Smart Statistical Models for Musical Data

Building a Database for Studying Harmony in Popular Music

John Ashley Burgoyne and Ichiro Fujinaga

McGill — Schulich School of Music / École de musique Schulich — CIRMMT Centre for Interdisciplinary Research in Music Media and Technology — DDMAL Distributed Digital Music Archives & Libraries Lab

Why is musicology still data-poor?

While sub-atomic physics is moving into a period of data scarcity, the reverse situation appears to be happening for music.... Contemporary music scholars have access to computational and database resources, comprehensive reference tools, high quality data acquisition methods, sophisticated modeling techniques, and other innovations that make it far easier to collect, analyze and interpret musically-pertinent evidence and artifacts.

1999

— David Huron, UC Berkeley

I'm sure that a closer relationship with music information science presents musicologists with an opportunity. The trouble is, we've been standing at this moment of opportunity for quite some time now.

2005

— Nicholas Cook, ISMIR

We use the song collection from the 2009 MIREX Audio Chord Detection subtask, which contains 210 songs (174 by the Beatles, and 18 by Queen and Zweieck, respectively), making it the largest audio-aligned chord test set used to date.

2010

— Matthias Mauch, PhD dissertaiton

Opportunities in automation

Traditional approaches to building databases in the humanities rely on humans to transcribe all pertinent data and metadata—and all of this expert labour is expensive!

Fully automated approaches are appealing for their ability to generate arbitrarily large databases, but they are too unreliable to use for serious research.

Human-assisted automation asks humans to correct errors while the database is being generated and uses this information to improve performance.

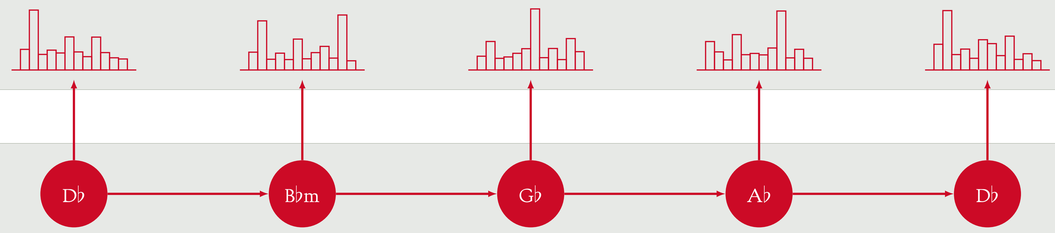
Musicological input is essential to building databases with human-assisted automation, even if some of the most interesting musicological questions must wait until the databases are built.

Transcribing chords from digital audio: Still unsolved!

- One of the major open problems in music information science is how to transcribe chords automatically from audio. It is—in theory—an easier task than transcribing a full score.
- No existing system can recognise chords reliably. The best-performing system currently, from Matthias Mauch's doctoral dissertation at Queen Mary, University of London (2010), achieves only 81 percent accuracy when chords are simplifed to major and minor only and considerably less when attempting to recognise complete chord symbols.

From chords to audio: Chroma vectors

- What are the aural properties of a chord?
- The standard assumption is that chords will produce a consistent **chroma vector** in audio. Chroma vectors represent the proportion of harmonic energy in each pitch class.
- The upper partials of the harmonic series confuse chroma vectors, especially for chords in first inversion.
- One paper (Harte 2005) has been published exploring one *Tonnetz*-inspired alternative to the chroma vector. We will explore others.



The Markov assumption: Can it be musical?

- In order to be computationally tractable, most models for musical audio make a **Markov assumption** that each chord can be predicted only from the immediately preceding chord and some predetermined contextual parameters.
- What context is necessary to represent tonal harmony as a Markov chain?
  - Key? (Burgoyne 2005 and others)
  - Metric position? (Mauch 2010)
  - Bass? (Mauch 2010)
  - Root? (area for exploration)
  - Intervals and voice-leading tendencies? (area for exploration)
- When the Markov chain for harmony and the expectations each chroma vector are combined into a single single statistical model, it is possible to adjust preliminary guesses for each chord to reflect the tonal context, just as humans do when making an analysis.

Engineering tricks

Discriminative training, as opposed to the more commonly-used generative training, concentrates the statistical power on modelling the music rather than the audio. Little other research in this area has explored its potential.

Augmented models use 'information geometry' to obtain the best predictions from across a family of more complicated models related to our basic model.

Billboard Hot 100

- We are working with a group of jazz musicians to transcribe 1 000 songs that appeared on the *Billboard Hot 100* chart between 1958 and 1991. It will be the largest database of its kind in the world.
- The songs were chosen using a sampling methodology that maximises the ability to draw conclusions about different time periods on the charts and levels of popularity.
- All transcriptions are double-keyed, include complete jazz notation for upper extensions, and are time-aligned with an audio recording.
- About half of the transcriptions have been completed to date, and we anticipate having all 1 000 songs available in early 2011.

Summary

We still need larger databases to realise the potential of computer-assisted musicology. With automated tools, we hope to grow the number of transcriptions availabe from the hundreds we have now to hundreds of thousands.

Social Sciences and Humanities Research Council of Canada / Conseil de recherches en sciences humaines du Canada — Canada