

з

5



# Audio Structure Analysis

- Audio features
- Cost measure and cost matrix
   self-similarity matrix
- Path extraction (pairwise similarity of segments)
- Global structure (clustering, grouping)

## Audio Structure Analysis

Müller/Kurth (EURASIP 2007) Rhodes/Casey (ISMIR 2007) Peeters (ISMIR 2007)

- Audio  $\rightsquigarrow$   $V:=(v^1,v^2,\ldots,v^N)$
- v<sup>n</sup> = 12-dimensional normalized chroma vector
- Local cost measure  $c: \mathbb{R}^{12} \times \mathbb{R}^{12} \to \mathbb{R}$

$$c(v^n, w^m) := 1 - \langle v^n, w^m \rangle$$

- $N \times N$  cost matrix  $C(n,m) := c(v^n,w^m)$ 
  - ~--- quadratic self-similarity matrix

## Audio Structure Analysis

#### Self-similarity matrix



## Audio Structure Analysis

#### Self-similarity matrix



## Audio Structure Analysis

#### Self-similarity matrix



#### Audio Structure Analysis



# Audio Structure Analysis

#### Self-similarity matrix

7

11





100

150

60

40

20

50

B.

 $A_2$ 

 $A_1$ 

200

Audio Structure Analysis



8



#### Audio Structure Analysis



# Similarity cluster

13

17



#### Audio Structure Analysis



#### Matrix Enhancement

#### Challenge: Presence of musical variations

- Fragmented paths and gaps
- Paths of poor quality
- Regions of constant (low) cost
- Curved paths

Idea: Enhancement of path structure

## Matrix Enhancement

Shostakovich Waltz 2, Jazz Suite No. 2 (Chailly)



# Matrix Enhancement

Idea: Usage of contextual information (Foote 1999)

$$C_L(n,m) := \frac{1}{L} \sum_{\ell=0}^{L-1} c(v_{n+\ell}, v_{m+\ell})$$

Comparison of entire sequences

- L = length of sequences
- $C_L$  = enhanced cost matrix

→ smoothing effect

14











#### **Transposition Invariance**

Goto (ICASSP 2003)

- Cyclically shift chroma vectors in one sequence
- Compare shifted sequence with original sequence Perform for each of the twelve shifts a separate
- structure analysis
- Combine the results

#### **Transposition Invariance**

Goto (ICASSP 2003)

- · Cyclically shift chroma vectors in one sequence
- Compare shifted sequence with original sequence Perform for each of the twelve shifts a separate
- structure analysis Combine the results

Müller/Clausen (ISMIR 2007)

- Integrate all cyclic information in one transposition-invariant self-similarity matrix
- Perform one joint structure analysis

# 50

## **Transposition Invariance**

Example: Zager & Evans "In The Year 2525"



# **Transposition Invariance**





51

53

49

## **Transposition Invariance**

Example: Zager & Evans "In The Year 2525"



## **Transposition Invariance**









 300
 0.9

 40
 0.8

 50
 0.8

 50
 0.6

 260
 0.6

 250
 0.6

 260
 0.6

 260
 0.4

 150
 0.3

 160
 0.3

 100
 0.7

 80
 0.1



71

Transposition-invariant self-similarity matrix

- Timbre, dynamics, tempo
- Musical key ~~> cyclic chroma shifts
- Major/minor
- Differences at note level / improvisations



## Conclusions: Audio Structure Analysis

Challenge: Hierarchical structure of music



Rhodes/Casey (ISMIR 2007)

System: SmartMusicKiosk (Goto)



75

# System: SyncPlayer/AudioStructure

