CSC475 Music Information Retrieval

Computational Ethnomusicology

George Tzanetakis

University of Victoria

2014

Introduction

University of Victoria Computer Science

Definition

The main task of *ethnomusicology* is to explain music and music making with reference to the social, but in terms of the musical factors involved in performance and appreciation
- J. Blacking, 1979.

Definition

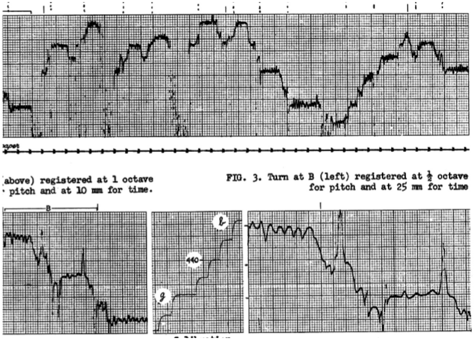
Computational Ethnomusicology is the design, development, and usage of computer tools that have the potential to assist in ethnomusicological research. - Tzanetakis et al, 2007

Motivation

- Existing work in MIR has focused on popular music (for example recommendation systems) or on classical music (for example score following and query-by-humming).
- Large amounts of recordings that are practically impossible to access without better tools
- The majority of music is not written down but the majority of musicological research is based on notated scores.
- The development of inexpesnive sensors and actuators have enabled detailed analysis and synthesis of performance gestures used to generate music in addition to the music itself.

Bartok, Seeger and the Mellograph

The origins of ethnomusicology can be traced to the hand written transcription of Hungarian folk songs done by Bela Bartok. Charles Seeger (the father of Pete Seeger) pioneered the study of ethnomusicology and the use of the Melograph, an electro-mechanic machine for the “transcription” of non-notated music.



Bartok, Seeger, and the Mellograph

Audio recordings were also very important in the development of ethnomusicology as a field: *"The only true notations are the sound-tracks on the record itself"*, Bartok.

Bela Bartok

Charles Seeger

Mellograph (pitch - time graph, volume - time graph)

Why MIR in CE ?

- Detect and analyze information that would be difficult and time-consuming to do by hand (onsets, pitch contours etc).
- Scale to large data sets.
- Provide visual represntations that can serve as “scores” for music that is not notated.
- Leverage and adapt the large amount of existing techniques and algorithms that have been developed (mostly for Western music).

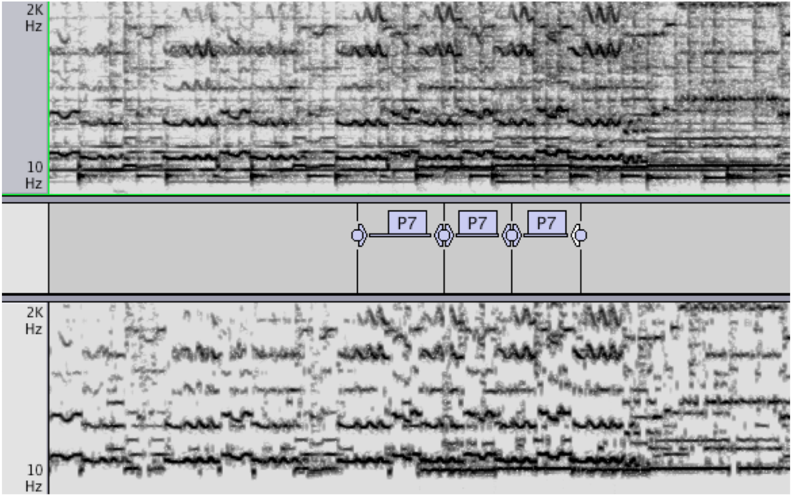
Some older work in CE

Ethnomusicologist James Kippen and computer scientist Bernard Bel developed grammar rules for tabla playing using a master-student paradigm. Lieberman (1970) manually transcribed large number of gamelan recordings and then used computer techniques to calculate statistics about the pitch material and search for patterns. Schloss (1987) did early work on the computer transcription of percussive music and proposed the "digital melograph".

Guidelines for CE Research

- Collaboration with music scholars, not just engineers
- Large collections
- Design of domain-specific and culture-specific techniques

Source separation and pattern detection of Duduk



CompMusic (2011-)

Computational Models for the Discovery of the World's Music, is a research project funded by the Europen Research Council and coordinated by Xavier Serra from the Music Technology Group of the Universitat Pompeu Fabra in Barcelona. The project runs from 2011 to 2016.
Five music cultures:

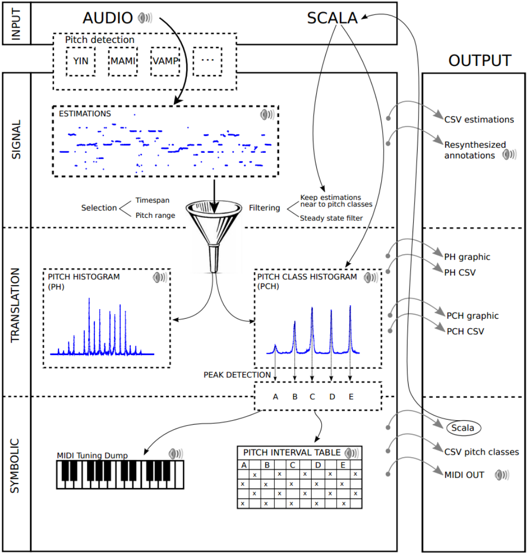
- Hidustani (North India)
- Carnatic (Sound India)
- Turkish-makam (Turkey)
- Arab-Andalusian (Maghreb)
- Beijing Opera (China)

Tasks

University of Victoria Computer Science

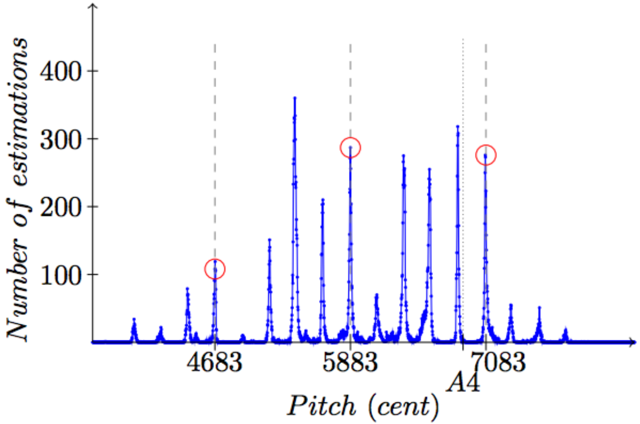
- Pitch Contour
  - Tuning and pitch analysis
  - Tonic detection
  - Melodic motifs
  - Melodic stability
- Rhythm (typically onsets)
  - Rhythm similarity and style classification
  - Domain specific algorithm (for example clave-aware)
  - Microtiming visualization and analysis
- Transcription
- Access, retrieval, browsing
  (for example Dunya from CompMusic)

Example 1: Tarsos

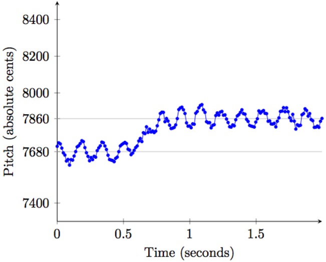


Tarsos: Pitch Histogram

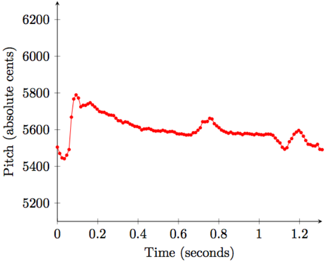
Indonesian Sleandro scale



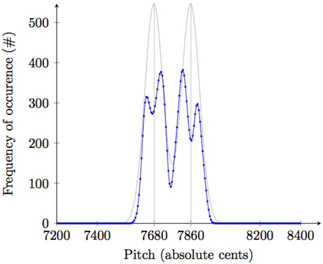
Tarsos: vibrato analysis

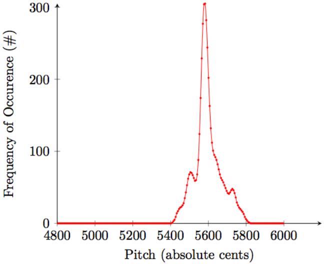


(a) Western vibrato melograph.



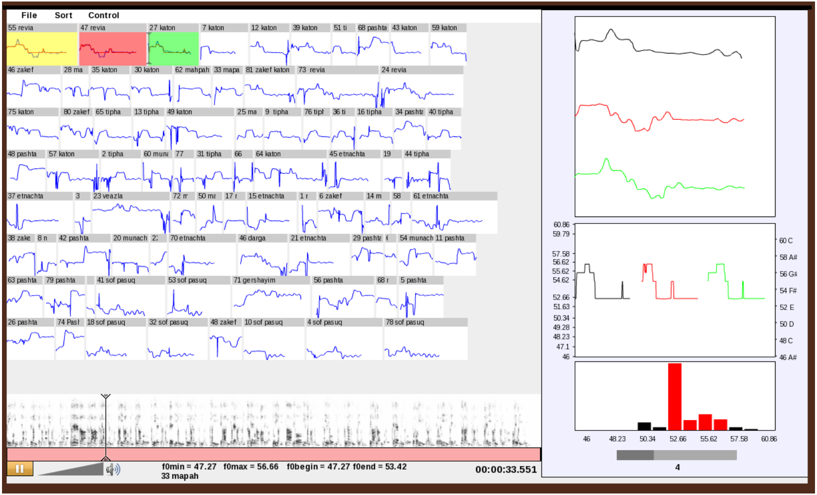
(b) Indian raga pitch gesture.





Example 3: Cantillion

Web interface for syntagmatic and paradigmatic analysis.
It enables non-linear listening and viewing similarly to a score



Cantillion: Torah Chant

University of Victoria
Computer Science

וַיֹּ֣אמֶר אֱלֹהִ֗ים יִקָּו֨וּ הַמַּ֜יִם

Genesis 1:9

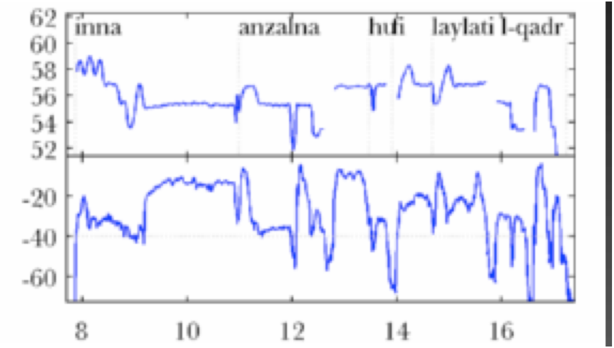
And God said "Let the waters be collected"

Letters in black
Vowel points in red
Cantillation signs in green

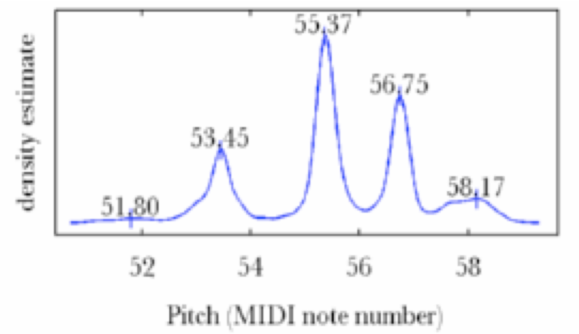
Cantillion: Melodic gestures

Segmentation into melodic gestures is done based on the text by an expert (Dr. Daniel P. Biro from the School of Music, University of Victoria).

Monophonic pitch extraction



Recording-specific scale



Cantillion: Hypothesis

Quantifying variation

The recitation of a particular piece of text from the Torah depends on the surrounding music culture. More specifically we will examine Marocco and Hungary.

Melodic similarity

Similarity between melodic fragments is calculated based on their associated pitch contour quantized using the recording-derived scale. Dynamic time warping (DTW) is used for calculating the alignment score.

Cantillion: Mean Average Precision

Ta'am (Morocco)	Average Precision (Morocco)	Ta'am (Hungary)	Average Precision (Hungary)
sofpasuq	0.550	sofpasuq	0.994
katon	0.399	revia	0.967
tipha	0.306	etnachta	0.945
mapah	0.299	pashta	0.683
pashta	0.269	tipha	0.673
revia	0.245	katon	0.562
etnachta	0.234	mapah	0.550
zakef	0.206	merha	0.530
merha	0.158	zakef	0.231
munach	0.147	munach	0.179
kadma	0.036	kadma	0.040

Table 1. Mean average precision for different te'amim based on the alignment distances.

Cantillion: Distance Histograms

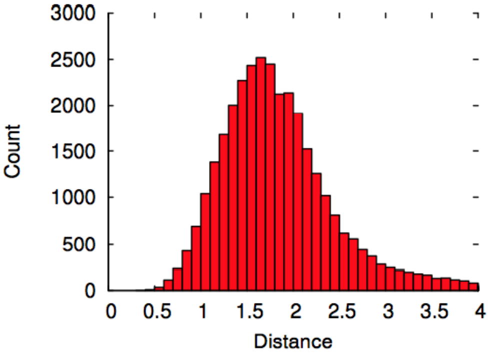


Figure 3. Distribution of distances between unrelated segments.

Distance Histograms

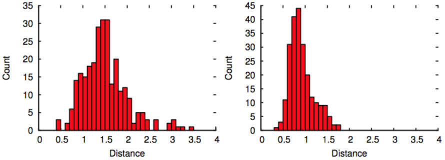


Figure 4. Distribution of distances between renditions of the *tipha* in the Moroccan interpretation (left) and the Hungarian interpretation (right).

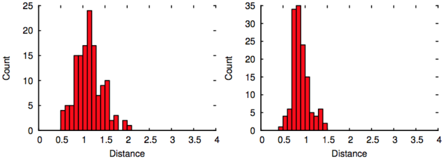


Figure 5. Distribution of distances between renditions of the *sof pasuq* in the Moroccan interpretation (left) and the Hungarian interpretation (right).

Musical Time

Theoretically rhythm is described as :

- Discrete beats
- Integer subdivisions (8th note, 16th notes)

In practice, *microtiming* plays an important role:

- Expressive deviations from exact timing
- “Swing”, “human fee”, “groove”, “playing ahead”, “playing behind”, “pushing the rhythm”
- Not just random human imperfection

Example 3: Instantaneous tempo curves

Timing analysis of two performances of a composition for the Kazakh dombra (time-stretched for alignment as they have different durations) by Matt Wright:

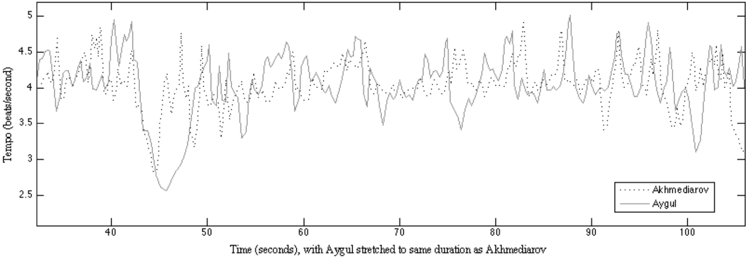


Figure 1. Instantaneous tempo curves of two performances of a composition for the Kazakh *dombra* (time-stretched for alignment as they have different durations).

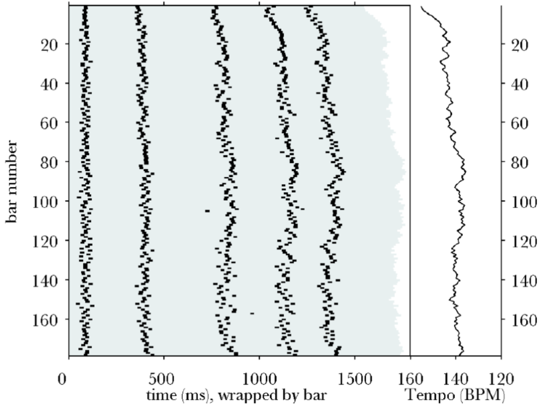
Clave

Clave literally means key. It can refer to:

- The instrument itself
- The part played by that instrument
- The metric structure expressed by the clave part. (e.g. this drum enters on the third note of the clave)

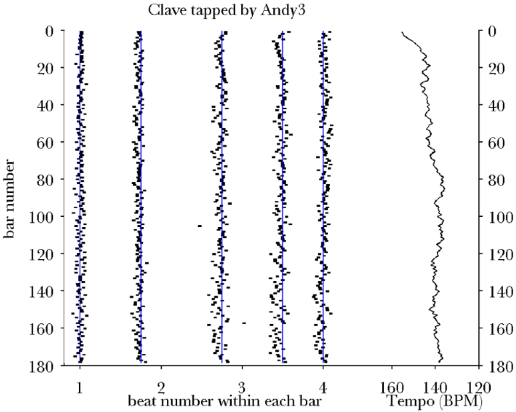
Example 4: Visualizing micro-timing

Original tap data and associated tempo curve for Afrocuban recording with underlying clave pattern.



Example 4: Visualizing micro-timing

Bar-wrapping uses time stretching to visually separate timing changes due to increasing/decreasing tempo and micro-timing changes at each theoretical beat location. Each theoretical beat location is a line.



Beat Tracking using Dynamic Programming

Dan Ellis in 2007 proposed the use of dynamic programming for beat tracking. Beat locations should generally correspond to onsets and the set of beats should reflect a locally-constant inter-beat interval. These constraints map to the local match and transition cost of dynamic programming. Beat tracking is performed by first estimating a global tempo and then using dynamic programming to find the best-scoring set of beat times that reflect the tempos as well as the corresponding moments of high onset strength.

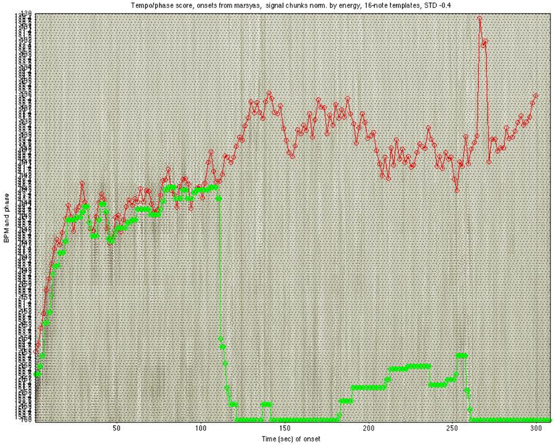
Template-based beat tracking using DP

University of Victoria Computer Science

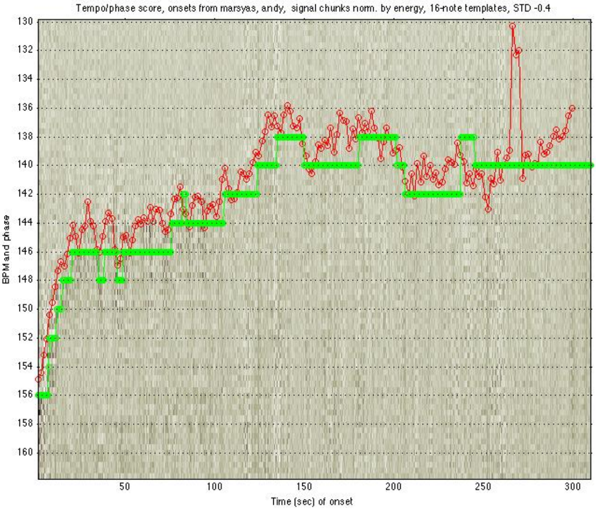
Main idea is to use domain specific knowledge i.e the clave.

- Consider each detected onset time as a potential note of the clave pattern.
- Exhaustively consider every possible tempo (and clave rotation) at each onset by cross-correlating each of a set of clave-pattern templates against an onset strength envelope signal beginning at each detected onset.
- Interpret each cross-correlation result as a score for the corresponding tempo (and clave rotation) hypothesis.
- Connect the local tempo and phase estimates to provide a smooth tempo curve and deal with errors in onset detection, using dynamic programming

Classic Dynamic Programming Beat Tracking

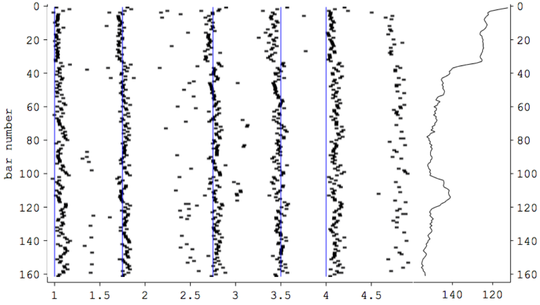


Template-based Dynamic Programming Beat Tracking



Example 4: Visualizing micro-timing

Automatic analysis. Notice the slight consistent delay of the tap data for the last theoretical clave beat location.



Future Direction 1: NIME for CE

Music is created by humans interacting with a wide assortment of devices we call musical instruments. A lot of interesting information can be obtained by analyzing the gestures used to create the music. Using either direct or indirect sensors and computer analysis it is possible to extract performance information that is not readily available from a recording. Such information can be used to study timing and dynamic nuances of expert performers, playing technique, and performer identity among other topics. There is little work explore instruments from around the world and even for Western instruments most of it has focused on the piano.

Example: Sensing for the Africal Gyil

Trail et all (NIME 2012)

Future Direction 2: Pedagogy

Learning traditional world instruments can be a challenge as there is no established pedagogy such as books and exercises. Expert players that can serve as mentors are few and in some cases disappearing. Computational ethnomusicology is ultimately about using computer to better understand music and music making. This understanding can be leveraged to create culture-specific and instrument-specific tools for music pedagogy. For example pitch analysis can be used to teach micro-tuning and pattern detection and time-stretching techniques can be used to isolate a particular phrase and practice it at a slower tempo.