Cent Filter Banks and its Relevance to Carnatic Music

Padi Sarala, Akshay Ananthapadmanabhan and Hema A.Murthy

3rd CompMusic Workshop

Indian Institute of Technology
Madras, India

Date: December 13, 2013

Outline of the presentation

- Importance of tonic with respect to Carnatic music
- Introduction to Cent filter banks
- Applications of cent filter banks:
  - Song identification in a concert
  - Motif recognition
  - Mridangam stroke recognition
- Experimental results
- Demo:Segmentation of concert into items for archival

Importance of Tonic with respect to Carnatic music

Tonic:

- In carnatic music, each singer performs the concert with respect to a reference called Tonic.
- The tonic is chosen by the performer and accompanying instruments are tuned to the same tonic.

Drone or Tambura:

- Generally in any concert, tonic is fixed and it is maintained throughout the concert using an instrument called the drone.
- The function of drone is to preserve the tonic throughout the concert.
- Tonic ranges from 160Hz to 250 Hz for female singers and 100Hz to 175Hz to male singers.[a b]

---

[a] Ashwin Bellur, Vignesh Ishwar, Xavier Serra, and Hema A. Murthy. *"A knowledge based signal processing approach to tonic identification in indian classical music"*. In *International CompMusic Wokshop*, 2012.

[b] Justin Salamon, Sankalp Gulati and Xavier Serra. **"A Multipitch Approach to Tonic Identification in Indian Classical Music"** , In Proc. of ISMIR 2012

Motivation for Cent Filter Bank Energy Feature

Sno	Carnatic music swara	Label	Frequency ratio	Carnatic music scale for different tonic			
				138	156	198	210
1	Shadja (Tonic)	S	1.0	138	156	198	210
2	Shuddha rishaba	R1	(16/15)	147.20	166.40	211.20	224
3	Chatushruthi rishaba	R2	(9/8)	155.250	175.50	222.75	236.25
4	Shatshruthi rishaba	R3	(6/5)	165.250	187.20	237.60	252
3	Shuddha gAndhara	G1	(9/8)	155.250	175.50	222.75	236.25
4	ShAdhArana gAndhara	G2	(6/5)	165.60	187.20	237.60	252
5	Anthara gAndhara	G3	(5/4)	172.50	195.0	247.5	262.5
6	Shuddha madhyama	M1	(4/3)	184.0	208.0	264.0	280
7	Prati madhyama	M2	(17/12)	195.50	221.0	280.5	297.5
8	Panchama	P	(3/2)	207.00	234.0	297.0	315
9	Shuddha daivatha	D1	(8/5)	220.80	249.60	316.8	336
10	Chatushruthi daivatha	D2	(5/3)	230.00	260.0	330.0	350
11	Shatshruthi daivatha	D3	(9/5)	248.40	280.80	356.4	378
10	Shuddha nishAdha	N1	(5/3)	230.0	260.0	330.0	350
11	Kaisika nishAdha	N2	(9/5)	248.40	280.80	356.4	378
12	KAkali nishAdha	N3	(15/8)	258.75	292.50	371.25	393.75

Table: Carnatic music swaras and their frequency ratios.

Melody in CM:

- CM is based on the twelve semitone scales and frequencies of semitones depends on the tonic.
- Melody is made up of set of notes. These set of notes in CM is defined with respect to the tonic.
- Table shows the frequencies corresponding to twelve semitones for four singers, each with a different tonic.
- Frequencies of semitones vary with respect to tonic.

Mel and Cent Filter banks

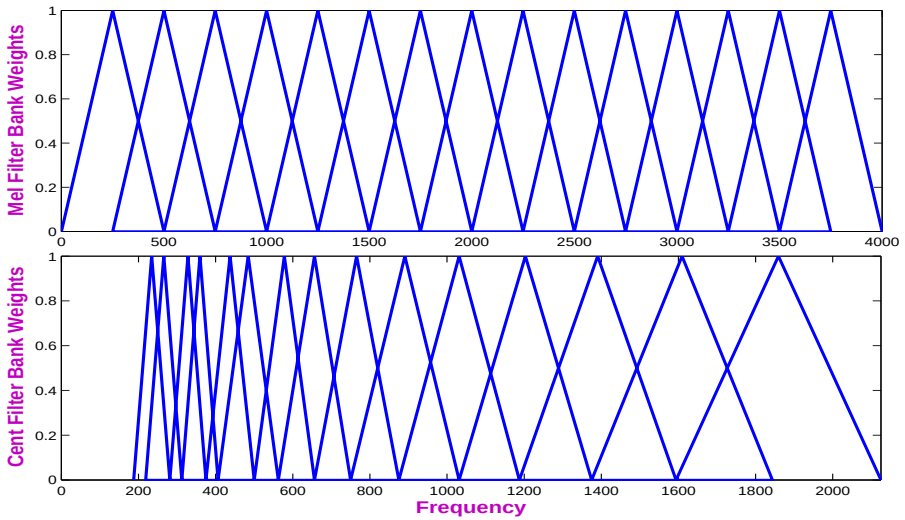


Figure: Filter banks of Mel scale and Cent scale.

Mel Scale

$$Mel\ Scale = 2595 \cdot \log_{10}\left(1 + \frac{f}{700}\right) \tag{1}$$

Cent Scale

$$Cent\ Scale = 1200 \cdot \log_{2}\left(\frac{f}{tonic}\right) \tag{2}$$

CentFilter Bank Energy Feature Extraction (1)

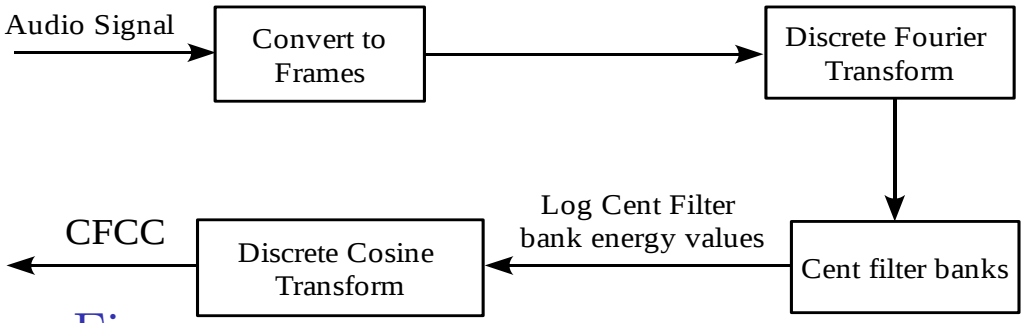


Figure: Cent Filter bank energy feature extraction.

Cent Filter Bank Extraction:

- The audio signal is divided into frames.
- The short-time Discrete Fourier Transform (DFT) is computed for each frame.
- The power spectrum is then multiplied by a bank of filters that are spaced uniformly in the tonic normalised cent scale. The cent scale is defined as:

$$cent = 1200 \cdot \log_2 \left(\frac{f}{tonic} \right) \tag{3}$$

- The energy in each filter is computed.
- Discrete Cosine Transform (DCT-II) of log filter bank energies is computed to get cepstral coefficients.

Applications of Cent filter banks

Cent filter bank based cepstral coefficients are applied for different music processing tasks like:

- Song identification in a carnatic music concert.
- Motif recognition in an Alapana.
- Mridangam stroke recognition in ThaniAvarthanam.

Song Identification in a concert

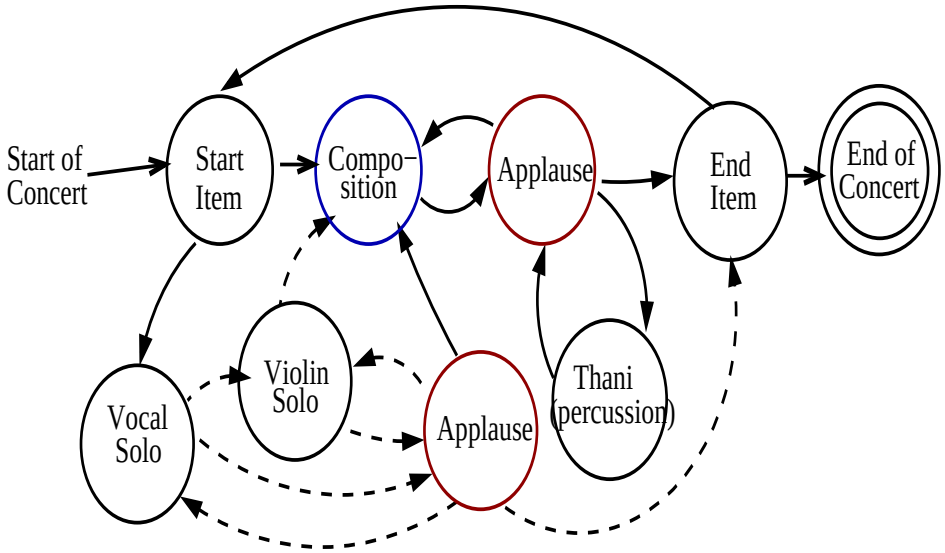


Figure: *General structure of a concert in carnatic music.*

Importance of Song:

- Composition segments are performed with respect to a raga.
- Locating these song segments in a concert is very much useful for musicians.
- Song segments can be used further for finding the number of items in a concert.

Experimental Evaluation

Singer Name	No. of Concerts	Duration (Hrs)	No. of Applause	Different Tonic
Male 1	4	12	89	158,148,146,138
Female 1	4	11	81	210, 208
Male 2	5	14	69	145, 148,150,156
Female 2	1	3	16	198
Male 3	4	12	113	145,148
Female 3	1	3	15	199
Male 4	26	71	525	140,138,145
Male 5	5	14	62	138,140

Table: *Database used for study, different Tonic values identified for each singer using pitch histograms.*

Database Used for the Study:

- 50 live recordings of male and female singers are taken for experiments. .
- All concerts are vocal and the total number of applauses are 990.
- It can be observed that even for a given singer the tonic varies across concerts.

Experimental Setup

Building the Models:

- From male (female) recordings 3 segments are randomly chosen for each class.
- MFCC, ChromaFCC and CFCC features are used to build 32 mixture GMM models for 4 classes namely *Vocal*, *Violin*, *ThaniAvarthanam*, and *Song*.

Segmentation of a Concert:

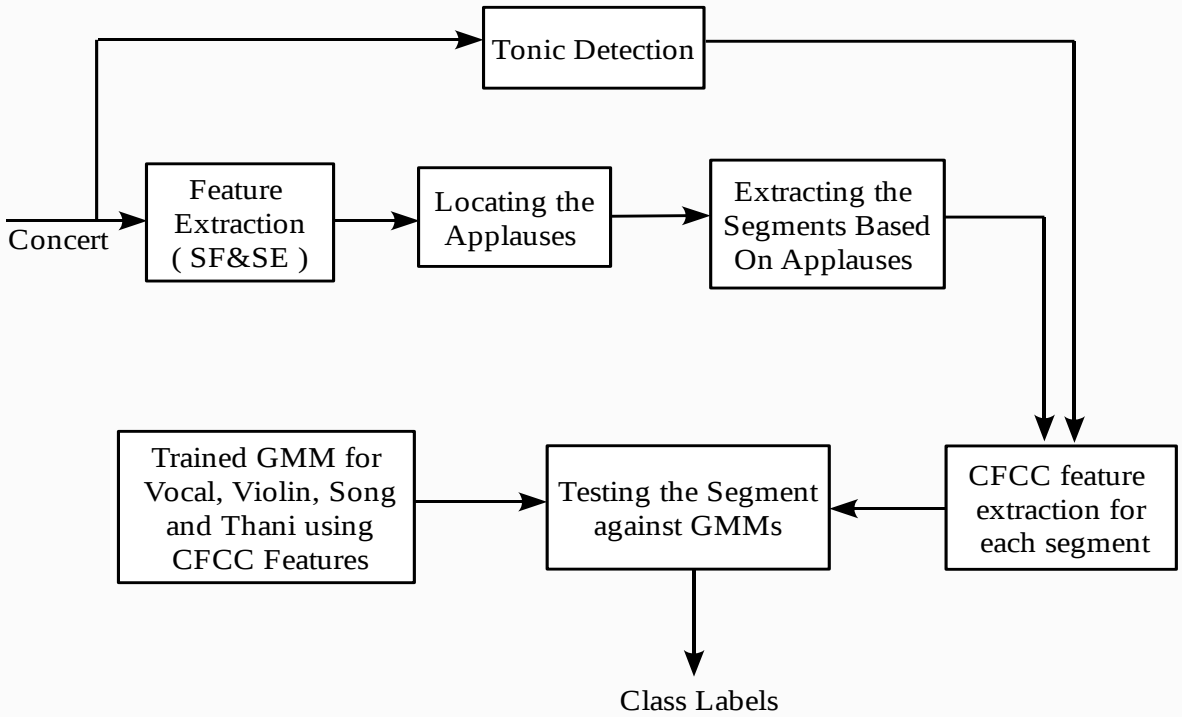


Figure: Segmenting the concert into Vocal, Violin, Song using CFCC features by building GMMs.

Experimental Results

Model	MFCC	ChromaFCC	CFCC
Male singers	78%	60%	90%
Female singers	92%	70%	97%

Table: *Main song identification performance using MFCC, ChromaFCC and CFCC.*

Segmentation Results:

Cent filter bank based cepstral coefficients better captures the notes positions compared with that of Chroma and MFCC features.[a]

[a] Padi Sarala and Hema A. Murthy. ***"Cent Filter Banks and its Relevance to Identifying the Main Song in a Carnatic Music"***. In Proc. of *CMMR, Marseille*, 2013.

Motif Recognition

Motif:

- Motif defines the characteristics of *Raga*.
- Motif can be thought of sequence of notes that are unique to a *Raga*.
- Pitch information is used for Motif recognition.[a]

[a] Vignesh Ishwar, Ashwin Bellur, , Xavier Serra, and Hema A. Murthy. ***"Motivic Analysis and its Relevance to Raga Identification"***. In *International CompMusic Wokshop*, 2012.

Database Used

Raga Name	Phrases labelled	Instances
Bhairavi	Phrase 1	70
	Phrase 2	51
Kambhoji	Phrase 1	104
	Phrase 2	48
	Phrase 3	45
Sankarabharanam	Phrase 1	81
	Phrase 2	51
	Phrase 3	98
Kalyani	Phrase 1	52
Varali	Phrase 1	52

Table: *Total number of phrases for each Raga*

Name of the Feature	Classification Accuracy
MFCC	55%
Pitch	65%
Chroma	63%
CQT	67%
CFCC	73%

Table: *Motif recognition accuracy*

Motif Recognition Results

Pitch:

Raga	Va1	Sh1	Ko1	Sh2	Bi1	Kl1	Ko2	Sh3	Bi2	Ko3
Va1	**43**	2	1	3	0	3	0	0	0	0
Sh1	0	**16**	38	3	0	8	5	10	0	0
Ko1	0	4	**91**	1	0	0	0	4	0	2
Sh2	0	3	6	**39**	0	0	1	1	0	0
Bi1	0	1	0	0	**62**	0	4	1	1	1
Kl1	0	10	16	7	0	**13**	0	3	0	2
Ko2	0	0	1	0	0	0	**42**	0	1	1
Sh3	0	13	38	3	0	2	2	**34**	1	3
Bi2	0	0	0	0	3	0	0	0	**46**	2
Ko3	0	0	2	0	0	1	0	0	2	**40**

Table: *Confusion Matrix for Motif Recognition using HMMs.*

CFCC:

Raga	Va1	Sh1	Ko1	Sh2	Bi1	Kl1	Ko2	Sh3	Bi2	Ko3
Va1	**39**	0	0	2	0	0	0	0	9	1
Sh1	4	**48**	5	4	1	5	1	8	4	0
Ko1	0	11	**74**	2	0	2	0	14	0	0
Sh2	1	2	2	**41**	1	1	0	1	1	0
Bi1	1	3	0	0	**50**	0	11	0	2	2
Kl1	0	2	1	3	0	**37**	0	2	6	0
Ko2	1	0	1	0	2	0	**42**	0	0	1
Sh3	1	17	6	4	1	4	1	**55**	4	4
Bi2	1	0	0	3	1	1	0	0	**44**	0
Ko3	0	0	0	0	0	0	1	0	0	**43**

Table: *Confusion Matrix for Motif Recognition using HMMs.*

Mridangam Stroke Recognition

Mridangam:

- Mridangam is a primary percussion instrument used in Carnatic music.
- Transcribing mridangam strokes is useful for students.
- Transcription can also provide information for musicians to practice instruments.
- Automatic transcription allow us to keep track of other musical traditions.

Characteristics of strokes:

$\underbrace{\textit{Bheem} \quad \textit{Cha} \quad \textit{Dheem} \quad \textit{Dhin}}_{\Downarrow}$ $\underbrace{\textit{Num} \quad \textit{Ta} \quad \textit{Tha} \quad \textit{Thi} \quad \textit{Tham} \quad \textit{Thom}}_{\Downarrow}$

Tuned to Tonic | Not Tuned to Tonic

Implementation of Mridangam stroke Recognition

A MULTIPITCH APPROACH TO TONIC IDENTIFICATION IN INDIAN CLASSICAL MUSIC Justin Salamon, Sankalp Gulati and Xavier Serra

Features used:

- Constant Q-Tranform (CQT) features with NMF activations are used for mridangam recognition.[a]
- Cent filter bank based cepstral coeffieicents.
- Feature extraction includes 6 octaves.

[a] Akshay, Juan P.Bello, Raghav Krishnan and Hema A. Murthy. ***"Tonic-Independent Stroke Transcription of the Mridangam"***. Accepted for AES 53rd International Conference, London, 2014.

Experimental Evaluation

Database Used:

Tonic	Stroke Instances
B	1325
C	1129
C#	1197
D	916
D#	1495
E	1100

Table: *Number of stroke instances for each tonic.*

Mridangam Recognition accuracy:

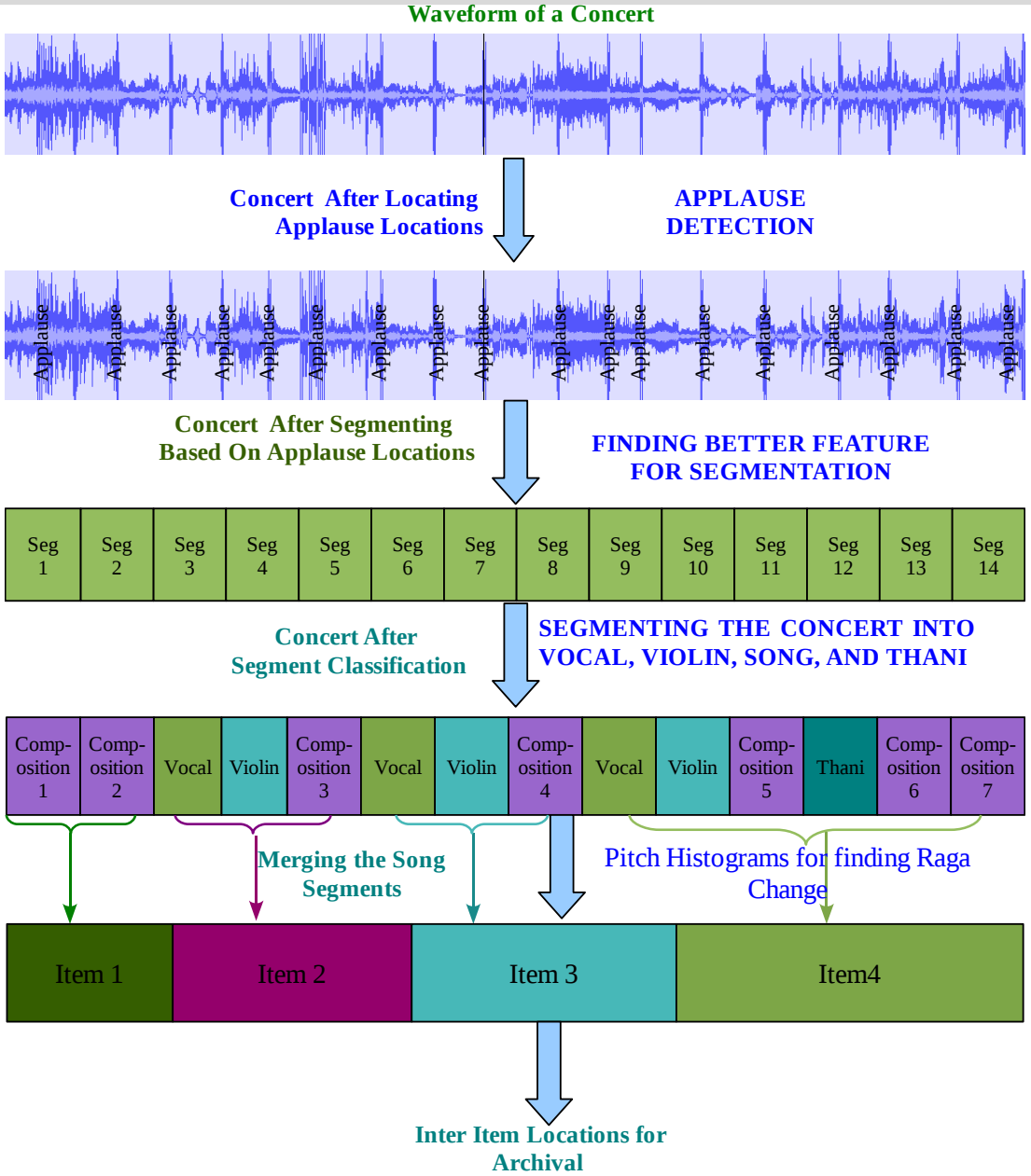
Name of the Feature	Classification Accuracy (%)
CQT	74
CFCC	77

Table: *Mridangam stroke recognition (with all tonics).*

Name of the Feature	Classification Accuracy (%)
CQT	62
CFCC	66

Table: *Mridangam stroke recognition (tonic invariant).*

Demo: Segmentation of a concert into Items for archival



THANK YOU