



# Towards Mobile Holographic Storytelling at Historic Sites

Jessica L. Bitter and Ulrike Spierling<sup>(✉)</sup>

Hochschule RheinMain, Unter den Eichen 5, 65195 Wiesbaden, Germany  
{jessicalaura.bitter,ulrike.spierling}@hs-rm.de

**Abstract.** Augmented Reality (AR) promises to enhance motivation and learning of museum visitors, whereby head-mounted displays (HMD) provide a greater presence and immersion than handheld devices (HHD). To assess this, we transferred interactive AR stories from an existing HHD prototype to the MS HoloLens 2. The app lets museum visitors of a Roman fort ‘meet’ video ghosts to learn about life at their historic location. We compared two alternative ways to trigger spatial videos. The evaluation shows that, while limitations occur when reusing media once produced for 2D screens on HMDs, the approach is worth focusing on. The HMD enhances the immersion even with 2D videos placed in space, thus granting virtual ghosts more ‘presence’ in the user experience.

**Keywords:** Augmented reality · Head-mounted display · Presence · Immersion

## 1 Introduction

Museums are places for learning. They exhibit cultural artefacts, or natural and fossil finds, thus conveying factual heritage knowledge. In addition, they enable real-life experiences, because exposed objects and sites appear as present ‘witnesses’ of the past. In the last several years, museums have begun to use emerging technologies such as Augmented Reality (AR) to enhance physical exhibits with virtual information, or to provide pure virtual realities with storytelling.

We focus on AR not only for its possibilities to add digital information, but mainly for its features to enrich an otherwise non-interactive site with human lifelike experiences, displaying dramatic storytelling in real space. Also, with this goal in mind, the related project “Spirit”, which preceded this study, presented a prototype of complex location-based and spatial interactive drama in a Roman fort with hand-held devices (HHD) like off-the-shelf tablets [1]. This app allowed to meet the ‘ghosts’ of roman soldiers and historic inhabitants of the site and to witness an exemplary drama of typical conflicts and life right on the spot within the mural remains. When combined GPS and tablet camera recognized a building (by image comparison with a prepared reference), filmed characters appeared as silhouette videos in front of the sighted backdrop (Fig. 1). The app addressed the human desire to imagine life right within the visited area, thus connecting with the past. The project evaluation looked at the influence of several design aspects

that would increase this aspired feeling of ‘presence’ [2]. One supporting aspect for presence was a kind of spatial staging of virtual scenes in the real environment, by letting filmed characters appear at different angles around the user. These were triggered by the gyroscopic sensor of the handheld device, which recognized rotations like panning to the left and right. Meanwhile, presence and immersion as a means for experiential learning are expected to move a big step forward with novel hardware, such as head-mounted displays (HMD) with integrated stereo vision and improved tracking features to recognize the spatial surroundings of users. In this regard, the Microsoft “HoloLens 2” serves as a precursor for future more consumer-accessible devices.

In this paper, we present the result of our first attempt to transfer the existing content of the “Spirit” project to being experienced on a HoloLens 2. We partly build upon theoretical elaborations expressed by Liu et al. [3], concerning the impact of using general interaction patterns from HHD-kinds of devices on an HMD. Therefore, we reuse (with permission) the previous given HHD media content and its basic concept for the HoloLens 2, with the initial goal to reach a most similar interaction pattern and experience on the HMD. Previously, due to the limitations of the handheld hardware, single video actions in a complex spatial scene could only be triggered separately by the gyroscope for appearing on the tablet screen in succession at different recognized orientations. The HoloLens instead – with its tracking option of spatial anchors – allows to place video objects directly into the environment. With these improved possibilities, we also looked at the general problem of how to best perceive linear filmic content in 3D space and compared two ways to trigger the ghost videos. Further, we are interested in the prospects of using 2D videos in 3D space. In the following, after discussing related work, we first explain these concepts, followed by core findings of the qualitative evaluation of the user experience.

## 2 State of the Art

In the field of AR in museums, most reported projects focus on either HHD or HMD. Geronikolakis and Papagiannakis noticed the lack of multiplatform applications. They tried to close that gap by providing a plugin for Unity3D with which the loading of an SDK and preparing of an approach can be sped up [4].

Handheld AR devices like a smartphone are currently the first choice to establish AR in a museum, as users are familiar with these and can bring their own. Usability is a main design issue to enhance the overall experience [5]. Along these lines, Choi et al. perceive HMDs as disadvantageous and not yet mature enough [6]. They argue that poor accessibility, insufficient image quality, inconvenient form factors and basic usability are reasons why HMDs (and in this case the HoloLens) are not yet ready to compete with handheld AR devices like a Smartphone.

On the contrary, Hou [7] and Sugiura et al. [8] perceive the usability of HMD devices (the HoloLens) as good. In their view, HMD seems to be a promising technology to enhance the visitor’s experience in a museum or exhibition, especially encouraging the learning in a museum environment. An innovative example is provided by Kim et al. [9], who developed a mobile tour guide that indicates points of interest on a HMD, while using a connected HHD to display a map with GPS and keywords.

Liu et al. [3] argued that there are not enough experiences of creating “same” content for both kinds of devices at once. They compared different features of tablets and the HoloLens 2, proposing how the HoloLens could partially adopt interaction patterns of a tablet. They hypothesized how the same AR app [1] would be transferable to the HoloLens under the prerequisites of the hardware constraints and possibilities.

O’Dwyer et al. [10] implemented a volumetric filmed actor in an AR museum narrative app that could be perceived on either a tablet or a HoloLens 1. While the evaluation compared the immersive effects of both devices, the linear presentation of one minute length could only be watched at one stationary location. Users could walk around the virtual actor during his monologue, with no further interaction or navigation.

It is not yet fully explored how to best use time-based media such as film storytelling together with 3D immersive technologies. Especially for learning apps, it is essential that users can perceive the content in time, while interacting in space. Hence, previous work in 360-degree filming can be relevant here. Pillai et al. [11] state that it is necessary to implement visual cues to guide users through filmic experiences on an HMD. Especially when the experience is also the first contact with a VR/AR HMD device, it is to be expected that users are more focused on exploring the new technology than following a story, according to Syrett et al. [12]. Gödde et al. [13] add that especially first-time users need an orientation phase, before they can focus on the video content. In addition, after every cut in the filmed story, they need to orient themselves anew to find the location of the action. Visual cues like waving do not always work when users look in an opposite direction.

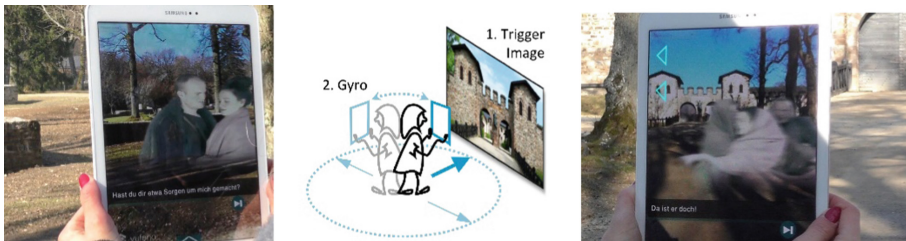
### **3 Transferring an AR Storytelling Experience from Tablet to HoloLens 2**

We followed a ‘design research’ approach [14] in the sense of building a prototype to understand the principles of a concept. We implemented the previously existing content of the tablet museum app into a new app for the MS HoloLens 2. Following Liu et al. [3], the goal was to try to keep as much of the “original” as possible, with the hypothesis that already only by using the HMD, there would be an increase in the feeling of ‘presence’ as part of the user experience. The original app already tried to implement a pseudo-spatial experience by letting the user explore the environment around the main point of view. While this was expected to be greatly enhanced with the HMD that enables ‘real’ spatial coordinate recognition, there was doubt concerning the believability of the 2-dimensional videos displaying the ghosts (as discussed in [3]).

#### **3.1 Interaction Pattern of the Given Tablet App**

The given HMD AR app aimed at creating an educational interactive drama experience in an outdoor museum. Visitors should be able to find locations at which lifelike ‘ghosts from the past’ appear around the user’s position, enacting scenes that represent their life back in 233 AD at the Roman fort that is now the museum [2]. GPS was used for pre-selecting an area, and image marker recognition could compare current camera images with a dataset of prepared views to trigger a first video scene that displayed a recorded

film scene. The whole story was divided into snippets that could then subsequently be triggered by the user through turning to the left or to the right, following UI arrows used as hints (Fig. 1). By means of the gyroscope of the hand-held device, these turns triggered the following scenes that were staged around the user. The result was an impression of a pseudo 3D world in which the user could look around. Technically, there was no spatial model underlying, as all videos were triggered in relation to the screen. After triggering, the cut-out video characters stayed within their initially captured image section of the depicted site to create the illusion of keeping their position in the environment, viewed through the metaphorical window of the tablet screen.



**Fig. 1.** AR storytelling tablet app, “Encounter/AR” mode. After triggering the first scene (right image), arrows indicated that turning left revealed the next scene (triggered by the gyroscope).

The story was a fictional re-enactment of possible actions within the buildings of the Roman fort. Apart from this “Encounter” mode providing the dramatic story in AR, there were two more interaction modes, without AR, that helped users navigate between the different locations and read some meta-information (Fig. 2). For a complete experience, these modes are also necessary between the AR locations, in order to be usable. The “Search” mode provided a GPS-enhanced map navigation and a possibility to search for buildings with a given stencil outline. The “Read” mode added factual information that fitted the dramatic storyline. For reading and GPS navigation, the tablet could be held horizontally in a relaxed pose.



**Fig. 2.** AR storytelling tablet app. Left and middle: “Search” mode with map and stencil. Right: “Read” mode, opened via the bar menu at the bottom [2].

The design of the “Encounter” mode was optimized for adjusting the aspired spatial user experience to the limitations of the tablet hardware used at this time. For example,

the cut-out video characters were presented in mono vision on top of the camera image of the tablet. There was no depth of field and no stereo vision possible. There was no occlusion feature that would allow video characters to be placed behind real objects. To see characters ‘around’ the user, these had to be triggered as a new video scene at any updated viewing angle to appear on the screen coordinates, as there was no 3D world model. Accordingly, as the HoloLens 2 would offer to program improved adapted features without many of those restrictions, the question was whether the feeling of presence would be increased in spite of this initial 2D design.

### 3.2 Transfer to HoloLens 2

The same media material produced for the tablet app was also used in the HoloLens application. However, to keep the initial steps simple and to mainly focus on the multi-direction interaction pattern of the “Encounter” (without “Search” and “Read”), we aimed at an office-only prototype for user testing. (Also, the COVID-19 pandemic posed project restrictions concerning museums during the research.) We used poster images of the original Roman fort buildings as image targets, which could be recognized by the HoloLens app and then used as a trigger to start the videos (Fig. 3, left).

Following Liu et al. [3], the aim was to stay as true to the original functions as possible while implementing for a different device with different unique interaction styles. In the “Encounter” mode, after finding an image target, the HoloLens triggers the first video of a sequence of several videos that make up the entire scene. The ghost videos are now possible to be placed in a three-dimensional virtual space around the user, instead of only on a screen. To keep the ghosts always facing the user’s camera, each video rotates on its y-axis in reaction to user movements (“billboard” behavior).



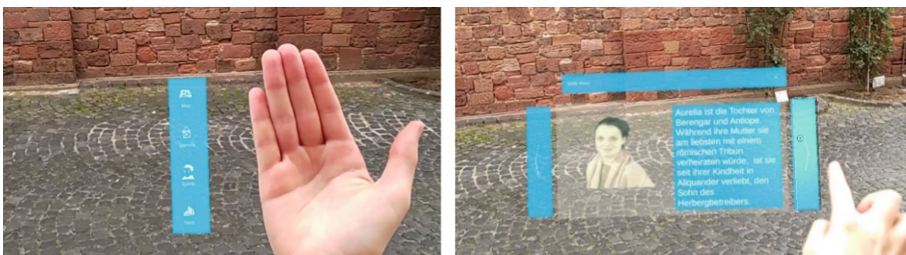
**Fig. 3.** Image capture of views through MS HoloLens 2 (left: in the office hallway, triggered by the poster image for evaluation, right: outdoor test).

As it was designed in the tablet app, the ghosts in the HoloLens app should also appear to the left and right of the user’s main view direction after the first image target recognition. For user tests with the HoloLens, we prototyped two alternative approaches to trigger the following video scenes:

- *Approach 1, “user follows story”*. All following videos of one scene start in direct succession after their previous video in a given order. In order to see them located to the left and right, the user has to follow the spatial sound of the voices and to turn around accordingly.
- *Approach 2, “user triggers story”*. The second approach uses the head rotation of the user to trigger the following videos only when the user finds them and looks at them. This approach is similar to the tablet version, in which also the turning of the tablet (gyroscope sensor) triggered the next videos. Apart from the necessity to trigger, the order is fixed.

As the videos in this second approach wait for the user to be triggered, the affordance emerged to provide hints where to look for the next video. Whereas on-screen arrows (left/right) have been used in the original tablet app for this purpose (Fig. 1), in our approach we decided to place hints more directly into the spatial environment. We placed 3D particle emitters, which appear as ‘dust’ or ‘mist’ vanishing in the air, at the same positions as the ghost videos. The particles match the turquoise fairy dust currently appearing in the beginning of the existing videos. They appear after the preceding video has stopped playing, either to the left or to the right, staying in place and bubbling until the user finds them.

For the remaining modes “Search” and “Read” including their screen menu, the transfer to the HMD was not as adequately possible. We decided to avoid any gaze-based interaction with the typical HoloLens menus, because gaze plays an important role in the AR story part of the app (looking around). However, the hands were still free to conceive a hand menu that offers most of the features of the tablet app (Fig. 4). This menu can be navigated with pointing or voice commands. However, both interaction styles were from the start expected to be less convenient to use than with the tablet. Hence, here is an obvious limitation of the HMD hardware, compared to a handheld device that easily integrates modes of reading in a convenient pose and position. We consider finding alternative solutions for general interaction a necessity in future work.



**Fig. 4.** The hand menu in the HoloLens 2 app (outdoor test)

## 4 User Evaluation

The implemented prototype was given to sixteen testers for evaluation. As we have been facing the COVID-19 pandemic for a long time, this evaluation was conducted remotely

without a possible supporting or observing researcher. We prepared a package for distribution with hardware and instructions, and asked the subjects to fill a questionnaire. The test users also had to mount prepared image marker printouts as backdrop posters to their walls at home to make it work (compare Fig. 3, left). This way, the setup was also location-independent (in contrast to a real outdoor scenario).

As a first step in iterative design, we aimed at only a few users to get quick results for improvement. According to Jakob Nielsen, only five users are enough to find 85% of the usability problems in a first prototype [15]. Naturally, this is not a quantitative evaluation, however, to show some distribution of opinions within the group of 16, we present lists of their results in numbers below.

For the purpose to present the effect of the transfer on the presence experience, we focus on the following guiding questions for the result presentation:

- Which of the two approaches contributes more to the understanding and experience?
- How is the presence and immersion perceived, if 2D videos are used with an HMD?

Out of 16 users, 10 subjects were female and 6 were male, the average age amounted to 33.25 years. 7 of them have watched online video presentations of the previous tablet AR app before; 3 have even used it. A majority of 12 testers have never used the MS HoloLens 2 or other HMD devices before.

The questionnaire has both a Likert scale, as well as freeform text input space. Here, the testers were asked to explain their ratings, which is at this point of development the most important part of the survey. The idea was to identify possible variations or consensus in their opinions (Table 1).

**Table 1.** Likert scale score. 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree.

Category	1	2	3	4	5	$\bar{x}$
After understanding how to interact, approach 1 was easy to use			1	4	11	4.6
The overall experience in approach 1 is good		1	1	9	5	4.1
After understanding how to interact, approach 2 was easy to use		1	2	6	7	4.2
The overall experience in approach 2 is good		1	1	8	6	4.2
Approach 1 is easier to use than approach 2	1	1	5	7	2	3.5
The distance between user and ghosts is good		4	4	2	6	3.6
The places and angles where the ghosts appear are appropriate	2	1	4	4	5	3.6
The ghosts seem present with me in space	1	2	3	6	4	3.6

In Table 1, we can see that the usability of approach 1 is slightly better rated (4.6) than its experience (4.1). Thus, the usability of approach 1 (“easy to use”) is mostly agreed on as good, with a tendency to “strongly agree”, whereas the overall experience is just good. The usability and the experience of approach 2 are mostly agreed as good. In

direct comparison with each other, we can see that the mean of 3.5 indicates only a slight tendency that the usability of approach 1 is perceived as superior to approach 2. Five test users answered “neutral”, which means that the difference between the approaches is not too big for a rough third of the testers.

Most users commented on their rating that it has felt natural to turn around to search for the ghosts. Thanks to the guiding through spatial sound in approach 1, they found the ghost videos effortlessly. However, the users also pointed out that they missed the beginnings of the videos because these started sequentially, before they had fully turned around. It created a feeling of being pulled through the story like in a movie, which has not given much time to process its content. For the alternative approach, however, in which a video would not start before the users’ gaze triggers it, some testers found it irritating having to look around in search for the ghosts. Those mostly had problems hearing the attracting sound, so that they did not have any guiding indicators, before they finally could only spot indicating particles after they turned around far enough. Additionally, three users pointed out that it would have been necessary to explain beforehand how their movements would influence the start of videos.

These findings suggest that the first approach, where the content does not ‘wait’ for the user, is not perfectly suited for a learning environment, although usability may be slightly easier. Having to rush to the next scene contravenes getting all the available information and remembering it later. Concerning the second approach, the feedback suggests that more importance must be given to guiding the users to minimize irritation and maximize the learning factor through the content. One possible redesign idea is to use the first approach, but add some non-essential footage to the scene beginnings in order to delay the start of each video’s core content in a way that enables a good synchronization between the average user reaction times and the storytelling.

Concerning the presence of the ghosts, the Likert scale score shows that the users’ tendency points vaguely towards a good presence and placement in the 3D environment, as all three related questions result in a mean of 3.6. This feedback gets more understandable when looking at the qualitative comments. Six users wrote that the ghosts were too close to them. This effect may be a result of the instance that the test has been performed indoors in small private homes. The possible distances have been restricted due to walls and furniture. This would be different in a museum space, but needs further attention. For example, in a different experiment, we tested the app outdoors in front of real walls (Fig. 3 right, and Fig. 4). While this setting offers more space, new challenges occur like bright sunlight affecting the feeling of presence of the ghosts. This will be addressed and evaluated in future work.

The comments about the place and angles of the ghosts confirm that it is of utmost importance to offer a seamless guiding system for the user. Four users did not understand the necessity of the spatial placement of the ghosts. Two of them even had difficulties to understand that they needed to look around at all, an effect that was also reported for the original tablet app [2]. This shows that the learning process can only take place if the functions of the app are clear to users.

Lastly, all of the users stated that it was clear to them that the videos are 2-dimensional, without complaining about this. One user explicitly emphasized that it did not disturb the impression of ghosts. Moreover, two users mentioned they found the

presence and immersion of the ghosts higher than on a screen, regardless that the videos were only 2D. We conclude that even 2D videos hold great potential for a presence experience, if placed in a 3D environment, contradicting the hypothetical concern of Liu et al. [3] that the videos could potentially be discarded as a “fake” impression. Still, we expect that the presence and immersion would even be enhanced greatly when using novel approaches such as volumetric filming in the future.

## 5 Conclusion

In this paper, we described a prototypical test to enhance AR presence experiences for increasing learning and motivation in a museum, through displaying AR stories of fictional contemporary witnesses in a more holographic way. We described the transfer of an existing concept and AR storytelling content from a tablet to the HoloLens 2. We conclude that it would indeed be advised to develop some new more specific interaction patterns for the HoloLens to ideally match the novel features of the HMD hardware. Still, it saved time that the videos and texts were already produced in a previous app, but for future content creation, assets could be made adaptable to better fit each 3D medium. While the original tablet app always waited for users to trigger videos in a complex scene, we implemented two approaches on the HoloLens. In the first approach, users follow the ongoing story, and in the second, the story continuation depends on the user triggering every step. Both approaches deserve further investigation as both have their specific advantages, given that identified small usability issues would be fixed for both. Lastly, spatial audio appeared to be an important feature for orientation and for locating the relevant action in the real environment, which must be further explored in future design studies.

We see the potential of these approaches for a new learning environment in museums through present and immersive characters or contemporary witnesses. Considering the fact that a full 3D impression – for example by volumetric filming – would be expensive, we found that the aspired feeling of presence and immersion was better than anticipated with only the 2D videos from the tablet app. In the future, we will work on overcoming the restrictions found in the current app, as well as on investigating the limitations faced in our first outdoor test.

**Acknowledgements.** This work has been funded (in part) by the German Federal Ministry of Education and Research (BMBF), funding program Forschung an Fachhochschulen, contract number 13FH181PX8. We would like to thank Yu Liu and Linda Rau for their inspiration and implementation advice. All video assets were produced within the project “Spirit” [2].

## References

1. Kampa, A., Spierling, U.: Smart Authoring for location-based augmented reality storytelling applications. In: Eibl, M., Gaedke, M. (eds.) INFORMATIK 2017. Gesellschaft für Informatik, Bonn, pp. 915–922 (2017). [https://doi.org/10.18420/in2017\\_93](https://doi.org/10.18420/in2017_93)

2. Spierling, U., Winzer, P., Massarczyk, E.: Experiencing the presence of historical stories with location-based augmented reality. In: Nunes, N., Oakley, I., Nisi, V. (eds.) ICIDS 2017. LNCS, vol. 10690, pp. 49–62. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-71027-3\\_5](https://doi.org/10.1007/978-3-319-71027-3_5)
3. Liu, Y., Spierling, U., Rau, L., Dörner, R.: Handheld vs. head-mounted ar interaction patterns for museums or guided tours. In: Shaghaghi, N., Lamberti, F., Beams, B., Shariatmadari, R., Amer, A. (eds.) INTETAIN 2020. LNICSSITE, vol. 377, pp. 229–242. Springer, Cham (2021). [https://doi.org/10.1007/978-3-030-76426-5\\_15](https://doi.org/10.1007/978-3-030-76426-5_15)
4. Geronikolakis, E., Papagiannakis, G.: An XR rapid prototyping framework for interoperability across the reality spectrum, ArXiv, abs/2101.01771 (2021)
5. Chan, B.Y., Ismail, Z.I.B.A., Jack, L.P., Asli, M.F.: Augmented reality mobile application. A feasibility study in a local national museum. J. Phys. Conf. Ser. **1358**. 12th Seminar on Science and Technology 2–3 October 2018, Kota Kinabalu, Sabah, Malaysia, IOP Publishing Ltd. (2019). <https://doi.org/10.1088/1742-6596/1358/1/012057>
6. Choi, H., Kim, Y.R., Kim, G.J.: Presence, immersion and usability of mobile augmented reality. In: Chen, J.Y.C., Fragomeni, G. (eds.) HCII 2019. LNCS, vol. 11574, pp. 3–15. Springer, Cham (2019). [https://doi.org/10.1007/978-3-030-21607-8\\_1](https://doi.org/10.1007/978-3-030-21607-8_1)
7. Hou, W.: Augmented reality museum visiting application based on the Microsoft HoloLens. J. Phys. Conf. Ser. **1237**, 052018 (2019). <https://doi.org/10.1088/1742-6596/1237/5/052018>
8. Sugiura, A., Kitama, T., Toyoura, M., Mao, X.: The use of augmented reality technology in medical specimen museum tours. Anat. Sci. Educ. **12**(5), 561–571 (2018). <https://doi.org/10.1002/ase.1822>
9. Kim, D., Seo, D., Yoo, B., Ko, H.: Development and evaluation of mobile tour guide using wearable and hand-held devices. In: Kurosu, M. (ed.) HCI 2016. LNCS, vol. 9733, pp. 285–296. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-39513-5\\_27](https://doi.org/10.1007/978-3-319-39513-5_27)
10. O’Dwyer, N., Zerman, E., Young, G.W., Smolic, A., Dunne, S., Shenton, H.: Volumetric video in augmented reality applications for museological narratives. J. Comput. Cult. Heritage (JOCCH) **14**, 1–20 (2021). <https://doi.org/10.1145/3425400>
11. Pillai, J.S., Ismail, A., Charles, H.P.: Grammar of VR storytelling: visual cues. In: Proceedings of the Virtual Reality International Conference-Laval Virtual, pp. 1–4 (2017). <https://doi.org/10.1145/3110292.3110300>
12. Syrett, H., Calvi, L., van Gisbergen, M.: The oculus rift film experience: a case study on understanding films in a head mounted display. In: Poppe, R., Meyer, J.-J., Veltkamp, R., Dastani, M. (eds.) INTETAIN 2016. LNICSSITE, vol. 178, pp. 197–208. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-49616-0\\_19](https://doi.org/10.1007/978-3-319-49616-0_19)
13. Gödde, M., Gabler, F., Siegmund, D., Braun, A.: Cinematic narration in VR – rethinking film conventions for 360 degrees. In: Chen, J.Y.C., Fragomeni, G. (eds.) VAMR 2018. LNCS, vol. 10910, pp. 184–201. Springer, Cham (2018). [https://doi.org/10.1007/978-3-319-91584-5\\_15](https://doi.org/10.1007/978-3-319-91584-5_15)
14. Hevner, A., Chatterjee, S.: Design Research in Information Systems: Theory and Practice. Springer, New York (2010). <https://doi.org/10.1007/978-1-4419-5653-8>
15. Nielsen, J., Landauer, T.K.: A mathematical model of the finding of usability problems. In: Proceedings of ACM INTERCHI’93 Conference, Amsterdam, The Netherlands, pp. 206–213 (1993). <https://doi.org/10.1145/169059.169166>