MIRrors: Music Information Research reflects on its future

Special issue foreword

Perfecto Herrera-Boyer · Fabien Gouyon

The automatic processing of music information has been the focus of an increasing number of researchers, companies and institutions in recent years. The International Symposium on Music Information Retrieval held in Plymouth, MA, in 2000 can be informally considered the foundational event, where many of the participants with backgrounds in consolidated disciplines such as computer music, speech processing, artificial intelligence, musicology, information science or music cognition converged to tackle new problems posed by the increasing digitalization of music data. The very nature of Music Information Research (MIR) is, hence, multidisciplinary. This, added to its short life, hinders its recognition as an identifiable discipline with a solid and healthily expanding core of research. The aim to bring this issue to light and to provide insights on how to overcome it is the binding link among the papers of this volume.

Most articles included in this special issue originated as peer-reviewed conference papers (but were considerably expanded or reworked since then) and were publicly presented and debated in a special session of the International Society for Music Information Retrieval Conference, held in Porto in October 2012.[1] The call for participation to this special issue stressed our requirement that contributions had a reflective nature about the past of MIR, combined with realistic perspectives on the future of specific topics, and on the future of the discipline as a whole. We were looking for original, challenging and thought-provoking papers able to connect the past with the present and the future of the discipline. We cultivated the metaphor that this special issue was going to act like a mirror; a mirror that, by properly

---

[1] http://ismir2012.ismir.net

P. Herrera-Boyer (✉)
Music Technology Group, Universitat Pompeu Fabra, Barcelona, Spain
e-mail: perfecto.herrera@upf.edu

F. Gouyon
INESC TEC, Porto, Portugal
e-mail: fgouyon@inescporto.pt

distorting the past (because it could not provide but a biased and filtered view) could be reflecting a possible picture of the future of our discipline.

Rather than provide a comprehensive overview of challenges on the path of MIR towards affirmation as a research discipline (which is the purpose of recently published documents as e.g. the Roadmap for Music Information Research [1]),[2] this special issue takes a deeper look at a selected number of issues relevant to the three essential MIR facets: (i) **music information**, (ii) the **humans** that create, search for, or use that information, and (iii) an **interaction scenario** that facilitates the connection between humans and information.

Under this view, MIR could be a particular case of human-computer interaction (HCI). As several papers in this special issue remark, it is startling to discover how loose the connection between MIR and HCI has been until now. Unlike in HCI, the **human** has been poorly addressed, when not disengaged, in MIR. In addition, several contributing papers illustrate that there are specific long-tradition disciplines that could provide models and methods to make MIR research user-centered.

The **interaction scenario** has also been problematic. The lion's share of MIR research traditionally follows bottom-up strategies, from the signal (audio, symbols, tags, etc.) to some more abstract concept. This implies a strong focus on very specific and isolated pieces of an —elusive— general picture. More pragmatic considerations such as how research results will be used, by whom, and in which real-world context—i.e. which interaction scenario—, are seldom addressed, or with a relatively naive view. Unfortunately, MIR research has not yet defined a clear methodology and experimental design procedures for assessing its worth, other than simplistic performance metrics. Addressing this issue implies fundamental methodological considerations regarding what "improvements" mean in practice and how to address them systematically. That is, considerations such as e.g. how evaluation should guide research, how to effectively build on the legacy from others and from the past, how to properly and effectively interface with neighboring disciplines and bodies of knowledge with different approaches to music information and human behavior and cognition.

We finally mention what has probably been, until now, the core of MIR: **music information**, specifically its extraction and representation. This has meant the development of algorithms capable of converting data into information and, sometimes, even into knowledge. Leveraging the conversion of information into knowledge should be the main goal of MIR systems. But, for achieving this goal, both the entity that processes and uses such information, and the natural and efficient ways to do so, have to be properly understood and characterized. Even though signal processing and pattern recognition techniques are somewhat limited, they have been put to great use in representing musical signals and discovering musical patterns, and have fundamentally changed how people interact with music. The challenge is to reshape MIR not with less music information but with greater emphasis on the human and interaction scenario, and on combining all three facets together.

Building upon systematic and meticulous reviews of the existing MIR legacy, the contributions to this special issue propose original and challenging research paths

---

overarching the three facets of MIR. They raise fundamental, timely and relevant questions appropriate for the general MIR researcher.

Which evaluation methodologies would best contribute to improvements? This is a central question in MIR, where quantitative measurements of algorithm performance is critical. In *"Evaluation in Music Information Retrieval"*, Urbano, Schedl and Serra discuss different aspects of Information Retrieval Evaluation that have been overlooked and need special attention in MIR, considering experimental validity, reliability and efficiency. Evaluation experiments produce large amounts of numbers and plots, but there is a lack of proper interpretation and discussion due in part to the lack of public and standardized resources, and in part to the lack of proper statistical and methodological training in many engineering and computer science curricula. Special attention is given on how to improve existing MIR systems by fully and properly interpreting the data generated in experiments.

How should music genre recognition systems be evaluated? If we consider the wealth of publications on the topic of music genre recognition (MGR) in the past decade, this is a particularly relevant specification of the previous general question regarding evaluation. A methodological perspective is adopted by Sturm in *"Classification accuracy is not enough: On the evaluation of music genre recognition systems"*. His thorough review of the majority of the research undertaken on this problem concludes that neither classification accuracy, recall and precision, nor confusion tables, necessarily reflect the capacity of a system to recognize genre in musical signals. The paper advocates that an evaluation of system behavior at the level of the music is required to usefully address the fundamental problems of MGR, and many other music information research tasks. He warns about confusing a musical concept (like genre) with certain magnitudes computed from the signal. Even though covariances between them could exist, they do not grant equivalence. Sturm's meta-analysis on the genre recognition problem motivates the development of a richer experimental and conceptual toolbox for evaluating any system designed to intelligently extract information from music signals.

Which are the relevant challenges in automatic music transcription? Focusing on another well-known MIR research topic, in *"Automatic music transcription: Challenges and future directions"*, Benetos, Dixon, Giannoulis, Kirchhoff and Klapuri call for advancing the current state of audio music transcription systems by either making use of more information (which could take the form of high-level models reflecting the musical conventions or instrument acoustics relevant to the piece in question), or by taking advantage of explicit user input in order to select algorithms, set parameters, or resolve ambiguities (i.e., semi-automatic transcription). Another promising direction for further research is the combination of multiple processing principles (as particular cases of ensemble learning techniques), such as different algorithms with complementary properties which estimate a particular feature, or algorithms which extract various types of musical information, such as the key, metrical structure, and instrument identities, and feed this information into a model that provides a context for the note detection process. As most of the authors in this special issue acknowledge, in order to enabling progress in these directions, expertise from a range of disciplines is needed, such as e.g. musicology, acoustics, audio engineering, cognitive science, artificial intelligence, and computing. In such a scenario, the sharing of code and data between researchers becomes increasingly important.

How can research reproducibility be fostered and what does MIR have to gain from it? This is a topic that is further developed and refined by Page, Fields, de Roure, Crawford and Downie in *"Capturing the workflows of Music Information Retrieval for repeatability and reuse".* It is amazing the amount of knowledge that could have been amassed in the lifetime of our discipline if we, as a community, would have cared more about sharing, repeatability, reuse and heritage of our domain-specific knowledge. A considerable amount of that is encoded as software, and this paper traces the legacy that has survived up until the present time. According to the authors, interoperability has not been —and should not be— achieved through the adoption of a single portal, toolkit, or programming language. Plurality of systems and the different approaches they embody is as important in avoiding skewed research and results as the plurality of datasets. Rather than a single complete system the proposal aims for making possible, from reusable and interoperable components, the quick assembly of any number of specific, targeted, applications. Whilst historically, the MIR community has produced a wide variety of tooling to meet its needs, many of which have been surveyed in the article, there needs to be recognition of the increasing dominance of a single platform and vigilance for the problems this may cause. Workflows not only provide a platform for principled reuse, but are also the building blocks for Research Objects, and through these come the opportunity to conduct MIR research in new transparent, reusable, repurposable, and repeatable ways.

Which methodologies are needed in the design of music features? The design and development of ad-hoc features representing multiples dimensions, facets and particularities of music has a long and well-established tradition in MIR. Hand-crafted musical knowledge-motivated design of features is, according to Humphrey, Bello and LeCun, not sustainable as it paves the way for shallow and limited architectures that usually fail to capture long-term musical structure. In *"Feature learning and deep architectures: New directions for music informatics"* the authors advocate for a paradigm shift that involves deep learning, a promising research path that has emerged from the machine learning community. Deep architectures are able to discover relevant features and abstraction levels provided enough richness is contained in the input data, and the authors present, in addition to a discussion on the way deep learning might change our conceptions about features (and not just our feature-related typical tasks), some short case studies illustrating its advantages.

Which methodologies from cognition and neuroscience would benefit MIR, and why? MIR is often considered an excessively self-referential discipline. The difficulties to get published in general scientific journals or high-impact conferences have been informally documented in our email lists and in personal communications.[3] According to Aucouturier and Bigand in the somehow Socratic discussion *"Seven problems that keep MIR from attracting the interest of cognition and neuroscience"*, MIR faces an opportunity to break such a trend, provided it realizes what it has to gain from neuroscience and cognition. In this way, with an appropriate framing and focus, new research opportunities should be waiting in the research agenda of those disciplines. The authors advocate for the music cognition and

---

[3]Who doesn't have a funny rejection story to share?

music neuroscience communities to take advantage of tools developed by MIR researchers which make it possible to study the mechanisms of music information processing and to test hypotheses on what type of signal characteristics are important for early auditory processing.

How and why should the user of MIR systems be considered in MIR research agendas? Two complementary papers have directly addressed "users" in MIR: *"Toward an understanding of the history and impact of user studies in music information retrieval"* and *"The neglected user in Music Information Retrieval research"*. In the former, Lee and Cunningham summarize and criticize the way MIR user studies have been conducted and published, and what impact these studies have had on the field. There are many factors impeding a strong impact for such studies: a lack of findability due to the scattered patterns of publication, weak connections among scholars, dominance of small scaled studies that are difficult to generalize, and the disconnection between researchers conducting user studies and system developers/evaluation task designers. Perharps most remarkable is the finding that there are many frequently-cited studies that, because they have no connection or presentation in the usual MIR channels, have exerted a null impact on the community. Different perspective and aims are pursued by Schedl, Flexer and Urbano in the latter paper. They analyze user-studies taking computer science and psychology points of view. Whereas the former centers on the problems of user modeling, machine learning, and evaluation, the latter is mainly concerned with proper experimental design and interpretation of the results of an experiment. Considering the lack of proper research on user-centric systems and the limitations, usefulness and costs of evaluating them, the authors stress the need for multi-faceted, personalized, social- and contextual-aware systems that are evaluated with a rich set of user-satisfaction criteria and the proper methodological rigour that grant the refutation or confirmation of hypotheses.

We hope this selection of well-crafted papers will get the attention they deserve. They provide deep, committed, and inspiring directions towards a bright future for MIR. We challenge our readers to take the time to go through them and to adapt their research activity according to the issues raised. If each reader can find one valuable piece of information to guide their future research, we believe the authors' efforts and our own in compiling this special issue will be worthwhile.

Reference

1. Serra, X., Magas, M., Benetos, E., Chudy, M., Dixon, S., Flexer, A., Gómez, E., Gouyon, F., Herrera, P., Jordà, S., Paytuvi, O., Peeters, G., Schlüter, J., Vinet, H., Widmer, G. (2013). Roadmap for Music Information ReSearch. In: Peeters, G. (Ed.), *Creative Commons BY-NC-ND 3.0 license*. ISBN: 978-2-9540351-1-6.