CompMusic research progress: halfway results

Xavier Serra
Music Technology Group
Universitat Pompeu Fabra, Barcelona
Index

• Introduction
• Research corpora
• Research topics addressed
  ▪ Tonic identification
  ▪ Intonation analysis
  ▪ Melodic analysis
  ▪ Rhythmic analysis
  ▪ Score-audio alignment
  ▪ Ontologies
  ▪ Lyrics-music analysis
  ▪ Music exploration
Some numbers

- Started on July 2011 and will finish on July 2016
- 2.5 million euros of funding from ERC
- 30 researchers plus several collaborators
- 60 publications (one special issue in JNMR)
- 3 workshops
- Large impact in MIR community
From traditional data processing methodologies to knowledge-driven approaches
Problem

Melody

Rhythm
Data

T.M. Krishna is at the forefront of classical vocalists in India today and is known in the musical world as a classicist. His tremendous stage presence, his awe-inspiring voice, his great scholarship and his ability to transcend cultural borders makes him the greatest ambassador of Carnatic classical music. He contributes to Carnatic music in numerous ways other than performing. He has started and is involved in many organisations whose work spreads across the whole spectrum of Carnatic music including research, archival and documentaries, taking Carnatic music to various parts of society and to smaller towns and villages, conducting festivals focused at the youth, supporting artists from rural south India who need opportunities to widen their horizons and spotting youth talent and giving them opportunities. He conducts creativity workshops for corporates and has lectured in various organisations like IIM Bangalore, IIT Madras, CII, and Harvard university. He is an author who has co-authored 'Voices within' a book dedicated to some of the greats of Carnatic music and contributes regularly to various journals and newspapers. He has also always been in the forefront to raise very many sensitive issues that affects Carnatic music over many years. A multifaceted personality.

CD 1
1 Tharunam Edamma - D.K. Jayaraman
   background vocals: Sukanya Sankararaman, Balaji Shankar
   lead vocals: D.K. Jayaraman
   mridangam: J. Vaidyanathan
   violin: R. K. Shriramkumar
   recording of: Tharunam Edamma
   composer: Syama Sastri
   lyricist: Syama Sastri

2 Parvathi Pathim
   background vocals: Sukanya Sankararaman, Balaji Shankar
   lead vocals: D.K. Jayaraman
   mridangam: J. Vaidyanathan
   violin: R. K. Shriramkumar
   recording of: Parvathi Pathim
   composer: Muthusvamy Dikshitar
   lyricist: Muthusvamy Dikshitar

3 Chenithodi Theve
   background vocals: Sukanya Sankararaman, Balaji Shankar
   lead vocals: D.K. Jayaraman
   mridangam: J. Vaidyanathan
   violin: R. K. Shriramkumar
   recording of: Chenithodi Theve

قدام الأصباهن
توضيح الميزان
الصنعة 1. شغل الأصباهن، زجل
يا تؤكتب الأخلاء، قلبي هو فواك
توضيح
هو فواك يا لضن
قد خحز الأذان
يا ولد ينطن
رقتاً بحن فواك ودش جداه
الصنعة 2.، اصباهن، توضيح
من محتواي سقط
كأنز في الحضان
يا نازل النذالي
من دعمي
يا نشيا الغزال
يا فطع الجمال
يا غضب فتغذ
يا جذب وناري
يا فطع فتغذ
مذب darling
من نشيا سقط
Methodology
Evaluation
Hindustani  Arab-Andalusian  Carnatic  Research corpora  Hindustani  Turkish makam  Beijing Opera
Audio Recordings

Carnatic
- 248 CDs
- 397 hours

Hindustani
- 233 CDs
- 271 hours

Makam - Turkey
- 225 CDs
- 259 hours

Beijing Opera
- 48 CDs
- 46 hours

Arab-Andalusian
- 56 hours of personal recordings
Research corpora

- **Purpose:** Develop methodologies with which to extract musically meaningful features from audio music recordings related to melody and rhythm.
- **Coverage:** Audio recordings, plus appropriate accompanying information, covering the varieties of melodies and rhythms present in each musical culture.
- **Completeness:** Editorial metadata and the descriptive information accompanying each audio recording.
- **Quality:** The audio has to be well recorded and the accompanying information has to be accurate.
- **Reusability:** Corpus has to be available for the research community to use.
Audio features

metadata:
  version:
    essentia: "1.2.2"
pitch_confidence:
  mean: 0.725613296032
  skew: -1.01243495941
spectral_flatness_db:
  mean: 0.189157426357
spectral_rms:
  mean: 0.0014992490178
spectral_strongpeak:
  mean: 0.62891064697
  var: 0.741435945034
mfcc:
  mean: [143.017318726, -38.4358482361, 7.8873262, -4.917526245, 159.187026978]
  var: [122.917526245, 159.187026978]
onset_rate: [1.70000004768]
tristimulus:
  mean: [0.29845097661, 0.602976679802]
Editorial metadata
Expert information
## Data status

<table>
<thead>
<tr>
<th></th>
<th>Carnatic</th>
<th>Hindustani</th>
<th>Makam-Turkey</th>
<th>Beijing Opera</th>
<th>Arab-Andalusian</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Audio Recordings</strong></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td><strong>Editorial Metadata</strong></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td><strong>Lyrics</strong></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️ (Partially)</td>
</tr>
<tr>
<td><strong>Expert Information</strong></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td><strong>Wikipedia/dbpedia</strong></td>
<td>✔️</td>
<td>✔️</td>
<td>Few</td>
<td>Baidu baike</td>
<td>Very Few</td>
</tr>
<tr>
<td><strong>Community Information</strong></td>
<td>✔️</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>Facebook, Youtube</td>
</tr>
<tr>
<td><strong>Audio Features</strong></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td><strong>Music Notation</strong></td>
<td>Not digital</td>
<td>Few &amp; Not machine readable</td>
<td>✔️</td>
<td>Few &amp; Not machine readable</td>
<td>Very few &amp; Not digital</td>
</tr>
</tbody>
</table>
Tonic identification

- Upper Pa (5th)
- Lead voice
- Sa (tonic)
- Lower Pa (5th)

Multipitch Histogram:

- $\rho_2$
- $\rho_4$
- $\rho_3$
- $\rho_5$

Normalized salience vs Frequency bins (1 bin = 10 cents), Ref: 55Hz
Intonation
Motivic analysis

Discovery Induction
Extraction

Matching Retrieval

Transform

Characterization

N dimensions
Rhythm analysis
Score to audio alignment
Score to audio alignment

Uşşak Saz Semâî

Section Level
Score to audio alignment

Uşşak Saz Semâi

Note
Level
Ontologies
Chinese is a **tonal language**

<table>
<thead>
<tr>
<th>妈</th>
<th>麻</th>
<th>马</th>
<th>骂</th>
<th>吗</th>
</tr>
</thead>
<tbody>
<tr>
<td>ma1</td>
<td>ma2</td>
<td>ma3</td>
<td>ma4</td>
<td>ma</td>
</tr>
<tr>
<td>“mum”</td>
<td>“hemp”</td>
<td>“horse”</td>
<td>“to scold”</td>
<td>“?”</td>
</tr>
</tbody>
</table>

Hypothesis: performers try to **reflect tone contours in the melody** in order to make lyrics understandable.
Music exploration
Conclusion

- Already half way through!!!

- Problem definition ✔
- Research corpora ✔
- Methodology ✔
- Evaluation ✔

- Now we just have to finish our research!!!
Thanks !!