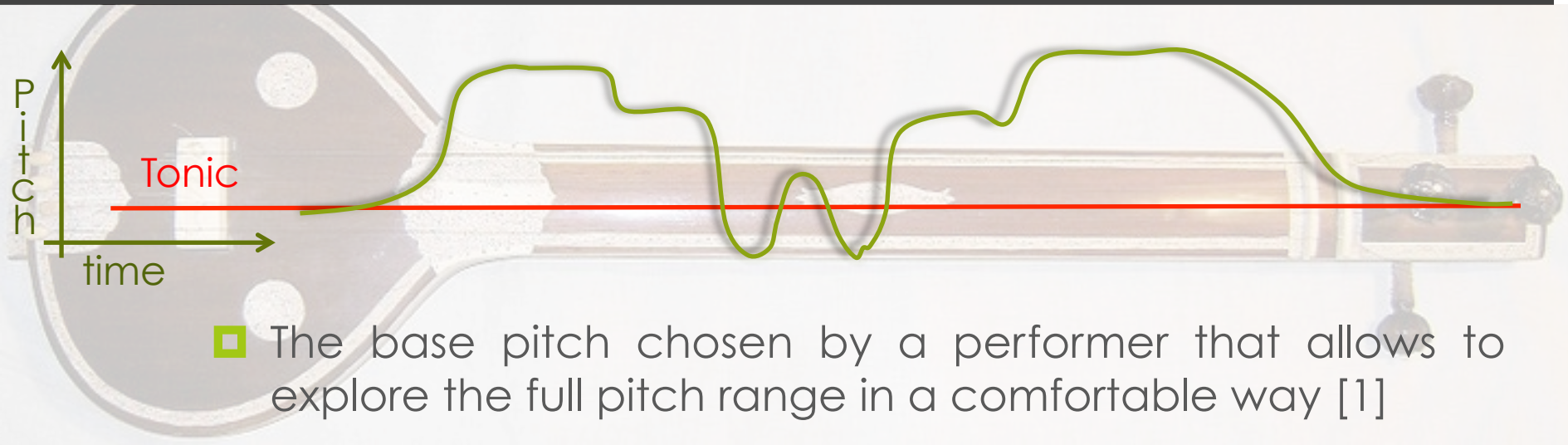
Tonic Identification System for Hindustani and Carnatic Music

Sankalp Gulati, Justin Salamon and Xavier Serra

Music Technology Group
Universitat Pompeu Fabra

{sankalp.gulati, justin.salamon, xavier.serra}@upf.edu

Introduction: Tonic in Indian art music



- The base pitch chosen by a performer that allows to explore the full pitch range in a comfortable way [1]
- Anchored as 'Sa' swar in a performance (mostly)
- All the other notes used in the raga exposition derive their meaning in relation to this pitch value
- All other accompanying instruments are tuned using this pitch as reference

Role of Drone Instrument

- Performer and audience needs to hear this pitch throughout the concert
- Reinforces the tonic and establishes all harmonic and melodic relationships

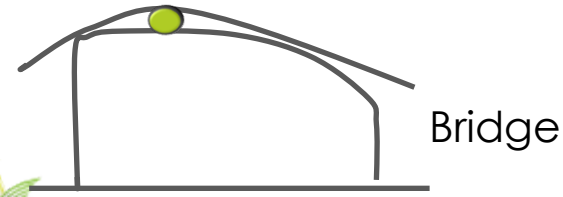
Sitar

Surpeti or Shrutibox

Electronic Tanpura

Tanpura

Introduction: Tonal structure of Tanpura



- Four strings
- Tunings
  - Sa-Sa'-Sa'-Pa
  - Sa-Sa'-Sa'-Ma
  - Sa-Sa'-Sa'-Ni
- Special bridge with thread inserted (Jvari)
  - Violate Helmholtz law [2]
- Rich overtones [1]

Introduction: Goals and Motivation

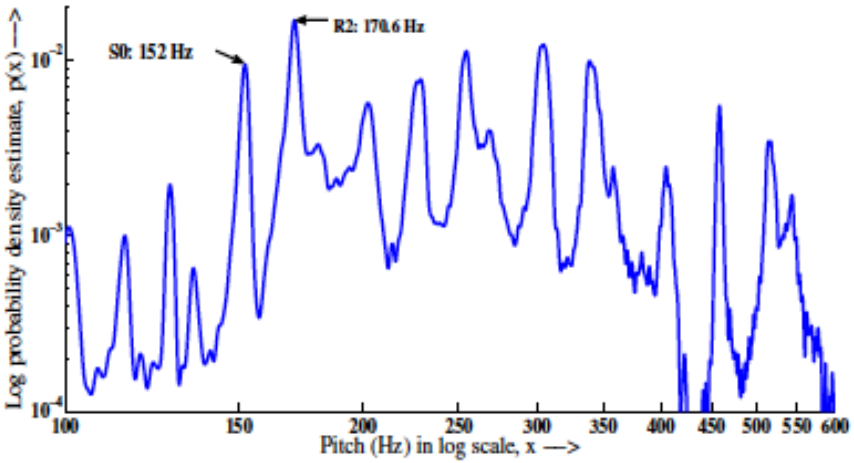
- Automatic labeling of the tonic in large databases of Indian art music
- Devise a system for identification of
  - Tonic pitch for vocal excerpts
  - Tonic pitch class profile for instrumental excerpts
- Use all the available data (audio + metadata) to achieve maximum accuracy
- Confidence measure for each output from the system

Introduction: Goals and Motivation

- Fundamental information
- Tonic identification: crucial input for:
  - Intonation analysis
  - Raga recognition
  - Melodic motivic analysis

Relevant work: Tonic Identification

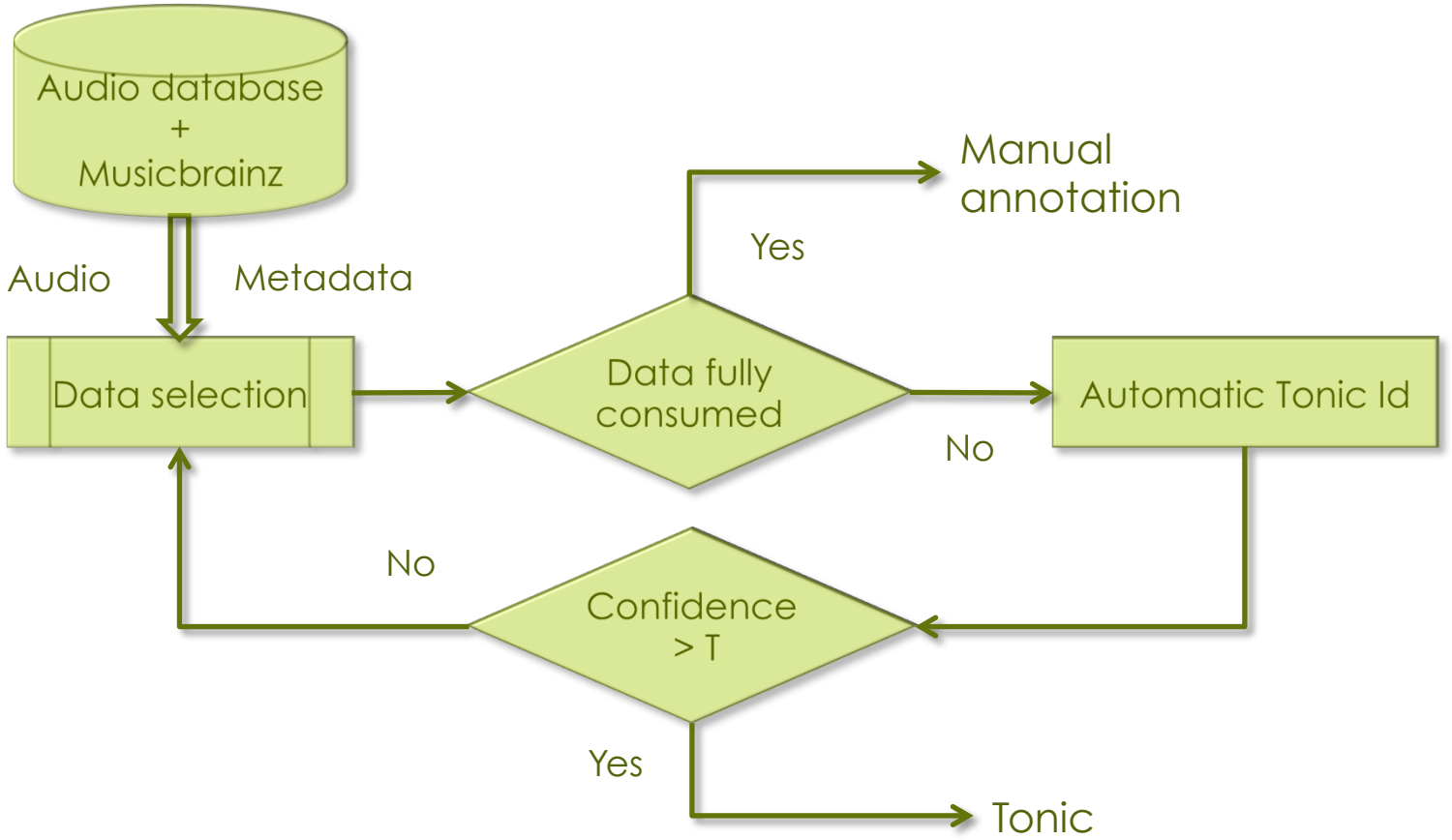
- Very little work done in the past
- Based on melody [4,5]
- Ranjani et al. take advantage of melodic characteristics of Carnatic music [4]



Relevant work: Summary

- Utilized only the melodic aspects
- Used monophonic pitch trackers for heterophonic data
- Limited diversity in database
  - Special raga categories, aalap sections, solo vocal recordings
- Unexplored aspects:
  - Utilizing background audio content comprising drone sound
  - Taking advantage of different types of available data, like audio and metadata
  - Evaluation on diverse database

Methodology: System Overview

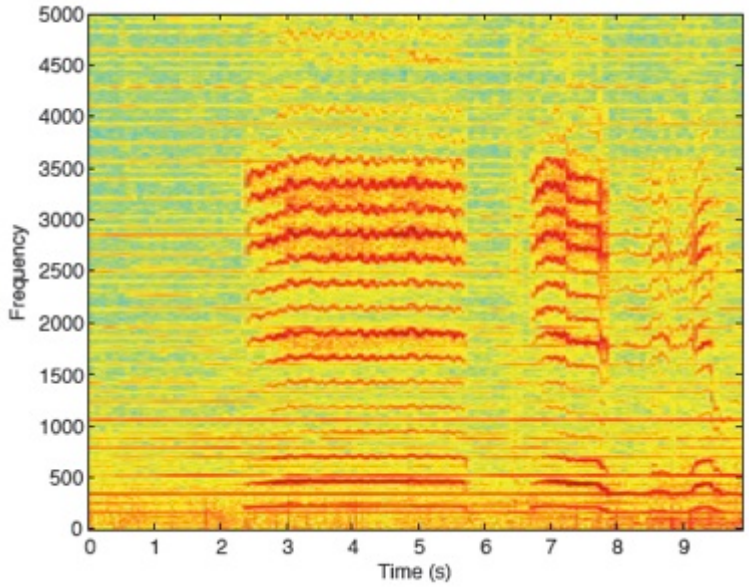


Methodology: System Overview

- Culture specific characteristics for tonic identification
  - Presence of drone*
  - Culture specific melodic characteristics
  - Raga knowledge
  - Melodic Motifs
- Use variable amount of data that is sufficient enough to identify tonic with maximum confidence.
  - Audio data
  - Metadata (Male/Female, Hindustani/Carnatic, Raga etc.)

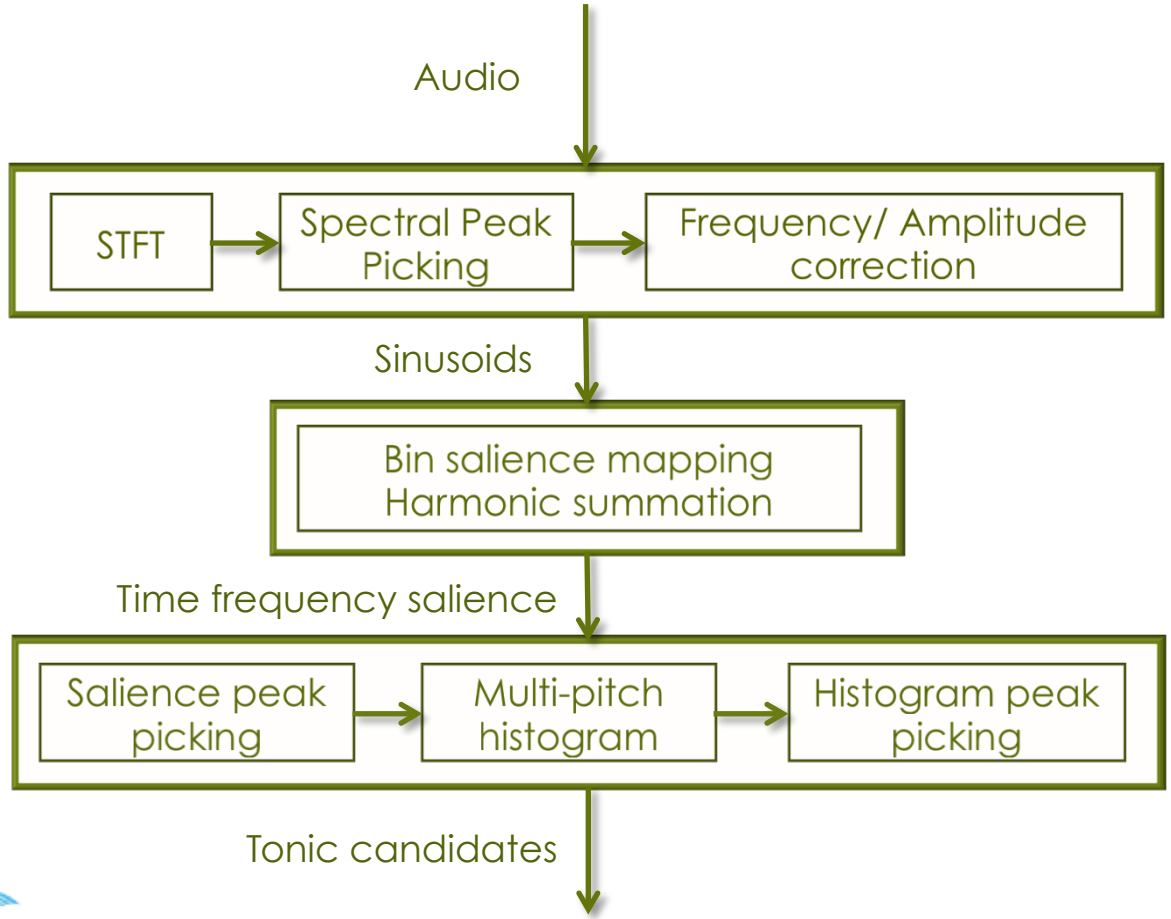
Methodology: Tonic Identification

- Audio example:
- Utilizing drone sound
- Chroma or multi-pitch analysis



Multi-pitch Analysis [7]

Tonic Identification: Signal Processing



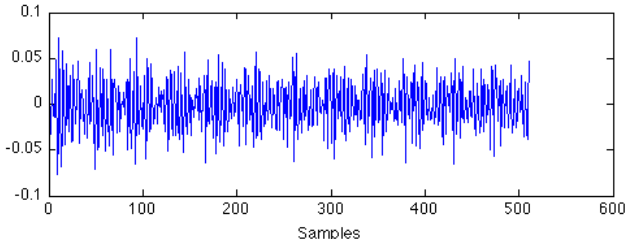
Sinusoid Extraction

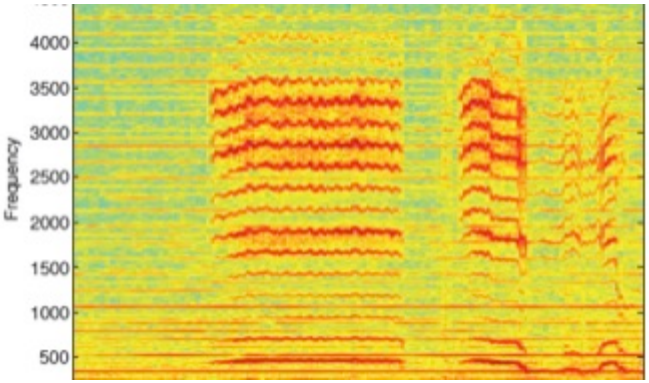
Pitch Salience computation

Tonic candidate generation

Tonic Identification: Signal Processing

STFT

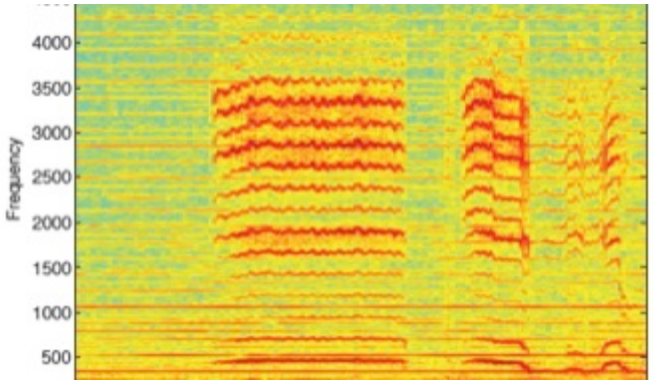


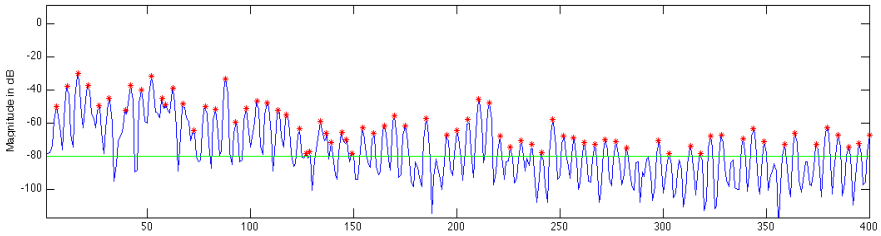


- STFT
  - Hop size: 11 ms
  - Window length: 46 ms
  - Window type: hamming
  - FFT = 8192 points

Tonic Identification: Signal Processing

Spectral Peak Picking

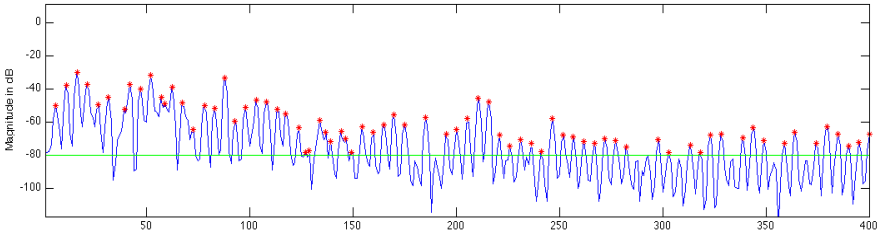




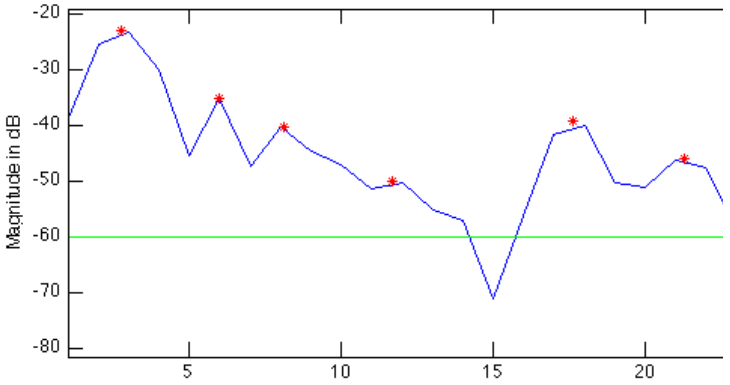
- Spectral peak picking
  - Absolute threshold: -60 dB
  - Relative threshold: -40 dB

Tonic Identification: Signal Processing

Frequency/ Amplitude correction

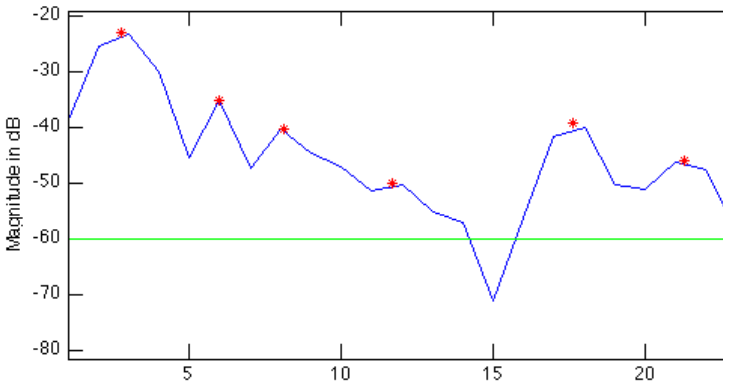


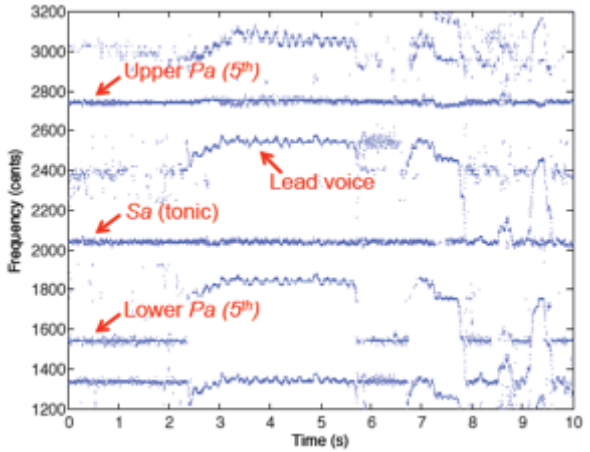
- Frequency/Amplitude correction
  - Parabolic interpolation



Tonic Identification: Signal Processing

Bin salience mapping
Harmonic summation

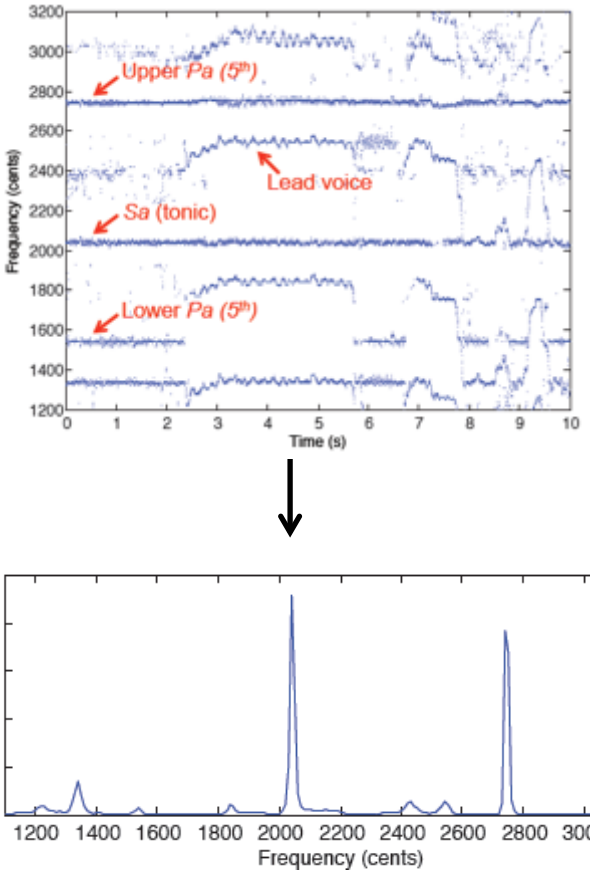




- Harmonic summation [7]
  - Spectrum considered: 55-7200 Hz
  - Frequency range: 55-1760 Hz
  - Base frequency: 55 Hz
  - Bin resolution: 10 cents per bin (120 per octave)
  - N octaves: 5
  - Maximum harmonics: 20
  - Alpha: 1
  - Beta: 0.8
  - Square cosine window across 50 cents

Tonic Identification: Signal Processing

Multi-pitch histogram



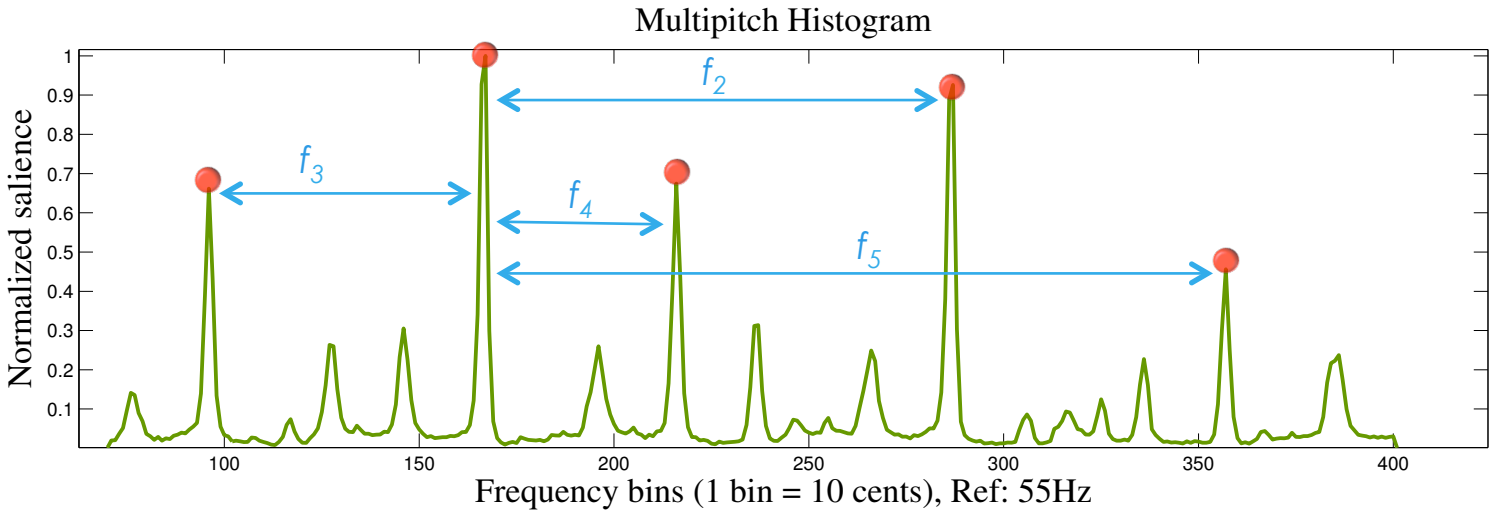
- Tonic candidate generation
  - Number of salience peaks per frame: 5
  - Frequency range: 110-550 Hz
  - After candidate selection salience is no longer considered!!!!

Tonic Identification : Two sub-tasks

- Caters to both vocal and instrumental excerpts
  - Identify tonic pitch class (PC) using multi-pitch histogram
  - Estimate the correct octave using predominant melody
- Use predominant melody extraction approach proposed by Justin Salamon et al. [6]
- Tonic PCP
  - Peak Picking + Machine learning
- Tonic octave estimation
  - Rule based method + Classification based approach

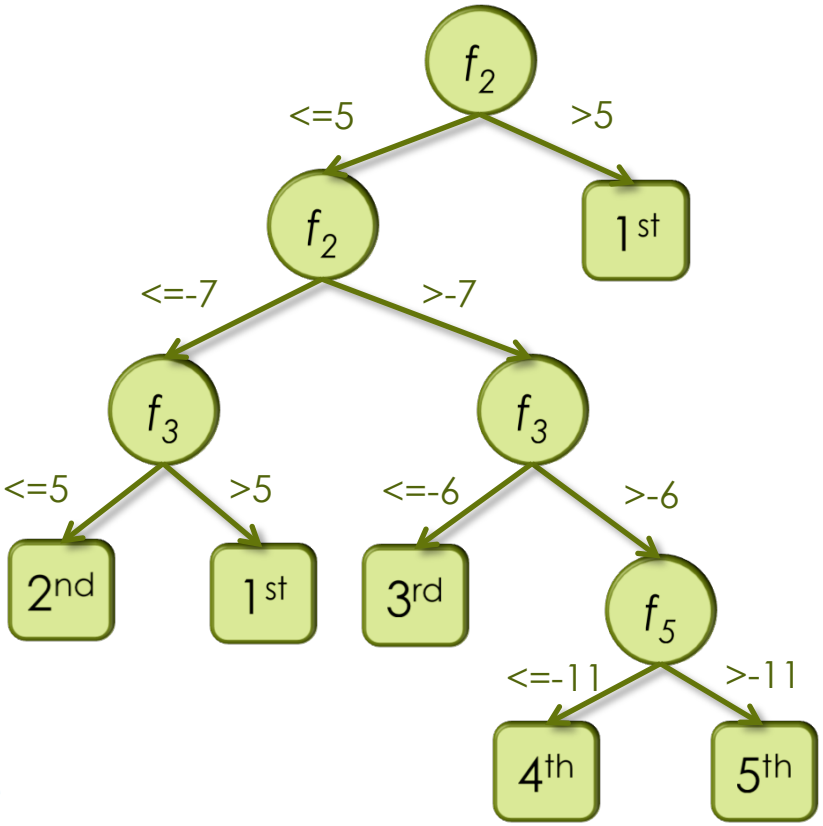
Tonic Identification : PC identification

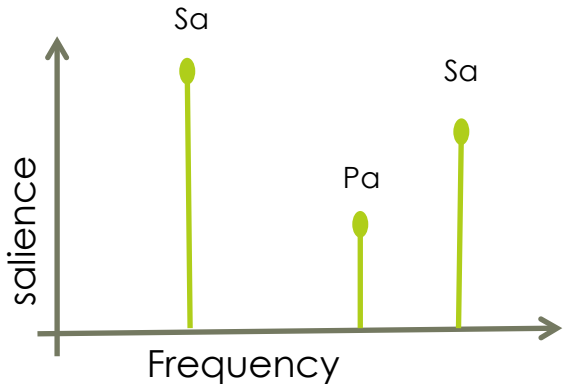
- Classification based template learning
- Two kind of class mappings
  - Rank of the highest tonic PC
  - Highest peak as Tonic or Non tonic
- Feature extracted # 20 (f_1-f_{10}, a_1-a_{10})

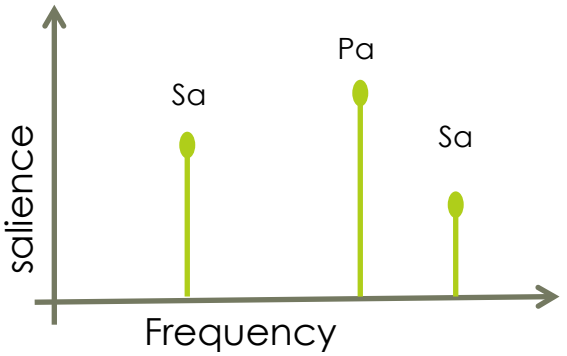


Tonic Identification : PC identification

- Decision Tree:

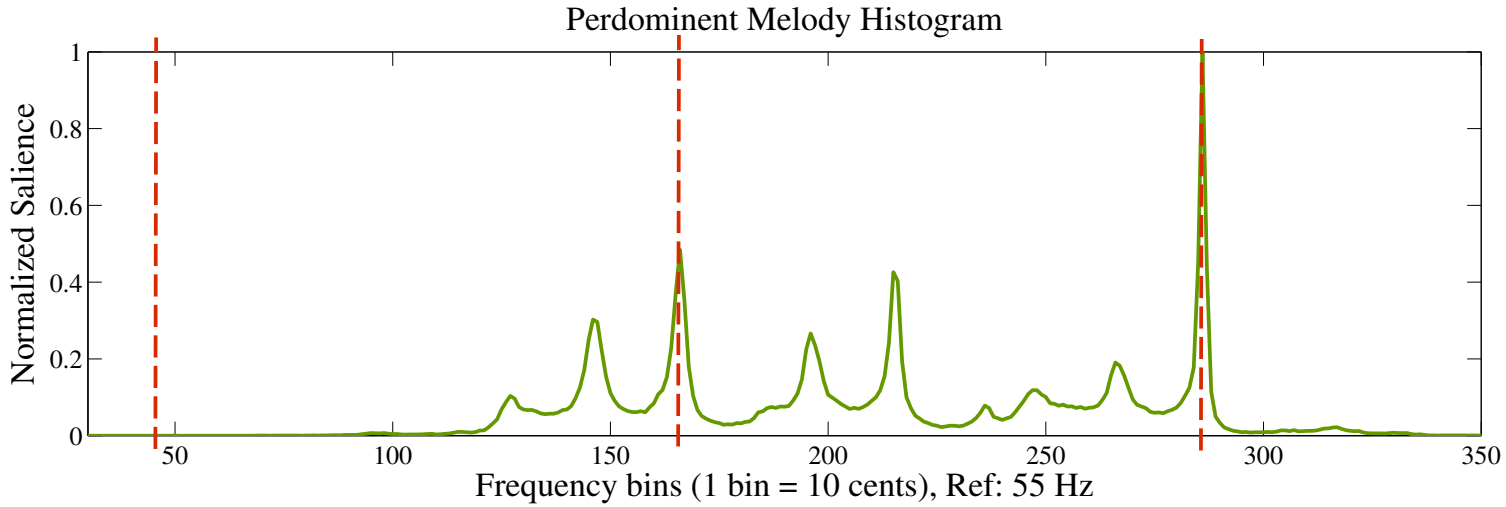






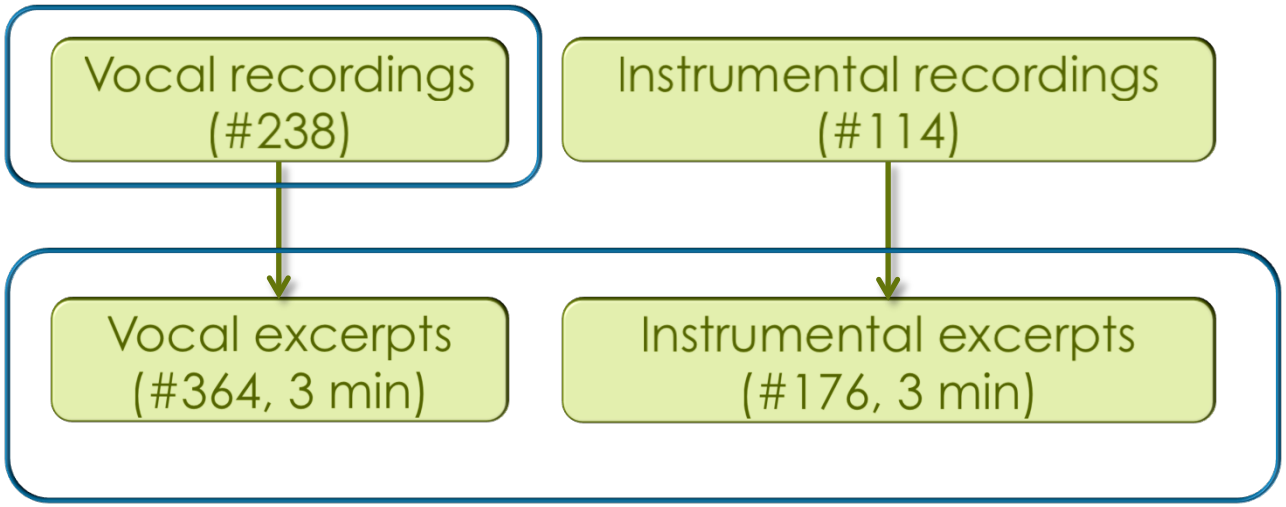
Tonic Identification : Octave Identification

- Tonic octave
  - Rule based method
  - Classification based approach
- 25 Features: a_1-a_{25}



Evaluation: Database

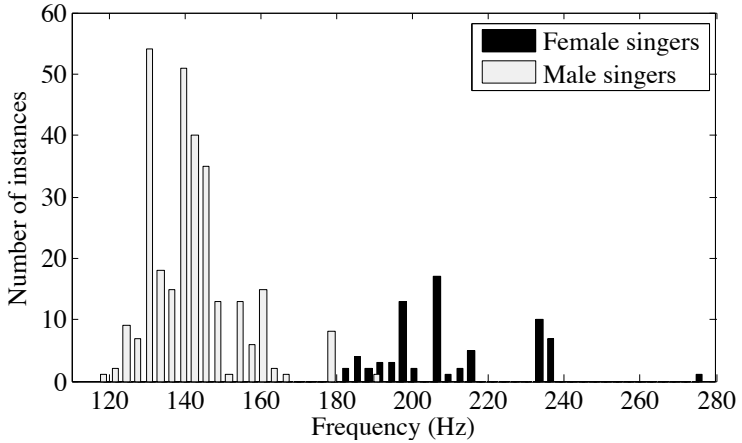
- Subset of CompMusic database (>300 Cds) [3]



Approach 2: #540, 3min (PCP) + 238, full recordings (Octave)

Evaluation: Database

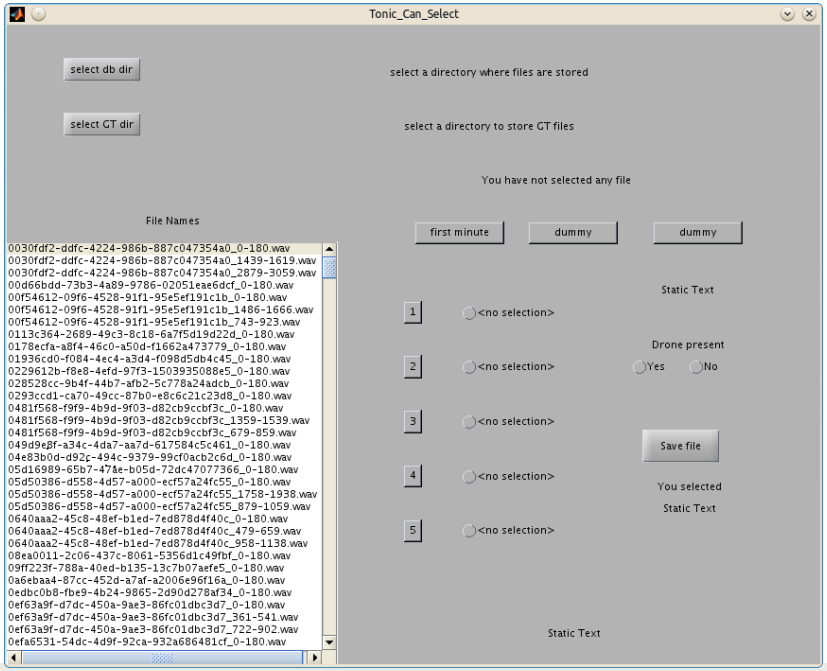
- Tonic distribution



- Statistics (for 364 vocal excerpts)
  - Male (80 %), Female (20%), Hindustani (38%), Carnatic (62%), Unique artist (#36)
- Statistics (for 540 vocal and instrumental excerpts)
  - Hindustani (36%), Carnatic (64%), Unique artist (#55)

Evaluation: Annotations

- Annotations done by the author
- Extracted 5 tonic candidates from multi-pitch histograms between 110-370 Hz
- Matlab GUI to speed up the annotation procedure



Evaluation: Accuracy measures

- Output correct within 50 cents of the ground truth
- 10 fold cross validation + rule based classification
- Weka: data mining tool
- Feature selection: CfsSubsetEval (features > 80% folds)
- Classifier: J48 decision tree
- Performs better than
  - SVM-polynomial kernel (6% difference in accuracy)
  - K* classifier (5% difference in accuracy)

Results

Approach\(%)	Map	#folds	Class EQ	# Features	Tonic pitch	Tonic PCP	5th	4th	Other
AP1_EXP1	-	-	-	-	-	85	10.7	0.93	3.3
AP1_EXP2	M1	1	no	1, S2	-	**93.7**	1.48	8.9	0.9
AP1_EXP3	M1	10	no	4, S3	-	92.9	1.9	3.5	1.7
AP1_EXP4	M1	10	yes	4, S4	-	74.2	11	7.6	6.7
AP1_EXP5	M2	1	no	1, S2	-	91	3.3	3	2.7
AP1_EXP6	M2	10	no	2, S5	-	91.8	2.2	3	3
AP1_EXP7	M2	10	yes	2, S5	-	87.8	4.2	4	3.9

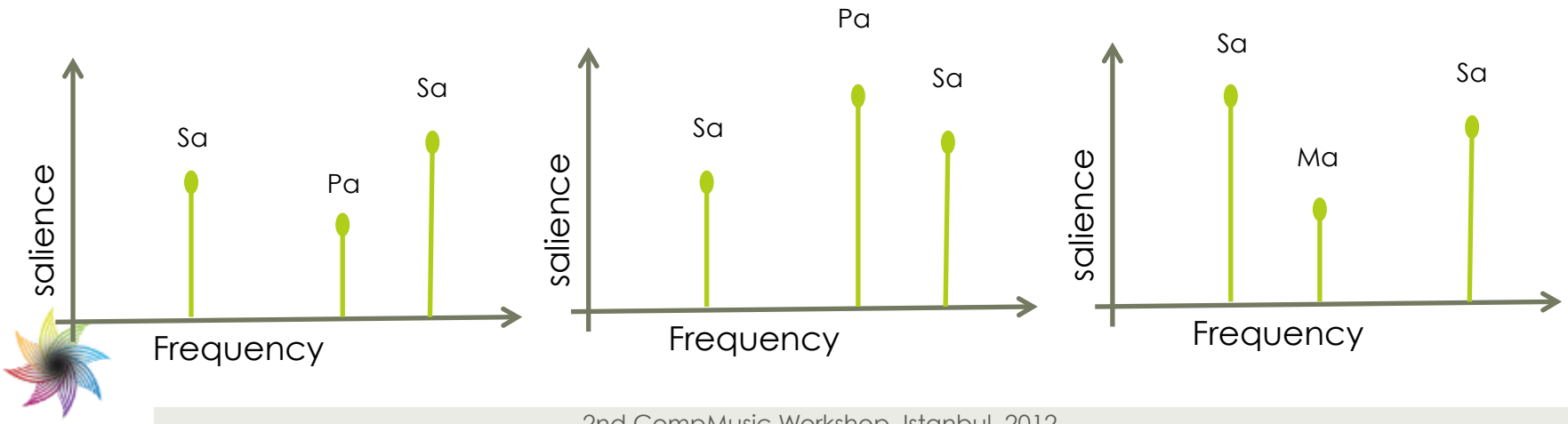
- M1 : tonic PCP rank, M2 : highest peak tonic or non-tonic
- S1: $[f_2, f_3, f_5]$, S2: $[f_2]$, S3: $[f_2, f_4, f_6, a_5]$, S4: $[f_2, f_3, a_3, a_5]$, S5: $[f_2, f_3]$

Results

- Approach 2, Octave identification
  - Rule based approach – 99 %
  - Classification based approach – 100%

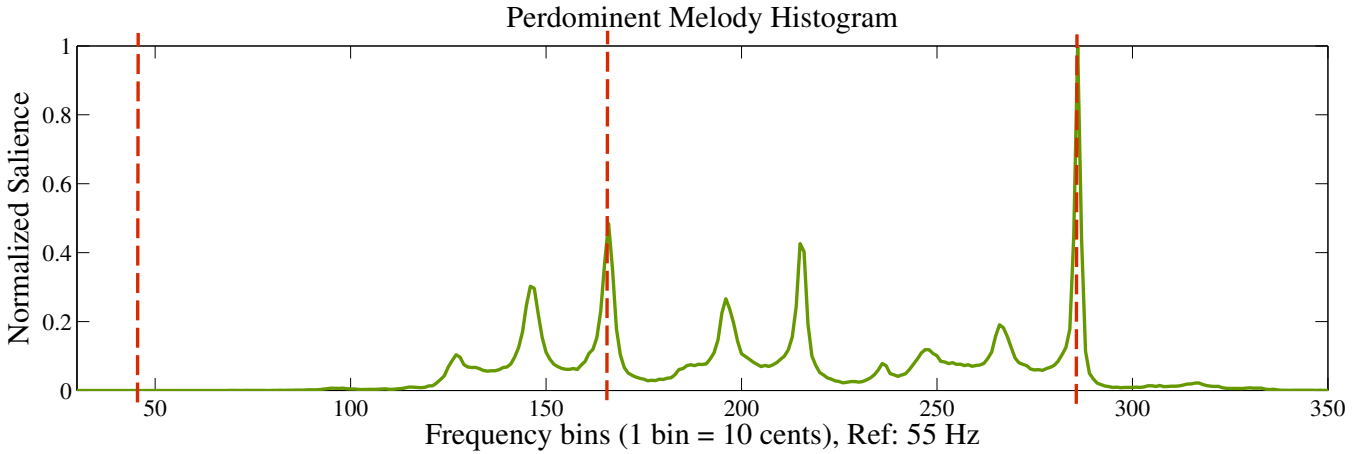
Discussion: PCP Identification

- AP-1: Performance for male singers (95%), female singers (88%)
- Error cases
  - Mostly Ma tuning songs
  - More female singers
- Sensitive to selected frequency range for tonic candidates, a range of 110-370 Hz works optimal



Discussion : Octave Identification

- Challenges faced by rule based approach
  - Hindustani musicians go roughly -500 cents below tonic
  - Carnatic musicians generally don't go that below tonic
  - Melody estimation errors at low frequency
  - Concept of Madhyam shruti



Conclusions and Future Work

- Drone sound in the background provides an important cue for the identification of tonic and can be utilized to automatically perform this task
- System should be fed with more information to differentiate between '*Pa*' and '*Ma*' tuning
- Future Work:
  - Exploring melodic characteristics for tonic identification
  - Deeper analysis of confidence measure concept
  - Study influence of cultural background on human performance for this task

REFERENCES

1. B. C. Deva. The Music of India: A Scientific Study. Munshiram Manoharlal Publishers, Delhi, 1980.
2. C. V. Raman. On some Indian stringed instruments. In Indian Association for the Cultivation of Science, volume 33, pages 29-33, 1921.
3. X. Serra. A multicultural approach in music information research. In 12th Int. Soc. for Music Info. Retrieval Conf., Miami, USA, Oct. 2011.
4. R. Sengupta, N. Dey, D. Nag, A. K. Datta, and A. Mukerjee, "Automatic Tonic (SA) Detection Algorithm in Indian Classical Vocal Music," in *National Symposium on Acoustics*, 2005, pp. 1-5.
5. T. V. Ranjani, H.G.; Arthi, S.; Sreenivas, "Carnatic music analysis: Shadja, swara identification and rAga verification in AlApana using stochastic models," Applications of Signal Processing to Audio and Acoustics (WASPAA), IEEE Workshop, pp. 29-32, 2011.
6. J. Salamon and E. G´omez. Melody extraction from polyphonic music signals using pitch contour characteristics. IEEE Transactions on Audio, Speech, and Language Processing, 20(6):1759–1770, Aug. 2012.
7. J. Salamon, E. G´omez, and J. Bonada. Sinusoid extraction and salience function design for predominant melody estimation. In Proc. 14th Int. Conf. on Digital Audio Effects (DAFX-11), pages 73–80, Paris, France, Sep. 2011.

Thank you

Questions?