



# From AI, creativity and music to IoT, HCI, musical instrument design and audio interaction: a journey in sound

Alan Chamberlain<sup>1</sup> · Adrian Hazzard<sup>1</sup> · Elizabeth Kelly<sup>1</sup> · Mads Bødker<sup>2</sup> · Maria Kallionpää<sup>3</sup>

Published online: 1 April 2021

© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2021

This introduction brings together a range of research, a majority of which was presented at the Audio Mostly conference hosted at the University of Nottingham. The conference brings together a range of researchers, industry, designers and educators to discuss all things audio-related, the main focus of the conference is to further understand the different ways in which we can interact with audio-based technologies. This special issue is testament to the international and interdisciplinary nature of the growing Audio Mostly conference, which brings together a range of experts from across the world.

A recurrent theme in this collection is the turn to the creative appropriation of sonic and musical interactions with computers. From the design of sound interactions with physical objects, the development and evaluation of new digital instruments and their expressive capabilities, to basic and explorative research about sound perception and the semantics of sonic experiences, the volume emphasizes the importance of sound in interaction. Sound offers not only numerous creative and expressive possibilities; it also suggests new opportunities for the exploration of new kinds of interactions with ubiquitous and pervasive computers. As our attentional ability seems to be increasingly overextended by a multitude of visual displays and affordances, sound might provide avenues for more ambient and less intrusive forms of interaction.

In many ways this special issue also a celebration of the end of the FAST project (Fusing Semantic and Audio Technologies for Intelligent Music Production and Consumption). The Engineering and Physical Research Council Funded project allowed researchers to work together in a truly radical way,

across disciplines, using co-production approaches with global partners. This Introduction to the Special Issue also serves as a small bibliography representing the work that the Editors of this issue published in relation to that project. We hope that it serves as a guide for anyone wanting to further understand some of the contemporary challenges, which have emerged in the areas relating to Audio-based Design, HCI and AI.

As with so many areas of research these days the rapid development of technology in many respects has led a state of instability in terms of understanding and forecasting the future of research. We hope that this range of papers will lead the reader on a journey through this research which will inspire and provide a ‘snapshot’ of the latest research in this area as we move into an age of autonomous, intelligent and experiential music-based technologies.

Many thanks from the Editors this issue.

Alan Chamberlain, Adrian Hazzard, Elizabeth Kelly,  
Mads Bødker & Maria Kallionpää.

## 1 Papers in this collection

**Cliffe, L. et al. “Materialising Contexts: Virtual Soundscapes for Real World Exploration”.**

Using a practice-based research through design methodology, the paper details the development of an interactive Audio Augmented Reality sound installation at the National Science and Media Museum in Bradford, UK. In the project, the designers assign archival sound sources to a vintage radio receiver from the museum’s collection with the aim of extending the visitors interest and engagement in the artifacts on exhibition. Their interaction model is described, showing how the exhibit and its acoustic zones created an *aura* for the audience to explore. A follow up qualitative study of users interacting with the system shows how the spatial extension through sound of the object of an exhibition can lead to improved engagement. By animating the audience to move

---

✉ Alan Chamberlain  
Alan.Chamberlain@nottingham.ac.uk

<sup>1</sup> University of Nottingham, Nottingham, UK

<sup>2</sup> Copenhagen Business School, Copenhagen, Denmark

<sup>3</sup> Hong Kong Baptist University, Kowloon Tong, Hong Kong

around the object and discover various archival recordings of voice and music, the project, among other things, shows how using sound can trigger the initial interest in an object as well as intensely focused listening and it should encourage museums and archives to use sound as a prominent modality for creating engaging presentations of their artifacts and exhibits.

**Cunningham, S. et al. “Supervised Machine Learning for Audio Emotion Recognition: Enhancing Film Sound Design using Audio Features, Regression Models and Artificial Neural Networks”.**

In this paper Cunningham offers some insight to audio emotion recognition (AER). The paper is interested in the multifaceted role audio plays in sound design for film. In relation to this Cunningham sets out that further work is needed to better analyze, understand and classify the affective qualities of non-musical audio, and under explored area in comparison to music emotion recognition. Consequently, this paper speaks to those interested in music information retrieval, machine learning and sound design in general. Cunningham’s approach is to employ and compare two machine learning approaches on the standard International Affective Digitized Sound (IADS) data set. Interestingly Cunningham’s findings contrast previous work on music emotion recognition, highlighting the need for future work in this area to better understand how machine learning methods can support a broad range of audio recognition that incorporates musical and non-musical audio.

**Engeln, N. and Groh, R. “CoHEARence of Audible Shapes - A Qualitative User-Study for Coherent Visual Audio Design with Resynthesized Shapes”.**

The paper investigates coherence between visual and auditive stimuli. In particular, the paper suggests ways of studying visual notation. An aim of the article is to inform the design of digital instruments, suggesting that creativity can be supported through the design of artifacts such as virtual instruments that do not imitate their physical, hardware counterparts. Based on a smaller qualitative study to explore how participants would visually render auditory stimuli in the form of various re-synthesized shapes in the form of simple frequencies and envelopes, the results from a follow up study are analyzed using concepts from semiotics. Findings include the propensity of the participants to make *iconic* links in their rendering, e.g., by illustrating envelopes as lines and curves/shapes for frequencies. However, some participants favored more metaphorical renderings such as “like whirling a rope” or similar. The paper proposes an initial syntax for ordering the transformations of sound into shapes or visual stimuli and provides recommendations for further research into the design of less ‘skeuomorphic’ tools and interfaces for sound design and musical creativity.

**Gibson, D. and Polfreman, R. “Analyzing Journeys in Sound: Usability of Graphical Interpolators for Sound Design”.**

This paper is concerned with graphical interpolation systems for sound synthesis user interfaces. The authors note little formal evaluation regards the effectiveness of these graphical interfaces has been previously conducted. This paper reports on a comparative usability study of three 2D graphical interpolators of increasing visual complexity (i.e. from nothing to a set of overlapping nodes), where participants try to mimic heard synthesis examples using each interface. Usability questionnaires employ five evaluation metrics, namely time, speed, distance, accuracy and satisfaction, alongside and mouse/trackpad traces of the interface interaction are used for analysis. The authors observed three distinct phases of use across all interfaces. The questionnaires revealed no clear distinction between the ‘usability’ of each visual design, but nonetheless there were distinct differences in use, with the more visually complex interfaces seeing less random exploration, increased accuracy and prolonged exploration.

**Grimaud A.M. & Eerola, T. “EmoteControl: An Interactive System for Real-Time Control of Emotional Expression in Music”.**

The article discusses musical expression and the link to emotions research. Developing and evaluating a digital system for non-musicians to interactively explore expressive cues in musical performances, qualitative studies on the user experience as well as formal usability evaluations are performed. Reflecting on the use of the software and possible extensions on mobile platforms and in browser-based applications, the authors suggest that fields such as music education, emotions research, gaming, and communications might benefit from easy-to-use and tailorable systems such as EmoteControl, but also reflect on possible therapeutic applications and how developmental psychology studies could be relevant application a.

**Iber, M. et al. “Auditory Augmented Process Monitoring for Cyber Physical Production Systems”.**

In this paper the authors take on the challenge of process monitoring support for workers in manufacturing settings, specifically how they might monitor the operational phases or errors in manufacturing equipment via auditory feedback. The authors long term aim is to develop a set of noise canceling headphones that can also broadcast real-time process monitoring and communication between workers. This paper offers a step towards that goal, exposing knowledge about how machine auditory emissions can be captured and processed to deliver real-time feedback via two proof of concept approaches situated in two exemplar settings. They engage with matters of machine learning for audio analysis, creative sonification based on physiological and musical models, and

workers implicit knowledge and understanding.

**Page, D.L. “Music & Sound-Scapes of Our Everyday Lives: Music & Sound-Making, Meaning-Making, Self-Making”.**

In this auto-ethnographic study, the author investigates his music-making practice and Self through the process of creating a DIY musical artifact. The impetus behind this investigation was the author’s desire to interrogate his experience of a more authentic connection with one form of music-making (with acoustic instruments) over another (digital, virtual). In Part 1, the author investigates music-making practice. In Part 2, he reinterprets his personal Music and Sound-making practice, including a reflective account of how his music-making practice evolved and a rigorous examination of his values. In Part 3, the author gives detailed insight into his ‘Music & Sound-making Centred Research Study Practice’, which arose out of his auto-ethnographic investigation, including the discrimination process in his practice.

**Richan, E. and Rouat, J. “A proposal and Evaluation of New Timbre Visualisation Methods for Audio Sample Browsers”.**

Modern sample libraries can contain thousands and thousands of artificial or recorded sound samples. To find particular samples, one usually uses keywords and subcategories, and then listens to the clips one by one. This paper focuses on the possibilities of facilitating the daunting task of finding optimal sound samples from such extensive libraries with the help of visual labels: the authors designed a sound sample browser which lets users visually label sounds using textural images. A series of studies will be presented, focusing on the effects of using shape, color, or texture as labels in a search task. This paper presents the results and an in-depth analysis of the study conducted: the findings demonstrated that, whereas a visual shape improves task performance, color and texture do not have a significant effect on it. The paper suggests new methods for generating textural labels and proposes future research potentials on the topic, both in face-to-face as well as online circumstances.

**Salselas, I. et al. “Sound Design Inducing Attention in the Context of Audiovisual Immersive Environments”.**

In this paper the authors explore notions relating to sound design and its application in immersive settings. By doing this they are able to review the literature to provide a range of understandings of the design, application and use of sound design to understand the ways in which this might be used to focus the users’ attention in such environments. This paper is important as it enables the reader to understand the links between sound, narrative and agency in immersive experiences. This was a particularly interesting paper which fitted

well with the conference theme, A Journey in Sound, and was of interest to many who attended the conference with an interest in VR and immersion. It will be interesting to see the development of these ideas as immersive experiences become more accessible to a wider range of people.

**Turchet, L. et al. “Touching the audience: Musical Haptic Wearables for Augmented and Participatory Live Music Performances”.**

There has been increasing interest in haptic wearable devices targeting music performers: for example, such solutions may enhance the communication between performing musicians and a sound technician in challenging light conditions or otherwise limited visibility in a concert space. Moreover, they may offer an improved accessibility for music making for the visually impaired. The authors of this paper present musical haptic wearables for audiences (MHWAs): these are devices specifically designed for enhancing the audience’s listening experience in a live-concert setup by offering tactile feedback corresponding to the musical impulses. On top of delivering haptic stimuli and recording physiological and gestural data of the participants, the MHWAs were designed to provide new capabilities for creative participation. The MHWAs were tested in two different live-performance situations that will be discussed in detail. Both test performances suggested that, although the experience was not homogeneous across participants, MHWAs have the potential to enrich the live music listening experience in terms of arousal, valence, enjoyment, and engagement.

**Vander Wilt, D. and Farbood, M. “A New Approach to Creating and Deploying Audio Description for Live Theater”.**

The authors propose a framework for automated audio description of live theater events using freely available tools. Each performance of a theatrical work is typically unique, particularly in relation to the timing of specific events and cues, making the automation of accessibility media such as audio description a particular challenge. The authors tackle this using a simple and cost-effective method. Specifically, they compare the mel-frequency cepstrum coefficient of a reference recording (i.e., an audio recording of a previous performance of the production) to an audio input of a live performance of the work using an existing audio time warping algorithm to synchronize the reference recording to the unfolding live performance. The manipulated reference recording is used to trigger pre-prepared audio description assets for the audience members. The authors have also factored in special case provisions for unexpected interruptions to the performance not present in the reference recording and the unique characteristics of performances with and

without music. The framework – via a software implementation – was tested with two separate ‘shows’ twice using different recordings for both the reference and ‘live’ input, both demonstrating close synchronization.

**Williams, P. and Overholt, D. “Design and Evaluation of a Digitally Active Drum”.**

Various products of newly developed technologies have gained vast popularity among musicians, as they enable them to reach towards new technical and aesthetical horizons and to maximize their potential as performers. However, despite of the vast selection of commercially available Electronic Percussion (EP) solutions, many musicians still prefer to use traditional acoustic drums. This paper presents Digitally Active Drum (DAD), which is an acoustic-electronic hybrid instrument combining the enhanced possibilities of a digital solution to the performance techniques and physicality of an acoustic drum. The paper showcases the prototype that was tested by five expert drummers and two co-performers in different instrumental combinations.

**Worthy, P. et al. “Personal and Ubiquitous Computing Musical Agency and an Ecological Perspective of DMIs: Collective Embodiment in Third Wave HCI”.**

Understanding the meaning of embodiment, and in particular sharing a collective notion of embodiment, has become a

key concern for the Third Wave HCI community. Researchers in the Digital Musical Instrument (DMI) field have come to recognize the importance of designing instruments which support musical agency and to develop a view of the inter-relationship between musician, instrument, context and audience as an ecology. This article details the design and development of a DMI informed by participatory design where musicians were involved in the design process. Through this process, the authors developed a practical understanding of musical agency and the ecology of inter-relationship between musician, instrument, context and audience, which provide insights into the nature of embodiment and meaning-making of relevance to Third Wave HCI researchers and the larger HCI community.

**Acknowledgements** This work was supported by the Engineering and Physical Sciences Research Council [grant number EP/L019981/1] project Fusing Semantic and Audio Technologies for Intelligent Music Production and Consumption. We would especially like to thank the Department of Music and School of Computer Science at the University of Nottingham for their support in helping host the conference.

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.