Computational Analysis of Intonation in Indian Art Music

Gopala Krishna Koduri[1], Joan Serrà[2], Xavier Serra[1]

[1]Music Technology Group, Universitat Pompeu Fabra, Barcelona, Spain.

[2]Artificial Intelligence Research Institute (IIIA-CSIC), Bellaterra, Barcelona, Spain.

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

Topics

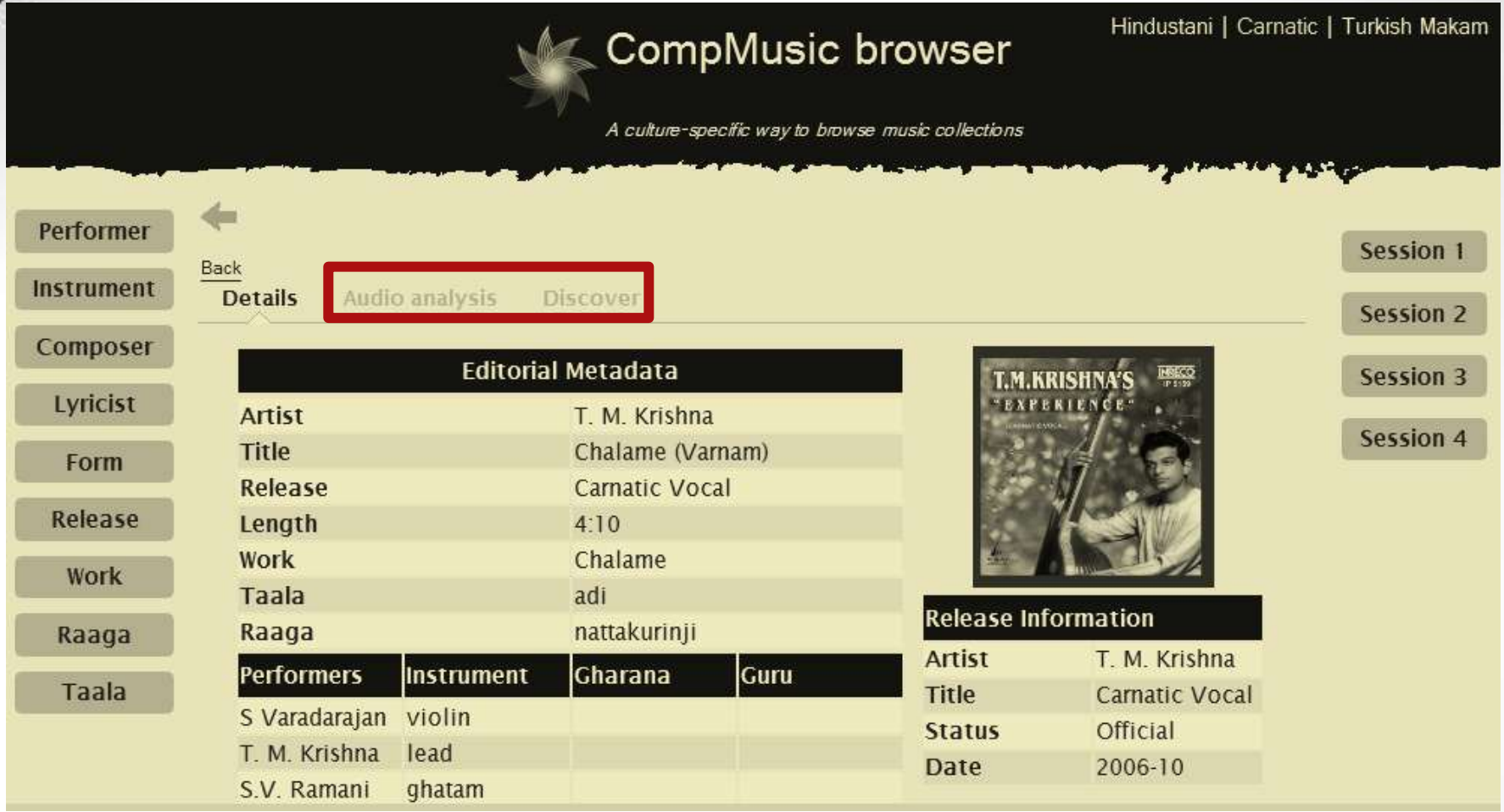
- What is intonation?
- Context and purpose
- Histogram parametrization
- Work in progress
  - Swara-based histograms
  - Pattern analysis
- Discussion & conclusions

compmusic

What is Intonation?

- Pitches used by a performer in a given performance.
- Our context: Pitch variations within a swara.
  - Eg: Ga_2 in Darbar and Nayaki 
- Relevant references
  - **Intonation**: Levy (1982), Belle et al (2009), Swathi et al (2009)
  - **Tuning**: Krishnaswamy (2003), J. Serrà et al (2011)

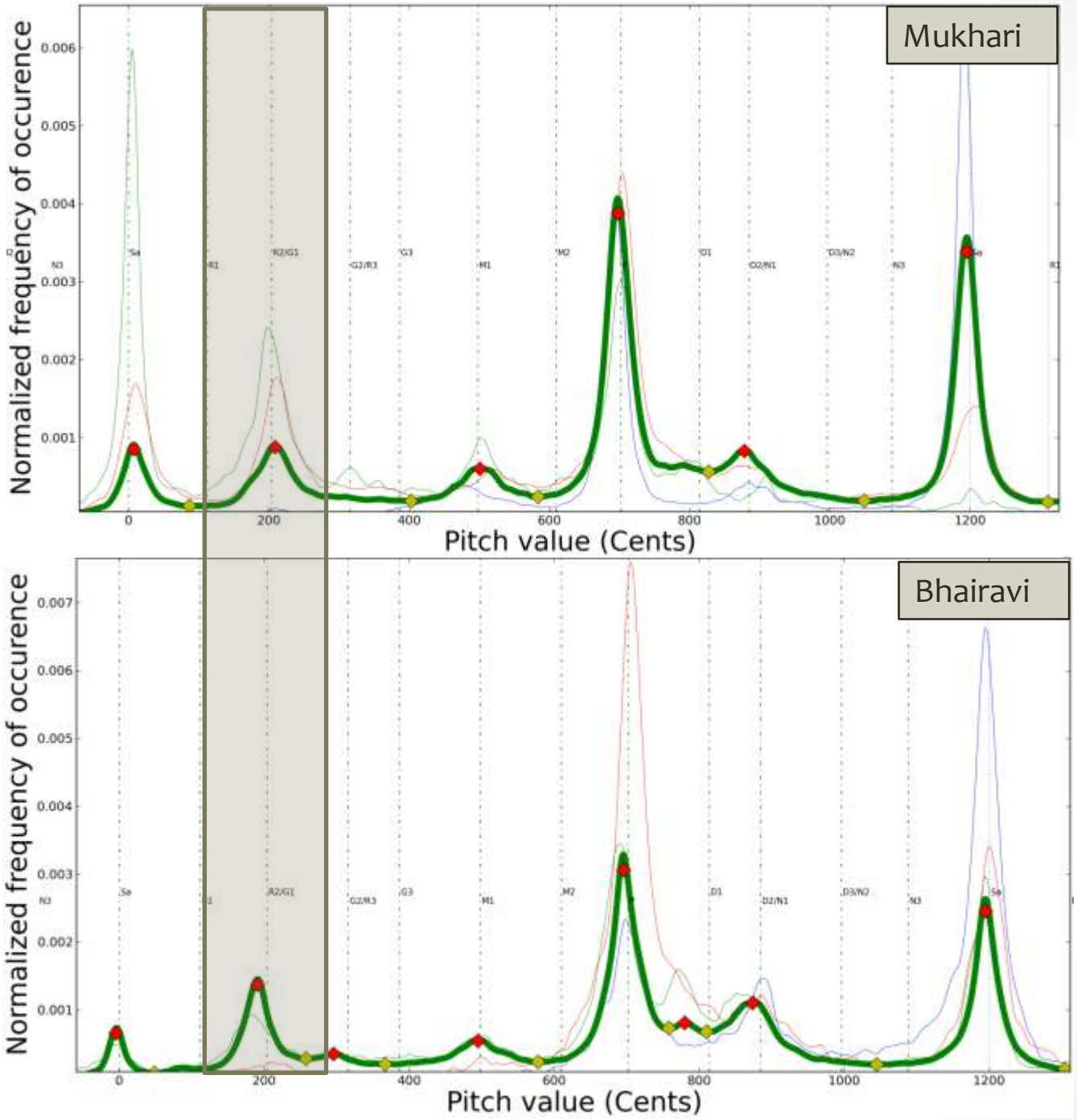
compmusic



- Melodic analysis
  - **Intonation profile,** motives, structure and low-level features.
- Rhythmic analysis
- Metadata
- Web-data

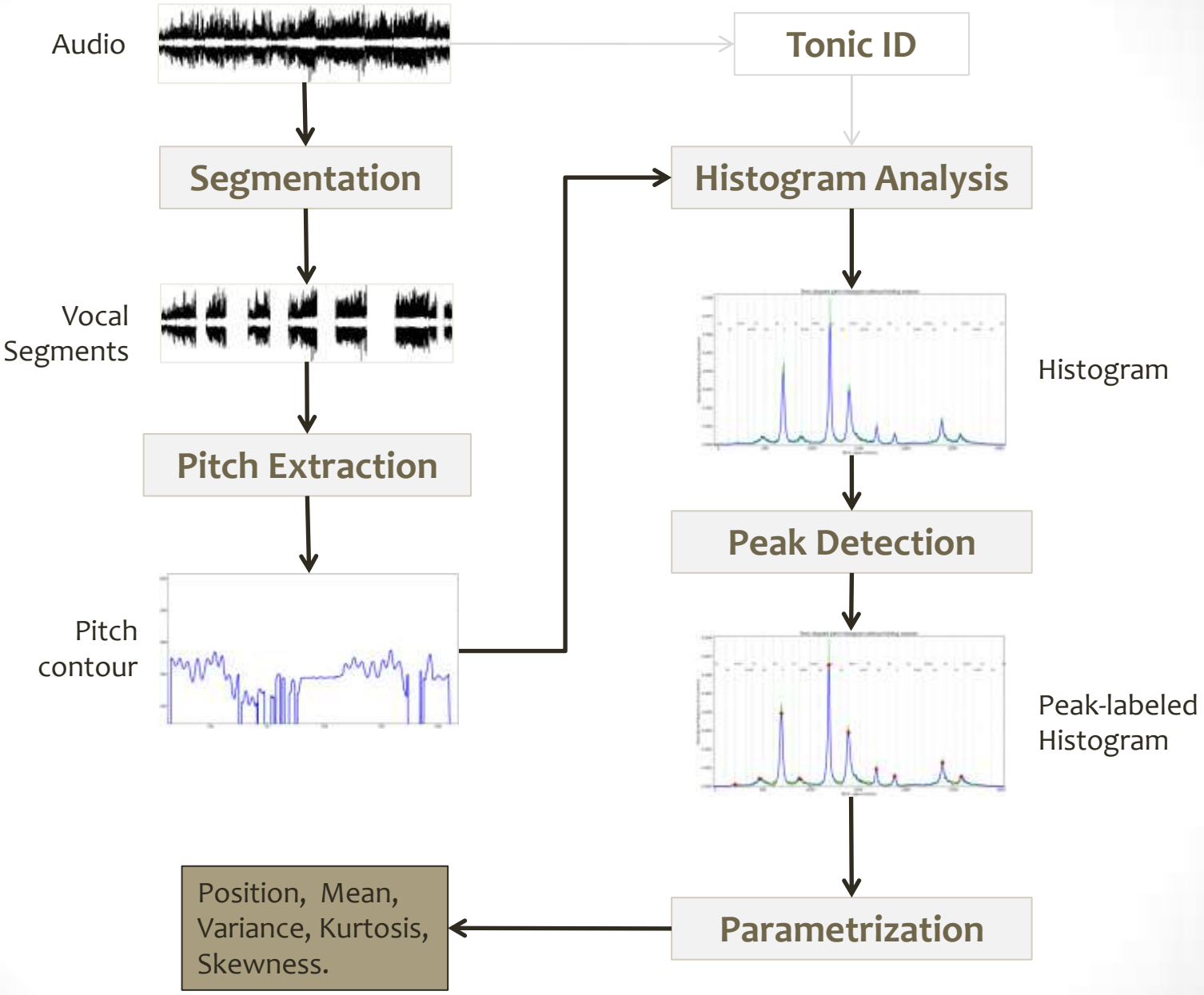
compmusic

Histogram analysis – Goal!



compmusic

Overview of Method

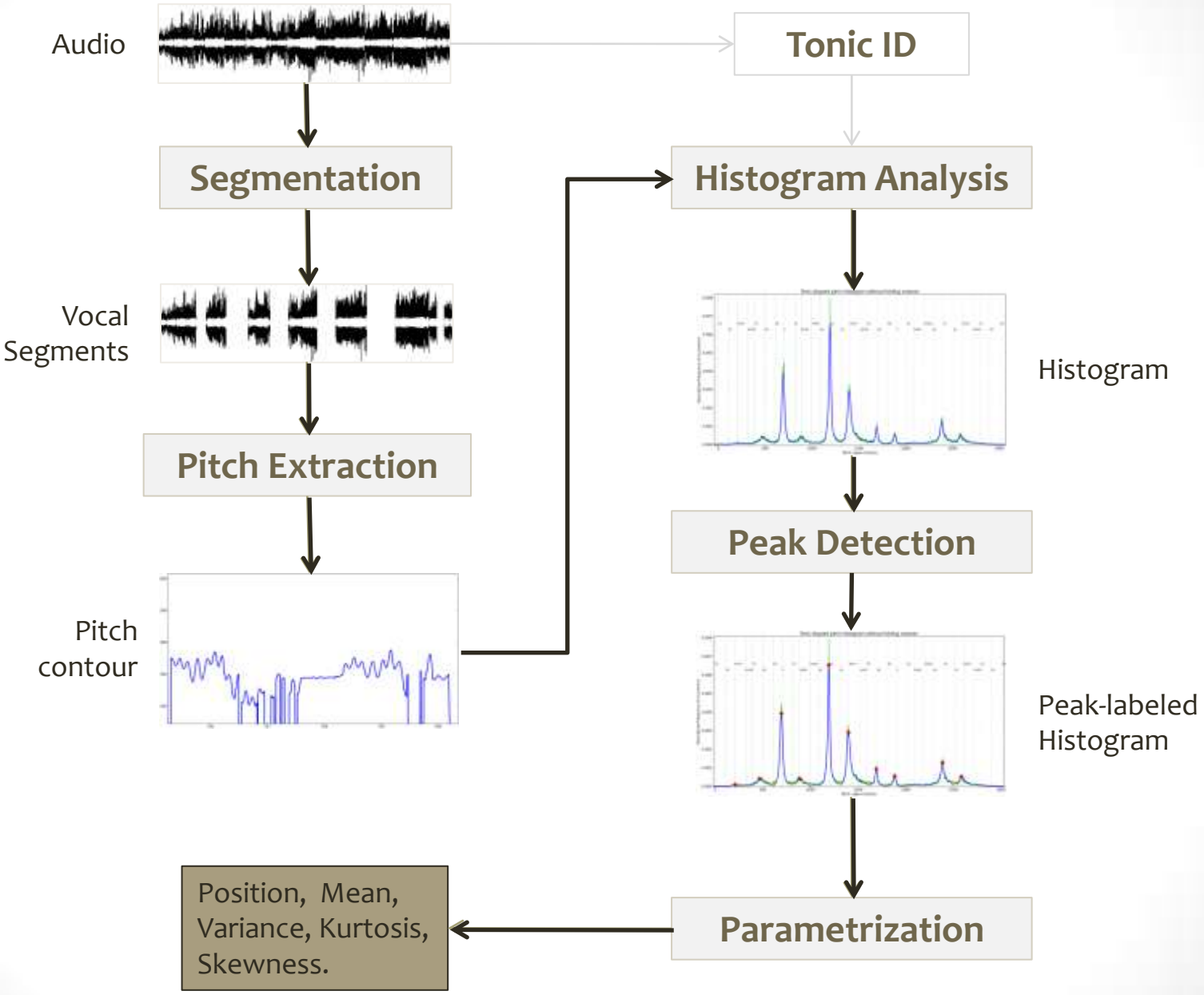


compmusic

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)

compmusic

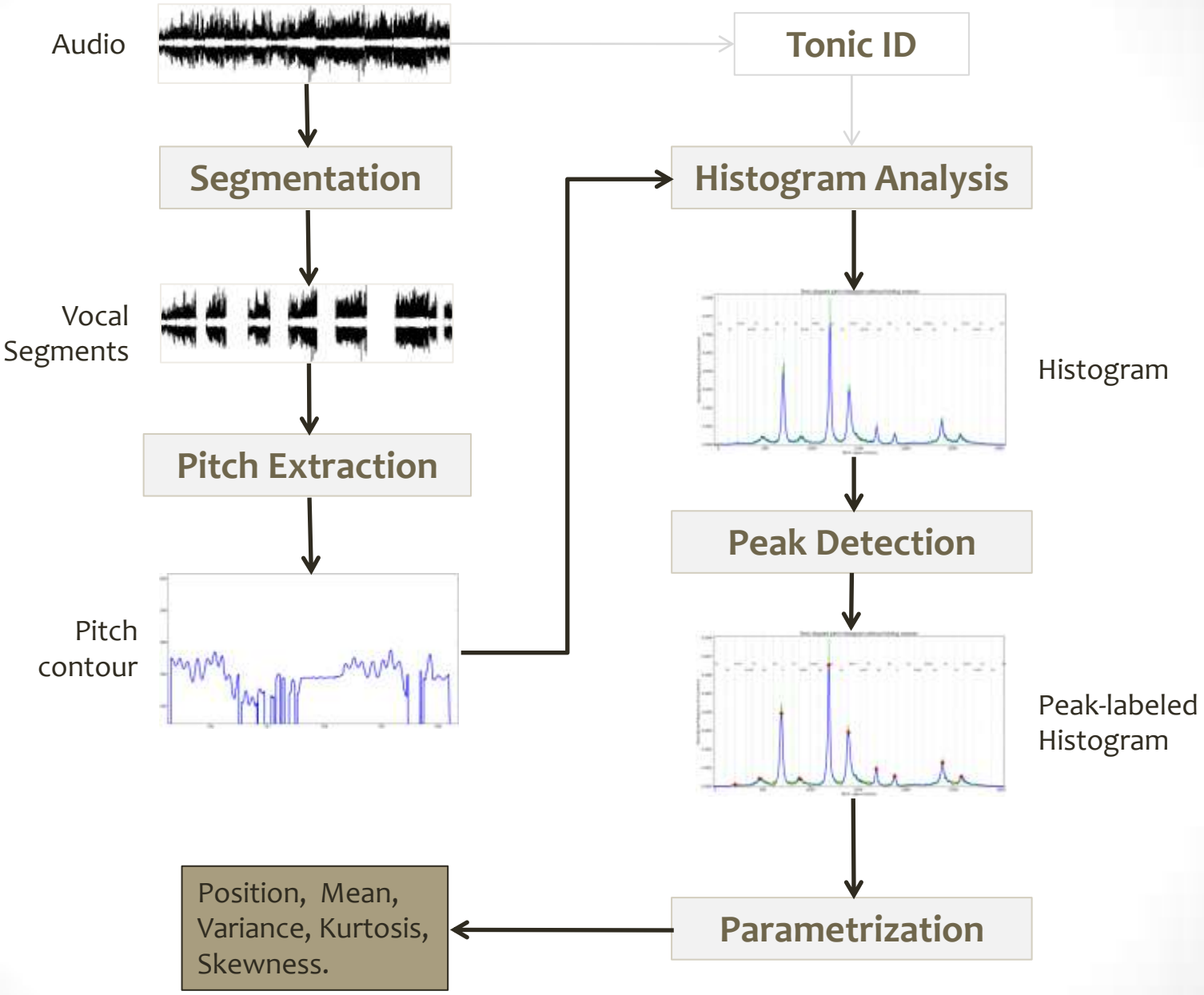

Audio
Tonic ID
Segmentation
Histogram Analysis
Vocal Segments
Histogram
Pitch Extraction
Peak Detection
Pitch contour
Peak-labeled Histogram
Position, Mean, Variance, Kurtosis, Skewness.
Parametrization

compmusic

Pitch Extraction

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

compmusic


Audio
Tonic ID
Segmentation
Histogram Analysis
Vocal Segments
Histogram
Pitch Extraction
Peak Detection
Pitch contour
Peak-labeled Histogram
Position, Mean, Variance, Kurtosis, Skewness.
Parametrization

compmusic

Histogram Analysis

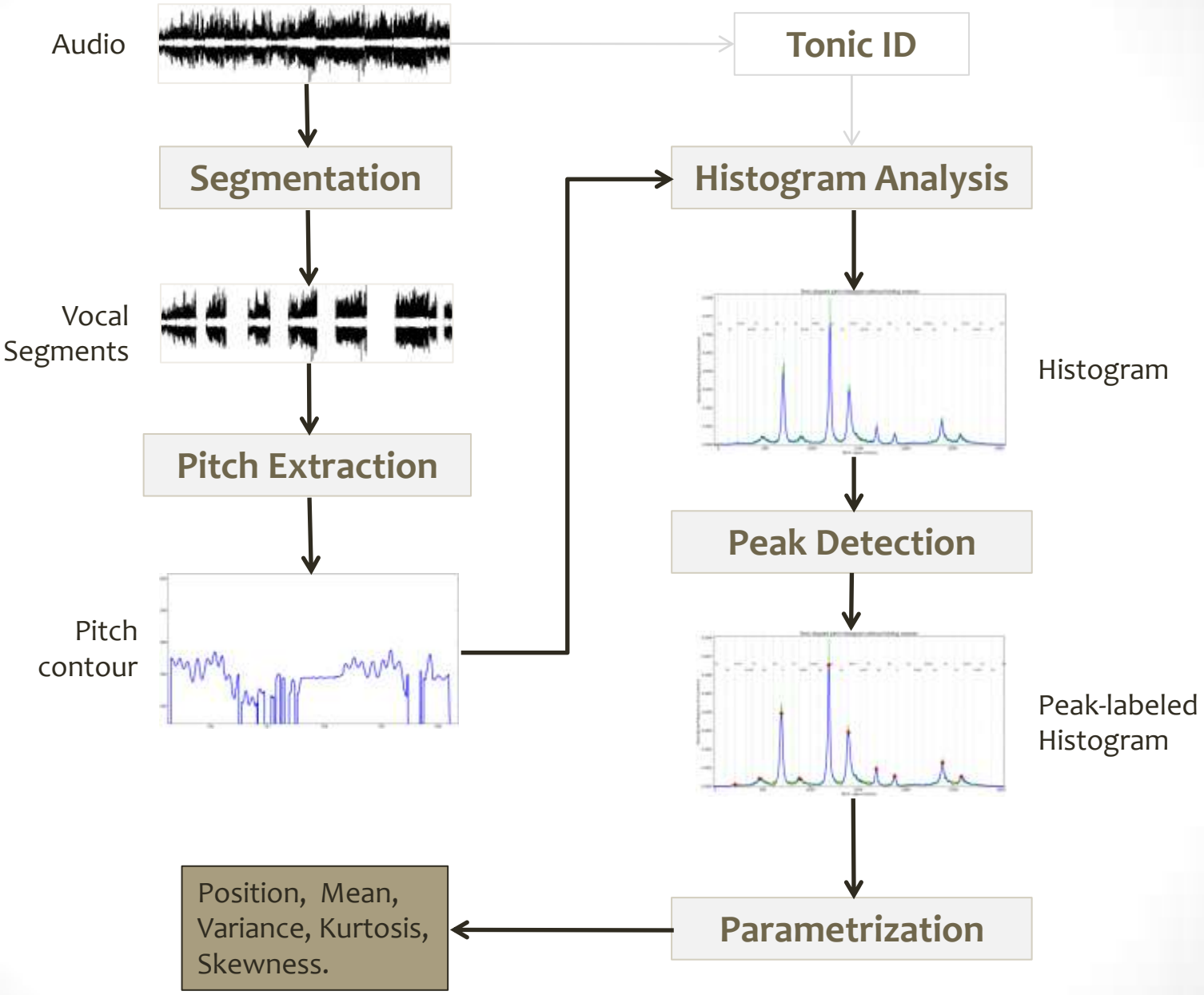
- … bin resolution!
- Histogram

$$H_k = \sum_{n=1}^{N} m_k$$

H_k is k^{th} bin count, N is fnumber 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

compmusic

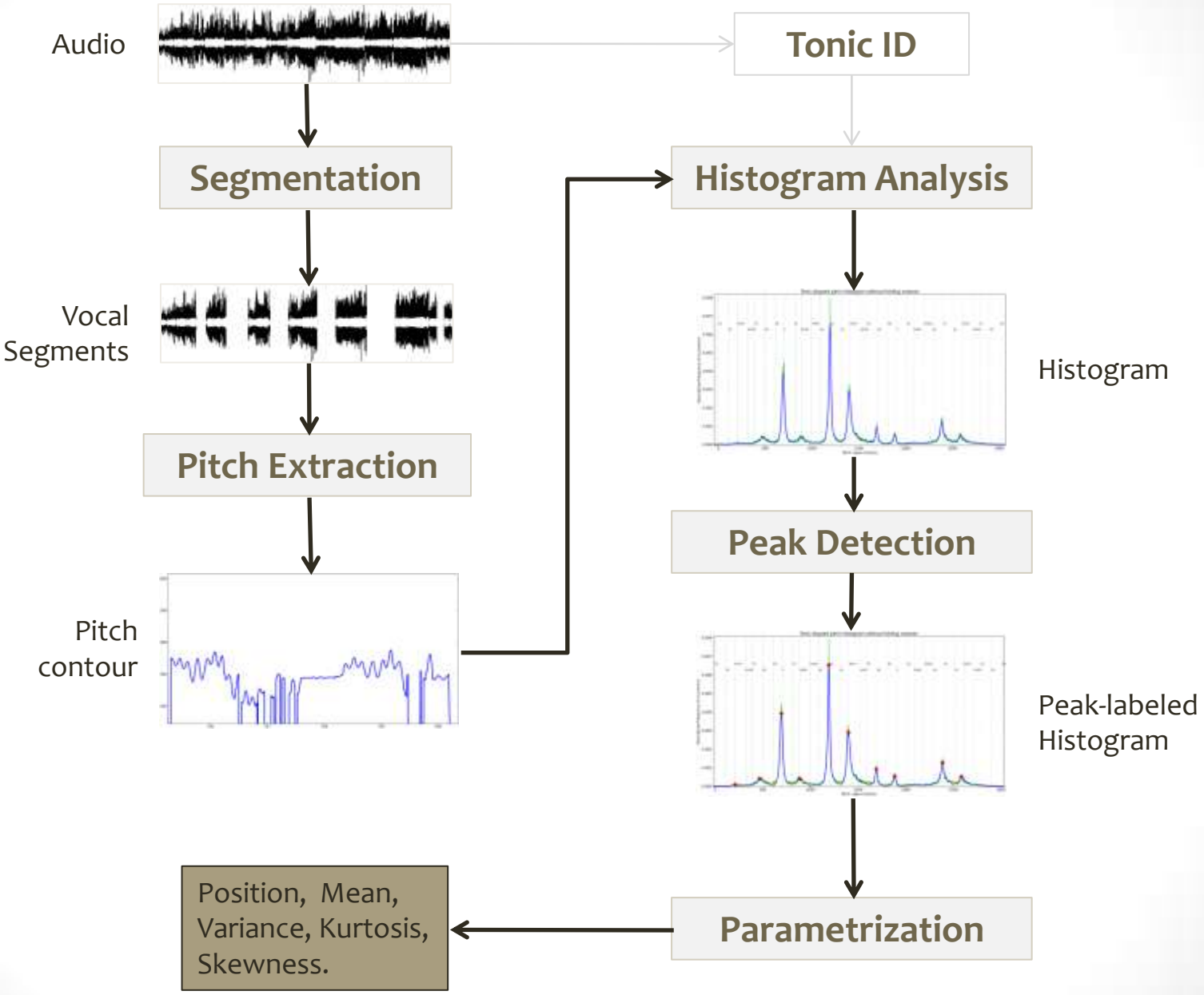

Audio
Tonic ID
Segmentation
Histogram Analysis
Vocal Segments
Histogram
Pitch Extraction
Peak Detection
Pitch contour
Peak-labeled Histogram
Position, Mean, Variance, Kurtosis, Skewness.
Parametrization

compmusic

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

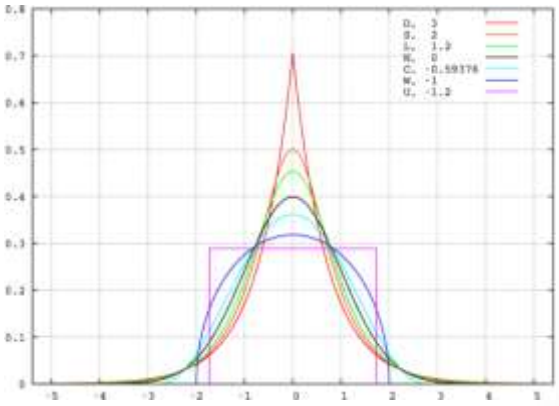
compmusic

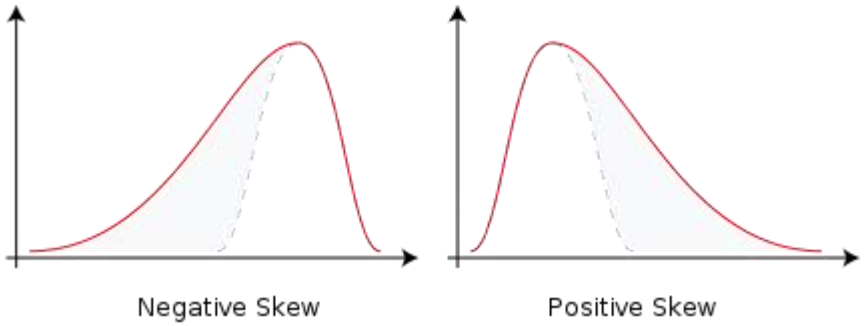

Audio
Tonic ID
Segmentation
Histogram Analysis
Vocal Segments
Histogram
Pitch Extraction
Peak Detection
Pitch contour
Peak-labeled Histogram
Position, Mean, Variance, Kurtosis, Skewness.
Parametrization

compmusic

Parametrization

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





compmusic

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!

compmusic

Results (1/3)

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

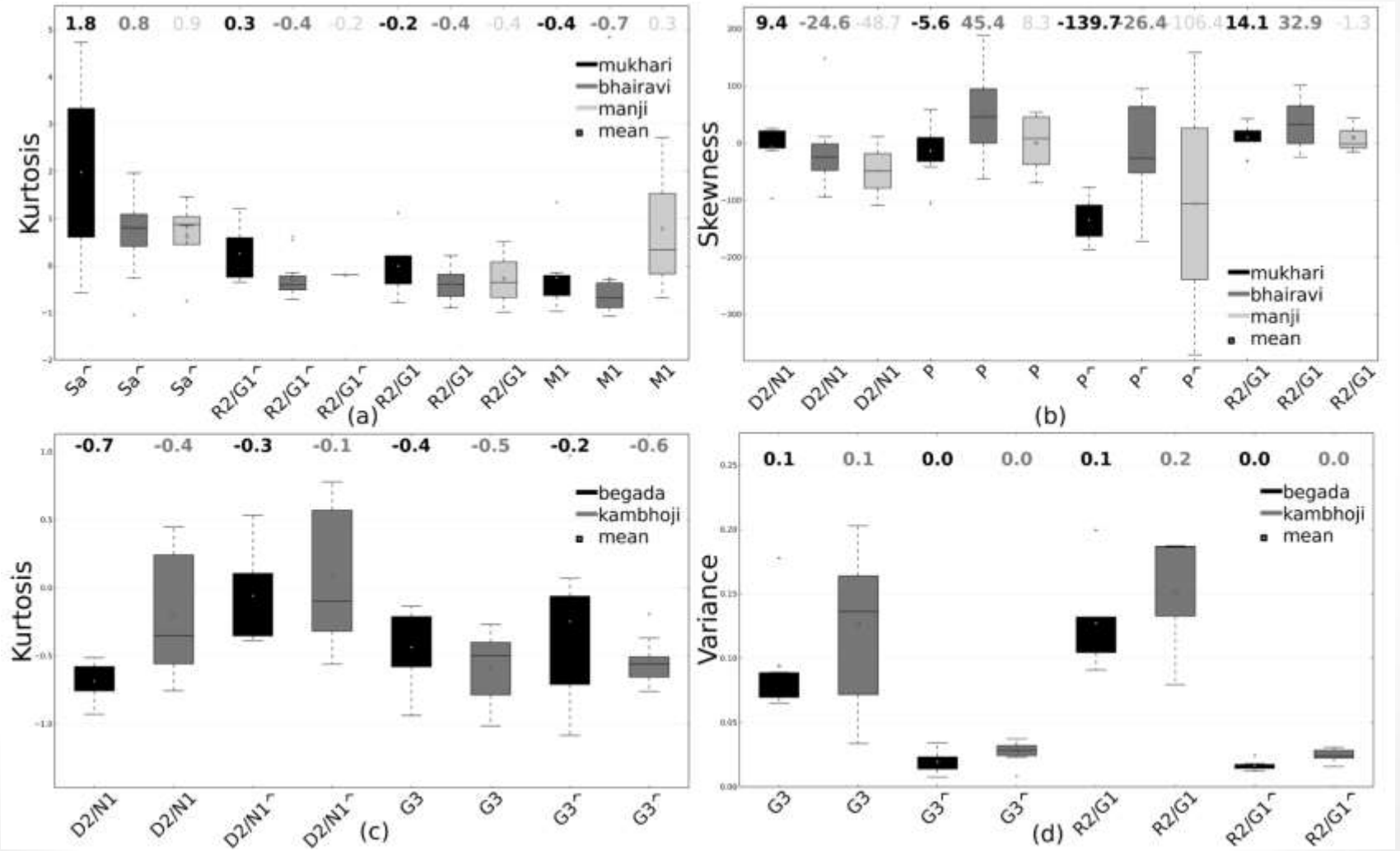
Table 1. Results of an exploration raaga classification test with 42 recordings in 3 raagas using different classifiers. The random baseline accuracy in this case is 28.57%.

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

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.

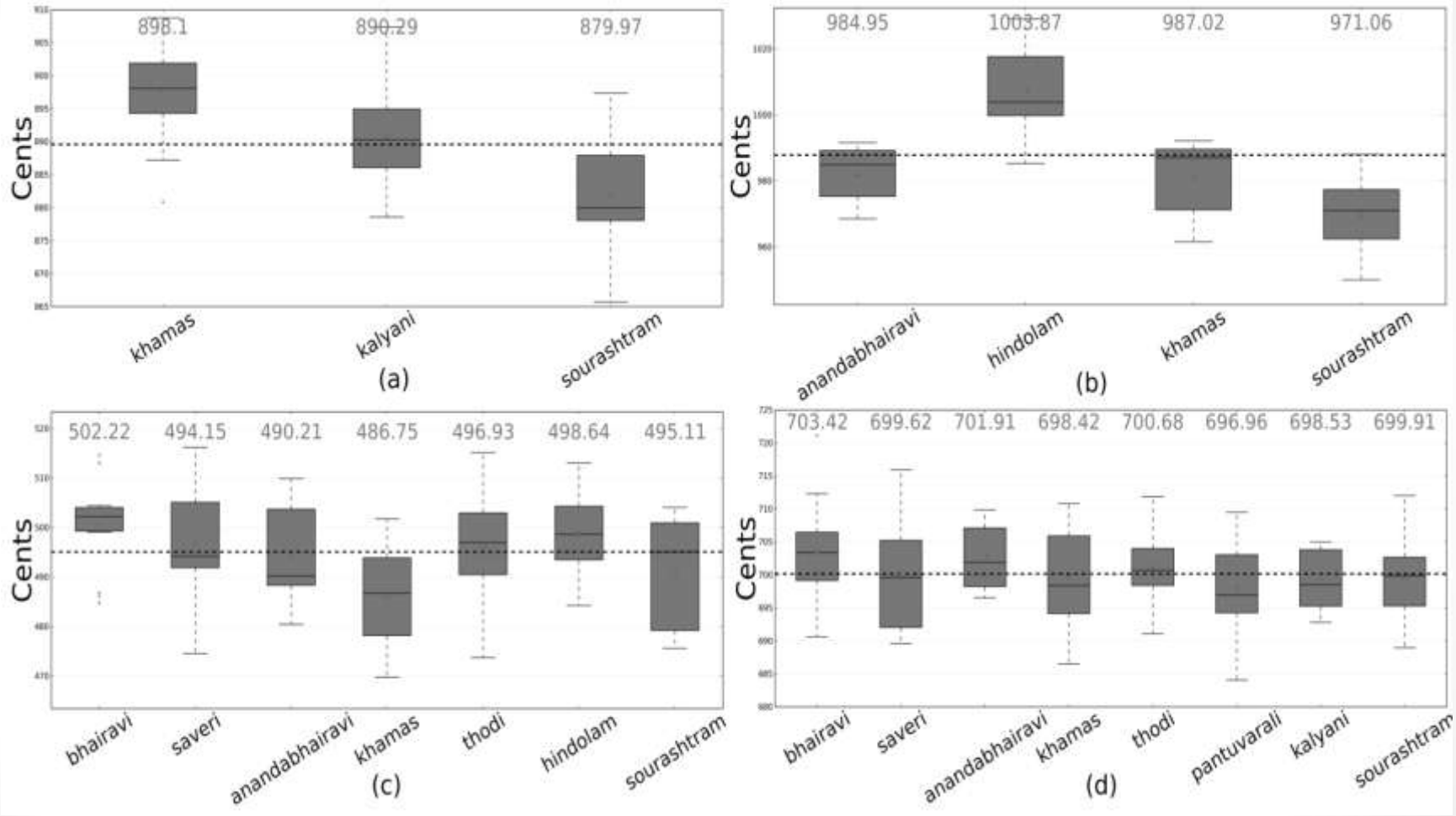
compmusic

Results (2/3)



compmusic

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.

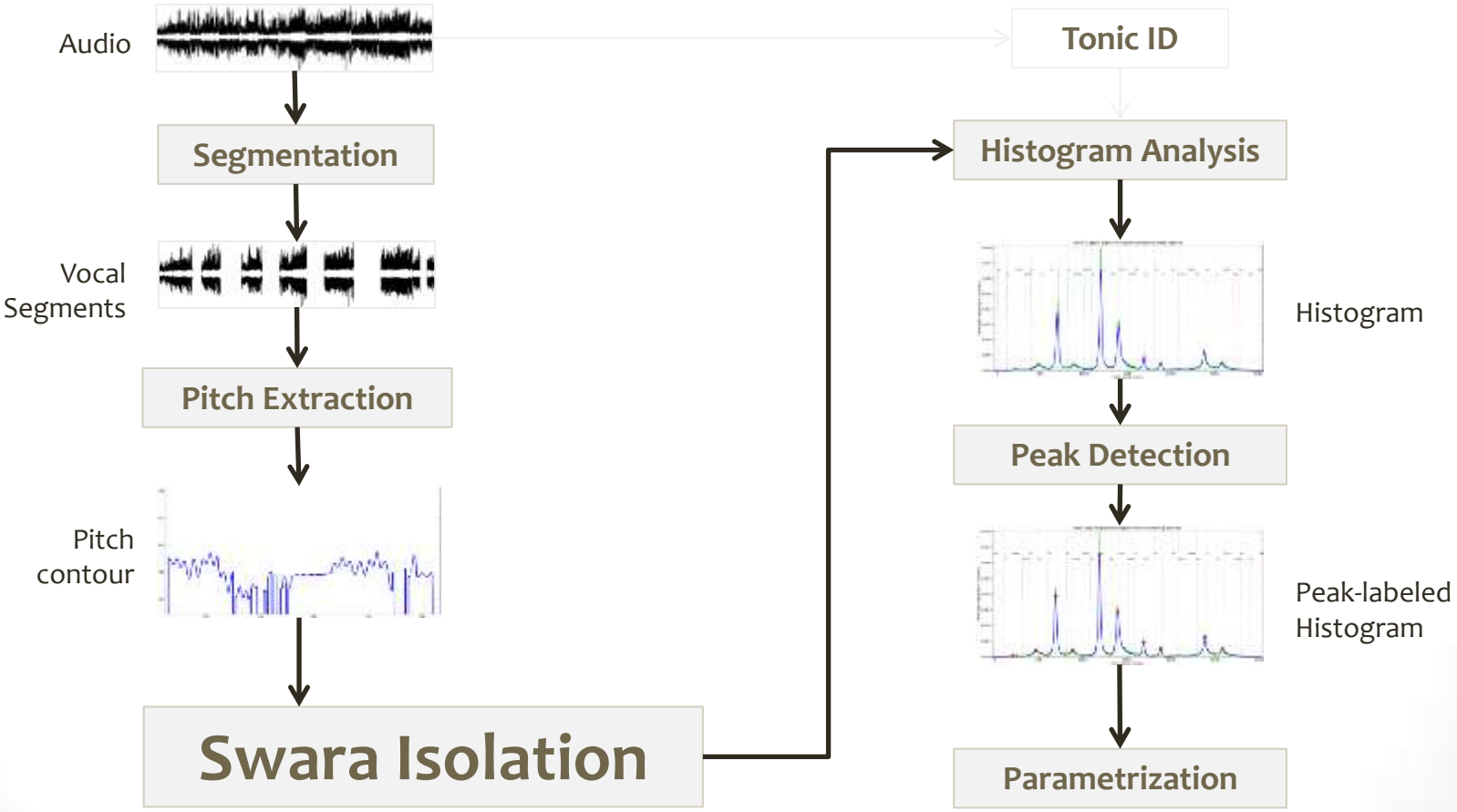
compmusic

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

compmusic

Swara Isolation [Work in progress]



compmusic

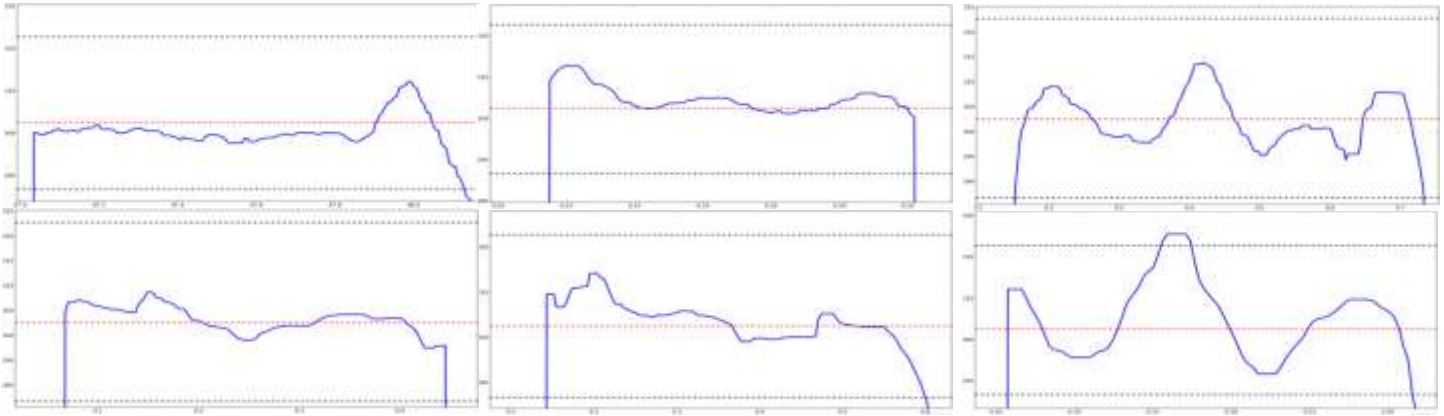
Swara Isolation [Ideas]

- 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

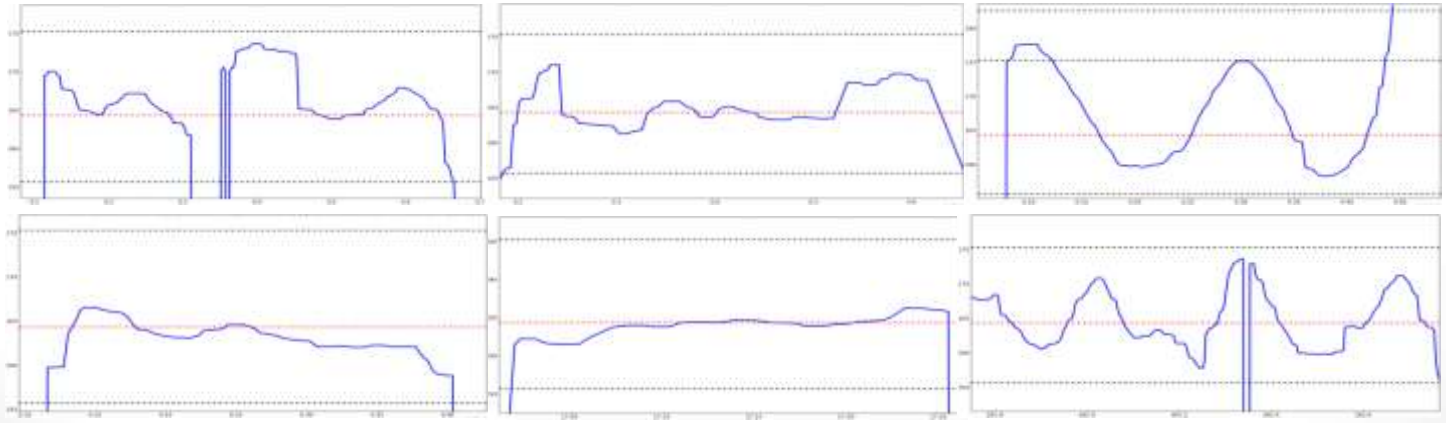
compmusic

Pattern Analysis [Work in progress]

R_2 in Mukhari



R_2 in Bhairavi



compmusic

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

compmusic

Questions & Discussion

Ideas and brain-storming during tea-session are most welcome!!

Computational Analysis of Intonation in Indian Art Music

Gopala Krishna Koduri[1], Joan Serrà[2], Xavier Serra[1]

[1]Music Technology Group, Universitat Pompeu Fabra, Barcelona, Spain.

[2]Artificial Intelligence Research Institute (IIIA-CSIC), Bellaterra, Barcelona, Spain.

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