Hierarchical approximate self-similarity

Structure detection in music

Christophe Rhodes, Michael Casey

Goldsmiths, University of London

Monday 24th September

Outline

Motivation

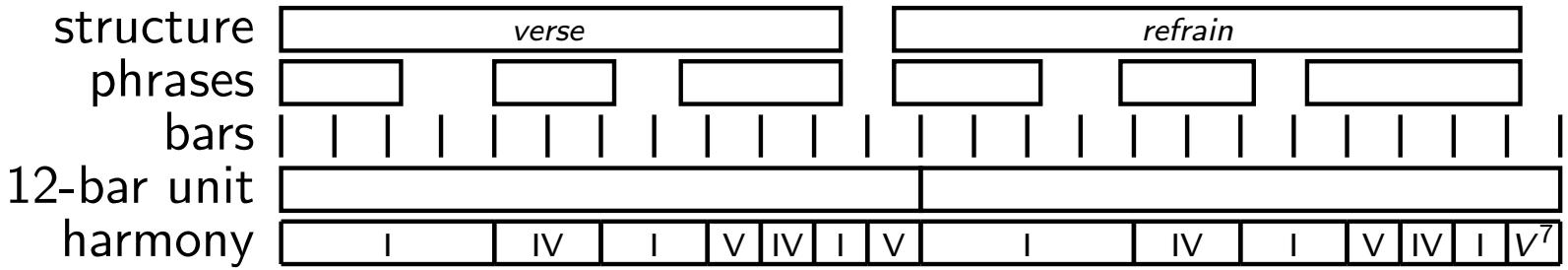
- Finding literal or approximate musical repetitions with little *a priori* knowledge;
- Inferring hierarchical arrangement of repetitions;
- Simple summarization of inferred structure;
- Agnostic to form of musical data.
- Derive attributes of performance (knowing score and repeats)

Motivation

- Finding literal or approximate musical repetitions with little *a priori* knowledge;
- Inferring hierarchical arrangement of repetitions;
- Simple summarization of inferred structure;
- Agnostic to form of musical data.
- Derive attributes of performance (knowing score and repeats)

Hierarchical structure

Hypothetical 12-bar blues:



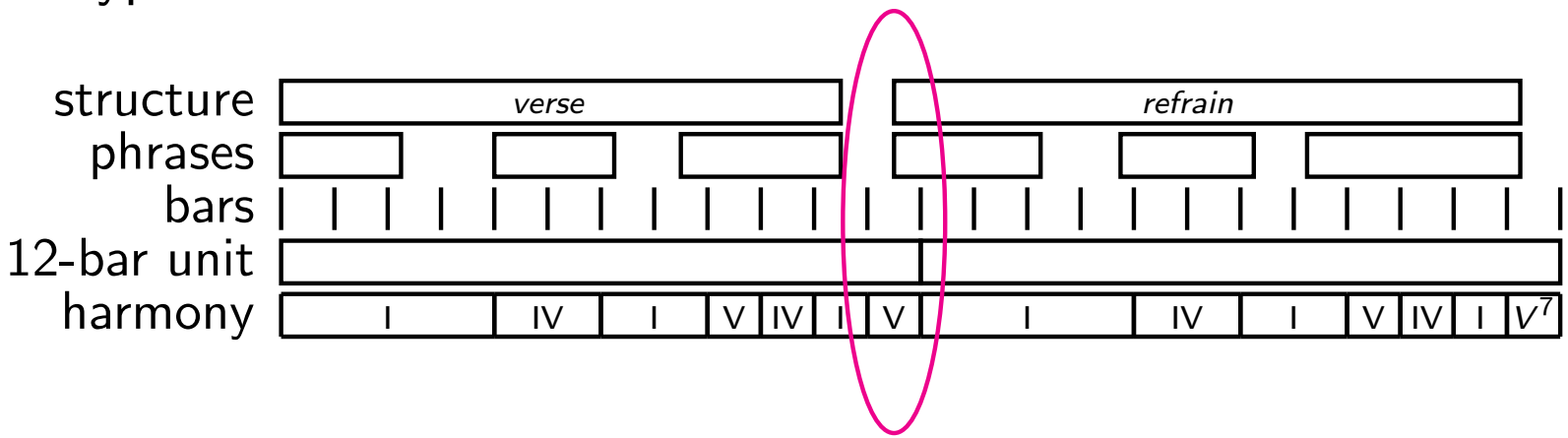
Two separate hierarchies:

- structure $\longrightarrow$ phrases;
- 12-bar unit $\longrightarrow$ harmony $\longrightarrow$ bars.

(not a realistic example!)

Hierarchical structure

Hypothetical 12-bar blues:



Two separate hierarchies:

- structure $\longrightarrow$ phrases;
- 12-bar unit $\longrightarrow$ harmony $\longrightarrow$ bars.

(not a realistic example!)

Outline

Approach

- Generate tree of hierarchical pairwise matching relationships
- Convert to map of regions related by (transitive) similarity relations
- Summarize map of regions

Approximations

Deal with sequential musical data. ("Strings", but maybe over an arbitrarily large or continuous alphabet.)

- Structures of interest are hierarchically arranged;
- Normalizeable (and meaningful) distance measure between sequence elements;
- Known length (time) scale of interest.

Hierarchy assumption

Boundaries of high-level (large extent) structures are not crossed by smaller units of that structure.

- Explicitly looking for hierarchical structure
- Design search such that non-hierarchical structure is not found

Advantages:

- No longer suffer from 'transitive closure' problem
- Search is fairly fast

Disadvantage:

- Sensitive to accuracy of high-level placement

Other Approximations

Error rate:

Define some error (or distance) between characters in the alphabet.

- 0 for the same character, 1 for a mismatch;
- cosine distance; ...

Strings 'match' if they have an average error below some threshold.
Constant (or monotonic) threshold allows search pruning.

Minimum length:

Some length scale below which we are not interested in matching.

Other Approximations

Error rate:

Define some error (or distance) between characters in the alphabet.

- 0 for the same character, 1 for a mismatch;
- cosine distance; ...

Strings 'match' if they have an average error below some threshold.
Constant (or monotonic) threshold allows search pruning.

Minimum length:

Some length scale below which we are not interested in matching.

Summarization

We obtain a tree of pairwise similarity relations over the input sequence. Generate summary labels by:

- converting tree into regions with the same similarity relations.
  - (this step depends on the hierarchical assumption)
- generating labels until no region which is either
  - larger than the minimum length, or
  - involved in a match with another region

  remains unlabelled.

Outline

Chopin, Op. 7 No. 2

$\|: \mathcal{A} :\|$ $\|: \mathcal{B} :\|$ $\mathcal{C}$ $\|: \mathcal{D} :\|$ *d.c.*

A A B B C D D A

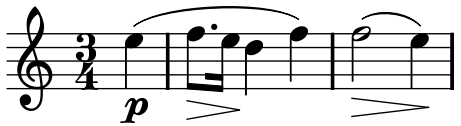
a b a b c b c b d e d e d a b

Chopin, Op. 7 No. 2

$\|: \mathcal{A} :\|$ $\|: \mathcal{B} :\|$ $\mathcal{C}$ $\|: \mathcal{D} :\|$ *d.c.*

A A B B C D D A

a b a b c b c b d e d e d a b

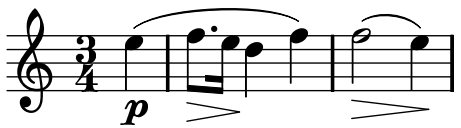
a:  ...

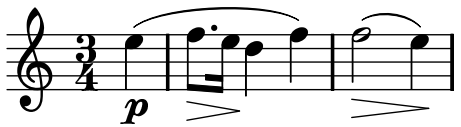
Chopin, Op. 7 No. 2

$\|: \mathcal{A} :\|$ $\|: \mathcal{B} :\|$ $\mathcal{C}$ $\|: \mathcal{D} :\|$ *d.c.*

A A B B C D D A

a b a b c b c b d e d e d a b

a:  ...

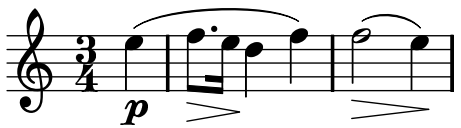
b:  ...

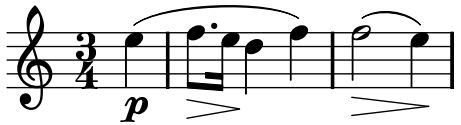
Chopin, Op. 7 No. 2

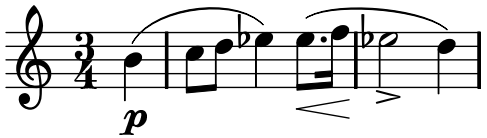
$\|: \mathcal{A} :\|$ $\|: \mathcal{B} :\|$ $\mathcal{C}$ $\|: \mathcal{D} :\|$ *d.c.*

A A B B C D D A

a b a b c b c b d e d e d a b

a:  ...

b:  ...

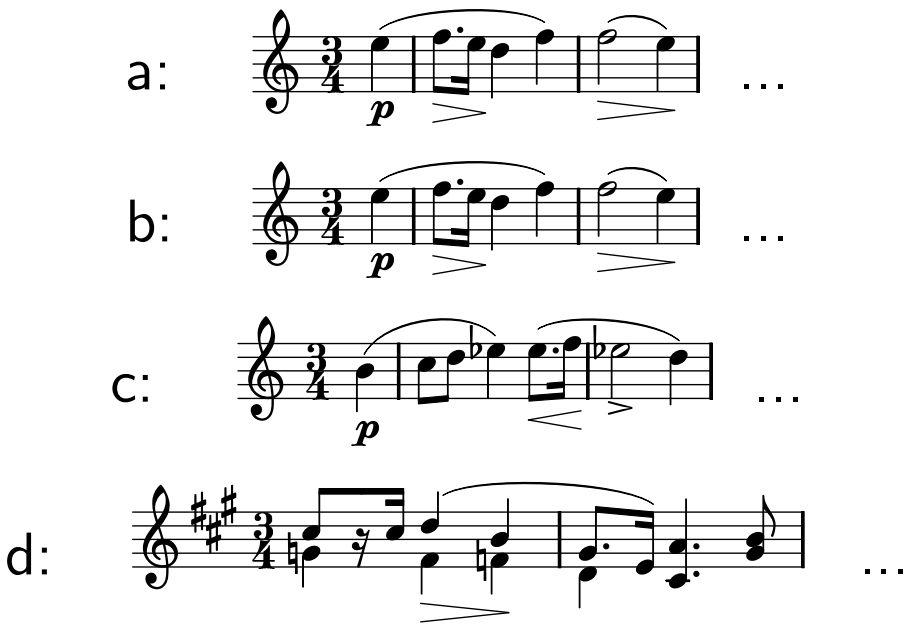
c:  ...

Chopin, Op. 7 No. 2

$\|: \mathcal{A} :\|$ $\|: \mathcal{B} :\|$ $\mathcal{C}$ $\|: \mathcal{D} :\|$ *d.c.*

A A B B C D D A

a b a b c b c b d e d e d a b

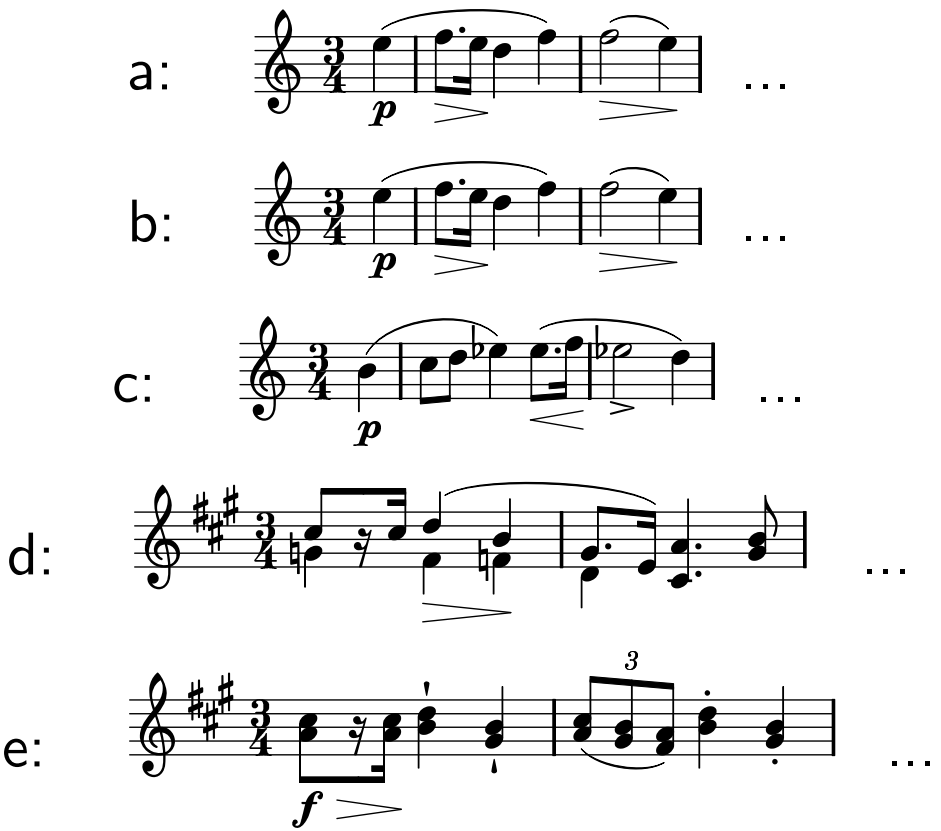


Chopin, Op. 7 No. 2

$\|: \mathcal{A} :\|$ $\|: \mathcal{B} :\|$ $\mathcal{C}$ $\|: \mathcal{D} :\|$ *d.c.*

A A B B C D D A

a b a b c b c b d e d e d a b



Chopin, Op. 7 No. 2: Results

27 recordings, 1930–2005. Ground truth from CHARM project.

Recording	Algorithmic labels	Ground Truth	Errors
Rubinstein (1939) Pobłolcka (1999) François (1956)	AABCCDA AABBCCDAB ABA	AABCDDA AABBCDDAB ABCDA	
Luisada (1990)	AABBCCA	AABBCDDA	(c)
Smith (1975)	ABABCBCBDDEABF	AABBCDDA	(b),(d)
Ts'ong (1993)	ABCCDDEA	AABBCDDA	(a)

- nine labellings completely* correct
- most of the rest are close; common discrepancies:
  - (d) labelling silence at end of track;
  - (c) missing one segment;
  - (b) labelling substructure;
  - (a) other.

(Provided ground truth incorrect in four cases)

Outline

Conclusions

- Can perform analysis of repetitions on expressive performances of classical music
- Explicit use of hierarchical assumption
- Given knowledge of score, better methods exist
- Extensible to other forms of sequential musical data

Future Work:

- Handle time warping efficiently
- Incremental variation of error threshold

Thanks:

- Craig Sapp, Raphael Clifford
- EPSRC GR/S84750/01