Music Similarity: Concepts, Cognition and Computation

Anja Volk, Elaine Chew, Elizabeth Hellmuth Margulis & Christina Anagnostopoulou

Published online: 29 Sep 2016.

Submit your article to this journal ↗

Article views: 72

View related articles ↗

CrossMark View Crossmark data ↗

Music Similarity: Concepts, Cognition and Computation

Similarity is fundamental to our experience of the world (Goldstone & Son, 2005), and to our experience of music in particular. Accordingly, modelling music similarity is crucial for researching musical structures and cognitive processes involved in the human engagement with music within the areas of Musicology, Music Theory and Music Cognition. Moreover, the computational modelling of music similarity has become a crucial need for music research, industry and consumers over the last few decades. The dramatic increase in the digitization of music calls for the development of computational methods in Music Information Retrieval (MIR), such as content-based querying and retrieval, automatic music classification, music recommendation and digital rights management. Music similarity is a fundamental topic involved in these different aspects of music information processing. Modelling similarity has become a major challenge in various areas of Computer Science, such as in Multimedia Retrieval, Data Mining and Bioinformatics. The importance of *domain specific* similarity functions to be employed in search engines has been stressed (Skopal & Bustos, 2011). In the domain of music, similarity is a highly context-dependent notion and poses serious challenges for computational modelling. At the same time, the need for computing tools to study music similarity is crucial for music scientists, who study similarity relations in the listening process, in composition and improvisation, and through the analysis of musical scores and performances.

This special issue dedicated to the topic of music similarity is a follow-up to the workshop 'Music Similarity: Concepts, Cognition and Computation', which was held at the Lorentz Center, International center for scientific workshops in Leiden, the Netherlands, 19–25 January 2015. The Lorentz workshop brought together 55 experts on music similarity from Computer Science, Musicology and Music Cognition in order to discuss cross-disciplinary strategies on the theoretical and computational modelling of music similarity. Main areas addressed during the workshop considered similarity computing, perception and cognition of similarity, similarity in music content (music analysis) and similarity in music expression (for information on the programme of the workshop, including presentations of participants, see Volk, Chew, Margulis, & Anagnostopoulou, 2015). Participants discussed achievements, challenges and future goals of music similarity research for different aspects such as pattern extraction and similarity, processes of categorization, similarity across modalities, similarity and expressive performance, large-scale music similarity, similarity and style, salience and variation, similarity in composition, similarity in timbre, harmonic similarity, similarity and memory, and evaluation and similarity.

The papers of this issue span the main areas of the workshop, namely similarity computing, perception and cognition of similarity, similarity in music content (music analysis) and similarity in music expression. The paper by Godøy et al. provides a framework for thinking about similarity between sound and movement. Why and in what ways might we perceive events in one of these modalities to be similar to events in the other? Godøy and colleagues argue that the cross-modal nature of many mental representations means that mental images of sounds often carry traces of the movements required to produce them, or of movements that are typically associated with them. They propose a continuum of sound-motor relationships, with direct relationships of this sort at one end, and more metaphoric connections at the other. Better understanding of perceived similarities between sound and music could lead to improved tools for cross-modal mapping in multimedia environments for art generation, and ultimately to music search engines that rely on gesture as part of the interface.

The paper by Boot, Volk and de Haas investigates the role of repeated patterns for the computational modelling of melodic similarity. More specifically, the paper focuses on the relevance of repeated melodic patterns for modelling similarity and compression in a retrieval setting for a dataset of 360 Dutch folk songs. Musicologists have suggested that shared patterns between melodies provide strong cues for the similarity of folk songs belonging to the same tune family. The paper proposes a framework to use these patterns for compression and classification in tune families, and compares the classification accuracy achieved by musicologist-annotated patterns to those found by pattern discovery algorithms. The superior compression ratio and retrieval accuracy of the annotated patterns show that state-of-the-art automatic pattern discovery still lags expert annotations in this retrieval setting. The experiment confirms that shared patterns indeed provide strong cues for melodic similarity and can thus be used successfully for compression and similarity estimation in computational approaches.

Flexer and Grill's paper considers the problem of limited inter-rater agreement in the modelling of music similarity. The paper focuses on two representative tasks in MIR, namely that of Audio Music Similarity and Retrieval (AMS) and Music Structural Segmentation (MSS). The former task is based on the modelling of similarity between music pieces, and the latter on assessments of similarity (or dissimilarity, where boundaries are concerned) within pieces. The paper tracks and analyses several years' information on the performance of these tasks in the Music Information Retrieval Evaluation eXchange competition, part of the International Society for Music Information Retrieval annual conference. Flexer and Grill posit that, when evaluation is based on multiple annotations or human judgements, an automated system's performance cannot exceed the average agreement among the annotations or evaluators. This performance ceiling is observed to have been reached early in the history of the AMS task, while there is still scope for improvement for the MSS task. Various strategies are proposed for improving the limiting upper bound.

Devaney's paper examines inter- vs. intra-singer similarity and variation in accompanied and unaccompanied vocal recordings of Schubert's 'Ave Maria'. Experiments investigate whether individual singers' recordings display sufficient self-consistency to allow a Support Vector Machine classifier to accurately identify singers when considering vocal pitch, timing, dynamics and timbre. The performance of the discriminative computational model is compared to that of human listeners. Reasons for inaccurate identification include violation of the expected balance between inter- and intra-singer variability, the salience of inter-singer variability vs. intra-singer similarity and the relevance of performance parameters.

Harrison, Musil and Müllensiefen's paper addresses melodic discrimination tests, which are based on similarity comparisons between melodies. Melodic discrimination tests are used to assess musical abilities of listeners. During the test, listeners are presented with similar versions of an unfamiliar melody, and have to decide whether or not the melodies are the same. Hence, these discrimination tests are intertwined with the notion of similarity. Melodic discrimination tests have usually been constructed and analysed using classical test theory, which is however not appropriate for optimizing test efficiency to ensure that each item contributes optimally to test performance. Moreover, classical test theory provides only limited information regarding construct validity, hence the question of how the test scores relate to the underlying construct of interest. Therefore, the paper uses methods of item response modelling to address test efficiency and construct validity for melodic discrimination tasks and introduces an explicit cognitive model for melodic discrimination tasks. The paper shows that item difficulty can be predicted by melodic similarity and complexity, in accordance with the proposed cognitive model.

The five papers exemplify that a high degree of methodological sophistication is required for achieving rich and valid results in modelling music similarity, a concept that has been described as *fuzzy* (McFee, Barrington, & Lanckriet, 2010), as *elusive* (Berenzweig, Logan, Ellis, & Whitman, 2003), a *cold-start problem* (Wang et al., 2005) or a *huge challenge* (Downie, Byrd, & Crawford, 2009) in MIR. Capturing different aspects of similarity such as music and motion, repeated patterns in folk songs, agreement in the human assessment of similarity between different annotators, or similarity in singers' performances of the same piece, the papers demonstrate that mutually informing perspectives on similarity are required for providing a better overall understanding of music similarity. The specific tasks of the annual Music Information Retrieval Evaluation eXchange (MIREX) at ISMIR constitute a fragmentation of music similarity into *different kinds of similarity*, which has also been introduced in Cognitive Science as an answer to the inherent complexity of similarity, suggesting that there is not just *one kind of similarity* (Medin, Goldstone, & Gentner, 1993; Smith, 1989). While this constitutes a necessary step in achieving scientifically rigorous approaches to a complex phenomenon, Godøy et al.'s article on similarity in sound and motion also points to future research that needs to overcome this fragmentation into isolated aspects of similarity, such as by addressing cross-modal aspects. For achieving comprehensive perspectives on music similarity that inform each other productively, instead of describing merely different kinds of similarity, we need to leverage musicology knowledge for building more musically sophisticated MIR technologies, and foster collaborations between MIR and music psychology. The papers of this issue provide examples in this direction by combining computational modelling with intricate musicological and/or empirical research investigating the human assessment of similarity.

The guest editors would like to thank the Lorentz center in Leiden for providing an excellent venue to discuss music similarity with international experts, the workshop participants who submitted papers to this special issue, reviewers who dedicated their time to the review process, the editor in chief of the Journal of New Music Research, Alan Marsden, for advice on all matters and the staff at Taylor and Francis that contributed to making the issue possible.

References

Berenzweig, A., Logan, B., Ellis, E. P. W., & Whitman, B. (2003). A large-scale evaluation of acoustic and subjective music similarity measures. *Proceedings of the 4th International Society for Music Information Retrieval Conference* (pp. 99–105). Baltimore, MD.

Downie, S., Byrd, D., & Crawford, T. (2009), Ten years of ISMIR: Reflections on challenges and opportunities. *Proceedings of the 10th International Society for Music Information Retrieval Conference* (pp. 13–18). Kobe, Japan.

Goldstone, R. L., & Son, J. (2005). Similarity. In K. Holyoak & R. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 13–36). Cambridge: Cambridge University Press.

McFee, B., Barrington, L., Lanckriet, G. (2010). Learning similarity from collaborative filters. In *Proceedings of the 11th International Society for Music Information Retrieval Conference* (pp. 345–350). Utrecht, The Netherlands.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100*, 254–278.

Skopal, T., & Bustos, B. (2011), On nonmetric similarity search problems in complex domains. *ACM Computing Surveys*, 43, 34:1–34:50.

Smith, L. B. (1989). From global similarities to kinds of similarities: The construction of dimensions in development. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 146–178). Cambridge: Cambridge University Press.

Volk, A., Chew, E., Margulis, L., & Anagnostopoulou, C. (2015, January). *Music similarity: Concepts, cognition and computation*. Lorentz Center Workshop, Retrieved from http://bit.ly/29tUpfk

Wang, J. C., Lee, H. S., Wang, H. M., & Jeng, S. K. (2005), Learning the Similarity of Audio Music in Bag-of- Frames Representation from Tagged Music Data, In Proceedings of the 6th International Society for Music Information Retrieval Conference, London, UK 85–90.

Anja Volk
Utrecht University
E-mail: a.volk@uu.nl

Elaine Chew
Queen Mary University of London

Elizabeth Hellmuth Margulis
University of Arkansas

Christina Anagnostopoulou
University of Athens