Special Section on Music Data Mining

Article *in* IEEE Transactions on Multimedia · August 2014

DOI: 10.1109/TMM.2014.2325693

Mitsunori Ogihara
University of Miami

George Tzanetakis
University of Victoria

Guest Editorial: Special Section on Music Data Mining

I. Music Data Mining

MUSIC has been an important application area for data mining and machine learning techniques for many years. Music data mining is an interdisciplinary area that studies computational methods for understanding and delivering music data and is a topic of growing importance with large commercial relevance and substantial potential. It attracts researchers not only from computer science, electrical engineering, and musicology but also library science and psychology [1].

During the last few years there has been a dramatic shift in how music is produced, distributed and consumed. A combination of advances in digital storage, audio compression as well as significant increases in network bandwidth has made digital music distribution a reality. Portable music players, computers and smart phones frequently contain personal collections of thousands of music tracks. Digital stores in which users can purchase music contain millions of tracks that can be easily downloaded.

The research area of music data mining has gradually evolved during this time period in order to address the challenge of effectively accessing and interacting with these increasing large collections of music and associated data such as styles, artists, lyrics and music reviews. The algorithms and systems developed frequently employ sophisticated and advanced data mining and machine learning techniques in their attempt to better capture the frequently elusive relevant music information.

Recent advancements in music listening technologies, in particular, the Internet-based music communities, radio stations and music stores, have introduced several new interesting aspects to the area, such as multimodal analysis of music data, community-based labeling of music, and listening pattern analysis. The introduction has made the area an exciting research ground and there is a strong and emergent need to publicize this area in multimedia literature. In addition, music data is of significant scale with diverse information sources and aspects. Many of the relevant aspects such as user behaviors and web-based information retrieval have direct analogies in other multimedia content items including images, videos, and pictures. As a result, the techniques and tools developed in music data mining can be relevant and useful in general to other areas of multimedia research.

II. The Special Section

The special section provides a leading forum for timely, in-depth presentation of recent advances in algorithms, theory, and applications in music data mining. The selected papers underwent a rigorous refereeing and revision process.

Digital Object Identifier 10.1109/TMM.2014.2325693

Efficient and intelligent music information retrieval (MIR) is a very important topic in music data mining [2]. A key step in MIR is to learn feature representations from available music data. Recently unsupervised feature learning techniques (e.g., sparse coding and deep neural networks) have been used to construct audio codewords, leading to the bag-of-frames (BoF) model and a term-document style representation of music. The paper by Su *et al.* presents a systematic empirical study that compares BoF variants on three MIR tasks (music genre classification, predominant instrument recognition, and audio tagging) by considering different options on codebook learning and encoding, feature pooling, term weighting, power normalization, and dimension reduction. The study improves the understanding of the BoF model and the analogy between traditional text words and audio codewords, and also provides several interesting insights in applying the model in MIR.

Music recommendation is receiving great attention recently and is becoming prevalent due to the problem of information overload. Its main goal is to provide users with the music pieces that they are likely to enjoy. A good music recommender should have a good balance between relevance (e.g., the recommendations should be similar to the existing relevant items) and novelty (e.g., the recommendations should have some variety and diversity). The paper by Lee and Lee proposes a music recommendation system based on collaborative filtering using dynamically promoted experts to provide recommendations that are novel yet relevant to users. In particular, singular value decomposition (SVD) and K-means clustering are first applied on the user-item matrix to group music pieces into different clusters (or music domains). Then a user is identified as an expert for a music cluster if his/her listening behaviors are concentrated in that cluster. For a target user, different users are selected dynamically as experts for the music clusters on which the user is a novice to provide final recommendations. Experiments demonstrate that the proposed system achieves a good balance between novelty (using music clusters) and relevance (with the help of experts).

A challenging problem in music data mining is to model expressive dynamics (e.g., variations in tempo and articulation) and discover expression patterns. Recently many data mining techniques, in particular, feature learning techniques have been used to identify useful patterns that are related to the musical expressive performance. In their paper, Grachten and Krebs perform an empirical study to evaluate the utility of several unsupervised feature learning methods in modeling the note intensities of performed music. They use a note centric representation that includes both harmonic and rhythmic characteristics of music contexts and apply different algorithms including non-negative matrix factorization (NMF), principal component analysis (PCA), and Restricted Boltzmann machines (RBM) to learn features. The learned features are evaluated in predicting the note intensities using linear regression models.

Music is of rich structure with its hierarchical and multi-faced organization of basic elements. Hence inferring and annotating the structural properties of music pieces is an important problem in music data mining as they play important roles in music indexing, similarity search and information retrieval. In their paper, Smith, Chew, and Chuan investigate a large corpus of annotated recordings and study the relationship between the acoustic properties of recordings and the boundary indications of listeners. Their results demonstrate a strong correlation between the acoustic boundaries, estimated by musical features reflecting timbre, harmony, key, rhythm and tempo, and the annotated boundary positions and the strength of the relationship is moderately affected by the musical feature representation. In another work, Serra *et al.* present an unsupervised approach based on structure features and time series similarity to automatically detect the temporal locations of segment boundaries (i.e., novelty and boundary detection) and to assess segment similarities and repetitions (i.e., classification/grouping). The structure features jointly consider both local and global aspects by measuring the relations between each time frame/window with all other times frames/window in a music piece and they are used to obtain reliable segment boundaries. The obtained segment boundaries are then used for music structure labeling based on time series similarity measures.

Acknowledgment

We are very indebted to the reviewers who reviewed the papers very carefully. We would also like to thank all the authors who submitted their papers to the special section. Special thanks to Dr. M. van der Schaar and Ms. R. Wollman for their great help and support in organizing the section.

Tao Li
School of Computer Science
Florida International University
Miami, FL 33199 USA
taoli@cs.fiu.edu

Mitsunori Ogihara
Department of Computer Science
University of Miami
Coral Gables, FL 33146 USA
ogihara@cs.miami.edu

George Tzanetakis
Department of Computer Science
University of Victoria
Victoria, BC Canada
gtzan@cs.uvic.ca

References

[1] T. Li, M. Ogihara, and G. Tzanetakis, *Music Data Mining*, ser. Chapman & Hall/CRC Data Mining and Knowledge Discovery Series. Boca Raton, FL, USA: CRC Press, 2012.

[2] T. Li and M. Ogihara, "Toward intelligent music information retrieval," *IEEE Trans. Multimedia*, vol. 8, no. 3, pp. 564–574, Jun. 2006.

Tao Li received the Ph.D. degree in computer science from the University of Rochester, Rochester, NY, USA, in 2004. He is currently a Professor in the School of Computer Science, Florida International University, Miami, FL, USA. His primary research interests include data mining, machine learning, bioinformatics, and music information retrieval. He is on the editorial board for IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE), *ACM Transactions on Knowledge Discovery from Data* (ACM TKDD), and *Knowledge and Information System Journal* (KAIS). He is a recipient of NSF CAREER Award in 2006 and multiple IBM Faculty Research Awards (2005, 2007, and 2008). In 2009, he received FIU's Excellence in Research and Creativities Award. In 2010, he received IBM Scalable Data Analytics Innovation Award.

Mitsunori Ogihara received the Ph.D. in information sciences from the Tokyo Institute of Technology, Tokyo, Japan, in 1993. He is currently a Professor of Computer Science at the University of Miami, Miami, FL, USA, and a Director of Data Mining in the Center for Computational Science at the University of Miami, Miami, FL, USA. He is on the editorial board for the journals *Theory of Computing Systems* and *International Journal of Foundations of Computer Science*. He is a Distinguished Scientist member of the Association for Computing Machinery. He is a recipient of the NSF CAREER Award in 1997.

George Tzanetakis (S'98-M'02-SM'11) received the B.Sc. degree in computer science from the University of Crete, Heraklion, Greece, and the M.A. and Ph.D. degrees in computer science from Princeton University, Princeton, NJ, USA. His Ph.D. work involved the automatic content analysis of audio signals with specific emphasis on processing large music collections. In 2003, he was a Postdoctoral Fellow at Carnegie-Mellon University, Pittsburgh, PA, USA, working on query-by-humming systems, polyphonic audio-score alignment, and video retrieval. He is an Associate Professor of Computer Science (also cross-listed in Music and Electrical and Computer Engineering) at the University of Victoria, Victoria, BC, Canada, where he holds a Canada Research Chair (Tier II) in the Computer Analysis of Audio and Music. He is the principal designer and developer of the open source Marsyas audio processing software framework (http://marsyas.info). His research deals with all stages of audio content analysis such as analysis, feature extraction, segmentation, and classification, with specific focus on music information retrieval (MIR). His work on musical genre classification is frequently cited. He received an IEEE Signal Processing Society Young Author Award in 2004.