd argue to confront users with rather conservative interaction principles with a clear and reliable mental model, if plain mediation and education objectives are favoured.

As we found AR-related interaction design concepts and standards missing, we intend to continue and further investigate aspects of user experience design in visualisation and interaction in Augmented Reality.

## References

- [DCB\*08] DAMALA A., CUBAUD P., BATIONO A., HOULIER P., MARCHAL I.: Bridging the gap between the digital and the physical: design and evaluation of a mobile augmented reality guide for the museum visit. In *Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts* (2008), DIMEA '08, pp. 120–127. 1
- [DPM07] DAMALA A., PASCAL H., MARCHAL I.: Crafting the mobile augmented reality museum guide, 2007. 3
- [EBW\*13] ENGELKE T., BECKER M., WUEST H., KEIL J., KUIJPER A.: Mobile browser D a generic architecture for rapid ar-multi-level development. *Expert Systems with Applications* 40, 7 (2013), 2704 – 2714. 2
- [FZB\*05] FÖCKLER P., ZEIDLER T., BROMBACH B., BRUNS E., BIMBER O.: Phoneguide: museum guidance supported by on-device object recognition on mobile phones. In *Proceedings of the 4th international conference on Mobile and ubiquitous multimedia* (2005), MUM '05, ACM, pp. 3–10. 2
- [KPR\*13] KEIL J., PUJOL L., ROUSSOU M., ENGELKE T., SCHMITT M., BOCKHOLT U., ELEFTHERATOU S.: A digital look at physical museum exhibits: Designing personalized stories with handheld augmented reality in museums. In *Digital Heritage International Congress (DigitalHeritage), 2013* (2013), vol. 2, pp. 685–688. 2
- [KZE\*13] KEIL J., ZOELLNER M., ENGELKE T., WIENTAPPER F., SCHMITT M.: Controlling and filtering information density with spatial interaction techniques via handheld augmented reality. In *Proceedings of the 15th International Conference on Human-Computer Interaction* (2013), HCI '13. 3
- [LH03] LEHN D. V., HEATH C.: Displacing the object: mobile technologies and interpretive resources. In *Proc. ICHIM 03* (2003). 3
- [LWD12] LEE M. J., WANG Y., DUH H. B.-L.: Ar ux design: Applying aeiou to handheld augmented reality browser. In *Mixed and Augmented Reality (ISMAR-AMH), 2012 IEEE International Symposium on* (2012), pp. 99–100. 2
- [MMM12] MADSEN C. B., MADSEN J. B., MORRISON A.: Aspects of what makes or breaks a museum ar experience. *2012 IEEE International Symposium on Mixed and Augmented Reality - Arts, Media, and Humanities (ISMAR-AMH)* (2012), 91–92. 2
- [MMT\*08] MIYASHITA T., MEIER P., TACHIKAWA T., ORLIC S., EBLE T., SCHOLZ V., GAPEL A., GERL O., ARNAUDOV S., LIEBERKNECHT S.: An augmented reality museum guide. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality* (2008), ISMAR '08, IEEE Computer Society, pp. 103–106. 2
- [PSAP\*05] PAPAGIANNAKIS G., SCHERTENLEIB S., AREVALO-POIZAT B. O. M., MAGNENAT-THALMANN N., STODDART A., THALMANN D.: Mixing virtual and real scenes in the site of ancient pompeii. *Computer Animation and Virtual Worlds* 16, 1 (2005), 11–24. 2
- [SW07] SCHMALSTIEG D., WAGNER D.: Experiences with handheld augmented reality. In *Proceedings of the 6th IEEE and ACM International Symposium on Mixed and Augmented Reality* (November 2007), ISMAR 2007, pp. 3–18. 2
- [VKK\*12] VAYANOU M., KARVOUNIS M., KYRIAKIDI M., KATIFORI A., MANOLA N., ROUSSOU M., IOANNIDIS Y.: Towards personalized storytelling for museum visits. In *Proceedings of 6th International Workshop on Personalized Access, Profile Management, and Context Awareness in Databases* (2012), PersDB 2012. 2
- [WAS\*10] WITHER J., ALLEN R., SAMANTA V., HEMANUS J., TSAI Y.-T., AZUMA R., CARTER W., HINMAN R., KORAH T.: The westwood experience: Connecting story to locations via mixed reality. In *Mixed and Augmented Reality - Arts, Media, and Humanities (ISMAR-AMH), 2010 IEEE International Symposium On* (oct. 2010), pp. 39–46. 2
- [WWK11] WIENTAPPER F., WUEST H., KUIJPER A.: Reconstruction and accurate alignment of feature maps for augmented reality. In *3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2011 International Conference on* (2011), pp. 140–147. 2