Computational Analysis of Intonation in Indian Art Music

Gopala Krishna Koduri\textsuperscript{1}, Joan Serrà\textsuperscript{2}, Xavier Serra\textsuperscript{1}

\textsuperscript{1}Music Technology Group, Universitat Pompeu Fabra, Barcelona, Spain.
\textsuperscript{2}Artificial Intelligence Research Institute (IIIA-CSIC), Bellaterra, Barcelona, Spain.

\textsuperscript{2}\textsuperscript{nd} CompMusic Workshop, Bahçeşehir Üniversitesi, Beşiktaş, İstanbul, Turkey
Topics

• What is intonation?
• Context and purpose
• Histogram parametrization
• Work in progress
  • Swara-based histograms
  • Pattern analysis
• Discussion & conclusions
What is Intonation?

• Pitches used by a performer in a given performance.
• Our context: Pitch variations within a swara.
  • Eg: Ga₂ in Darbar and Nayaki
• Relevant references
  • Tuning: Krishnaswamy (2003), J. Serrà et al (2011)
• Melodic analysis
  • **Intonation profile**, motives, structure and low-level features.
• Rhythmic analysis
• Metadata
• Web-data
Histogram analysis – Goal!

Normalized frequency of occurrence

Pitch value (Cents)

Mukhari

Bhairavi
Overview of Method

Audio → Tonic ID

Segmentation → Histogram Analysis

Pitch Extraction → Peak Detection

Vocal Segments

Pitch contour

Histogram

Peak-labeled Histogram

Position, Mean, Variance, Kurtosis, Skewness.
Segmentation

- Why just vocal segments?
- Segment classes
  - Vocal (solo/mix)
  - Violin (solo/mix with percussion)
  - Percussion (solo)
- Support vector machine model
  - Trained on 300 minutes audio data
  - Features: MFCCs, pitch confidence, spectral flatness, flux, rms, rolloff, strongpeak, zcr and tristimulus
- Accuracy: 96% (10-fold cross-validation test)
Audio

Vocal Segments

Segmentation

Pitch Extraction

Pitch contour

Tonic ID

Histogram Analysis

Peak Detection

Peak-labeled Histogram

Parametrization

Position, Mean, Variance, Kurtosis, Skewness.
Pitch Extraction

- Violin interference
  - Filling in the gaps
  - Mimicking with time-lag.
- Multi-pitch analysis
  - Predominant melody extraction
- Combination with YIN
  - Why?
Overview

Audio → Vocal Segments → Pitch Extraction → Tonic ID

Segmentation → Histogram Analysis

Histogram → Peak Detection

Peak-labeled Histogram → Parametrization

Position, Mean, Variance, Kurtosis, Skewness.
Histogram Analysis

• ... bin resolution!
• Histogram

\[ H_k = \sum_{n=1}^{N} m_k \]

\( H_k \) is \( k^{th} \) bin count, \( N \) is number of pitch values, 
\( m_k = 1 \) if \( c_k \leq P(n) \leq c_{k+1} \) and \( m_k = 0 \) otherwise. \( P \) is array of pitch values and \( c_k, c_{k+1} \) are bounds of \( k^{th} \) bin.

• Purpose of average histogram
  • Reliability of peak estimation in single histogram
Overview

Audio → Segmentation → Vocal Segments → Pitch Extraction → Pitch contour

Vocal Segments → Pitch Extraction

Pitch Extraction → Peak Detection → Histogram Analysis

Peak Detection → Histogram Analysis

Histogram Analysis → Parametrization

Parametrization

Position, Mean, Variance, Kurtosis, Skewness.
Peak Detection

- Pitch contour smoothed using a Gaussian kernel
- $D_p$ and $L_p$: Depth and look-ahead parameters
- Valleys are deeper than $D_p$
- Peaks are local maxima
  - Locality: $L_p$ bins ahead.
- Average histogram
  - $D_p$ and $L_p$ set to higher values
- Histogram of a single recording
  - $D_p$ and $L_p$ set to lower values
Audio → Segmentation → Vocal Segments → Pitch Extraction → Tonic ID

Pitch Extraction → Histogram Analysis → Peak Detection → Peak-labeled Histogram → Position, Mean, Variance, Kurtosis, Skewness → Parametrization
Parametrization

- Distribution bounds
- Calculate the parameters
  - Position
  - Mean
  - Variance
  - Kurtosis (Peakedness)
  - Skewness (Slantedness)
Data

• Subset of CompMusic Carnatic dataset
  • 16 raagas, 170 recordings (at least 5 per raaga), 35 vocalists

• Task 1: Explorative raaga recognition task
  • 3 raagas, 42 recordings
  • 2 raagas, 26 recordings

• Task 2: Distinguishing allied raagas
  • 3 sets, 7 raagas, 60 recordings

• Task 3: Analysis of peak positions
  • All recordings from the subset!
Results (1/3)

<table>
<thead>
<tr>
<th>Features/Classifier</th>
<th>Naive Bayes</th>
<th>1-Nearest Neigh.</th>
<th>SVM</th>
<th>Logistic Regression</th>
<th>Random Forest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean and Height</td>
<td>63.43%</td>
<td>56.67%</td>
<td>61.81%</td>
<td>56.33%</td>
<td>64.62%</td>
</tr>
<tr>
<td>All parameters combined</td>
<td>63.76%</td>
<td>68.90%</td>
<td>65.19%</td>
<td>68.86%</td>
<td>70.71%</td>
</tr>
</tbody>
</table>

**Table 1.** Results of an exploration raaga classification test with 42 recordings in 3 raagas using different classifiers. The random baseline accuracy in this case is 28.57%.

<table>
<thead>
<tr>
<th>Features/Classifier</th>
<th>Naive Bayes</th>
<th>1-Nearest Neigh.</th>
<th>SVM</th>
<th>Logistic Regression</th>
<th>Random Forest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean and Height</td>
<td>39.6%</td>
<td>39.85%</td>
<td>41.25%</td>
<td>43.65%</td>
<td>48.85%</td>
</tr>
<tr>
<td>All parameters combined</td>
<td>58.05%</td>
<td>67.6%</td>
<td>74.25%</td>
<td>77.45%</td>
<td>74.45%</td>
</tr>
</tbody>
</table>

**Table 2.** Results of an exploration raaga classification test with 26 recordings in 2 raagas using different classifiers. The random baseline accuracy is 20% in this case.
Results (2/3)
Results (3/3)

Peak positions of (a). D2, (b). N2, (c). M1 and (d). P. The dashed line shows the mean of the corresponding swara obtained from the general template.
Back to the browser!

- As a similarity measure for raagas
  - Characteristics of common swaras
- Evolution of raagas
  - Composed sections
- As a similarity measure for artists & schools
  - Especially, the improvised sections
Swara Isolation [Work in progress]

Audio → Segmentation → Pitch Extraction → Tonic ID → Histogram Analysis → Peak Detection → Parametrization → Swara Isolation

Vocal Segments

Pitch contour

Histogram

Peak-labeled Histogram
Swara Isolation [Ideas]

- Why?
  - Discard the irrelevant/non-contextual pitch values
    - How do we discriminate??
  - Much clearer distributions
- Moving window & mean frequency
- Histogram per swara
  - Multiple peaks indicating the ‘contribution’ or ‘interaction’ of other swaras
Pattern Analysis [Work in progress]

R₂ in Mukhari

R₂ in Bhairavi
Pattern Analysis [Ideas]

- Why?
  - Patterns as atomic units of description
  - Similarity measures directly involving patterns
- Dictionary of patterns
  - All gamaka patterns on all swaras?
  - Just the characteristic gamakas?
  - Phrases instead of swaras (hierarchical)?
- Scale-invariant pattern matching techniques
Questions & Discussion

Ideas and brain-storming during tea-session are most welcome!!
Computational Analysis of Intonation in Indian Art Music

Gopala Krishna Koduri¹, Joan Serrà², Xavier Serra¹

¹Music Technology Group, Universitat Pompeu Fabra, Barcelona, Spain.
²Artificial Intelligence Research Institute (IIIA-CSIC), Bellaterra, Barcelona, Spain.

2nd CompMusic Workshop, Bahçeşehir Üniversitesi, Beşiktaş, İstanbul, Turkey