Automatically extracting harmony from recorded music – and what to do with it

Matthias Mauch – Symposium on Music Information Retrieval @ University of Utrecht, February 2012
Audio Chord Transcription 1

- DBN models musical context [1][2]
- bass, key, metric position
- 2012 state of the art adaptation: Ni et al. [3]
Audio Chord Transcription 1

- DBN models for musical context
- bass, key, metric position
- 2012 state-of-the-art adaptation: Ni et al. [3]

Matthias Mauch
Audio Chord Transcription 1

- DBN models musical context [1][2]
- bass, key, metric position
- 2012 state of the art adaptation: Ni et al. [3]
averaging features across repeated song segments [4]

non-systematic noise is attenuated \(\Rightarrow\) better results

Figure 6.1: "It Won't Be Long". Effect of using the repetition information (see discussion in Section 6.4.1): comparing the fully-automatic autobeat-autoseg method to the baseline method that does not use repetition information. (a) the two bottom bars are black at times where the chord has been recognised correctly (using MIREX-style evaluation) over the whole song. The two top bars display the manual segmentation (for reference) and the automatic segmentation used to obtain the autobeat-autoseg results. (b) manually-annotated and automatically-extracted chord sequences for an excerpt of the song.
Audio Chord Transcription 2

- averaging features across repeated song segments [4]
- non-systematic noise is attenuated ⇒ better results

![Graph showing song-wise improvement in RCO for the methods using segmentation cues over the respective baseline methods. The lower part of the figures shows the performance difference per song, and the upper part summarises the same information in a histogram. Using autobeat-autoseg improves performance on 71% of songs compared to autobeat-noseg (Figure 6.7a); manbeat-manseg improves RCO scores for 63% of songs compared to the manbeat-noseg method (Figure 6.7b).]

Our main hypothesis—that repetition cues can be used to improve chord transcription—has been confirmed with statistical significance. Next, we consider the difference between manual and automatic beat extraction to find out whether the more practicable automatic approach yields significantly lower results.

Matthias Mauch
Chordino & NNLS Chroma

- “NNLS Chroma” [5] Vamp plugin (e.g. for Sonic Visualiser)
- Chordino – a basic chord estimator
... now what to do with it?
... now what to do with it?
... now what to do with it?

Researchers

Chordino

Musicians & Learners

General Public

Matthias Mauch
... now what to do with it?

Chordino  SongPrompter  (Yanno)  Songle.jp  Driver’s Seat

Researchers

Musicians & Learners

General Public

Matthias Mauch
SongPrompter

... 
Verse:  
Bm  G  D  A  
Once you were my love, now just a friend,  
Bm  G  D  A  
What a cruel thing to pretend.  
Bm  G  D  A  
A mistake I made, there was a price to pay.  
Bm  G  D  A  
In tears you walked away.  

Verse:  
When I see you hand in hand with some other  
I slowly go insane.  
Memories of the way we used to be ...  
Oh God, please stop the pain.  

Chorus:  
D  G  Em  A  
Oh, once in a life time  
D/F#  G  A  
Nothing can last forever  
D/F#  G  
I know it's not too late  
A7  F#/A#  
Would you let this be our fate?  
Bm  G  Asus4  
I know you'd be right but please stay  
A  
Don't walk away  

Instrumental:  
Bm  G  D  A  Bm  G  A  

Chorus:  
Oh, once in a life time  
Nothing can last forever  
I know it's not too late  
Would you let this be our fate?  
I know you'd be right but please stay.  
Don't walk away.  

...
SongPrompter

![Audio waveforms and spectrograms with lyrics]

Matthias Mauch
SongPrompter

- automatic alignment works best with speech and chord features [6]
- visual display from automatic alignment
  - lyrics, segmentation and chords
- audio playback
- original audio
- auto-extracted bass and drum track
SongPrompter

- automatic alignment works best with speech and chord features [6]
- visual display from automatic alignment
  - lyrics, segmentation and chords
- audio playback
- original audio
- auto-extracted bass and drum track
SongPrompter

Matthias Mauch
Yanno

- Stowell’s research [7]: schools use YouTube in music lessons to play along to.
- how can music technology help?
Yanno

Matthias Mauch
Matthias Mauch
Songle Web Service
Songle Web Service [8]

- adding interaction
- engaging user experience
- insights through automatic annotations
- anyone can contribute – it’s social!
- use for MIR research
- crowd-sourcing more training data
- exposure to broader audience

Matthias Mauch
http://songle.jp/
Songle

Error Correction Interface: Chords
Last.fm already have genre tags, similarity

we want a complement: intuitively understandable audio features

‘harmonic creativity’ (structural change [9])

noisiness, energy, rhythmic regularity, ...

Spotify app based on Last.fm audio API
Driver’s Seat

Audio Feature API

Spotify ID API

Extraction

Audio

Spotify Apps API

Matthias Mauch
Driver’s Seat
Application-centred MIR – what’s in it?

- Research impact – exposure to the world.
- Crowd-sourcing feedback – training data.
- Can show what people really want.
- Very expensive to do well.
- Research shift: from music processing towards interaction and implementation.
Discussion

- Is the development of large proof-of-concept applications too costly? Better invest time in basic research?
- Few companies use today’s MIR methods. Should we be focusing on perception?
- Any other questions?
References

“Automatically extracting harmony from recorded music – and what to do with it”


Matthias Mauch