An Approach for Linking Score and Audio Recordings in Makam Music in Turkey

Sertan Şentürk\textsuperscript{1} André Holzapfel\textsuperscript{2} Xavier Serra\textsuperscript{1}

\textsuperscript{1}Universitat Pompeu Fabra
\textsuperscript{2}Bahçeşehir University

\{name.surname\}@upf.edu
Outline

• Introduction
• Background
• Methodology
• Initial Experiments
• Conclusion
Introduction and Motivation

- Audio recordings and scores are highly valuable sources of information to study music
  - Different representations of a music piece
  - Provide complementary information
- Linking relevant representations from both sources is beneficial
- Section linking is an important step to assess the structural organization of the performance of a piece.
  - Complementary to other computational tasks
  - Analysis might be focused in the section level
  - Provide a deeper insight on various properties of music
Makam Music in Turkey

- At least 17 intervals in an octave
- Different tunings (*ahenk*)
- A score representation based on extending Western notation
  - Typically follows Arel-Ezgi-Uzdilek theory
  - Generally devoid of the expressive elements such as embellishments, heterophony etc...
  - Most are transcriptions of the performances with the intent to help people study the music piece
Related Computational Research

- Verse-Chorus Detection
Related Computational Research

• Cover Song Detection

Figure 3. CRPs for the song *Day Tripper* as performed by The Beatles, taken as song $X$, versus two different songs, taken as song $Y$. These are a cover made by the group Ocean Colour Scene (a) and the song *I’ve Got a Crush on You* as performed by Frank Sinatra (b). Parameters are $m = 9$, $\tau = 1$, and $\kappa = 0.08$.

(Serrà et. al, 2009)
Related Computational Research

- Image Processing
Methodology

Information Sources
- Audio
- Score
* Metadata

Descriptor Extraction
- Pitch Contours
- Pitch Histogram

Candidate Estimation
- Section Candidates

Hierarchical Linking
- Linked Sections

Hierarchical Linking

Additional Metadata
- Section info (order of the sections) (leading role of each voice)
- Theoretical metadata (key signature of the major, modal centers of the minor, intervals of the measure)
Information Sources

- **Audio**
- **Score**
- **Metadata**

**Information Sources**
- MusicBrainz
- Audio Recordings

**Descriptor Extraction**
- Tuning & Audio Pitch Contour
- Synthetic Pitch Contours

**Candidate Estimation**
- Morphological Operations
- Thresholding & Structural Component Analysis
- Hough Transformation
- Analytic Geometry Operations

**Hierarchical Linking**

---

**Additional Metadata**
- Section info
  - (order of the sections)
  - (starting note of each section)
- Theoretical metadata
  - (key signature of the makam)
  - (modal centers of the makam)
  - (intervals of the makam)

**MusicBrainz**
- Artists
- Works
  - (makam)
  - (usual)

**Audio Recordings**
- Recording of
Candidate Estimation

Information Sources
- MusicBrainz
  - Artists
    - Recording of works
- Additional Metadata
  - Composed
  - Lyric of works

Descriptor Extraction
- Tuning & Audio Pitch Contour
- Synthetic Pitch Contours

Candidate Estimation
- Section Candidates

Morphological Operations
- Thresholding & Structural Component Analysis
- Hough Transformation
- Analytic Geometry Operations

Hierarchical Linking
Hierarchical Linking

Information Sources
- MusicBrainz
  - Artists
  - Works
    - (makam)
    - (forms)
    - (usul)
- Additional Metadata
  - Section info
    - (order of sections)
    - (starting note of each section)
  - Theoretical metadata
    - (key signature of the makam)
    - (mode centers of this makam)
    - (intervals of the makam)

Descriptor Extraction
- Audio Recordings
- Symbolic Notation(s)

Candidate Estimation
- Tuning & Audio Pitch Contour
- Synthetic Pitch Contours

Hierarchical Linking
- Linked Sections
Implementation

• Symbolic notation in symbTr format (Karaosmanoğlu, 2012)
• Makam Toolbox (Bozkurt, 2008) used for tuning analysis and f0 estimation
  • Makam Toolbox uses YIN (Cheveigné & Kawahara, 2002) and applies post-processing
• Main framework implemented in MATLAB
Dataset

• “Instrumental Pieces Played with the Ney” collection from neyzen.com. The musicians:
  • Look at a few versions of the same piece, pick a preferred one
  • Check the score and make corrections, if necessary;
  • Perform the piece while referring to the score

<table>
<thead>
<tr>
<th>Composition</th>
<th>Composer</th>
<th>Structure</th>
<th>#Events in Score</th>
<th># Sections in Recording</th>
<th>Neyzen / Ney</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hicaz Saz Semai</td>
<td>Muhittin Erev</td>
<td>4 Hane</td>
<td>265</td>
<td>8, 8</td>
<td>Salih Bilgin/Kız</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 Teslim</td>
<td></td>
<td></td>
<td>Salih Bilgin/Mansur</td>
</tr>
<tr>
<td>Hicaz Saz Semai</td>
<td>Kul Mehmet</td>
<td>4 Hane</td>
<td>592</td>
<td>4, 4</td>
<td>Salih Bilgin/Kız</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Salih Bilgin/Mansur</td>
</tr>
<tr>
<td>Hicayni Peşrev</td>
<td>Kul Mehmet</td>
<td>4 Hane</td>
<td>592</td>
<td>4, 4</td>
<td>Salih Bilgin/Kız</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Salih Bilgin/Mansur</td>
</tr>
<tr>
<td>Hüseynî Saz Semai</td>
<td>Lavgacı Andon</td>
<td>4 Hane</td>
<td>307</td>
<td>8, 8</td>
<td>Salih Bilgin/Kız</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 Teslim</td>
<td></td>
<td></td>
<td>Salih Bilgin/Mansur</td>
</tr>
<tr>
<td>Rast Saz Semai</td>
<td>Osman Bey</td>
<td>4 Hane</td>
<td>323</td>
<td>8, 8</td>
<td>Salih Bilgin/Kız</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 Teslim</td>
<td></td>
<td></td>
<td>Salih Bilgin/Mansur</td>
</tr>
<tr>
<td>Uşşak Saz Semai</td>
<td>Salih Dede</td>
<td>4 Hane</td>
<td>429</td>
<td>8, 8, 12 (teslim repetition)</td>
<td>Salih Bilgin/Kız</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 Teslim</td>
<td></td>
<td></td>
<td>Salih Bilgin/Mansur</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Volkan Yılmaz/Müstahsen</td>
</tr>
</tbody>
</table>
# Results: Pre-Hierarchical

<table>
<thead>
<tr>
<th>Pre-Hierarchical Linking</th>
<th>Candidate Estimated</th>
<th>Candidate Not Estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Link Present</td>
<td>72</td>
<td>12</td>
</tr>
<tr>
<td>Link Not Present</td>
<td>13</td>
<td>-</td>
</tr>
</tbody>
</table>

72 / 84 correct links
## Results: Pre-Hierarchical Linking

<table>
<thead>
<tr>
<th>Pre-Hierarchical Linking</th>
<th>Candidate Estimated</th>
<th>Candidate Not Estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Link Present</td>
<td>72</td>
<td>8 + 4</td>
</tr>
<tr>
<td>Link Not Present</td>
<td>14</td>
<td>-</td>
</tr>
</tbody>
</table>

72 / 84 correct links

*Karar* detection failed in 2 pieces!
Manual correction...
Results: Post-Hierarchical Linking

<table>
<thead>
<tr>
<th>Post-Hierarchical Linking</th>
<th>Candidate Estimated</th>
<th>Candidate Not Estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Link Present</td>
<td>84</td>
<td>0</td>
</tr>
<tr>
<td>Link Not Present</td>
<td>0</td>
<td>-</td>
</tr>
</tbody>
</table>

84 / 84 correct links!

Note: No claim of statistical significance...
Discussion & Conclusion

• The results might be regarded as “proof-of-the-concept”
  • Currently adding more data from commercial recordings with various instruments

• Implemented system is pretty fast:
  • ~ 5-7 seconds on Ubuntu 64-bit computer with 3.33GHz Intel processor and 13.5 GB ram

• Initial experiments on audio-score alignment
  • Qualitatively (i.e. listening) adequate but needs more work

• Plan to carry comparative studies on section matching
  • Generalized Hough Transform, Geodesics, Time Series Analysis...
References


