Acoustic experiences for cultural heritage sites: a pilot experiment on spontaneous visitors' interest

Vincenzo Norman Vitale[1], Marco Olivieri[2], Antonio Origlia[1], Niccolò Pretto[3], Antonio Rodà[3], and Francesco Cutugno[1]

[1] Department of Electrical Engineering and Information Technology, Federico II University of Naples, Naples, Italy {vincenzonorman.vitale, antonio.origlia, cutugno }@unina.it

[2] Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy marco5.olivieri@mail.polimi.it

[3] Department of Information Engineering, University of Padova, Padova, Italy {niccolo.pretto, antonio.roda}@dei.unipd.it

Abstract. Providing technologies to support the visiting experience in cultural venues of artistic value is an important issue that needs to be addressed by considering the delicate nature of the places. Architectural heritage and visual arts are two valuable examples: the most sensible choice for augmenting the comprehension and the experience concerning this kind of cultural heritage is through audio cues (e.g., using audio guides). This work describes a pilot experiment to evaluate the impact of soundscapes and environmental acoustics reconstruction in the visitors' experience. We proposed, to the visitors, an audioguide integrated with an optional experience presenting a comparison between a choir recorded in an anechoic chamber and the same choir as it would have sounded in the Church of the San Martino Charterhouse, using a preliminary acoustic survey. Experiments were conducted with real visitors to maximise the ecology of the collected data and results show that people are significantly interested towards the proposed experience, motivating further efforts to improve the quality and depth of the contents provided with this strategy.

Keywords: Cultural Heritage Experience · Audioguide · Environmental Acoustic Reconstruction · Impulse Response Measurement

1 Introduction

In recent years, new technologies offer unprecedented opportunities to promote and improve understanding of cultural heritage by engaging visitors in new personalised experiences. There are several approaches to enhance the visitors' experience (e.g., smartphone applications, interactive multimedia installations [2, 1], virtual, augmented [29] or mixed realty [22] applications, projections [9]), based on several interaction methods, such as sensor-based, device-based, tangible, collaborative, multimodal, and hybrid interaction methods [4]. From a design point of view, however, the main focus must not be represented by the technology itself, but by the place, the collection or the object that needs to be promoted [23]. In this sense, technological intervention must be designed to be integrated in cultural experience, accompanying it without ever interfering with it. Often the environment where the artefact is exhibited or the site we want to enhance cannot be altered for preservation or respect reasons. In such cases, personal and/or wearable devices (such as smartphones or VR headset) represent a good opportunity for intervention.

Our case study, the church of the San Martino Charterhouse in Naples, has this constraint and, furthermore, it is a visually rich environment that should not be superimposed with other visually impacting contents. It also has an important peculiarity: a sound box under the wooden choir, designed to enhance the singing voices of the monks. This echo chamber is connected to the choir through three grated holes in the floor and its effect cannot be easily perceived during the visit without disturbing the other visitors, as it would require sounds to be played in the environment. Considering these requirements, an augmented audioguide based on smartphone was designed in order to enhance the visit of the Charterhouse. The first prototype of the audioguide includes a partial reconstruction of the acoustic profile of the church. A full reconstruction of the environment response to acoustics is, of course, a complex and delicate task so, a preliminary investigation is presented, in this work, to evaluate the spontaneous interest that would arise in the Charterhouse visitors, towards a sound comparisons experience complementing the standard narrative approach. Tests were conducted in the real environment by proposing the audioguide experience to visitors who spontaneously came to visit the Charterhouse to maximise data ecology and provide a good indicator to establish if the effort of performing a complete acoustic survey of the church environment meets the interest of the people.

After the following section concerning related works, the case study is presented in Section 3. Section 4 provides the description of the audioguide, while Section 5 describes how the audio scene was created. Then, the test and its results are presented. Lastly, Section 7 presents our conclusions.

2 Related Works

The expression *smart tourism* describes the increasing reliance of tourism destinations, tourists and related industries, on emerging Information and Communication Technologies (ICT)[5, 14] that allow for massive data collection and value extraction. So, visitors will enjoy a wide range of personalised services, experiences and more, aimed at cultural heritage promotion. Indeed, with the spreading of smartphone and Internet of Things (IoT) devices [6], cultural sites can easily collect information at finer grades, reducing the risk to interfere with visitors. Moreover, cultural sites can leverage this information to offer personalised and context aware experiences. This way is possible to create an integrated and personal "lifelong" visitor model providing a starting point for lifelong experiences personalisation [16]. In this direction several efforts have been made, for user model definition [10] aiming at the integration of pre, during and post visit experiences [18, 11]. Anyway to build such models, we need some information about visitors and visits. The collection process must respect visitors privacy, avoiding any interference with visits and related experiences. So these systems have been devised to "accompanies the visitor and augments her overall museum experience" [28], being designed like informative, not obtrusive, frames surrounding the interaction with the exhibit in cultural sites. The audio-guide perfectly reflect the above description, it is one of the most diffused information sources for visitors in cultural heritage sites, accompanying visitors in their experiences without obstructing the visual channel. Anyway, the classic audio-guide concept has been improved in order to enable data collection, at different grades and interaction levels. Some location-aware mobile audio-guides have been deployed in Hecht museum, at University of Haifa [17, 19], allowing visitors to receive the right information at the right place. On the other hand, those audio-guide produced information about visitor behaviour inside the museum, allowing curators to evaluate the effectiveness of exhibits, on the basis of indicators like *Attraction Power* and *Holding Power* [19]. A further version of a mobile audio guide has been integrated with an eye-tracker [20], providing the user with the possibility to receive information about an exhibit and/or a specific part of it, relieving him/her with the necessity to request specific information,

Concerning the acoustic experience, the first concept to consider is the *auralisation*, the audio equivalent of 3D visualisation, which enables quantifiable acoustic properties of buildings, sites and landscapes captured via measurements or recreated through computer based modelling, to form the basis of an audio reconstruction and presentation of a space. For acoustic heritage it helps to build a more multi-sensory picture of our past, and the experience of being present within it [21]. These reconstructions create a "acoustical photography", which is preserved for posterity (becoming precious in case of alteration of the original space) [13] and could be used in studio sound processing for musical productions or 3D games. Computer modelling is the only possible solution in case of missing building, as in [26]. The most common method for collecting the acoustic properties of a building is the extraction of room impulse responses (RIRs) from on-site acoustic measurements. Several techniques could be used, however, one of the most robust and used is based on Exponential Sine Sweep [12, 15].

3 Case Study

The San Martino Charterhouse is a monumental monastery, built to meet the specific requirements of the carthusian monastic rule, based on the benedictine motto *ora et labora*. Built at the beginning of the XIV century but renovated several times, this impressive and stunning monastic complex hosted the works of generations of craftsmen and artists from all over Europe. After losing its religious function, it has been converted in a museum in 1867 because it could in itself well represent the art, craft and architecture in Naples through the centuries. Since then, it has become a sort of museum of the city of Naples, it started to acquire several collections of very different items (nativity scenes, coaches, pottery, paintings, drawings, coats of arms, etc.), all of them connected in different ways to the history of the city. The Charterhouse, thus, represents a sort of meta-museum, constituted by the Charterhouse itself plus all the collections it hosts. Considering the complexity of this *communication system*, designing well-integrated technologies to enhance the quality of the experience of such a rich place is clearly a challenge.

While the venue itself is large, in this work we concentrated on the church and its immediately connected environment, which represent a sub-visit concentrating on the most artistically rich parts of the Charterhouse. Specifically, the application presents contents about four points of interest (POI) :

1 “Parlatorio” (Parlor): a room to receive external visitors.
2 “Sala del Capitolo” (Chapter Hall): the decision room for monastic order representatives.
3 “Coro dei Padri” (Fathers’ Choir): choir for the monks.
4 “Sala del Tesoro Nuovo” (New Treasure Hall): the room in which the monastic order kept their belongings.

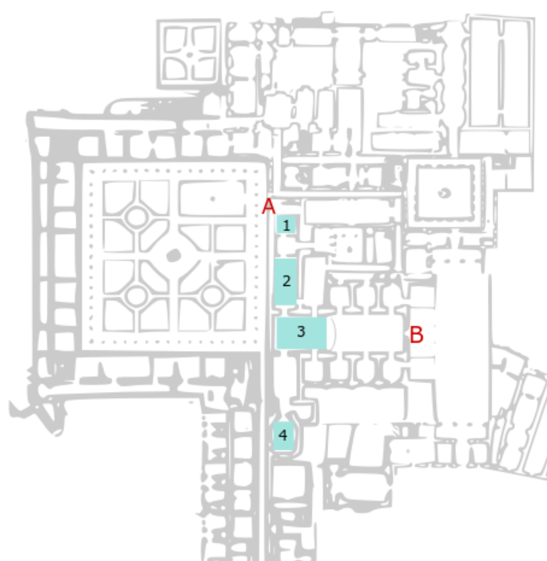

Fig. 1. A map of San Martino Charterhouse in Naples

Fig. 2. A view of the choir

4 Audio guide

The audioguide "*Caruso*" [8] is an Android smartphone application offering a location aware information layer to cultural sites visitors. It has has been proposed for the first time during the Or.C.HE.STRA project [3], as an interactive personal audioguide, for an outdoor augmented reality (AR) experience. In its first version, it offered information about monuments surrounding the visitor, along with 3D soundscapes, for improved cultural enjoyment [7, 8]. During the CHROME project, the application has been completely rewritten and extended, in order to provide a platform to integrate the most recent technologies developed for 3D sound experience and to be deployed in indoor scenarios. Using different kinds of sensors (e.g. Bluetooth beacons, RFID tags, etc.) and instrumented/technological exhibits, Caruso locates the visitor inside the cultural site, presenting contents to the visitors, consistently with the context.

The indoor positioning system for Caruso, is based on Bluetooth beacons technology. In order to identify a particular room or a part of it, one or more beacons have been placed in every monitored space. So, Caruso searches for signals issued by monitored Bluetooth beacons and, with an ARMA RSSI filter, it determines whether a visitor is in a monitored area for which contents are available. Caruso automatically provides the visitor with the appropriate audio content as he/she enters the area of interest to let people know they are entering a relevant area without forcing them to look at the smartphone, which is an action we wanted to avoid as a design choice. As done in some previous works [19], for different purposes, every time a visitor reaches/leave an area of interest Caruso updates a report. In our specific case, the report will be updated also on every interaction with multimedia contents, by reporting, for instance, every time an audio content starts or stops.

5 Acoustic measurements

The architectural complexity of the church of the San Martino Charterhouse and the inaccessibility of the echo chamber impede the creation of a precise physical model of the church. Therefore, the room impulse response (*RIR*) was preferred for reconstructing the acoustic of the site. This method considers a room as a black box system and assumes it is linear and time invariant (*LTI*, figure 3). By introducing an input signal $x(t)$, we obtain an output $y(t)$, which represents the signal perceived by a listener in the room and this is the sum of the generated noise, $n(t)$, and the deterministic function of the input signal:

$$y(t) = n(t) + F[x(t)] \tag{1}$$

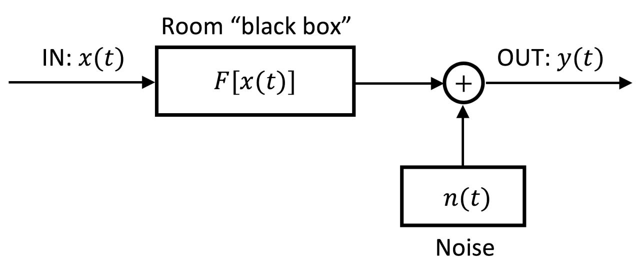

Fig. 3. LTI system

Therefore, it is possible to study the function $F[x(t)]$ as the convolution between the input signal and the impulse response $h(t)$, which is the response of the system after being excited by the Dirac delta:

$$y(t) = n(t) + x(t) \otimes h(t) \tag{2}$$

As said in the related work section there are several ways to measure the acoustic response of a room. Most of them uses speakers to emit a particular input signal $x(t)$ generated by the computer (Figure 4): this is not an impulse signal because the speakers, having limited power, cannot emit a Dirac delta so as to obtain the impulse response, therefore usually a long signal is reproduced which diffuses energy over time while the microphone records the response of the room, $y(t)$. It is then necessary a further processing step to obtain the impulse response $h(t)$ through an appropriate deconvolution technique.

Nowadays it is possible to have speakers with an almost completely flat frequency spectrum, so you can obtain very precise results. On the other hand, the use of these speakers introduces a non-linear distortion of the signal, so the room is not excited by the input signal $x(t)$ but by a signal already distorted $w(t)$. However, there are some electroacoustic measurement methods, which differ for the input signal $x(t)$ with which the environment is excited that are able to relax this problem.

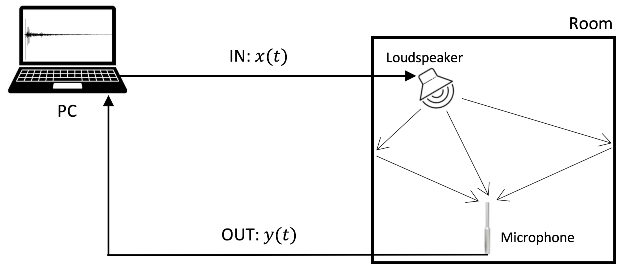

Fig. 4. RIR measurement

Among the various methods of measurements, such as the *MLS* (Maximum-Length-Sequence) or *TDS* (Time Delay Spectrometry), the Exponential Sine Sweep, *ESS*, is based on the exponential growth of the frequency sweep in the input signal $x(t)$ and allows to obtain the linear impulse response of the system separated from the impulse components corresponding to the harmonic distortions of the non-linear part introducing by the speakers [27, 15]. The main purpose of this method is the separation of the distortion peaks, due to non-linear components, from the acoustic response [12]. The underlying idea of this method is to generate as an input signal, $x(t)$, a sine sweep with an exponential variation of the frequency over time and to replace the circular deconvolution with a linear convolution over time. In order to follow this approach, we need to generate an inverse filter $\xi(t)$, that is able to incorporate the input signal $x(t)$ into the delayed function of the Dirac $\delta(t)$ and then perform the convolution between the output signal $y(t)$ with the inverse filter in order to obtain the acoustic response $h(t)$:

$$x(t) \otimes \xi(t) = \delta(t) \tag{3}$$

$$h(t) = y(t) \otimes \xi(t) \tag{4}$$

The following equation describes the input sine sweep signal as an exponential variation from f_1 and f_2 in a total time T:

$$x(t) = \sin\left[\frac{2\pi f_1 T}{\ln\left(\frac{f_2}{f_1}\right)}\left(e^{\frac{t}{T}\ln\left(\frac{f_2}{f_1}\right)} - 1\right)\right] \tag{5}$$

All the measurements were made using the plug-in suite Aurora[4], that in addition to the generation of the acoustic response, offers the possibility of

[4] pcfarina.eng.unipr.it/Aurora_XP (Last accessed: February 20th, 2020)

analysing and manipulating the wave and obtaining summary parameters for each band of octave frequencies. The following list describes the equipment used for the measurements:

- *Behringer ECM8000*: omni-directional microphone with XLR output and ultra-linear frequency response.
- *Genelec 8030B*: monitor with a fairly flat frequency response between $50Hz$ and $25kHz$. Tweeters of $19mm$ diameter and woofers of $130mm$ diameter.
- *Edirol UA101*: Roland USB sound card with 8 inputs including two XLRs and 8 jack outputs.
- *Samsung Ultrabook*: PC for measurements with Audition 3.0 (the version supported by the Aurora plug-in).
- Stands to support the microphone and the speaker.

Using Aurora, we generated the sine sweep signal with sampling frequency of 48000 Hz and 32 bit resolution, start and end frequency respectively at $22-22000$ Hz and duration of 15 seconds. In this way Aurora can automatically generates the correspond inverse filter, $\xi(t)$, required to obtain the impulse through convolution (4).

The measure was made with the microphone perpendicular to the floor and the tip pointing upwards with a height of approximately 160 cm from the floor in order to simulate the listener point. The height of the speaker with its support has remained fixed at 150 cm, in order to simulate the voice of a singer. The speaker stand was positioned in the choir section that is parallel to the high altar, whereas the microphone was positioned near to the ancient wooden bookrest in the centre of the choir.

The acquired signal was processed to achieve the response of the environment. Through the tools provided by Aurora the impulse responses were obtained by separating the distorted components from the linear impulse useful for the analysis.

6 Field Test

The experiment involved 45 visitors who spontaneously visited the San Martino Charterhouse. The experimenters provided them the audio guide "Caruso" and then they started the visit from one of entry points, A or B, shown in Figure 1. The experimenters provided a brief explanation about how to use the application and let the visitors roam freely in the environment, reducing the impact of the experience on natural behaviour. No visitors needed intervention from the experimenters and visited the environment with their own times. The target item, the audio 3D reconstruction, was presented as an optional content integrating the narrative explanation of the sound box, in the "Coro dei Padri". Also, the application let users decide which sounds to listen to, so that indicators of interest towards the full experience could be collected.

In Figure 5, the visitors distribution per age group is reported, while in Figure 6 the average visits' length per age group is reported. Collected data

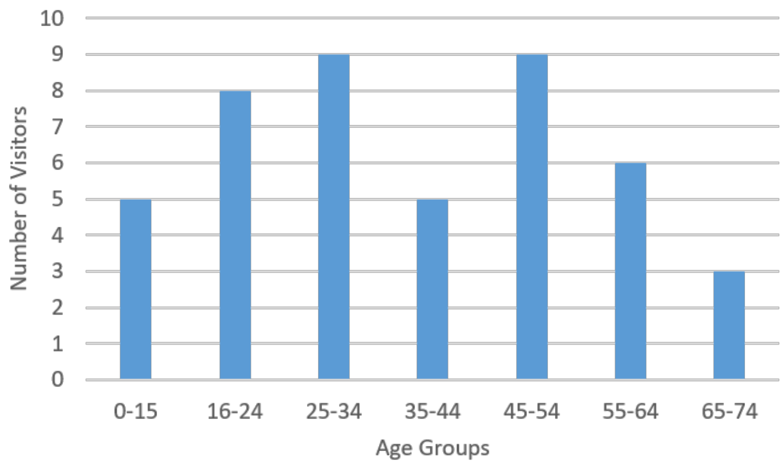

Fig. 5. Number of visitors per age group.

Did not listen	Listened 1	Listened 2
10	4	31

Table 1. Summary of the visitors' behaviour

show that, in general, the visitors were interested in the contents provided by the audioguide and, specifically, they showed interest in the experience provided by the sound box effect reconstruction: the majority of the visitors, in fact, listened to both the anechoic singing and the simulated version, as shown in Table 1. Furthermore, most of the people (89%) listened to both files for their entire length (15 seconds). This indicates a high interest, in accessing to the experience, proving its value with respect to the rest of the narrative. This motivates a further effort to improve the experiences through a more precise 3D acoustic reconstruction of the environment, aiming at enriching the narrative in the final audioguide. In general, the application design appears to be functional towards the complexity of the environment and meets the interest of the visitors without presenting contents that superimpose with the rich visual experience provided by the Charterhouse itself.

7 Conclusion

The introduction of cutting-edge technological elements in cultural heritage, brought a great number of opportunities to promote tourism, knowledge and cultural sites too. In this work we show that is possible leverage such technological elements, to promote new experiences, without interfering with channels

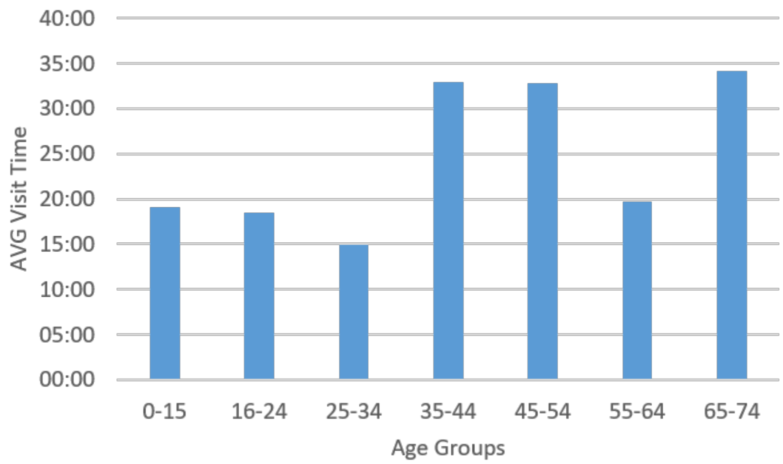

Fig. 6. Average visit time per age group

significant for visual arts, architecture and more. Moreover, we showed that soundscapes and environmental acoustics reconstruction have a great impact on visitors' experience. After this experimental study, many possible evolution can be envisioned. We clearly need to experiment on longer and different soundscapes. Furthermore, with a better indoor positioning system, we will be able to provide users with a better engaging experience, for example by improving the virtual acoustic reconstruction (necessarily based on an extended set of measurements). Finally, the measurement of the room impulse response could be improved using a microphone array [25, 24] instead of a single omni-directional microphone. This method allows indeed a time as well as a spatial separation of the elements of the impulse response.

Acknowledgments

This work is partly funded by the Italian PRIN project Cultural Heritage Resources Orienting Multimodal Experience (CHROME) #B52F15000450001.

References

1. Avanzini, F., Canazza, S., De Poli, G., Fantozzi, C., Micheloni, E., Pretto, N., Rodà, A., Gasparotto, S., Salemi, G.: Virtual reconstruction of an ancient greek pan flute. In: 13th Sound and Music Computing Conference, Proceedings (SMC 2016). pp. 41–46 (2016). https://doi.org/10.5281/zenodo.851179
2. Avanzini, F., Canazza, S., De Poli, G., Fantozzi, C., Pretto, N., Rodà, A., Angelini, I., Bettineschi, C., Deotto, G., Faresin, E., Menegazzi, A., Molin, G., Salemi, G., Zanovello, P.: Archaeology and virtual acoustics. a pan flute from ancient egypt. In:

Proceedings of the 12th International Conference on Sound and Music Computing. pp. 31–36 (2015)
3. Barile, F., Calandra, D.M., Caso, A., DAuria, D., Di Mauro, D., Cutugno, F., Rossi, S.: ICT solutions for the OR. C. HE. STRA project: from personalized selection to enhanced fruition of cultural heritage data. In: Signal-Image Technology and Internet-Based Systems (SITIS), 2014 Tenth International Conference on. pp. 501–507. IEEE (2014)
4. Bekele, M.K., Champion, E.: A comparison of immersive realities and interaction methods: Cultural learning in virtual heritage. Frontiers in Robotics and AI **6**, 91 (2019). https://doi.org/10.3389/frobt.2019.00091, https://www.frontiersin.org/article/10.3389/frobt.2019.00091
5. Buhalis, D., Amaranggana, A.: Smart tourism destinations enhancing tourism experience through personalisation of services. In: Information and communication technologies in tourism 2015, pp. 377–389. Springer (2015)
6. Chianese, A., Piccialli, F.: Improving user experience of cultural environment through IoT: the beauty or the truth case study. In: Intelligent interactive multimedia systems and services, pp. 11–20. Springer (2015)
7. D'Auria, D., Di Mauro, D., Calandra, D.M., Cutugno, F.: Caruso: Interactive headphones for a dynamic 3D audio application in the cultural heritage context. In: Information Reuse and Integration (IRI), 2014 IEEE 15th International Conference on. pp. 525–528. IEEE (2014)
8. D'Auria, D., Di Mauro, D., Calandra, D.M., Cutugno, F.: A 3D audio augmented reality system for a cultural heritage management and fruition. Journal of Digital Information Management **13**(4) (2015)
9. De Feo, E., Russo, A., Venier, E., Breschigliaro, M., Pretto, N.: Pamu: A multimedia park to enhance the renaissance walls of padua. DISEGNARECON **11**(21) (2018)
10. Dim, E., Kuflik, T., Reinhartz-Berger, I.: When user modeling intersects software engineering: the info-bead user modeling approach. User Modeling and User-Adapted Interaction **25**(3), 189–229 (2015)
11. Falk, J.H., Dierking, L.D.: The museum experience revisited. Routledge (2016)
12. Farina, A.: Advancements in impulse response measurements by sine sweeps. In: Audio Engineering Society Convention 122 (May 2007), http://www.aes.org/e-lib/browse.cfm?elib=14106
13. Farina, A., Ayalon, R.: Recording concert hall acoustics for posterity. In: Audio Engineering Society Conference: 24th International Conference: Multichannel Audio, The New Reality (Jun 2003), http://www.aes.org/e-lib/browse.cfm?elib=12277
14. Gretzel, U., Sigala, M., Xiang, Z., Koo, C.: Smart tourism: foundations and developments. Electronic Markets **25**(3), 179–188 (2015)
15. Holters, M., Corbach, T., Zölzer, U.: Impulse response measurement techniques and their applicability in the real world. In: Proceedings of the 12th International Conference on Digital Audio Effects (DAFx-09) (2009)
16. Kuflik, T., Kay, J., Kummerfeld, B.: Lifelong personalized museum experiences. In: Proceedings of Workshop on Pervasive User Modeling and Personalization (PUMP'10). pp. 9–16 (2010)
17. Kuflik, T., Lanir, J., Dim, E., Wecker, A., Corra, M., Zancanaro, M., Stock, O.: Indoor positioning: challenges and solutions for indoor cultural heritage sites. In: Proceedings of the 16th international conference on Intelligent user interfaces. pp. 375–378. ACM (2011)

18. Kuflik, T., Wecker, A.J., Lanir, J., Stock, O.: An integrative framework for extending the boundaries of the museum visit experience: linking the pre, during and post visit phases. Information Technology & Tourism **15**(1), 17–47 (2015)
19. Lanir, J., Kuflik, T., Dim, E., Wecker, A.J., Stock, O.: The influence of a location-aware mobile guide on museum visitors' behavior. Interacting with Computers **25**(6), 443–460 (2013)
20. Mokatren, M., Kuflik, T., Shimshoni, I.: Listen to what you look at: Combining an audio guide with a mobile eye tracker on the go. In: AI* CH@ AI* IA. pp. 2–9 (2016)
21. Murphy, D., Shelley, S., Foteinou, A., Brereton, J., Daffern, H.: Acoustic heritage and audio creativity: the creative application of sound in the representation, understanding and experience of past environments. Internet Archaeology **44** (2017). https://doi.org/10.11141/ia.44.12
22. Plecher, D.A., Wandinger, M., Klinker, G.: Mixed reality for cultural heritage. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). pp. 1618–1622 (March 2019). https://doi.org/10.1109/VR.2019.8797846
23. Pretto, N., Micheloni, E., Gasparotto, S., Fantozzi, C., De Poli, G., Canazza, S.: Technology-enhanced interaction with cultural heritage: an antique pan flute from egypt. ACM Journal on Computing and Cultural Heritage (JOCCH) **13**(2) (2020)
24. Salvati, D., Canazza, S., Rodà, A.: A sound localization based interface for real-time control of audio processing. pp. 177–184 (2011)
25. Salvati, D., Rodà, A., Canazza, S., Foresti, G.: A real-time system for multiple acoustic sources localization based on isp comparison (2010)
26. Segura-Garcia, J., Montagud-Climent, M., Mirasol-Menacho, S., Oleza-Simó, J.: Theatrical virtual acoustic rendering with head movement interaction. Artificial Intelligence for Engineering Design, Analysis and Manufacturing **33**(3), 359–368 (2019). https://doi.org/10.1017/S0890060419000192
27. Stan, G.B., Embrechts, J.J., Archambeau, D.: Comparison of different impulse response measurement techniques. Journal of the Audio Engineering Society **50**, 249–262 (04 2002)
28. Stock, O., Zancanaro, M., Busetta, P., Callaway, C., Krüger, A., Kruppa, M., Kuflik, T., Not, E., Rocchi, C.: Adaptive, intelligent presentation of information for the museum visitor in peach. User Modeling and User-Adapted Interaction **17**(3), 257–304 (2007)
29. Tscheu, F., Buhalis, D.: Augmented reality at cultural heritage sites. In: Information and Communication Technologies in Tourism 2016. pp. 607–619. Springer International Publishing, Cham (2016)