

Music Structure Analysis by Finding Repeated Parts

Jouni Paulus, Anssi Klapuri

Institute of Signal Processing
Tampere University of Technology
Tampere, Finland

Audio and Music Computing for Multimedia Workshop
27.10.2006





Outline

Introduction

Proposed Method

Problem formulation

Approach in practise

Evaluation

Summary



Introduction

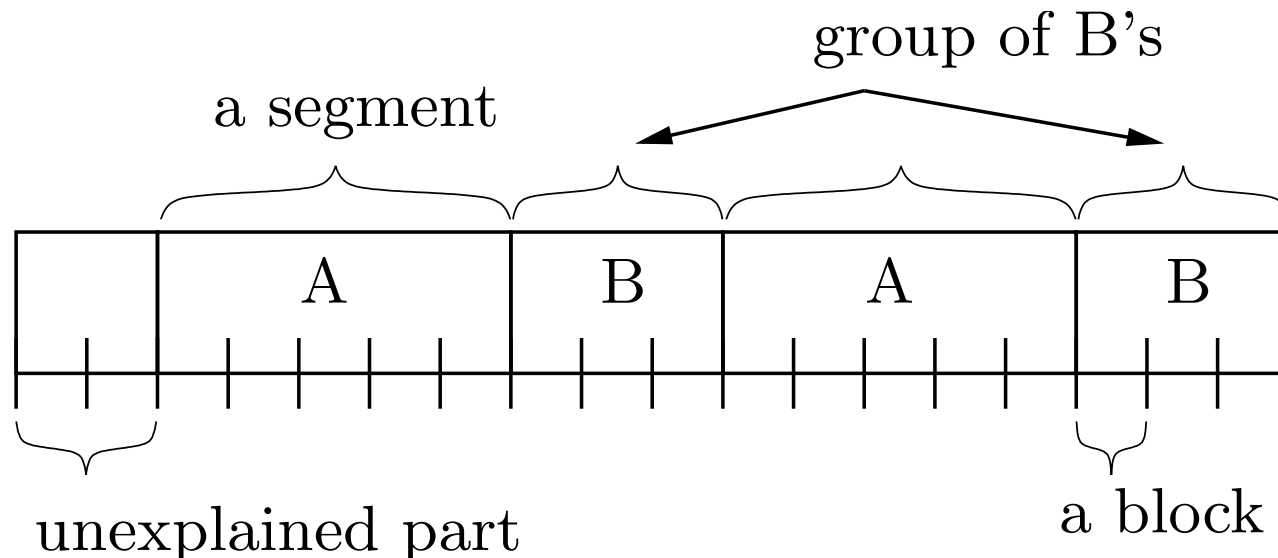
- Analysis of the structure of musical pieces.
 - Divide piece into distinguishable parts.
- Especially in pop music, structure relies on repetitions.
 - E.g., “intro, verse, chorus, verse, chorus, chorus”.
- Proposed method aims to locate the repeated parts.
 - Resulting description may not cover whole piece.
- Applications include
 - smart playback,
 - thumbnailing, and
 - musicological tools.





Definitions

- Musical piece presented as a sequence of *blocks* b_1, b_2, \dots, b_N .
- One or more consecutive blocks form a *segment* $s_k = b_i, b_{i+1}, \dots, b_j$.
 - An occurrence of a part.
- A group of non-overlapping segments, g_i .





Definitions, cont'd

- Segment groups have the properties
 - within-group dissimilarity, d_i , and
 - occurrences of a part should be similar
 - coverage, c_i .
 - the occurrences of a part cover some duration of the piece
- All segment groups form a set of *segment group candidates* $G = \{g_1, g_2, \dots, g_M\}$.
- Structural *explanation* $E \subset G$, such that the included groups do not overlap temporally.
- Select the best subset of the candidate groups.





A “good” explanation?

- Requirements
 - Segments in a group should resemble each others.
 - As much of the piece should be explained as possible.
 - The explanation should be as simple as possible.
- Calculating a cost for an explanation

$$C = \text{dissimilarity} + \alpha \text{ unexplained} + \beta \text{ complexity}, \quad (1)$$

with $\alpha, \beta \geq 0$.

- Evaluate above as

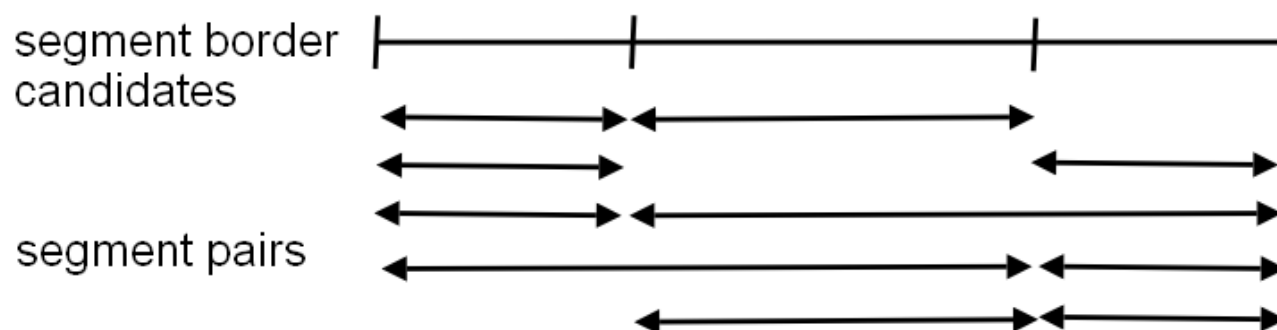
$$C(E) = \sum_{g_i \in E} d_i c_i + \alpha \left(1 - \sum_{g_i \in E} c_i\right) + \beta \log(1 + |E|). \quad (2)$$





From signal to segments

- Block division?
 - Utilise meter estimation. Block division from measure borders.
- Where a segment may start or end?
 - Given N locations, $O(N^4)$ non-overlapping segment pairs.
 - Not feasible to use all measure borders.
 - Create possible locations with novelty in general timbre.
- Create all non-overlapping segment pairs.





Segment groups

- Within-group dissimilarity?
 - Match two segments with DTW, using chroma (tonal descriptor) feature.
 - Average distance between all segment pairs in the group.
- Create candidate groups iteratively based on segment pairs.
 - Start from a pair, add non-overlapping segments one by one.
 - Heuristic rules to limit the number of groups (lengths of segments in a group, within-group dissimilarity, maximum number of segments in a group).





From groups to explanation

- Creating the best explanation is a combinatorial problem.
- Idea of the search algorithm:
 - Start from empty explanation,
 - iteratively, add one more group to explanation, and
 - evaluate the cost.
- Search space can be limited.
 - Estimate minimum cost that can be obtained by extending current explanation.
 - If the minimum is worse than best so far, no point to continue to that direction.





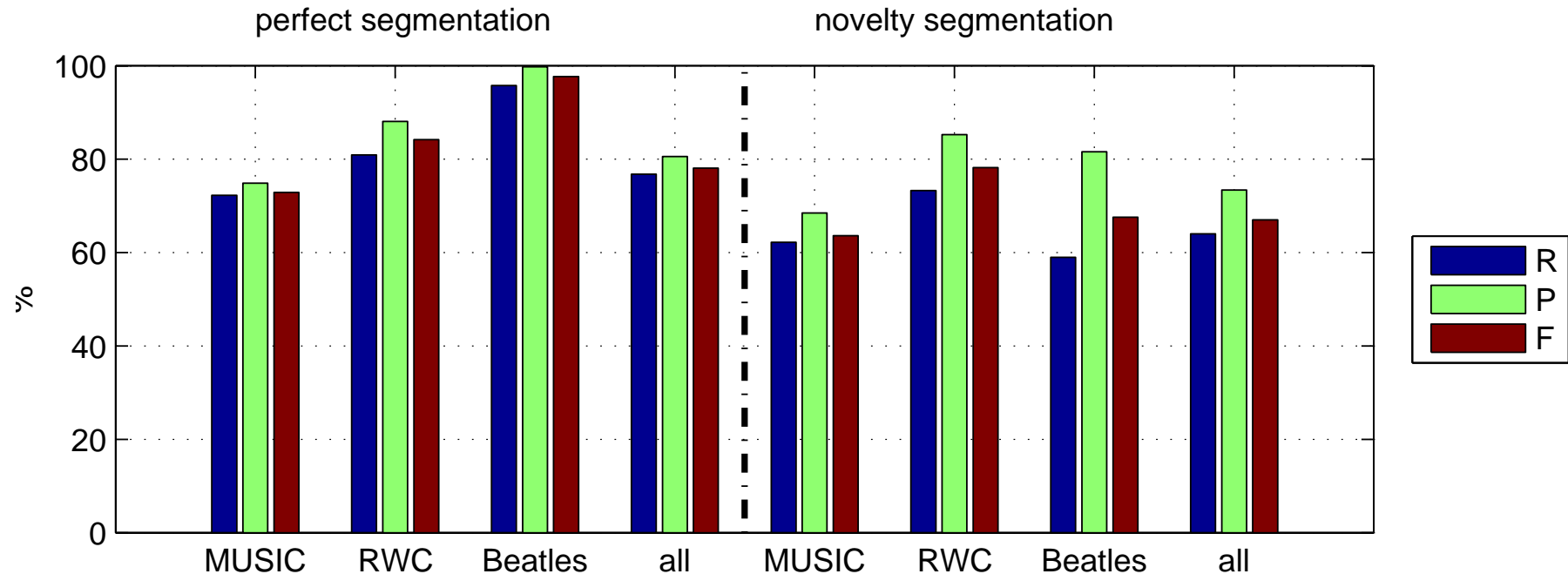
Material

- 50 popular music pieces, from 3 different collections
 - MUSIC
 - RWC-Pop
 - The Beatles
- Manual annotations of clear structural parts.
 - Main focus on repeated parts, but also solo etc. handled.
 - Some parts of the pieces may be left un-annotated.
- 3-fold cross-validation





Results



- R = recall rate, amount of annotated parts found
- P = precision rate, amount of found parts annotated
- F = harmonic F-measure from R and P





Summary

- Music structure analysis problem formulated in terms of parametric cost function.
- Proposed method relies on the repetitions of musical parts.
- Analysis system performance evaluated with real music pieces.
- Segmentation has quite large effect on performance.
- Difficult to find global values for parameters α and β so that method works for all music types.

