

LATE-BREAK SESSION ON MUSIC STRUCTURE ANALYSIS

Bruno Rocha

University of Amsterdam,
The Netherlands
b.m.machadorocha@uva.nl

Jordan B. L. Smith

Queen Mary University of
London, UK
jblsmith@
eecs.qmul.ac.uk

Geoffroy Peeters

IRCAM-CNRS STMS,
Paris, France
geoffroy.peeters@ircam.fr

Joe Cheri Ross

Indian Institute of Technology,
Bombay, India
joecheri@gmail.com

Oriol Nieto

New York University, USA
oriol@nyu.edu

Jan Van Balen

Utrecht University,
The Netherlands
j.m.h.vanbalen@uu.nl

1. INTRODUCTION

Music structure segmentation is a very relevant topic in music information retrieval as it is both an end in itself and a means to improve performance in other tasks [1]. The goal of this session was to discuss the challenges current algorithm designers and corpora annotators face, even though an agreement on the definition of music structure is hard to reach.

2. COURSE OF DISCUSSION

2.1 Participants

The session started with participants introducing themselves and describing their involvement in music structure research.

Rocha is developing a structure segmentation algorithm for electronic dance music in a short project involving the University of Amsterdam and Elephantcandy, a company that develops mobile audio applications.

Peeters has developed music structure estimation algorithms and progressively arrived to the conclusion that this task is ill-defined outside the context of a specific application; he then did a survey and a proposal to help describing what are the facets of music structure.

Smith helped develop and now maintains the SALAMI database. He is currently studying the link between these annotations and the recordings, and the differences between annotators.

Nieto developed the new music segmentation algorithm for The Echo Nest. He is now working on his dissertation, in which he aims to add various perceptual models to state-of-the-art content-based methods for music structure discovery.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.

© 2012 International Society for Music Information Retrieval

Van Balen works on the analysis of choruses and cognitive salience in popular music.

Benjamin Martin, Joe Cheri Ross, Gopala Koduri, Mickael Legoff, Maria Panteli, Katerina Kosta and Macarena Valera were also present at the session¹.

2.2 Annotation methodologies

The discussion started with a focus on annotation methodologies. Recently created corpora have incorporated labels that are multi-dimensional in that they distinguish between boundaries and groupings based on different musical factors, such as similarity or instrumentation. It was suggested that those developing analysis algorithms should attempt to do multi-dimensional analysis as well to take advantage of this ground truth. This approach could achieve higher accuracy, since the analysis task is more specific.

In this discussion, Moby's song "Natural Blues" was raised as an example of a piece of music whose design on a digital audio workstation can be easily deduced by the listener. It was pointed out that despite the multi-dimensionality of its construction, listeners manage to make discrete decisions regarding structural changes, which raises the question: do they make them consistently?

2.3 Perception of musical structure

This led to a discussion of the perception of musical structure, in which it was hypothesized that listeners perceive structural changes on the basis of musical cues that can be dependent on genre, culture or listening experience. It would be interesting to investigate results of annotations made by different subjects and explore the areas in which listeners tend to disagree the most. This could probably lead to the development of a cognitive model that would help improving the task of structural segmentation.

2.4 Annotation alignment

Should researchers align annotations with downbeats? Is it wrong to consider a pickup as a start of a section?

¹ There is a mailing list for music structure related discussions. The address is: structsegment_mir@googlegroups.com

These questions led to a new point of discussion. It was suggested that considering multiple ground-truths could solve this problem. When enough people are annotating music, it is as though a perceptual experiment is being conducted [2]. The ground truth obtained through this approach is perhaps the one that is most easily justifiable. The development of a framework conjugating different state-of-the-art techniques and checking for disagreements between them instead of using subjects was also proposed - something similar to what [3] did on the task of beat annotation.

2.5 Evaluation

On the topic of evaluation, it was pointed out it is important to have clearly defined annotation guidelines. Algorithm designers should set clear goals when building their methods: a) what layer are they approaching? b) what perception are they trying to mimic (the composer, the user, the musicologist)?

3. CONCLUSION

While structural analysis is emerging as a very important task in MIR, its inherent fuzziness remains a problem: definitions for what structure is and how it is annotated remain operational. Recent approaches to annotation [4-7] have responded to this by formalizing their methodology (to varying degrees), but a strong musicological or cognitive basis for these approaches has not been put forward.

4. REFERENCES

- [1] M. Casey, R. Veltkamp, M. Goto, M. Leman, C. Rhodes, and M. Slaney: "Content-based music information retrieval: current directions and future challenges," *Proceedings of the IEEE*, Vol. 96, No. 4, pp. 668-696, 2008.
- [2] M. Bruderer, M. McKinney, and A. Kohlrausch: "The perception of structural boundaries in melody lines of Western popular music," *Musicae Scientiae*, Vol. XIII, No. 2, pp. 273-313, 2009.
- [3] J. Zapata, A. Holzapfel, M. Davies, J. Oliveira, and F. Gouyon: "Assigning a confidence threshold on automatic beat annotation in large datasets," *Proceedings of the International Society for Music Information Retrieval Conference*, pp. 157-162, Porto, Portugal, 2012.
- [4] G. Peeters and E. Deruty: "Is music structure annotation multi-dimensional? A proposal for robust local music annotation," *Proceedings of International Workshop on Learning Semantics of Audio Signals*, Graz, Austria, 2009.
- [5] J. Smith, J. Burgoyne, I. Fujinaga, D. De Roure, and J. Downie: "Design and creation of a large-scale database of structural annotations," *Proceedings of the International Society for Music Information Retrieval Conference*, Miami, USA, 2011.
- [6] F. Bimbot, E. Deruty, S. Gabriel, and E. Vincent: "Semiotic structure labeling of music pieces: Concepts, methods and annotation conventions," *Proceedings of the International Society for Music Information Retrieval Conference*, pp. 235-240, Porto, Portugal, 2012.
- [7] J. Serrà, M. Müller, P. Grosche, and J. Arcos: "Unsupervised detection of music boundaries by time series structure features," *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 1613-1619, Toronto, Canada, 2012.
- [8] F. Bimbot, E. Deruty, S. Gabriel, and E. Vincent: "Methodology and resources for the structural segmentation of music pieces into autonomous and comparable blocks," *Proceedings of the International Society for Music Information Retrieval Conference*, Miami, USA, 2011.
- [9] G. Peeters and K. Fort: "Towards a (better) definition of the description of annotated MIR corpora," *Proceedings of the International Society for Music Information Retrieval Conference*, pp. 25-30, Porto, Portugal, 2012.