

MIREX 2011 “MUSIC STRUCTURE SEGMENTATION” TASK: IRCAMSUMMARY SUBMISSION

Geoffroy Peeters

Ircam Sound Analysis-Synthesis Team - CNRS STMS

ABSTRACT

This extended abstract details a submission to the Music Information Retrieval Evaluation eXchange (MIREX) 2011 for the “Structure Segmentation” task. The system named ircamsummary performs both structure estimation and summary generation. Detailed description of the system can be found in [3] [1] and [2]. We briefly summarized them below.

1. IRCAMSUMMARY IMPLEMENTATION

Ircamsummary is both a Matlab-compiled software, a pcode function or a C++ software and library running under Linux, Windows-XP and Mac-OS-X. It performs music structure estimation using the “state” or the “sequence” approach and music audio summary generation using various strategies. Ircamsummary uses the library ircambeat for beat/downbeat synchronous analysis.

2. IRCAMSUMMARY ALGORITHM DESCRIPTION

The flowchart of ircamsummary is represented in Figure 1.

Ircamsummary is a software allowing structure estimation both as “state” representation and “sequence” representation. It also allows the generation of music audio summary using various strategies.

Four independent sets of features are extracted from the signal: MFCC, Chroma features, Spectral Flatness/ Crest measures, Spectral Valley/ Contrast factors. They are processed separately in the remaining.

Various temporal modelling can then be applied to the instantaneous features. In the MIREX-2011 submission, a simple beat-synchronous mean-value modelling has been used. Other temporal modelling such as Dynamic features modelling or membership to hidden states are possible.

Four distance matrices are computed from the four feature sets. After processing (stretching) the four distance matrices are combined together using a weighted sum.

In the case of “state” representation (MIREX-2011 submission), the resulting distance matrix is used to perform a rough segmentation of the signal.

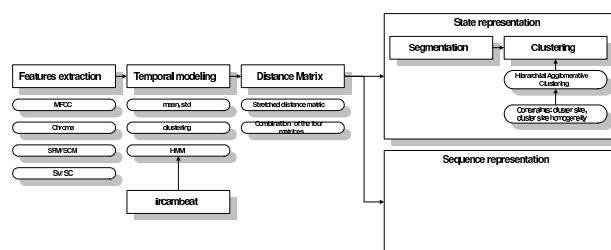


Figure 1. Flowchart of ircamsummary

The resulting segments are then progressively merged using a hierarchical agglomerative clustering. A specific distance is used in order to consider both segments similarity and dissimilarity while enhancing the diagonal (“sequence” approach). Two constraints are added for the linkage: a) cluster size and b) cluster size homogeneity.

3. MIREX-2011 RESULTS AND DISCUSSIONS

4. ACKNOWLEDGMENTS

This work was partly supported by “Quaero” Programme, funded by OSEO, French State agency for innovation

5. REFERENCES

- [1] G. Peeters. Deriving musical structures from signal analysis for music audio summary generation: Sequence and state approach. In U.K. Wilf, editor, *CMR 2003 (LNCS 2771)*, Lecture Notes in Computer Science, pages 142–165. Springer-Verlag Berlin Heidelberg 2004, 2004.
- [2] G. Peeters. Sequence representation of music structure using higher-order similarity matrix and maximum-likelihood approach. In *Proc. of ISMIR*, Vienna, Austria, 2007.
- [3] G. Peeters, A. Laburthe, and X. Rodet. Toward automatic music audio summary generation from signal analysis. In *Proc. of ISMIR*, pages 94–100, Paris, France, 2002.