Editorial

Special Issue on "Sound and Music Computing"

Tapio Lokki [1,*] (ID), **Meinard Müller** [2], **Stefania Serafin** [3] **and Vesa Välimäki** [4] (ID)

1   Department of Computer Science, Aalto University, 02150 Espoo, Finland
2   International Audio Laboratories Erlangen, Friedrich-Alexander Universität Erlangen-Nürnberg,
     91058 Erlangen, Germany; meinard.mueller@audiolabs-erlangen.de
3   Department of Architecture, Design and Media Technology, Aalborg University,
     2450 Copenhagen SV, Denmark; sts@create.aau.dk
4   Department of Signal Processing and Acoustics, Aalto University, 02150 Espoo, Finland;
     vesa.valimaki@aalto.fi
*   Correspondence: tapio.lokki@aalto.fi

check for
updates

1. Introduction

Sound and music computing is a young and highly multidisciplinary research field. It combines scientific, technological, and artistic methods to produce, model, and understand audio and sonic arts with the help of computers. Sound and music computing borrows methods, for example, from computer science, electrical engineering, mathematics, musicology, and psychology.

For this special issue, 44 manuscripts were submitted and were carefully reviewed. Finally, 29 high-quality articles were published, and we are very pleased with the outcome. Some of the articles are revised and extended versions of papers published earlier in related international conferences, such as in the 14th Sound and Music Computing Conference SMC-17 (Espoo, Finland), the 18th International Society for Music Information Retrieval Conference ISMIR-17 (Suzhou, China), or the 2017 New Interfaces for Musical Expression Conference NIME-17 (Copenhagen, Denmark).

This editorial briefly summarizes the published articles and guides you to read them in detail. The articles could be categorized in many ways, as such multidisciplinary field has a wide variety of topics. Here, we have organized the articles based on their application areas or special techniques applied in research. We hope that these articles will inspire researchers in sound and music computing to conduct more excellent research and spread the word about this vibrant, multidisciplinary field.

2. Sound and Music Computing Techniques

2.1. Audio Signal Processing

Cecchi et al. [1] have written the only review article for this special issue. Their long paper gives a complete overview of audio signal processing methods for the equalization of the loudspeaker-room response, which is a fundamental problem in sound reproduction. The increasing popularity of small mobile speakers having non-ideal properties makes this topic ever more important.

Necciari et al. [2] propose an improved auditory filter bank called *Audlet*, which allows perfect reconstruction. The new filter bank is compared with the gammatone filter bank, and its used in a single-channel audio source separation task is demonstrated.

Brandtsegg et al. [3] discuss approaches to real-time convolution with time-varying filters, which extends the convolution reverberation concept. For example, the sounds produced by two players can be convolved with each other to obtain exciting audio effects.

Damskägg and Välimäki [4] address a problem known as time-scale modification in which the objective is to temporally stretch or compress a given audio signal while preserving properties like pitch and timbre. To handle different signal characteristics, the main idea of the paper is to modify

the phase of the signal's time-frequency bins in an adaptive fashion using an implicit bin-wise fuzzy classification based on three classes (sinusoid, noise, transient).

Esqueda et al. [5] present virtual analogue models of the Lockhart and Serge wavefolders. The input–output relationship of both circuits was digitally modeled using the Lambert-W function. Aliasing distortion is ameliorated using a first-order antiderivative method. An earlier version of this paper received a best paper award at the SMC-17 conference.

2.2. Machine and Deep Learning

Deep learning is a hot topic also in sound and music computing. In this special issue, there are several articles applying deep learning techniques to various problems. The article by Wang et al. [6] describes an automatic music transcription algorithm combining deep learning and spectrogram factorization techniques. It is applied to a specific piano, and the results outperform the earlier methods in note-level polyphonic piano music transcription. Blaauw and Bonada [7] describe a singing synthesizer based on deep neural networks called the Neural Parametric Singing Synthesizer (NPSS), which can generate high-quality singing when a musical score and lyrics are given as the input. The NPSS can learn the timbre and expressive features of a singer from a small set of recordings. Lee et al. [8] discuss a learning approach based on convolutional neural networks (CNNs) to derive meaningful feature representations directly from the waveform of an audio signals (rather than using frame-based input representations such as the short-time Fourier transform). Such approaches are interesting in view of end-to-end music classifications tasks including genre classification and auto tagging. As one main contribution, the authors discuss the properties of the learned sample-level filters and show how their CNN-based learning approach behaves under certain downsampling and normalization effects.

Machine learning is also traditionally applied by many researchers. Green and Murphy [9] report on spatial analysis of binaural room impulse responses. The results of this article indicate that neural networks are able to detect the direction of the direct sound, but are less accurate at predicting the direction of arrival of the reflections, even in quite simple cases. More work on this topic is needed, to be able to study room acoustics with machine learning. Lovedee-Turner and Murphy [10] have collected a database of spatial sound recordings for the purpose of classification of acoustic scenes as well as the material for machine learning algorithms. To validate the database they also introduce a classifier that performs better than a traditional Mel-frequency-cepstral-coefficient classifier. The article by Pesek et al. [11] introduces algorithmic concepts for modeling and detecting recurrent patterns in symbolically encoded music. Given a monophonic symbolic representation of a piece of music, the algorithm outputs a hierarchical representation of melodic patterns using an unsupervised learning procedure without the need of hard-coded rules from music theory. Also the comprehensive article by Bountouridis [12] is concerned with pattern analysis of symbolic music representations. Inspired by multiple sequence alignment techniques that are well known in bioinformatics, the authors show how such methods can be adapted to symbolic music analysis. In particular, sequence alignment and retrieval techniques are used for measuring melodic similarities and for detecting musically interesting relations within song families. Carabez et al. [13] study a brain-computer interface, which consists of headphones and an electroencephalography- or EEG-based measurement system, which registers the user's brain activity. Using machine learning techniques, they demonstrate promising results on reading from the user's mind the direction of arrival of sound stimuli.

2.3. Automatic Transcription and Programming

Mcleod et al. [14] address a central problem in music information retrieval known as music transcription: given an audio recording of a piece of music, the goal is to extract symbolic note parameters such as note onsets and pitches. In this article, the authors focus on a-cappella music recordings with four singers (bass, tenor, alto, soprano). Combining an acoustic model based on probabilistic latent component analysis (PLCA) and a music language model based on Hidden Markov

Models (HMM), the authors present an approach for jointly tackling the problems of multi-pitch transcription as well as voice assignment.

Lazzarini [15] presents a new framework for unit generator development for the computer music language Csound, introducing the concept of unit generators and opcodes, and its centrality with regards to music programming languages in general, and Csound in specific.

3. Sound and Music Computing Applications

3.1. Sound Synthesis and User Control

Sound synthesis and control of novel and computer-based instruments are one of the main areas in sound and music computing. In Selfridge et al. [16] several physical models of objects swinging through air are presented. Listening tests showed that the models were rated as plausible as recordings. Such models are particularly interesting when used in real-time audio-visual simulations. This is a revised and extended version of a paper winning a best paper award at SMC-17 Conference. Michon et al. [17] present two original concepts: mobile device augmentation and hybrid instruments. Several tools, techniques, as well as thoughtful considerations and useful advices on how to design such instruments are presented. This paper is an extension of a paper that won the best paper award at the NIME-17 conference. The paper by MacRitchie and Milne [18] investigates four different pitch layouts on the computer screen, and finds how easy or difficult it is to play melodies on each of them. Their results lead to novel design rules for such musical instruments. Kelkar and Jensenius [19] asked people to listen to short melodies and move their hands as if their movement was creating the sound. The authors found that people tend to use one of six different mapping strategies. They also observed an interesting gender difference, as one of the strategies was more often used by women than by men.

3.2. Audio Mixing and Audio Coding

Wilson and Fazenda [20] present a method to generate automatic audio mixes. The study concerns three audio processing activities: level-balancing, stereo-panning, and equalization. The presented work will pave the way to automatize the work of audio engineers, especially in object-based audio broadcasting.

Jia et al. [21] propose an efficient, psychoacoustic coding method for multiple sound objects in a spatial audio scene. This technology can be applied to 3-D movies, spatial audio communication systems, and virtual classrooms.

3.3. Games and Virtual Reality

Hansen and Hiraga [22] introduce and evaluate *Music Puzzle*, which is an audio-based game. Interestingly, they tested the game with different user groups. People with hearing loss had problems in a game that used speech, but less with a game based on music. In contrast, people with low engagement in music performed worse in a music game. Based on this study the authors could explain the impact of hearing acuity and musical experience on focused listening of different sounds.

Schaerlaeken et al. [23] investigate the impact of playing for a virtual audience, both from the perspective of the player and the audience. The study highlights the use of immersive virtual environments as a research tool and a training assistant for musicians who are eager to learn how to cope with their anxiety in front of an audience.

Yiyu et al. [24] discuss an audio processor architecture, which is suitable for rendering a virtual acoustic environment using a finite-difference approach. Such a system can be useful for providing realistic acoustic experiences for gaming or virtual reality.

Puomio et al. [25] present a perceptual study on the effect of virtual sound source positions in spatial audio rendering using headphones with head-tracking. A listening test was conducted comparing optimized and non-optimized virtual loudspeaker setups in the simulations of a small

room and a concert hall. Their results suggest that the simulation of a small room benefits more from the optimization of virtual source positions than a large room.

3.4. Sonic Interaction, Musicology, and New Hardware

Verde et al. [26] investigate computational musicology for the study of tape music works, and existing computer vision techniques are applied to the analysis of such tracks.

Hayes and Stein [27] present an approach to incorporate environmental factors within the field of site-responsive sonic art using embedded audio and data processing techniques. The main focus is on the role of such systems within an ecosystemic framework, both in terms of incorporating systems of living organisms, as well as sonic interaction design.

In Yağanoğlu and Köse [28] a wearable vibration communication system for the deaf is presented. The wearable device proved to have a high success rate in localization tasks, which are problematic for deaf people.

Quintana-Suárez et al. [29] authored an article on a sensor device that enables to remotely monitor the activity and health of elderly people. Such technology is generally called Ambient Assisted Living, and this article in particular presents a low-cost acoustic sensor.

4. Conclusions

Overall the special issue shows the variety of topics researched by the sound and music community, ranging from spatial sound, sound processing, sonic interaction design, and music information retrieval to new interfaces for musical expression with applications in art, culture, gaming, and virtual and augmented reality.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cecchi, S.; Carini, A.; Spors, S. Room Response Equalization—A Review. *Appl. Sci.* **2018**, *8*, 16.
2. Necciari, T.; Holighaus, N.; Balazs, P.; Průša, Z.; Majdak, P.; Derrien, O. Audlet Filter Banks: A Versatile Analysis/Synthesis Framework Using Auditory Frequency Scales. *Appl. Sci.* **2018**, *8*, 96.
3. Brandtsegg, Ø.; Saue, S.; Lazzarini, V. Live Convolution with Time-Varying Filters. *Appl. Sci.* **2018**, *8*, 103.
4. Damskägg, E.P.; Välimäki, V. Audio Time Stretching Using Fuzzy Classification of Spectral Bins. *Appl. Sci.* **2017**, *7*, 1293.
5. Esqueda, F.; Pöntynen, H.; Parker, J.D.; Bilbao, S. Virtual Analog Models of the Lockhart and Serge Wavefolders. *Appl. Sci.* **2017**, *7*, 1328.
6. Wang, Q.; Zhou, R.; Yan, Y. A Two-Stage Approach to Note-Level Transcription of a Specific Piano. *Appl. Sci.* **2017**, *7*, 901.
7. Blaauw, M.; Bonada, J. A Neural Parametric Singing Synthesizer Modeling Timbre and Expression from Natural Songs. *Appl. Sci.* **2017**, *7*, 1313.
8. Lee, J.; Park, J.; Kim, K.L.; Nam, J. SampleCNN: End-to-End Deep Convolutional Neural Networks Using Very Small Filters for Music Classification. *Appl. Sci.* **2018**, *8*, 150.
9. Green, M.C.; Murphy, D. EigenScape: A Database of Spatial Acoustic Scene Recordings. *Appl. Sci.* **2017**, *7*, 1204.
10. Lovedee-Turner, M.; Murphy, D. Application of Machine Learning for the Spatial Analysis of Binaural Room Impulse Responses. *Appl. Sci.* **2018**, *8*, 105.
11. Pesek, M.; Leonardis, A.; Marolt, M. SymCHM—An Unsupervised Approach for Pattern Discovery in Symbolic Music with a Compositional Hierarchical Model. *Appl. Sci.* **2017**, *7*, 1135.
12. Bountouridis, D.; Brown, D.G.; Wiering, F.; Veltkamp, R.C. Melodic Similarity and Applications Using Biologically-Inspired Techniques. *Appl. Sci.* **2017**, *7*, 1242.
13. Carabez, E.; Sugi, M.; Nambu, I.; Wada, Y. Identifying Single Trial Event-Related Potentials in an Earphone-Based Auditory Brain-Computer Interface. *Appl. Sci.* **2017**, *7*, 1197.

14. McLeod, A.; Schramm, R.; Steedman, M.; Benetos, E. Automatic Transcription of Polyphonic Vocal Music. *App. Sci.* **2017**, *7*, 1285.

15. Lazzarini, V. Supporting an Object-Oriented Approach to Unit Generator Development: The Csound Plugin Opcode Framework. *Appl. Sci.* **2017**, *7*, 970.

16. Selfridge, R.; Moffat, D.; Reiss, J.D. Sound Synthesis of Objects Swinging through Air Using Physical Models. *Appl. Sci.* **2017**, *7*, 1177.

17. Michon, R.; Smith, J.O.; Wright, M.; Chafe, C.; Granzow, J.; Wang, G. Mobile Music, Sensors, Physical Modeling, and Digital Fabrication: Articulating the Augmented Mobile Instrument. *Appl. Sci.* **2017**, *7*, 1311.

18. MacRitchie, J.; Milne, A.J. Exploring the Effects of Pitch Layout on Learning a New Musical Instrument. *Appl. Sci.* **2017**, *7*, 1218.

19. Kelkar, T.; Jensenius, A.R. Analyzing Free-Hand Sound-Tracings of Melodic Phrases. *Appl. Sci.* **2018**, *8*, 135.

20. Wilson, A.; Fazenda, B.M. Populating the Mix Space: Parametric Methods for Generating Multitrack Audio Mixtures. *Appl. Sci.* **2017**, *7*, 1329.

21. Jia, M.; Zhang, J.; Bao, C.; Zheng, X. A Psychoacoustic-Based Multiple Audio Object Coding Approach via Intra-Object Sparsity. *Appl. Sci.* **2017**, *7*, 1301.

22. Hansen, K.F.; Hiraga, R. The Effects of Musical Experience and Hearing Loss on Solving an Audio-Based Gaming Task. *Appl. Sci.* **2017**, *7*, 1278.

23. Schaerlaeken, S.; Grandjean, D.; Glowinski, D. Playing for a Virtual Audience: The Impact of a Social Factor on Gestures, Sounds and Expressive Intents. *Appl. Sci.* **2017**, *7*, 1321.

24. Yiyu, T.; Inoguchi, Y.; Otani, M.; Iwaya, Y.; Tsuchiya, T. A Real-Time Sound Field Rendering Processor. *Appl. Sci.* **2018**, *8*, 35.

25. Puomio, O.; Pätynen, J.; Lokki, T. Optimization of Virtual Loudspeakers for Spatial Room Acoustics Reproduction with Headphones. *Appl. Sci.* **2017**, *7*, 1282.

26. Verde, S.; Pretto, N.; Milani, S.; Canazza, S. Stay True to the Sound of History: Philology, Phylogenetics and Information Engineering in Musicology. *Appl. Sci.* **2018**, *8*, 226.

27. Hayes, L.; Stein, J. Desert and Sonic Ecosystems: Incorporating Environmental Factors within Site-Responsive Sonic Art. *Appl. Sci.* **2018**, *8*, 111.

28. Yağanoğlu, M.; Köse, C. Wearable Vibration Based Computer Interaction and Communication System for Deaf. *Appl. Sci.* **2017**, *7*, 1296.

29. Quintana-Suárez, M.A.; Sánchez-Rodríguez, D.; Alonso-González, I.; Alonso-Hernández, J.B. A Low Cost Wireless Acoustic Sensor for Ambient Assisted Living Systems. *Appl. Sci.* **2017**, *7*, 877.