Multimedia Technologies for Enriched Music Performance, Production, and Consumption

Liem, Cynthia C.S.; Gómez , Emilia; Tzanetakis, George

DOI
10.1109/MMUL.2017.20

Publication date
2017

Document Version
Final published version

Published in
IEEE Multimedia

Multimedia Technologies for Enriched Music Performance, Production, and Consumption

Cynthia C.S. Liem
Delft University of Technology

Emilia Gómez
Universitat Pompeu Fabra

George Tzanetakis
University of Victoria

Internet access, mobile devices, social networks, and automated multimedia technologies enable sophisticated information analysis and access. They have radically changed how people find entertainment, discover new interests, and generally express themselves online—seemingly without any physical or social barriers. Furthermore, thanks to the increasing affordability of sensing, storage, and sharing, information is taking on increasingly rich and hybrid multimedia forms. We now are frequently confronted with multimodal information streams, which are co-occurring in various social consumption settings.

This phenomenon has also enabled opportunities in the music domain. In terms of music performance, musicians can exploit (live) analysis of and novel interaction mechanisms with musical data in multiple modalities, creating new forms of expression. For music production, sophisticated multimedia data analysis techniques can lead to more efficient and scalable workflows and richer and more accessible interfaces. In music consumption, music data containing embedded contextual and social information can lead to novel consumer experiences, stimulating greater music appreciation. This way, concerts turn into multimodal, multiperspective, and multilayered digital artifacts that can be easily explored, customized, personalized, (re)enjoyed, and shared among various types of users. Similar notions and opportunities hold for the consumption of general music recordings.

The goal of this special issue was to gather state-of-the-art research on multimedia methods and technologies aimed at enriching music performance, production, and consumption. In recent years, researchers in the multimedia community have started looking at music as not just an audio-only phenomenon but as something that should be studied in multimedia contexts.[1] Another aspect of interest to the multimedia community is that to become meaningful, music requires human listeners. As such, music encompasses both a multimodal representation and a human consumption context.

Research Challenges

When considering multimedia-related opportunities for enriching music performance, production, and consumption, various research challenges exist.

As noted, music information can reach beyond isolated audio and thus can be more appropriately modeled as multimodal information streams, including combinations of audio, video, images, music scores, text, and gestures.[2–6] Analyzing, describing, and indexing this information presents considerable research challenges: frequently, information that is

1070-986X/17/$33.00 © 2017 IEEE

Published by the IEEE Computer Society

Figure 1. Live visualization from the PHENICX project:[14] various types of synchronized music-related information during a symphonic performance.

salient or meaningful to humans is not trivially extracted from the signals.

Music performance and production usually involve human creators. This requires novel interaction mechanisms, vocabularies, and user interfaces to support creative processes for experts and novices alike.[7–11] At the same time, music consumption and appreciation fundamentally relies on human audience members. Audience engagement can improve through novel user experience paradigms, social networking, and other techniques targeted at improving audience immersion and inclusion. For example, in parallel to a musical timeline, supporting information can be provided that stimulates active listening at different expertise levels.[12,13] Figure 1 shows an example of how this can be done in a live setting through visualization.[14] Even more challenging is the case in which performers and audiences engage with one another through a digitally mediated environment.[15]

Beyond the concert hall, music consumption is also often a part of everyday (non-musical) activities.[16] In a situation of digital abundance, getting the right music to the right audience will require advances in user awareness, personalization, intent, context awareness, and automatic context adaptation. However, such topics have thus far been underrepresented in music information retrieval research.[17]

In This Issue

The four articles selected for this special issue cover a broad range of topics related to enriched music experiences.

In "Enriched Multimodal Representations of Music Performances: Online Access and Visualization," Esteban Maestre, Panos Papiotis, Marco Marchini, Quim Llimona, Oscar Mayor, Alfonso Pérez, and Marcelo Wanderley consider the acquisition and representation of multimodal music performance recordings. The authors present a case study on the acquisition of an annotated dataset of multimodal string quartet recordings, including audio, video, and gesture data. This includes the presentation of Repovizz, a system for online sharing and visualization of multimodal, time-aligned datastreams. The authors release their string quartet dataset to the research community and invite researchers to host and exchange their own multimodal music performance data through the Repovizz platform.

In "Toward Computer-Assisted Understanding of Dynamics in Symphonic Music,"

The model can be employed in musicology-oriented use cases in which differences between various performances of the same piece must be analyzed.

Maarten Grachten, Carlos Eduardo Cancino Chacón, Thassilo Gadermaier, and Gerhard Widmer study characteristics of recorded music performances—more specifically, of classical symphonic music. They propose a computational model for predicting expressive dynamics in ensemble performances through basis-function modeling. The basis functions represent information in the musical score. An additional challenge provided by the given use case is that the score reflects information for multiple players at once, but an overall dynamics assessment needs to be made. Different variants of the model are tested, showing that nonlinear modeling is beneficial for prediction. The model can, for example, be employed in musicology-oriented use cases in which differences between various performances of the same piece must be analyzed.

The article "Open Symphony: Creative Participation for Audiences of Live Music Performances," by Yongmeng Wu, Leshao Zhang, Nick Bryan-Kinns, and Mathieu Barthet, also considers music performance. However, in contrast to the previous article, which analyzes recorded performed music, this work pioneers opportunities for creative audience participation during a live music performance. The authors present a typology for participatory live music performance (PLMP) systems, together with the Open Symphony system, which lets audience members interactively and multimodally vote on different playing modes for musical ensemble members during a live improvisational performance. The qualitative and quantitative feedback obtained during two performance sessions reveals both challenges and opportunities for technology-enabled creative communication from audiences to performers and will be useful for informing the design of future interactive systems.

Finally, in "A Method and Toolkit for Digital Musical Instruments: Generating Ideas and Prototypes," Filipe Calegario, Marcelo Wanderley, Stephane Huot, Giordano Cabral, and Geber Ramalho focus on the physicality of music-making and on tools that help create new digital musical instruments. More specifically, they propose a design method for generating ideas on how to hold new digital musical instruments and control their sound. Furthermore, they present the Probatio physical prototyping toolkit. The modular toolkit involves two bases and five control element blocks, which can be employed with existing sound synthesizers. Initial user study responses give insight into the positive and negative aspects of the current prototype, suggestions for improvements, and possible use contexts for future prototypes.

The articles in this special issue demonstrate the diversity and breadth of research into enriched music experiences, particularly in the context of music performance. All articles contain pioneering work in the area and point toward interesting future lines of research in the field. Possible future work not covered in this issue was also touched on with the many topics presented in the "Research Challenges" section.

Enabling richer music experiences will naturally help inspire better experiences in the broader field of multimedia. As such, advances in the field will improve the overall quality of our interaction with multimedia objects. **MM**

Acknowledgments

We thank all reviewers who contributed to the selection of articles for this special issue.

References

1. C.C.S. Liem et al., "The Need for Music Information Retrieval with User-Centered and Multimodal Strategies," *Proc. 1st Int'l Workshop on Music Information Retrieval with User-Centered and Multimodal Strategies (MIRUM) at ACM Multimedia*, 2011, pp. 1–6.
2. A. Bazzica, C.C.S. Liem, and A. Hanjalic, "On Detecting the Playing/Non-Playing Activity of Musicians in Symphonic Music Videos," *Computer*

Vision and Image Understanding, Mar. 2016, pp. 188–204.
3. J. Lībeks and D. Turnbull, "You Can Judge an Artist by an Album Cover: Using Images for Music Annotation," *IEEE MultiMedia*, vol. 18, no. 4, 2011, pp. 30–37.
4. M. Müller et al., "Freischütz Digital: A Multimodal Scenario for Informed Music Processing," *Proc. Int'l Workshop on Image and Audio Analysis for Multimedia Interactive Services* (WIAMIS), 2013, pp. 1–4.
5. Á. Sarasúa, B. Caramiaux, and A. Tanaka, "Machine Learning of Personal Gesture Variation in Music Conducting," *Proc. 2016 CHI Conf. Human Factors in Computing Systems*, 2016, pp. 3428–3432.
6. A. Schindler and A. Rauber, "Harnessing Music-Related Visual Stereotypes for Music Information Retrieval," *ACM Trans. Intelligent Systems and Technology* (TIST), vol. 8, no. 2, 2016, article no. 20.
7. L. Dahl, "Triggering Sounds from Discrete Air Gestures: What Movement Feature Has the Best Timing?" *Proc. 14th Int'l Conf. New Interfaces for Musical Expression* (NIME), 2014, pp. 201–206.
8. K. Graham-Knight and G. Tzanetakis, "Adaptive Music Technology Using the Kinect," *Proc. 8th ACM Int'l Conf. Pervasive Technologies Related to Assistive Environments*, 2015, article no. 32.
9. P. Knees and K. Andersen, "Searching for Audio by Sketching Mental Images of Sound: A Brave New Idea for Audio Retrieval in Creative Music Production," *Proc. 2016 ACM Int'l Conf. Multimedia Retrieval* (ICMR), 2016, pp. 95–102.
10. P. Seetharaman and B. Pardo, "Audealize: Crowdsourced Audio Production Tools," *J. Audio Eng. Soc.*, vol. 64, no. 9, 2016, pp. 683–695.
11. G. Tzanetakis, "Natural Human-Computer Interaction with Musical Instruments," *Digital Tools for Computer Music Production and Distribution*, IGI Global, 2016, pp. 116–136.
12. M. Goto, "Active Music Listening Interfaces Based on Signal Processing," *Proc. 2007 IEEE Int'l Conf. Acoustics, Speech and Signal Processing* (ICASSP), vol. 4, 2007, pp. IV–1441.
13. M.S. Melenhorst and C.C.S. Liem, "Put the Concert Attendee in the Spotlight. A User-Centered Design and Development Approach for Classical Concert Applications," *Proc. 16th Conf. Int'l Soc. Music Information Retrieval* (ISMIR), 2015, pp. 800–806.
14. C.C.S. Liem, E. Gómez, and M. Schedl, "PHENICX: Innovating the Classical Music Experience," *Proc. 2015 IEEE Int'l Conf. Multimedia and Expo* (ICME), 2015; doi: 10.1109/ICMEW.2015.7169835.
15. E. Geelhoed et al., "Co-Present and Remote Audience Experiences: Intensity and Cohesion," *Multimedia Tools and Applications*, 2016, pp. 1–34.
16. A. Demetriou, M. Larson, and C.C.S. Liem, "Go with the Flow: When Listeners Use Music as Technology," *Proc. 17th Int'l Soc. Music Information Retrieval Conf.* (ISMIR), 2016.
17. M. Schedl, A. Flexer, and J. Urbano, "The Neglected User in Music Information Retrieval Research," *J. Intelligent Information Systems*, vol. 41, no. 3, 2013, pp. 523–539.

Cynthia C.S. Liem is an assistant professor in the Multimedia Computing Group at Delft University of Technology, The Netherlands. Her research interests include music and multimedia search and recommendation, with a special interest in multimodal analysis and dynamic and contextual characteristics of user engagement. Liem received her PhD in computer science from Delft University of Technology (partially sponsored through a Google European Doctoral Fellowship). Contact her at c.c.s.liem@tudelft.nl.

Emilia Gómez is an associate professor (Serra-Húnter and ICREA Academia Fellow) in the Department of Information and Communication Technologies at the Universitat Pompeu Fabra (UPF), Barcelona, where she leads the Music Information Research Lab in the Music Technology Group. Her research interests include the music information retrieval (MIR) field. She tries to understand the way humans describe music and emulate these descriptions through computational models dealing with large music collections. Gómez received her PhD in computer science at the UPF. She currently is president-elect of the International Society for Music Information Retrieval. Contact her at emilia.gomez@upf.edu.

George Tzanetakis is a professor in the Department of Computer Science with cross-listed appointments in ECE and Music at the University of Victoria, Canada. His research interests span all stages of audio content analysis, such as feature extraction, segmentation, and classification, with specific emphasis on music information retrieval. More recently, he has been exploring new interfaces for musical expression, music robotics, computational ethnomusicology, and computer-assisted music instrument tutoring. Tzanetakis received his PhD in computer science from Princeton University. Contact him at gtzan@uvic.ca.