Neuro-curation

A case study on the use of sonic enhancement of virtual museum exhibits

Duncan Williams
Acoustics Research Centre
University of Salford
Manchester, UK
d.a.h.williams@salford.ac.uk

Ian Daly &
Inas Y Y Al-Taie
School of Computer Science and Electronic Engineering
University of Essex
Colchester, UK
i.daly@essex.ac.uk,
iyyalt@essex.ac.uk

Paola Di Giuseppantonio Di Franco &
Michael Tymkiw
School of Art
University of Essex
Colchester, UK
pd17425@essex.ac.uk
mtymkiw@essex.ac.uk

ABSTRACT

For the past several years, museums have widely embraced virtual exhibits—certainly before COVID-19, but especially after the virus's outbreak, which has required cultural institutions to temporarily close their physical sites to audiences. Indeed, even once these institutions reopen and the world returns to a new normal, virtual exhibits will remain a defining feature of museums: partly as a means to expand audiences, and partly as a way to increase revenue generation. This paper describes a case study in which a variety of soundscapes were presented accompanying a number of VR objects from the British Museum, in order to determine whether there was any appreciable improvement in viewer engagement with different types of soundscape. Soundscapes were created using synthesis, combinations of foley style effects, spoken word narration, and musique concrete based on a palette drawn from the International Affective Sounds Database. Participants (N=95) were asked to rate their engagement in an online experiment - engagement was highest in the foley-style soundscape condition. This work has implications for future targeted soundscape design, in order to target individual engagement and to facilitate exhibit evaluation, a field we describe as "neuro-curation".

CCS CONCEPTS

• H.5.5 [Information Interfaces and Presentation]: Sound and Music Computing—Methodologies and techniques

KEYWORDS

Museum exhibits, curation, audio augmented reality, virtual reality, soundscaping

ACM Reference format:

Duncan Williams, Ian Daly, Inas YY Al-Taie, Paola Giuseppantonio Di Franco, Michael Tymkiw . Neuro-Curation: A case study on the use of sonic enhancement of virtual museum exhibits. In *Proceedings of Audio Mostly 2021. ACM, New York, NY, USA, 5 pages.*

1 Introduction

Sound has long been known to have the potential to give museum exhibits additional emotional power [1] as well as allowing visitors to generate different interpretative perspectives on an exhibit [2], [3]. Consequently, museums have begun to explore sound in order to produce more entertaining exhibitions [4]. The LISTEN system [5] is an audio augmented reality (AAR) system for museum exhibits which tracks visitor behaviour in order to adjust delivery of the audio content of the exhibition. For example, it tracks which artworks a visitor has viewed and how long they have visited for in order to assign a behavioural model to visitors. This is then used to adjust audio delivered to the visitor. The audio includes attractor sounds to attempt to draw visitors towards certain objects or artworks. It also uses 3D soundscapes that are based on the visitors movements to enhance the exhibit. However, while the LISTEN system is customised to the visitor it does not investigate how sound can be optimised for individual objects within the exhibit.

In 'The rough mile' [6] pre-recorded audio segments were used to augment an outdoor location in a way that provided a transformative audio landscape that aimed to use audio to reframe (rather than compliment) the local context. In essence AAR is used to play specific audio clips to visitors when they reach specific real-world locations in order to overlay a fictional narrative on the real world locations. Context aware automated audio guides proposed in [7] use a combination of context modelling and artwork detection to create custom museum audio guides matched to the objects the user is looking at.

Despite the explosion of virtual exhibits, relatively little attention has been paid to how sound may create a more engaging experience for audiences. That is, although museums have long included sound as parts of virtual part of real exhibits (e.g., through narrated tours or through the addition of music), there is a death dearth of knowledge about how the power of sound may be harnessed to influence the length of time spent looking at objects within a virtual exhibit; about the emotional response of audiences when viewing these objects; or about the level of attention or distraction experienced by visitors. In mainstream virtual worlds (e.g., video gaming or VR experiences), the immersive power of sound is well understood and exploited to maximize user experience. Thus, in the context of museum work, this information could provide powerful tools for museum and culture curators to improve exhibit design and maximize footfall..

This gap in knowledge represents a significant opportunity. After all, by better understanding how sound shapes an audience's experiences of virtual exhibits, museums can not only prolong the length of a visitor's engagement or pique their curiosity for more information. They can also use a visitor's more engaged, prolonged experiences in virtual exhibitions as a springboard for selling related products and services (e.g., books, posters, tickets).

As a preliminary step in developing a more nuanced understanding of sound's influence on how audiences experience virtual exhibits, we have conducted a series of online experiments with 95 participants. These experiments have shown that carefully designed soundscapes customised to specific exhibits can positively effect visitor engagement with virtual exhibits. Specifically, when soundscapes are tailored to specific objects, our recent study has shown that visitors report greater engagement with these objects and higher levels of interest in them. These findings suggest that tailored soundscapes have the potential not only to significantly boost visitor engagement with virtual exhibitions but and also to provide a foundation for potential commercial applications beyond museums. Some such beneficiaries could include other cultural institutions (such as sites of cultural heritage or theatres), as well as various institutions beyond the cultural sector that depend on the virtual display of objects or environments for pedagogical, sales, or marketing purposes (such as schools, retailers, or real estate developers)

2. Experiment

We conduct an online experiment to investigate 'sonic enhancement' of virtual museum exhibits. The goal is to investigate how sound effects visitor engagement within the context of a virtual museum exhibition.

Participants are presented with a series of paired 3D models of individual objects sourced from real museum exhibits and soundscapes. These pairings are pseudo-randomised across participants, and at least one pairing includes a "no sound" condition to act as a control. After presentation of each 3D model-soundscape pair, participants are asked to report their current felt affect, engagement, and sense of presence, and to reflect on how these were affected by the model-soundscape pairing. Standard psychology test batteries were used for this; Self-Assessment Manikins (SAM), Likert Scales (LS), and Presence Questionnaire (PQ) test.

Our online experiment was composed of a set of 6 trials. Two different types of trial were used. Trial type A presented an object-soundscape pair to the participant for 30s after which the participants were asked to report their current felt affect and engagement. In trial type B the object and soundscape were also presented for 30s after which participants were given the option to hear more information about the 3D object they were shown.

The 3D objects (see Figure 1) were presented on screen using the ThreeJS toolbox for Javascript. Participants were able to rotate the object, pan the camera around the object, and zoom into or out of the object using either the computer mouse or touchscreen controls depending on their device. Participants were given detailed instructions on how to control their view of the 3D object before the start of the experiment.

During the additional information section of the trial the 3D object remained on screen and participants were still able to interact with it. However, the soundscape was not played and instead a pre-recorded audio file was played to participants containing additional information about the object.

Three additional layers of audio information were available to participants and played sequentially in response to yes/no options presented to participants to request if they wanted extra information. Each additional information segment lasted approximately 30s and contained information about the construction, history, and cultural relevance of the objects. This project has been reviewed on behalf of the University of Essex Ethics Committee and Neuromodulation Committee and has been given approval (ERAMS reference: ETH1920-1530).

2.1 Soundscape Design

soundscape. For our purposes, we take a view on [8] which defines a soundscape as an "Acoustic environment as perceived or experienced and/or understood by a person or people, in context". In other words, we expect that our responses to a soundscape may change depending on individual experiences (memories) and expectations (including cultural expectations), as well as the manner of the presentation (e.g., if presented synchronously with visual content). In our context, working with VR and visual material from museums, we suggest that the world of sound-to-picture (film, television, and video games) has the most comparable view on soundscapes: here, we may find ecologically congruous soundscapes, as in the sound walk (i.e., real-world sounds recorded and played back as accurately as possible), hyper-real sounds, such as the special effects (or "Foley") world of film and video games, or combinations of real-world and synthesised sounds designed not to convey a specific perceptual feature of the scene or artefact in question, but instead to incite specific affective responses in the listener. Chion describes a remarkable property of sound-for-picture in this regard as *synchresis*, wherein highly artificial sounds are used synchronously with video to create a hyper-real response almost involuntarily in the viewer [9]. This effect is perhaps most akin to the world of traditional music, where music is intended to convey emotion to the listener that may be artificial (not the listener's own existing emotional state), though the intended and induced emotion may be different, again depending on the context, lived experience, and cultural expectations of the individual listener [10]. Inspired by this, we create four categories of soundscape. Thus we have a category of "real-world" museum ambience recordings, originally captured on location in various museum locations. The reader might imagine that these sound like they are in the foyer of a museum, or a relatively quiet gallery space. There are rustles, light chatter, long echoey footsteps, and other such incidental noises. Next, we create a "Foley" style soundtrack which features sound-effects that are inspired by the individual artefacts in question, but are not necessarily intended to convey realism. For example, an ancient gourd might be accompanied by the sounds of an outdoor environment, with occasional flowing water. This category borrows heavily from the practice of Foley in sound-to-picture work. Next, we borrow from the world of computer music with two further categories, the first a soundscape made using generative sound synthesis – in the western musical tradition we have a culture of this type of music in the work of ambient sound artists like Brian Eno [11]. Finally, we use the same principle to create a sonic collage in the tradition of musique concrete [12], which combines sound effects from the International Affective Database of Sounds [13], a repository of sound samples that have been pre-rated in terms of discrete emotional response.

A complete soundscape for each artefact was generated in each of these categories, along with a discrete recording of a male voice, made using a close microphone technique that most readers would be familiar with (a "radio DJ" effect) to capture narration about each artefact which could be triggered independently as part of the experiment, regardless of which of the soundscape types was to be played at any given point.

Figure 1: Original scan of marble head of Livia, G70 British Museum Online Collection. Model by Daniel Pett.

2.2 Evaluation

After each trial participants were asked to report their current felt affect on the valence arousal scales using the self-assessment manikin [8]. They were also asked to report their level of engagement with the object-soundscape pairing using a subset of the measuring presence questionnaire [9]. Participants were first asked to report their felt affect on the valence and arousal scales. These were presented in random order. Participants were then asked the 5 engagement questions, these were also presented in random order, Q1-Q5, as follows

- Engagement Q1 "How much did the auditory aspects of the display involve you?"
- Engagement Q2 "To what extent did you find the object visual features engaging?"
- Engagement Q3 "Did the audio or silence add to your experience of the object?"
- Engagement Q4 "How aware were you of events occurring in the real world around you?"
- Engagement Q5 "How engaged did you feel with the object and its accompanying audio content?"

3. Results

We first pre-process the dataset to remove participants who did not engage with the experiment. Lack of engagement is measured by looking at the variance of responses participants gave to the questions over all trials. Participants who gave the same answers to the questions on all trials (the variance of the answers was zero) are assumed to not be engaging with the experiment as intended. This subset of participants were removed from the dataset.

We then use a multivariate linear regression analysis pipeline to determine if there are any significant relationships between either the soundscapes or objects presented to participants and the responses given by participants to the affect or engagement questions.

We use linear models to measure the fit of the soundscape type and object to each of the participant responses. This allows us to evaluate the effect of changing the soundscape or the object presented to participants on their felt affect or level of engagement. We also note that our 5 engagement questions are likely to be highly correlated with one another. Therefore, we use principle component analysis (PCA) to identify a projection of the set of answers given by participants to these questions that better captures the key dimensions of changes in variance in responses given by participants to these questions [14].

Specifically, PCA identifies a translation matrix that translates the set of answers given by participants to these 5 questions into a set of 5 principal components, which are sorted in order of decreasing variance. We then repeat the linear regression analysis described above for each of the resulting principal components.
Post-hoc testing (paired *t*-tests) is then used to investigate all significant effects found via our regression analysis. This allow us to investigate which individual soundscape and/or object presentations to participants produces the observed significant responses.

We also investigate the effects of changes in soundscape type of object on whether participants ask to hear more information about the object. We hypothesise that simultaneous presentation of some types of soundscape with objects has the potential to increase participant engagement with, and interest in, the objects and, furthermore, that this will manifest in participants requesting to hear more about the object.

We first used the GLM to generalize the linear regression model that describe the linear relationship and to estimate the effect of the object types and soundscapes on the participant responses. In our results, we focused on the p-Value (typically ≤ 0.05) to determine the significance of our results. There is a significant relationship between the object types and soundscapes on the answer's participants give to the Q4 (soundscape: $p < 0.001$, p-Value for object type= 0.007) and Q5 (p-Value for soundscape= 0.00041356, p-Value for object type= 0.030461). If we adjust for multiple comparisons (multiply p-Values by 5, because we have 5 questions asking similar things), then we have significant results for Q4 (p-Value for soundscape= 0.00484625, p-Value for object type= 0.03638) and Q5 (p-Value for soundscape= 0.0020678).

After that, we used post-hoc testing to investigate which soundscapes and objects cause these differences. Figure 2 and 3 represent the effect of individual soundscape types and model types on results regarding to Engagement Q4, respectively. Figure 4 represents the effect of soundscape types on results regarding to Engagement Q5.

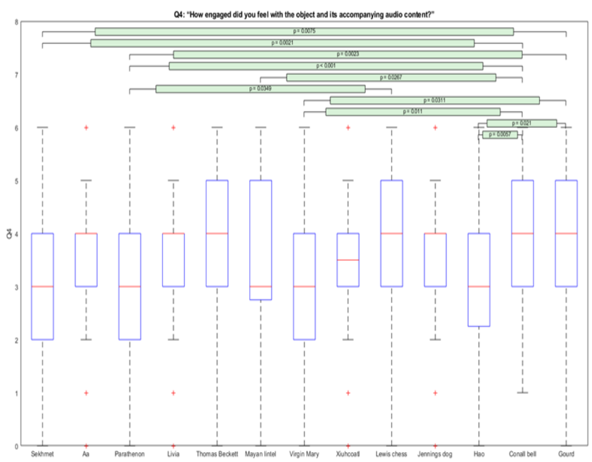

Figure 3. Distributions of answers given by participants to Engagement Q4 on the effect of individual soundscape types

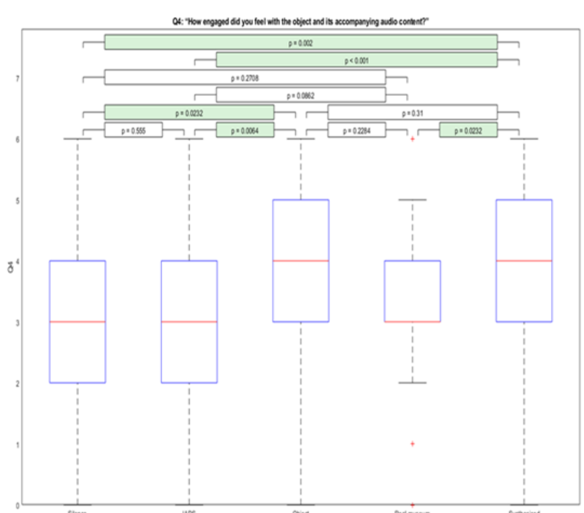

Figure 3. Distributions of answers given by participants to Engagement Q4 on the effect of individual model types

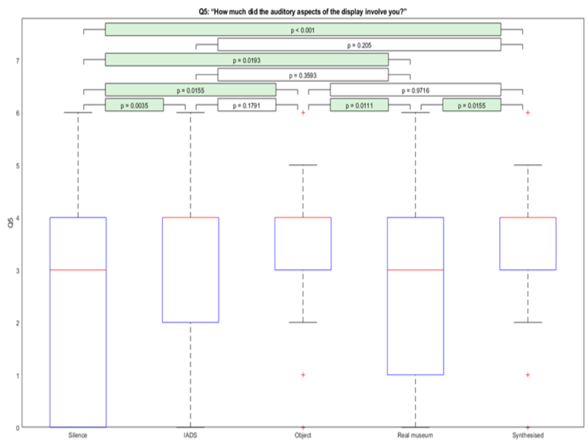

Figure 4. Represents the effect of soundscape types on results regarding to Engagement Q5

4. Conclusions

Our recent project specifically focussed on sound's effect in virtual exhibition environments, building on our previous studies aimed at understanding how new technologies effect perception and our understanding of heritage [15]–[17]. Our initial results are highly promising, demonstrating that soundscapes can significantly boost engagement with virtually exhibited objects whenever such soundscapes are tailored to the specific objects in question. There is a significant relationship between soundscape and engagement with the object-soundscape pairing. Specifically, object-inspired soundscapes and synthesised soundscapes increase engagement with the object-soundscape pairing.

Three of the soundscape types are significantly more engaging. Specifically, the IADS inspired sound, the object inspired sound, and the synthesised soundscapes were found to be more engaging than the real museum soundscapes and the silence. This suggests that, when participants are asked to reflect on just the soundscape, they find any type of soundscape more engaging than either silence of the sort of typical background sound they would normally hear in a museum setting.

Two distinct groups of objects were found in our analysis, a low engagement group of 6 objects and a high engagement group of 7 objects. When the analysis is repeated separately with each of these two groups a significantly larger effect of soundscape is found with the more interesting object group. In other words, when the objects are already interesting to the participants pairing those objects with a suitable soundscape acts to significantly enhance engagement with the object.

This suggests that if we are able to detect which objects are already somewhat interesting to a visitor to a museum we could act to boost that engagement further through delivery of an appropriate soundscape paired to the object.

The results of our current project show that virtual exhibits are more engaging for visitors when carefully designed soundscapes, tailored to the contents of the exhibit, are played to visitors. This suggests virtual exhibits could better engage visitors through careful design of soundscapes.

In the future we hope to build on our research exploring how sound may be used to modify the affective state of an individual [18], as well as our research which showed how state-of-the-art affective computing technology may be used to dynamically modulate sound and music over time to optimise an “affective trajectory” (a change in felt affect in an audience over time) [19], [20]. However, one of the most exciting aspects of this work is the set of questions it has raised for future investigation. For example, although we have proven that sound may heighten engagement, can museums tailor sounds to individual objects in less laborious ways (e.g., by drawing on AI or sound databanks)? Also, to what extent might a museum vary sounds during a virtual exhibit based on dynamic indicators of a spectator’s mood, length of stay, or other factors? By answering these and other questions in the next phase of our study, we seek to significantly nuance our preliminary knowledge and, in so doing, make our knowledge more targeted and applicable to real-world settings.

REFERENCES

[1] N. Bubaris, “Sound in museums – museums in sound,” *Curatorship*, pp. 391–402, 2014.

[2] T. Boon, “Music for Spaces: Music for Space,” *Journal of Sonic Studies*, vol. 8, 2014.

[3] M. Hutchison and L. Collins, “Translations: experiments in dialogic representation of cultural diversity in three museum sound installations,” *Museum Studies*, vol. 7, no. 2, 2009.

[4] L. Cliffe, J. Mansell, C. Greenhalgh, and A. Hazzard, “Materialising contexts: virtual soundscapes for real-world exploration,” *Personal and Ubiquitous Computing*, 2020.

[5] A. Zimmermann and A. Lorenz, “LISTEN: a user-adaptive audio-augmented museum guide,” *User Modeling and User-Adapted Interaction*, vol. 18, 2008.

[6] A. Hazzard, J. Spence, and S. Greenhalgh, Chris McGrath, “The Rough Mile: Reframing Location Through Locative Audio,” in *Proceedings of the 12th International Audio Mostly Conference on Augmented and Participatory Sound and Music Experiences*, 2017, pp. 1–8.

[7] L. Seidenari, C. Baecchi, T. Uricchio, A. Ferracani, M. Bertini, and A. del B. Bimbo, “Deep Artwork Detection and Retrieval for Automatic Context-Aware Audio Guides,” 2017.

[8] I. O. for Standardization, *ISO 12913-1: 2014 Acoustics—soundscape—part 1: definition and conceptual framework*. ISO Geneva, 2014.

[9] M. Chion and C. Gorbman, *Audio-vision : sound on screen*. New York, NY [etc.]: Columbia University Press, 1994.

[10] A. Gabrielsson, “Emotion perceived and emotion felt: Same or different?,” *Musicae Scientiae*, vol. 5, no. 1 suppl, pp. 123–147, 2002.

[11] B. Eno and P. Schmidt, “Oblique strategies,” *Opal, London*, 1978.

[12] P. Schaeffer *et al.*, *La musique concrète*. Presses universitaires de France, 1967.

[13] R. A. Stevenson and T. W. James, “Affective auditory stimuli: Characterization of the International Affective Digitized Sounds (IADS) by discrete emotional categories,” *Behavior research methods*, vol. 40, no. 1, pp. 315–321, 2008.

[14] I. T. Jolliffe, “A Note on the Use of Principal Components in Regression,” *Applied Statistics*, vol. 31, no. 3, p. 300, 1982, doi: 10.2307/2348005.

[15] P. D. G. Di Franco, C. Camporesi, F. Galeazzi, and M. Kallmann, “3D printing and immersive visualization for improved perception of ancient artifacts,” *Presence: Teleoperators and Virtual Environments*, vol. 24, no. 3, pp. 243–264, 2015.

[16] P. Di Giuseppantonio Di Franco, F. Galeazzi, and V. Vassallo, *Authenticity and cultural heritage in the age of 3D digital reproductions*, vol. 1. McDonald Institute, 2018.

[17] F. Galeazzi and P. Di Giuseppantonio Di Franco, “Theorising 3D Visualisation Systems in Archaeology: Towards more effective design, evaluations and life cycles,” 2017.

[18] D. Williams *et al.*, “A Perceptual and Affective Evaluation of an Affectively -Driven Engine for Video Game Soundtracking,” *ACM Computers in Entertainment*, Jun. 2016.

[19] I. Daly *et al.*, “Identifying music-induced emotions from EEG for use in brain-computer music interfacing,” in *Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on*, 2015, pp. 923–929.

[20] I. Daly *et al.*, “Personalised, Multi-modal, Affective State Detection for Hybrid Brain-Computer Music Interfacing,” *IEEE Transactions on Affective Computing*, Oct. 2017.