TISMIR special collection - Call for Papers

Multi-Modal Music Information Retrieval

Deadline for submissions - now extended to 01.09.2024

We have been delighted with the response to this collection and due to numerous requests for additional time, we are extending the deadline for all submissions to the 1st of September to allow a bit more time for teams to polish off their manuscripts and ensure high quality submissions for this collection.

Scope of the Special Collection

Data related to and associated with music can be retrieved from a variety of sources or modalities: audio tracks; digital scores; lyrics; video clips and concert recordings; artist photos and album covers; expert annotations and reviews; listener social tags from the Internet; and so on. Essentially, the ways humans deal with music are very diverse: we listen to it, read reviews, ask friends for recommendations, enjoy visual performances during concerts, dance and perform rituals, play musical instruments, or rearrange scores.

As such, it is hardly surprising that we have discovered multi-modal data to be so effective in a range of technical tasks that model human experience and expertise. Former studies have already confirmed that music classification scenarios may significantly benefit when several modalities are taken into account. Other works focused on cross-modal analysis, e.g., generating a missing modality from existing ones or aligning the information between different modalities.

The current upswing of disruptive artificial intelligence technologies, deep learning, and big data analytics is quickly changing the world we are living in, and inevitably impacts MIR research as well. Facilitating the ability to learn from very diverse data sources by means of these powerful approaches may not only bring the solutions to related applications to new levels of quality, robustness, and efficiency, but will also help to demonstrate and enhance the breadth and interconnected nature of music science research and the understanding of relationships between different kinds of musical data.

In this special collection, we invite papers on multi-modal systems in all their diversity. We particularly encourage under-explored repertoire, new connections between fields, and novel research areas. Contributions consisting of pure algorithmic improvements, empirical studies, theoretical discussions, surveys, guidelines for future research, and introductions of new data sets are all welcome, as the special collection will not only address multi-modal MIR, but also cover multi-perspective ideas, developments, and opinions from diverse scientific communities.

Sample Possible Topics

- State-of-the-art music classification or regression systems which are based on several modalities
- Deeper analysis of correlation between distinct modalities and features derived from them
- Presentation of new multi-modal data sets, including the possibility of formal analysis and theoretical discussion of practices for constructing better data sets in future
- Cross-modal analysis, e.g., with the goal of predicting a modality from another one
- Creative and generative AI systems which produce multiple modalities
- Explicit analysis of individual drawbacks and advantages of modalities for specific MIR tasks
- Approaches for training set selection and augmentation techniques for multi-modal classifier systems
- Applying transfer learning, large language models, and neural architecture search to multi-modal contexts
- Multi-modal perception, cognition, or neuroscience research
- Multi-objective evaluation of multi-modal MIR systems, e.g., not only focusing on the quality, but also on robustness, interpretability, or reduction of the environmental impact during the training of deep neural networks

Guest Editors

- **Igor Vatolkin** (lead) - Akademischer Rat (Assistant Professor) at the Department of Computer Science, RWTH Aachen University, Germany
- **Mark Gotham** - Assistant professor at the Department of Computer Science, Durham University, UK
- **Xiao Hu** - Associated professor at the University of Hong Kong
- **Cory McKay** - Professor of music and humanities at Marianopolis College, Canada
- **Rui Pedro Paiva** - Professor at the Department of Informatics Engineering of the University of Coimbra, Portugal

Submission Guidelines

Please, submit through https://transactions.ismir.net, and note in your cover letter that your paper is intended to be part of this Special Collection on Multi-Modal MIR.

Submissions should adhere to formatting guidelines of the TISMIR journal: https://transactions.ismir.net/about/submissions/. Specifically, articles must not be longer than 8,000 words in length, including referencing, citation and notes.

Please also note that if the paper extends or combines the authors' previously published research, it is expected that there is a significant novel contribution in the submission (as a rule of thumb, we would expect at least 50% of the underlying work - the ideas, concepts, methods, results, analysis and discussion - to be new).

In case you are considering submitting to this special issue, it would greatly help our planning if you let us know by replying to igor.vatolkin@rwth-aachen.de.