

Carnatic Music: A Computational Perspective

Hema A Murthy
Department of Computer Science and Engineering
IIT Madras
hema@cse.iitm.ac.in

e-mail: hema@cse.iitm.ac.in

December 13 2013



Outline

MIR and Carnatic Music

Carnatic Music Concert

Preliminaries

Tonic

Melodic cues

Processing the Drone

Gamakās in Carnatic Music

Cent filterbanks

Mridangam stroke transcription

Modes of the mridangam

Transcription

Pitch Extraction

Conclusions



Music Information Retrieval and Carnatic Music

- Indian classical music – rich repertoires, many traditions, many genres
- An oral tradition
- Well established teaching and learning practices
- Hardly archived and studied scientifically
 - Indian classical music is rich in Manodharma – improvisation
 - Difficult to analyse and represent using ideas from Western Music
- Objective: Enhance experience through MIR



Structure of a Carnatic Music Concert: A listener's perspective

- A concert is made up a sequence of items.
- Items correspond to different forms.
- Each form has a specific characteristic.
- *Rāgas* are seldom repeated (except in thematic concerts)
- One or more pieces in a concert are taken up for elaboration
 - **Kriti:**
 - an alaapana, a composition (pallavi, anupallavi, charanam), niraval, svaraprasthara, solo percussion
 - **Raagam Taanam Pallavi (RTP):**
 - RTP: an alaapana, a taanam, a composition (pallavi only), svaraprasthara, solo percussion.
 - Pallavi is rendered at different speeds with niraval.
 - Svaraprasthara may include multiple melodies.
 - Rhythmic cycles chosen – complex (e.g. Adi taalam (tisra nadai))
 - **Other types:** padham, jaavali, viruttam, slokam, varnam,

The Main item and RTP are generally the hallmarks of a concert.



Primary Aspects of Indian Classical Music from a Computational Perspective I

- Tonic – the base *sur* chosen by the musician to render the music.
 - Each musician has his/her own tonic
 - Bombay Jayashree: 220 Hz, T M Krishna: 140 Hz, MDR: 110 Hz.
- *Rāga* or melody
 - *Gamakās* – the inflection of notes
 - *Gamakās* are associated with phrases of *rāgas*¹
 - Exploration of *Ālāpana* through the relevant phrases
 - Phrases are rendered in different tempos
 - Phrases are derived from compositions – especially from the trinity
- Talas
 - Strong tradition of rhythm
 - Carnatic Music – kalai, nadai, jaati.
 - Mridangam stroke analysis
 - Segmentation of tani
 - Can we transcribe the same?
- Other Aspects 1: The concert is more like a conversation between the artist and the audience.



Primary Aspects of Indian Classical Music from a Computational Perspective II

- The appreciation by the audience is more a “here and now”²
- A concert is replete with applauses – can we use these to segment and archive them in terms of pieces?
- Other Aspects 2: Since motivic analysis is based on pitch – and most algorithms fall short – explore new algorithms for pitch based on phase.

¹ T M Krishna and Vignesh Ishwar, “Carnatic Music: Svvara, Gamaka, Motif and Raga Identity”, 2nd CompMusic Workshop, Istanbul, Turkey

² M V N Murthy, “Applause and Aesthetic Experience,”<http://compmusic.upf.edu/zh-hans/node/151>

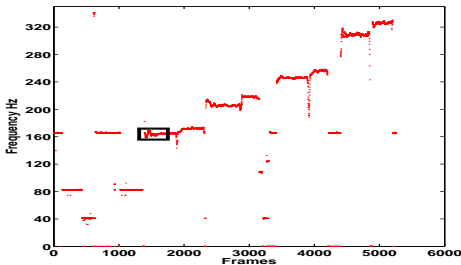
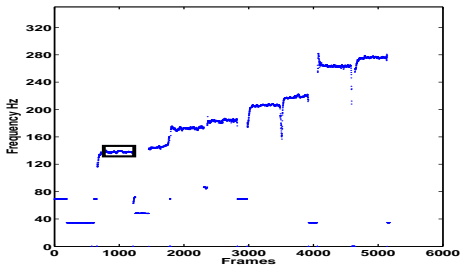


Tonic I

3.

- Tonic – A fundamental concept of Indian classical music
- Tonic – Pitch chosen by the performer to serve as reference
- The svara Sa in the middle octave range is the tonic
- Drone is played to establish tonic – Tanpura/Tambura
- Accompanying instruments also tune to the tonic
- Melodies defined relative to tonic

Tonic II



Tonic 1 🐼 Tonic 2 🐼 Drone 1 🐼 Drone 2 🐼 Alapana 🐼

³ Ashwin Bellur, "Automatic identification of tonic in Indian classical music," MS Thesis, IIT Madras, 2013



Approaches to identify tonic

- Music provides various cues to the listener about identity of the tonic
 - Cues can be divided into two broad classes
1. Melodic characteristics of the music ¹
 2. Tuning of the drone²

¹ S Arthi H G Ranjani and T V Sreenivas. Shadja, swara identification and raga verification in alapana using stochastic models. WASPAA 2011 pages 29-32, 2011.

² Salamon, J., S. Gulati, and X. Serra (2012). A multipitch approach to tonic identification in indian classical music. In Proc. of ISMIR, 157163

Cue 1 - Melodic characteristics of the music

- Mono pitch extracted from the audio – prominent pitch
- Pitch range: 40 to 700 Hz , window size = 133ms
- Histograms as primary representation
- Typical Histograms for Hindustani and Carnatic music (bin width 1 Hz)

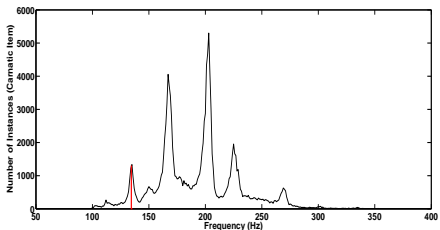


Figure: Carnatic Pitch Histogram

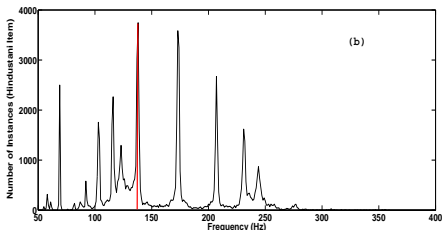


Figure: Hindustani Pitch Histogram



Pitch Histograms based processing

1. A peak indicating note Sa always present, not necessarily the tallest peak
2. Drone and percussion ensure a peak at Sa
3. Histogram envelope is almost continuous due to gamakas. (Example 🎧)
4. Fixed ratio between peaks representing svara Sa and Pa
5. Less inflected nature of Sa and Pa
6. Characteristics more prominent in Carnatic music



Cue 2 - Drone

- Determine tuning of the drone to identify tonic
- Drone omnipresent in Indian classical music
- Strings of the tambura tuned to indicate svara Sa
- Tuning of one of the strings varies depending on the raga being performed
- Attempt to develop fast tonic identification techniques with minimal data

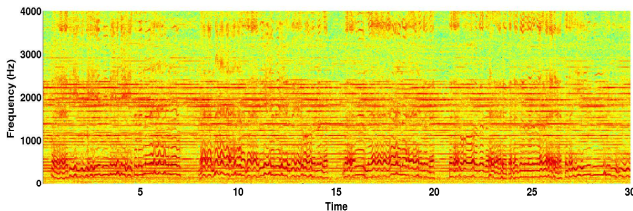


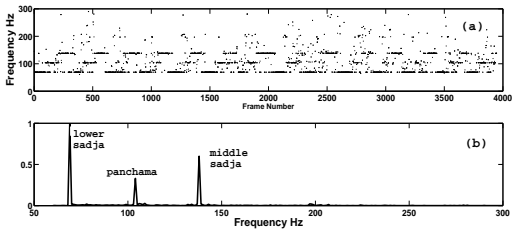
Figure: *Spectrogram of an excerpt of Carnatic music*

Example 1  Example 2  Example 3 



Determining tuning of the drone

- Drone omnipresent in the background
- Drone extracted in low energy regions
- Lead vocal frequencies predominantly occupy middle and upper octaves
- Drone frequently registers pitch values at the lower octave *Sadja/Sa*





Drone Prominent Frames

- Pick frames with drone as the prominent source
- A host of low level audio descriptors were employed
- Pitch estimated using a selected bag of frames using signal processing, dictionary learning methods (Non Negative Matrix Factorisation (NMF)).

Results:

- Performance almost as high as 90% on 1.5min of data when signal processing cue is used.
- Performance 98% with drone on 1.5mins of data.



Gamakās in Carnatic Music⁴

- Written forms of a *rāga*
 - ArOhana – sequence of notes in ascending order (generally starts from the tonic *Sa*).
 - AvarOhana – sequence of notes in descending order (generally starts from the tonic *Śa*).
- Vocal/Instrumental form of a *rāga*
 - Replete with *Gamakās*
 - Although specific placeholders – notes can meander quite significantly about the placeholder.
 - The “note uttered” and “note sung” need not have a one-to-one correspondence

⁴Vignesh Ishwar, Shrey Dutta, Ashwin Bellur and Hema A Murthy, “Motif spotting in an Alapana in Carnatic music,” ISMIR 2013.



Classification of *Gamakās* (SSP – Subbarama Dikshitar, Dasavidha *Gamakā* –as told to us by the experts)

- A finite number of meandering patterns are defined
 - kampitam, jaaru, vali, spuritham, ...
- The *Gamakās* are stitched together/one *gamakā* is overlayed on the other
- Other than *gamakās*, “brikhas” are also used.
- Standard set of melodic phrases are observed in every piece.



Characteristic Pitch and Spectrogram for some of the *Gamakās*

Some example *Gamakās* created by Vignesh Ishwar – tOdi

Kampitham:



Jaaru:



Odukkal:



Nokku:



Orikkai:





Gamakās in a given ālāpana

Vignesh tOdi *Ālāpana*:



KVN bhairavi *Ālāpana*:





Summary of pitch analysis on complete *ālāpanas*

- The musician does not seem to always stick to the grammar of the *gamakā*.
- Some definite phrases do exist
 - Most keen listeners identify *rāgas* without difficulty.
- A set of Time-Frequency (T-F) motifs?
 - T-F representation of pitch?
- Mapping from *rāga* to the language as defined by musicians required.
- Derive motifs for *rāgas* using “machine learning?”



Histograms of notes for different *rāgas* I

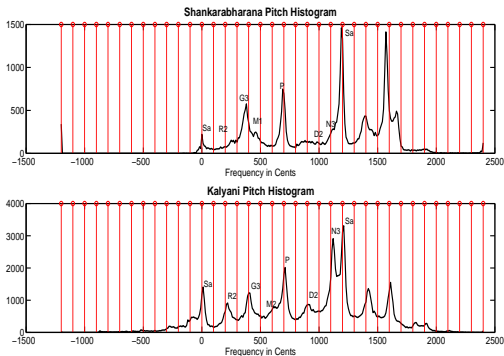


Figure: Pitch Histograms of Ragas Kalyani and Sankaraabharana



Histograms of notes for different *rāgas* II

- Significant band of frequencies around every peak that is also frequent.
- The two *rāgas* have only one note that is different – but their histograms are very different.
- Pitch contour is seamless and continuous.

Motifs in a *rāga*

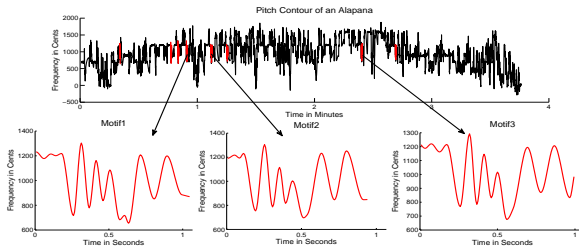


Figure: a) Motifs Interspersed in an Alapana; b) Magnified Motif

- The motif is repeated in different parts of the *Ālāpana*
- There are a number of motifs for each *rāga*
- Question: Motif as a query, can we locate it in an *Ālāpana*?
- *Ālāpana*: Long, erroneous pitch contour, computationally intensive.

Saddle points in tandem with Rough LCS used to locate motifs. Shrey Dutta will discuss this work.



A new feature for analysis of Indian Music I

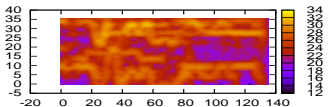
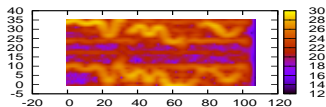
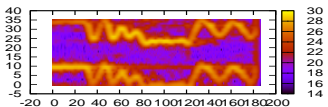
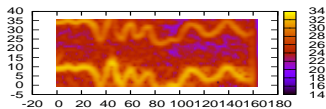
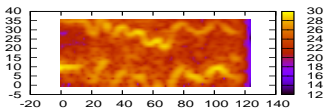
Tonic shift – causes a linear shift in cent scale.

A possible solution: cent filterbank banks

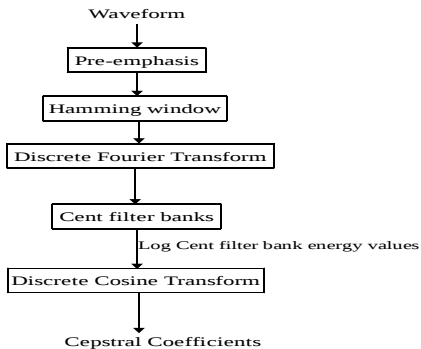
- Analysis of a concert in Indian Music depends on tonic.
- Motifs can span more than one octave.
- Cochlea – can be modeled by a bank of constant Q filters.
- For Indian music: Normalise by tonic – place filters uniformly on the cent scale:
Cent Filterbanks



Cent Filterbanks



Extraction of Cent filterbank features⁵





Cent Filterbanks

- Cent Filterbanks are similar to CQT (Constant Q transform) filters with a difference – they are tonic normalised.
- CQT is equivalent to Cent Filterbanks with a tonic of 2.
- Padi Sarala will talk more about cent filterbanks and their applications to segmentation motif recognition ⁶, stroke recognition ⁷ and archival ⁸

⁶ unpublished

⁷ Akshay Ananthapadmanabhan, Juan Bello, Raghava Krishnan and Hema A Murthy, "Tonic independent stroke transcription of the mridangam, AES, 2014

⁸ Padi Sarala and Hema A Murthy, "Inter and intra segmentation of Carnatic music recordings for archival," ISMIR 2013.



Repeating the experiments of C V Raman⁹.

- Strokes were played carefully
- The number of strokes is 8
- According to Raman there are 5 different modes – resonances
- Therefore 5 basis vectors were estimated in NMF

⁹ Akshay Ananthapadmanabhan, Ashwin Bellur and Hema A Murthy, “Modal analysis and transcription of strokes of the mridangam using non-negative matrix factorisation,” ICASSP 2013, Vancouver, Canada

Modes of Mridangam and Activations for different strokes

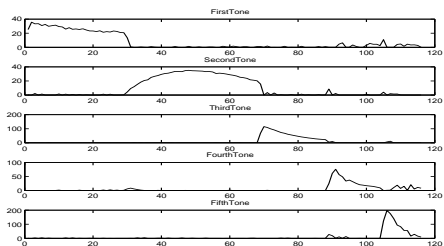


Figure: Modal Tones

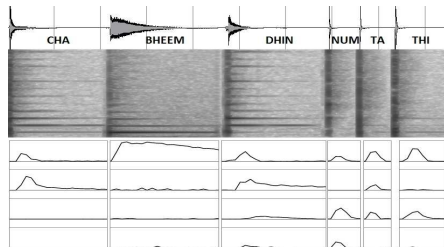


Figure: Strokes and their modes

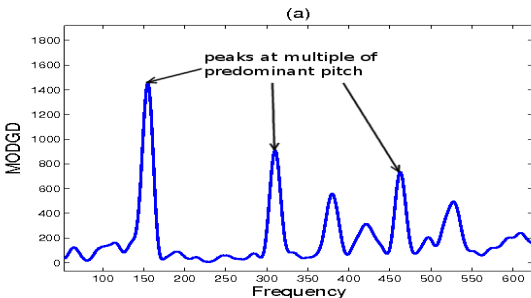


Transcription of strokes

- Training of HMMs using activations
- Testing on untranscribed data – use NMF activations to segment the strokes
- Performance evaluation
 - An accuracy of about 87-89% was obtained on a controlled experiment.
- Current work:
 - Tonic independence using CQT and cent filterbanks along with HMMs (unpublished) – very good results – Sarala/Akshay will talk about this.
 - Dr. Umayalpuram Sivaraman is working with us on transcribing data for his tanis – he has defined a number of new strokes – a task by itself.
 - Other upcoming artists are helping us transcribe various other artists.

Pitch extraction using modified group delay functions¹⁰

- A group delay based approach to pitch extraction.
- Flattened magnitude spectrum is a sinusoid.
- Estimate the frequency of the sinusoid using modified group delay functions.
- Results are encouraging – Rajeev Rajan will give details.



¹⁰ Rajeev Rajan and Hema A Murthy, "Group delay based melody monophony extraction from music," *ICASSP 2013*.



Conclusions

- Different Aspects of Carnatic Music Analysis addressed
- Tonic identification: Group delay based, NMF
- Motif discovery: Rough Longest Common Subsequence
- Segmentation of recordings of Carnatic Music
- Transcription of mridangam
- Future Work
 - Motif discovery using varnams, oneliners of songs
 - Does an *Alāpana* in CM have a rhythm?
 - Processing music uploaded into Music Brainz – convert many of the research ideas to enhance MIR
 - Multipitch pitch tracking using phase.
 - Cognitive signal processing and music – signal processing and machine learning used in tandem.