Sound and Music Computing at Aalborg University in Copenhagen

Stefania Serafin, Sofia Dahl, Amalia de Götzen, Cumhur Erkut, Dan Overholt, Hendrik Purwins and Bob Sturm
Aalborg University Copenhagen
sts@create.aau.dk

ABSTRACT

In this studio report we present the research and teaching activities of the Sound and Music Computing Group at Aalborg University Copenhagen. A new Master education in Sound and Music Computing which starts on September 2014 is introduced.

1. INTRODUCTION

Aalborg University Copenhagen has recently expanded the research and teaching activities in the field of Sound and Music Computing. Seven full time faculty members are now hired in different areas of Sound and Music Computing, and a new Master education is starting on September 2014. In this studio report, we first describe the research activities of the different faculty members. We then introduce the laboratories available, and we present the aims of the new master education.

2. MUSIC INFORMATION RETRIEVAL

Funded by a 2-year grant from the national Danish Research Fund, associate professor Bob L. Sturm has closely surveyed most of the literature regarding music genre recognition, and much about music emotion recognition and autotagging, to determine how such systems have been evaluated . For instance, we examined the 100 best classification accuracies for music genre recognition systems in the standard dataset GTZAN [1, 2]. It turns out that all these results, however, say nothing about which system is better than any other, and which, if any, is even addressing the problem of music genre recognition [3]. In short, the standard, systematic and rigorous approaches used in evaluation in music information retrieval (MIR) research for the past 20 years appear to not be *scientific* approaches to evaluation [4, 5, 2, 3]. Sturm has demonstrated that there are ways to evaluate MIR systems that are valid with respect to answering whether they are actually addressing the problems for which they are designed [6, 3, 7]. As a result, many new directions have emerged that can help complete the evaluation cycle that is missing in MIR research [8].

Hendrik Purwins works on statistical time series analysis starting with a (multi-dimensional) time-series as input, subject to: 1. preprocessing (filtering or transformation into the frequency or wavelet domain, or a domain spanned by learnt atomic functions), 2. probabilistic description of the time series, 4. statistical methods for recognition, visualization, and predictions of signatures in the time series, and 5. model evaluation. Currently he applies these methods to the comparison of theory/practice of intonation in Byzantine chant [9], decoding auditory attention to single voices in polyphonic music using EEG [10], automatic phrase continuation from drum loops, harmonic sequences and bass-guitar melodies [11], and sound texture resynthesis.

3. SONIC INTERACTION DESIGN

Sonic interaction design is an emerging field which was recently defined thanks to a successful COST action, as the study and exploitation of sound as one of the principal channels conveying information, meaning, and aesthetic/emotional qualities in interactive contexts. This field lies at the intersection of interaction design and sound and music computing [12]. At Aalborg University Copenhagen we examine sonic interaction from different angles, as shown in the following.

3.1 Sonic interactions for multimodal environments

Since we live in a multi sensorial world, sonic interactions become more meaningful when combined with simulations of other modalities. It is well known that sound can indeed complement, enhance or even substitute other senses.

Stefania Serafin's research explores sonic interactions when combined with other senses such as haptic and visual feedback. This includes the physics based simulation of multimodal interactions together with evaluation of the user experience.

Recent applications have been focused on the field of virtual reality, where the focus on auditory feedback has been rather limited when compared, for example, to the focus placed on visual feedback or even on haptic feedback .

Other applications have been in the field of cultural heritage, in order to use sonic interaction technologies to reconstruct and preserve musical instruments. In particular, thanks to a project supported by the Culture 2000 EU framework, we reconstructed and exhibited the devices and the music of the Rai Studio di Fonologia Musicale in Milan see article in [13]. In a recently funded project by the

Danish Sound Technology Network in collaboration with Volvo User Experience Lab and the Copenhagen Institute for Interaction Design, the possibilities to provide auditory and haptic feedback in cars is investigated, in order to reduce the high visual load provided to drivers.

An issue with physics based models and new interfaces for musical expression is the fact that most interfaces and software tools are built and used only by the instrument maker. In order to cope with this limitation, we are collaborating with Imogen Heap's team in order to combine the Gloves with our physics based musical instruments built as externals to the Max6 platform [14].

3.2 Music performance and sonic interaction design

The research of Amalia de Götzen encompasses two main areas: music performance and sonic interaction design. In the area of music performance her focus is on the perception of sound movement expressiveness [15]. Music and Emotion have been widely studied with reference to many musical aspects: tempo, timbre, melody, harmony, agogic etc. but in contemporary music also sound movement in space is used as an expressive musical parameter. This investigation ranges from listner perception to movement characteristics that impact expressiveness.

Her focus research area on interaction spans from sonic interaction of objects to foley representations of real world sounds [16, 17]. Both topics are related to expressiveness as well from two different perspectives: the coupling between the expressive gesture that produce a sound with the use of physical models to synthesize it and the analysis of foley sounds in order to produce expressive surrogates of everyday sounds.

3.3 Mobile sonic interaction design

As contemporary smartphones and tablets become ubiquitous, their nature changes from mere function to a tool for self- expression, creativity, or play. Mobile application development embraces this change, as evidenced by the popularity and profits of sound and music applications and games in various application stores. These application domains are attractive in education, especially in computer science, media technology, and interaction design. The main question is how to systematically integrate the rapidly evolving knowledge, know-how, tools, and techniques of mobile (audio) programming and mobile interaction design into university curricula. Accounts of our development and teaching experience are provided in [18], and applications are reported, e.g., in [19, 20]. This experience has been put into practice as the Mobile Lab of Sonic Interaction (MOLSI) (see below), run by Cumhur Erkut, who has joined AAU CPH in summer 2013.

During his previous work in Finland, Dr. Erkut has focused on mobile SMC in research, teaching and supervision. Among the applications developed, AudioReitit was an exercise in mobile sonic interaction for control and awareness of time. The iOS application facilitated open APIs (Helsinki Regional Transport and Google Weather) together with GPS and minimal GUI input, and used auditory display to provide information. The application has received two awards (Second prize in the category of "Innovative Interfaces" by Helsinki Regional Traffic (HSL) in 2011 and a special prize from the Helsinki City Data Center in December 2012).

Based on this experience, we note that both the public and the private sector strategically consider mobile applications as additional touch points besides the physical front and back offices. To help user involvement and participation in value creation, they also provide open data and APIs. In Denmark, for instance, there is a dedicated portal enlisting such APIs (http://digitaliser.dk). We therefore plan to harness this data and APIs for designing novel services and applications, equipped with advanced auditory displays in collaboration both with product sound design and service design.

4. NEW INTERFACES FOR MUSICAL EXPRESSION

Music, at its core, is a cultural phenomenon. In this context, Dan Overholts research explores both practical and theoretical aspects at the forefront of todays music interface technologies, facilitating expressive and collaborative musical activities. Just as language is a tool created by humans which in turn shapes and informs our thinking and actions, music is shaped by the tools and methods we use to produce it. In the Augmented Performance Lab, advanced explorations of musical expressivity are pursued through the creation of new instruments, interfaces and techniques.

One of Overholts recent projects centered on intercultural investigations, using technology to augment genre-mixing in musical performance. The international project was funded by the Danish Council for Research and Innovation, and involved partner researchers at U.C. Berkeleys Center for New Music and Audio Technologies (CNMAT). The project looked at how new musical instruments can be used in culturally collaborative music practices, examining the influence of music technology on different music styles under a microscope, and extending the ways in which interactive systems can affect human musical collaborations between cultures. The first scientific focus of the project was to improve existing methods and theories for creating such musical interfaces. The second focus was to explore the use of these in the development of new hybrid acoustic-electronic music instruments, in order to encourage intercultural collaborations [21]. The research enhances cross-cultural collaborative music-making by employing computer technology within interactive performance interfaces and environments.

Overholts other funded research projects include Hybrid Instruments [22], culturally inclusive physical therapy for the disabled using music and gaming as motivational tools [23], and Culturally Enhanced Augmented Realities (CultAR), an EU-funded (STREP FP7) project with partners in Finland, Austria,and Italy. The main objective of CultAR is to develop an integrated mobile full mixed reality platform combining markerless augmented reality, mobile 3D maps, tactile interfaces and directional audio for advanced context-aware personalized and digital cultural experience in urban environments. With CultAR, the goal is to build a hybrid space of ones urban surroundings and the related cultural information space, and strengthen the connection in a truly ubiquitous and personal manner. Overholts current research within CultAR is focused on cross-modal human perceptual sensitivity to audio / haptic spatialisation using bone conduction. This includes an eight-channel headband with bone conduction transducers capable of producing stimuli within both the haptic (sub-audio) and full-range audio frequencies, and techniques such as sensory substitution within augmented-reality soundscapes.

5. EMBODIED MUSIC COGNITION

Body movement is an essential part of most music performance. Not only do instrumentalists spend years in refining their playing movements in order to control minute details of the sound production, but their overall body movements are also conveying important information to the audience (c.f. [24, 25]). Furthermore, there is also converging evidence that our own body movements affect how we perceive music. For instance, our perception and preference of rhythm and tempo is linked to our bodies and how they move [26, 27].

Sofia Dahl's research investigates all these aspects of embodied music cognition and spans disciplines such as musicology, psychology, neuroscience, music performance, and music acoustics. A special focus of Dahl's research is rhythmic movements and their link to our perception and control of timing and tempo [28, 29]. Current research questions include how full body movements help musicians to control sound and timing during playing; how we perceive temporal changes; and when and how we experience music to have qualities that encourages us to move.

6. LABORATORIES

6.1 Multisensory experience lab

In the Multi-Sensory Experience Lab we research the integration of different senses by combining technology, computer simulations and user experience evaluation. The lab consists of three main spaces. The larger space is used for multimodal (audio-visual-haptic) simulations, and contains a motion capture system (16 cameras motion Optitrack system by NaturalPoint), a nVisor SX head mounted display and Oculus head mounted display, a 24 channels surround sound system (Dynaudio BM5A), and several devices for haptic feedback. In addition, the lab contains an anechoic chamber and a 64 speakers wavefield synthesis system.

6.2 Augmented cognition lab

The Augmented Cognition Lab is dedicated to the study of perception, cognition, affective states and aesthetic experience in digital and multimodal media and cognitive technologies. In particular involving complex stimuli and interactive displays. The lab counts with state of the art equipment for measuring brain activity (electroencephalography EEG) and several psychophysiological measurements devices and methods such as muscle activity (EMG), eye-tracking devices, heart and pulse rate measurements, computer vision and thermal imaging technology for recognition of affective states.

6.3 Mobile Lab of Sonic Interaction

The Mobile Lab of Sonic Interaction (MOLSI) is a capsule laboratory for the design and development of mobile applications. The apps that promote interaction with or through sound are the focus area of MOLSI. The MOLSI facilities include popular development tools for iOS (Objective-C), Android (Java), and cross-platform (C sharp with Xamarin), development certificates for on-device evaluation, creative coding platforms such as OpenFrameworks or Processing, sensors (Leap Motion, Kinect), and a capsule multichannel audio setup with Genelec 6010A and 5040A speakers.

6.4 Augmented performance lab

The Augmented Performance Lab (APL) is a performance space that contains 3D audio, depth cameras to record and process real-time musical performance either carried out on augmented instruments or new interfaces, including mobile devices. In a second room there is a control space that streams the performance data, analyzes it in real-time and simultaneously records both the control data and multichannel audio.

7. THE SMC EDUCATION

From September 1st, 2014, Aalborg University in Copenhagen will offer a Master of Science in Sound and Music Computing. The Master of Science is a 2-year, research-based, full-time study programme, set to 120 ECTS credits. Its mission is to train the next generation of professionals to push forward the sound and music technologies of the new information society. By combining practical and theoretical approaches in topics such as computational modeling, audio engineering, perception, cognition, and interactive systems, the programme gives the scientific and technological background needed to start a research or professional career. This program trains students on the technologies for the analysis, description, synthesis, transformation and production of sound and music, and on the technologies and processes that support sound and music creation.

8. CONCLUSIONS

In this studio report we have presented the research and education activities which take place in the Sound and Music Computing group at Aalborg University in Copenhagen.

More information can be found here:
http://media.aau.dk/smc

9. REFERENCES

[1] G. Tzanetakis and P. Cook, "Musical genre classification of audio signals," *IEEE Trans. Speech Audio Process.*, vol. 10, no. 5, pp. 293–302, July 2002.

[2] B. L. Sturm, "The GTZAN dataset: Its contents, its faults, their effects on evaluation, and its future use," *http://arxiv.org/abs/1306.1461*, 2013.

[3] ——, "The state of the art ten years after a state of the art: Future research in music information retrieval," *J. New Music Research (in press)*, 2014.

[4] ——, "A survey of evaluation in music genre recognition," in *Proc. Adaptive Multimedia Retrieval*, Oct. 2012.

[5] ——, "Classification accuracy is not enough: On the evaluation of music genre recognition systems," *J. Intell. Info. Systems*, vol. 41, no. 3, pp. 371–406, 2013.

[6] ——, "A simple method to determine if a music information retrieval system is a "horse"," *(submitted)*, 2014.

[7] F. Gouyon, B. L. Sturm, J. L. Oliveira, N. Hespanhol, and T. Langlois, "On evaluation in music autotagging research," *(submitted)*, 2013.

[8] J. Urbano, M. Schedl, and X. Serra, "Evaluation in music information retrieval," *J. Intell. Info. Systems*, vol. 41, no. 3, pp. 345–369, Dec. 2013.

[9] M. Panteli and H. Purwins, "A quantitative comparison of chrysanthine theory and performance practice of scale tuning, steps, and prominence of the octoechos in byzantine chant," *Journal of New Music Research*, vol. 42, no. 3, pp. 205–221, 2013.

[10] M. S. Treder, H. Purwins, D. Miklody, I. Sturm, and B. Blankertz, "Decoding auditory attention to instruments in polyphonic music using single-trial eeg classification," *Journal of neural engineering*, vol. 11, no. 2, p. 026009, 2014.

[11] S. Cherla, H. Purwins, and M. Marchini, "Automatic phrase continuation from guitar and bass guitar melodies," *Computer Music Journal*, vol. 37, no. 3, pp. 68–81, 2013.

[12] K. Franinovic and S. Serafin, *Sonic Interaction Design*. Mit Press, 2013.

[13] M. M. Novati and J. Dack, *The Studio Di Fonologia: a musical journey 1954-1983, update 2008-2012*. Ricordi, 2012.

[14] S. S. et al., "Controlling physically based virtual musical instruments using the gloves," in *Proc. New Interfaces for Musical Expression, 2014*, June 2014.

[15] A. de Götzen, E. Sikstrm, D. Korsgaard, S. Serafin, and F. Grani, *The grouping of sound movements as expressive gestures*. Springer, 2013.

[16] S. Delle Monache, D. Rocchesso, J. Qi, L. Buechley, and A. de Götzen, *Paper mechanisms for sonic interaction*. Association for Computing Machinery, 2012, pp. 61–68.

[17] A. de Götzen, E. Sikstrm, F. Grani, and S. Serafin, *Real, foley or synthetic? An evaluation of everyday walking sounds*. Logos Verlag Berlin, 2013, pp. 487–492.

[18] C. Erkut, A. Jylhä, and S. Serafin, "(and Sound) of SiMPE: Showcasing Outcomes of a Mobile Audio Programming Seminar," in *SimPE '13 in MobileHCI '13*, Munich, Germany, Aug. 2013.

[19] S. Baldan and S. Serafin, "Design and evaluation of Sonic Tennis: an audio-only iPhone game," in *SimPE '13 in MobileHCI '13*, Aug. 2013.

[20] C. Erkut, "Mobile Probes for Special Needs," in *SimPE '13 in MobileHCI '13*, Munich, Germany, Aug. 2013.

[21] D. Overholt and A. Freed, "Grafting Acoustic Instruments and Signal Processing: Creative Control and Augmented Expressivity," in *Abstract in 166th Meeting of the Acoustical Society of America*, Dec. 2013.

[22] D. Overholt, "Advancements in Violin-Related Human-Computer Interaction," in *Arts and Technology : Second International Conference, ArtsIT 2011, Esbjerg Denmark, December 2011, Revised Selected Papers. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering (Book 101)*, Sep. 2012.

[23] J. Larsen, D. Overholt, and T. Moeslund, "The Actuated Guitar : A platform enabling alternative interaction methods," in *Proceedings of the Sound and Music Computing Conference 2013*, Logos Verlag Berlin, Germany, Jul. 2013.

[24] S. Dahl, F. Bevilacqua, R. Bresin, M. Clayton, L. Leante, I. Poggi, and N. Rasamimanana, "Gestures in performance," in *Musical gestures: Sound, movement, and meaning*. New York: Routledge, 2009, pp. 36–68.

[25] S. Dahl and A. Friberg, "Visual perception of expressiveness in musicians' body movements," 2007.

[26] J. Phillips-Silver and L. J. Trainor, "Hearing what the body feels: Auditory encoding of rhythmic movement," *Cognition*, vol. 105, no. 3, pp. 533–546, 2007.

[27] S. Dahl, D. Huron, G. Brod, and E. Altenmüller, "Preferred dance tempo: Does sex or body morphology influence how we groove?" *Journal of New Music Research*, 2014.

[28] K. Petrini, F. E. Pollick, S. Dahl, P. McAleer, L. McKay, D. Rocchesso, C. H. Waadeland, S. Love, F. Avanzini, and A. Puce, "Action expertise reduces brain activity for audiovisual matching actions: An fmri study with expert drummers," *NeuroImage*, vol. 56, no. 3, pp. 1480–1492, 2011.

[29] S. Dahl, "Striking movements: A survey of motion analysis of percussionists," *Acoustical science and technology*, vol. 32, pp. 168–173, 2011.