Raga Identification of Carnatic music for music Information Retrieval

2 authors:

Rajeswari Sridhar
Anna University, Chennai
19 PUBLICATIONS   85 CITATIONS

SEE PROFILE

T.V Geetha
Anna University, Chennai
117 PUBLICATIONS   362 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   Cross Lingual Information Access View project

Project   Cross-Lingual Information Access View project

Raga Identification of Carnatic music for Music Information Retrieval

Rajeswari Sridhar[1], and T.V. Geetha[2]

[1, 2] Department of Computer Science and Engineering, Anna University, Chennai 600025, India

[1]Email: rajisridhar@annauniv.edu, [2]Email: tvgeedir@cs.annauniv.edu,

Abstract— In this work we propose a method to identify the raga of a Carnatic music signal. Raga is analogous to melody in Western music but is much richer and stronger in its representation. The main motive behind Raga identification is that it can be used as a good basis for music information retrieval of Carnatic music songs or Film songs based on Carnatic music. The input polyphonic music signal is analyzed and made to pass through a signal separation algorithm to separate the instrument and the vocal signal. After extracting the vocal signal we segment the vocal signal using our proposed segmentation algorithm. Using our proposed singer identification algorithm we determine the singer to know the fundamental frequency of the singer. The frequency components of the signal are then determined and we map these frequency components into the swara sequence and thereby determine the Raga of the particular song, which could be used to index the songs and further for retrieval based on the Raga.

*Index Terms—***Carnatic music, Raga, Music information retrieval**

I. INTRODUCTION

Music information retrieval is the need of the hour due to the availability of large amount of music on the Internet. Music information retrieval can be thought of as part of Multimedia information retrieval. A lot of work has been done in the other components of multimedia like text, video and the one that is yet to be fully developed is audio. Audio processing involves processing speech and music. In this paper we discuss music processing which could be used as the basis of music information retrieval using music characteristics. This paper is organized as follows: Section 2 discusses about the Characteristics of Carnatic music, Section 3 discusses on Existing work in Raga and melody identification, Section 4 discusses on the proposed System Architecture, Section 5 discusses on Results and Evaluation and Section 6 concludes the paper.

2. CHARACTERISTICS OF CARNATIC MUSIC

Indian music is broadly classified into South Indian Carnatic music and North Indian Hindustani music. Both the systems of music are rich in their own style and Carnatic music is much more complex in the way the notes are arranged and rendered [1]. Indian music in general is based on Raga and Talam. Raga and Talam can be thought of as equivalent to melody and rhythm in Western music. Raga is much more complex than melody and scale in western music. Raga can be thought of as the sequential arrangement of notes that is capable of invoking the emotion of a song. A note is called as a Swara in Carnatic music. Raga is characterized by Arohanam and Avarohanam. Arohanam defines the ascending order of the arrangement of notes and avarohanam is defined as the descending order arrangement of notes. A raga is classified into melakarta raga (parent) and janya raga (child). The various notes in Carnatic music are S, R, G, M, P, D, N which can be thought of as analogous to C, D, E, F, G, B, A. Every note has a frequency associated with it as given in [2]. A melakarta Ragam will have all the seven swaras and a janya ragam will have atleast 5 of the seven swaras [1]. Every song in Carnatic music has an associated Talam. There are essentially 175 talams as explained in [1]. In Carnatic music any song will be of time duration which is an integer multiple of a Talam. A talam is analogous to the beat of European music and the seven basic talams is discussed in [2]. It indicates the pacing of the music and the placement of syllables in the composition. In Carnatic music, the singer indicates the talam using hand gestures [2].

Usually songs, which are based on South Indian Carnatic music, will belong to one of the several Ragas or a combination of Ragas and one of the Talam. Hence it is a good idea to index songs, which are based on the Carnatic music characteristics using these components, Raga and Talam. Hence, in this work we have proposed an algorithm to determine the ragam and talam of a Carnatic music song and have proposed that these components could be used as features for indexing and later for retrieval.

3. EXISTING WORK

Literature work in Carnatic music retrieval is on a slow pace compared to Western music. Some work is being done in Swara identification [2] and Singer identification [3] of Carnatic music. In Hindustani music work has been done in identifying the Raga of Hindustani music [4]. In [4] the authors have created a HMM based on which they have identified two ragas of Hindustani music. The fundamental difference between Hindustani Raga pattern and Carnatic Raga pattern is that in Hindustani we have R1, R2 as against R1, R2, R3 in Carnatic. Similarly G, D, N all has three distinct frequencies in Carnatic music as compared to two frequencies in Hindustani [5]. This reduces the confusion in identifying the distinct frequencies in Hindustani music as compared to Carnatic music. The authors have not used polyphonic music signal and have assumed that the input music signal is a voice only signal. The fundamental frequency of the signal was also assumed and based on these features the raga identification process was done for two Hindustani ragas. On the western music aspect, melody retrieval is being performed by researchers. The one proposed by [6]

is based on identifying the change in frequency in the given query. The query is received in the form a humming tune and based on the rise and fall in the pitch of the received query, the melody pattern that matches with the query's rise and fall of pitch is retrieved. In the work proposed by [7] for melody retrieval the authors have identified melody based on features like distance measures and gestalt principles. The approach we have followed is based on low level signal features and we have identified the raga by considering a polyphonic signal as input to our system.

4. SYSTEM ARCHITECTURE

The proposed architecture of the system is given in Figure 1. The various modules of our system are signal separation, segmentation, feature extraction, frequency mapping based on singer identification and finally raga identification. The various blocks are explained in the following sections.
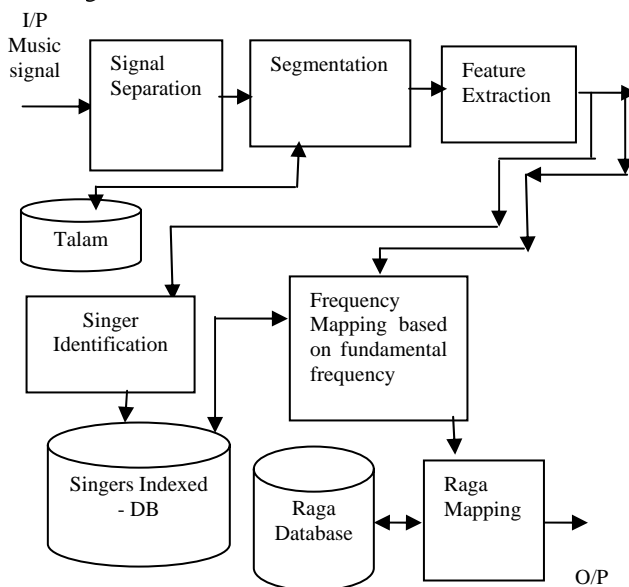

Figure 1: System Architecture.

4.1 Signal Separation

The performances of two signal separation algorithms proposed by [8, 9] have been compared and the authors of [10] have concluded that the one proposed by [8] is better suited for Carnatic music. The algorithm proposed by [8] is slightly modified to efficiently suit Carnatic music. The threshold value of the harmonic is chosen so that on repeated iterations the signal gets separated. During the first iteration the threshold value for the harmonic is as given by [9] and the signal is separated into voice and music. During the next iteration the threshold value is chosen depending on the extracted output of the first iteration. After the signal is separated into voice and instrument, the signal is segmented to identify the frequency components present in the signal for swara identification yielding to raga identification.

4.2 Segmentation

The segmentation algorithm we have proposed involves determination of onset and offset followed by a two level segmentation process. The modules are explained below:

4.2.1 Onset and Offset detection

Usually, a Carnatic song is associated with one Ragam. Sometimes, there is a special song which is normally sung in a combination of Ragams where, every 4 or 5 lines will be sung in one Ragam the next 4 or 5 lines in another ragam and so on. Such a song is called as a Ragamalika, which means composed of many Ragas. This is not a usual scenario and hence in our discussion we are trying to identify songs that are based on only one Ragam. Every song is associated with a Talams as indicated [2]. The Talam's first count starts at the beginning of the song and ends when the song completes. In other words, any song is an integral multiple of a pre-specified Talam. This gives an indication that in order to identify the swaras of the song it is a good idea to find out the beginning and the end of the Talam. This requires us to identify the onset and offset of a song. Onset refers to the beginning of a musical note or other sound, in which the amplitude rises from zero to an initial peak. Approaches to onset detection can operate in the time domain, frequency domain, phase domain, or complex domain, and include looking for the following changes:

1. Increases in spectral energy
2. Changes in spectral energy distribution (spectral flux) or phase
3. Changes in detected pitch
4. Spectral patterns recognisable by machine learning techniques such as neural networks.

Simpler techniques such as identifying increase in time-domain amplitude can typically lead to an unsatisfactorily high amount of false positives or false negatives [11]. The technique that we adopted is based on identifying change in spectral energy. Hence, we take a copy of the signal and convert it into frequency domain and observe the change in spectral energy. We then map this point to the time domain signal and mark that as the onset. We detect the offset in a manner similar to detecting onset where we identify the point at which the spectral energy starts dying as the offset. Both the onset and offset points are transformed to the time domain signal and then the time domain signal between the points of onset and offset is taken for the purpose of raga identification.

4.2.2 First Level Segmentation

After determining the onset and offset the signal between these two points are considered for the next level of segmentation. Our assumption is that, this segmented signal must be an integral multiple of one of the Talams as indicated in [2]. A database of 175 Talams with their time duration is maintained. This database is sorted based on decreasing frequency of usage of the talams. We try to map the talams time duration with the signal. A fitness function is written based on time duration to assign the

time duration of the Talam to the segmented input frame. The fitness function tries to match the segment against the duration of a talam which is retrieved from the database. The matching is based on the fact that the duration of segment must be an integral multiple of the talam duration. The talam that matches the segment input is identified as the Talam for the input song. The next level of segmentation is based on the identified Talam.

4.2.3 Second Level Segmentation

After obtaining the talam associated with a segment, the segment is divided into individual talam segments by observing the time duration in the Talam database. Any Talam has a fixed predetermined component pattern. Each component depending on whether it is a count – '|', thattu '0', veech '0', has varying duration. After identifying the Talam in the first level of segmentation the pattern of the identified Talam is known. Using this feature the segmented signal is further segmented at each count, Thattu or Veechu of the given Talam. At this point of segmentation we have individual segments which correspond to the Thattu, Veechu or count. In Carnatic music, for a count, Thattu, Veechu there can be one, two, or four swaras depending on the tempo in which the song is sung. That is, each component of a talam obtained after the second level of segmentation can be realized as one, two or four swaras. Hence each of our segments at this level of segmentation could correspond to either one, two or four swaras. This requires us to further segment the signal. Hence we segment each talam component obtained as a result of second level segmentation into four equal frames. This segmented signal is then used for the purpose of Raga identification.

4.3. Feature Extraction

After identifying the smallest possible segment we assume that this could correspond to a swara. We then use the HPS algorithm as explained in [12] to identify the frequency components in a single segment. In each segment there is a possibility for the presence of more than one frequency component. In order to determine a frequency component corresponding to a swara in each segment, we use energy as the basis for identifying the distinct frequency component in each segment. After identifying the frequency component present in each segment, we determine the one with the highest energy as the one that corresponds to the swara. We now have to map this frequency component to the corresponding swara. In order to perform this mapping, we need to identify the fundamental frequency of the signal. The fundamental frequency is identified by performing the process of singer identification.

We extract features pertaining to identifying the singer of the song as explained in [3]. The singer of the song is identified as explained in [3] and this is used as the basis for identifying the fundamental frequency of the song. The fundamental frequency is essential for Carnatic music analysis, since a singer can use a frequency in the range of 240 to 400 Hz as the fundamental frequency.

TABLE I.
SWARAS AND THEIR RATIOS WITH 'S'

Swara	Ratio	Swara	Ratio
S	1	M2	27/20
R1	32/31	P	3/2
R2	16/15	D1	128/81
R3	10/9	D2	8/5
G1	32/27	D3	5/3
G2	6/5	N1	16/9
G3	5/4	N2	9/5
M1	4/3	N3	15/8

The fundamental frequency normally corresponds to S and all other swaras of R, G, M, P, D, and N will depend on this S. The ratio of the other swaras relative to S is listed in Table I. Using literature on Carnatic music and normal singers singing pattern we have formulated a table of singers with their corresponding fundamental frequency. This is populated in the singer's database and the fundamental frequency of the singers is retrieved for Swara determination. Using Table I and the extracted frequency components raga identification is performed.

4.4 Frequency mapping and Raga determination

The swara is determined by identifying the singer's fundamental frequency and the dominant energy based frequency in each segment. Using the the ratio available in Table I the swara corresponding to every frequency is determined using the identified dominant frequency in every segment. We then identify the individual swaras available and use these swaras and compare it with a database of Raga to identify the raga of the given song. The Raga database is a table consisting of the name of the raga, the arohanam, avarohanam of raga in the form of swara components. We use a simple string matching to compare the identified swara pattern with the Raga database and determine the Raga of the input Carnatic song.

5. RESULTS & EVALUATION

The input signal is sampled at 44.1 KHz. We consider the identification of three melakarta Ragams for the purpose of evaluating this algorithm. For the purpose of Raga identification we have considered songs sung by like Nithyasree, M.S.Subalakshmi, Balamuralikrishna and Ilayaraja. From music knowledge we have identified their fundamental frequencies as 400 Hz, 320 Hz, 400 Hz and 240 Hz respectively. We considered songs that belong to melakarta ragas, Sankarabharanam, Kanakangi, Karaharapriya ragam and belonging to Adi talam or Roopaka Talam. We made the signal to pass through the signal separation algorithm, and segmentation algorithm. The result showing the segmentation points for one input is shown in Fig 2. This is the first level of segmentation where the protruding lines indicate the points of segmentation. After identifying the segmentation points we determine the frequency components using the HPS algorithm and tabulated the frequency values which have the dominant energy. The result is shown in Figure 3. The distinct frequencies obtained for the signal separated song

in Figure 2 is 400 600 427 420 417 554 418 420 418 625 623 494 416 417 418. Using the singer identification system, we determined that the singer is Nithyasree for the considered sample and then identified the ragam as Sankarabharanam. The details of the Raga identified based on singer and the number of songs for which the raga was identified is given in Figure 4
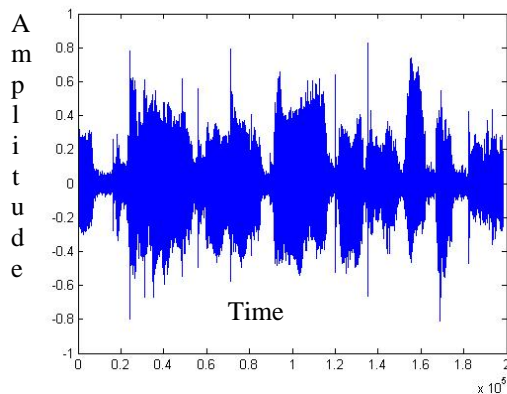

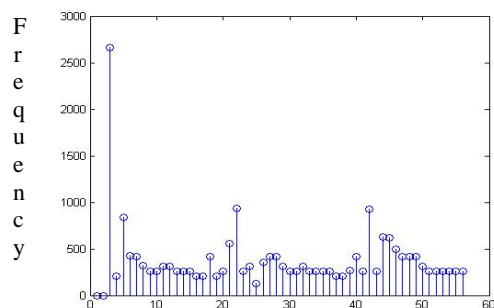Figure 2: Input signal separated and segmented
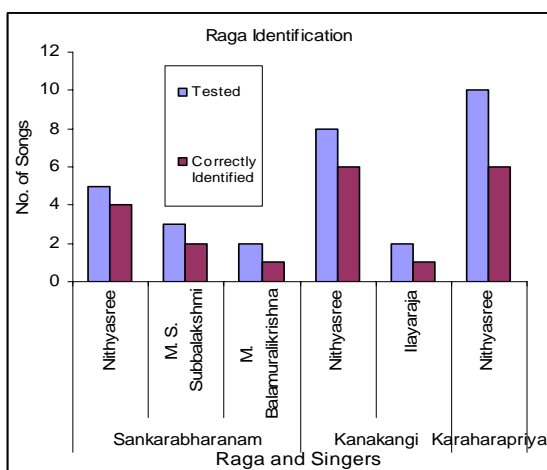

Figure 3: Frequency components

Figure 4: Raga identified based on singers

For the singer Nithyasree, the number of songs for which raga identified correctly was high compared to M.S. Subbalakshmi and Balamuralikrishna because of the fact that Nithyasree's fundamental frequency, which we have assumed is more correct compared to other singers.

For the ragam Karharapriya, Nithyasree's voice gave correct results for 6 songs out of the tested 10 songs. This is because of the fact that, those songs had a better frequency range compared to other songs and hence the result is wrong because of the assumption of the fundamental frequency. The same algorithm was run for Hindustani music and we observed that the algorithm gave correct results for two of the Hindustani Ragas. Since Hindustani music and Carnatic music's note indication is different as explained earlier, some combination which is not present in Hinudstani music is also likely to occur which is to be omitted.

6. CONCLUSION & FUTURE

This work is based on low level signal features for Raga identification. We have tested for three Ragas belonging to two Talams sung by three singers. The system has to be tested for different singers, all melakarta Ragas and ten most frequently occurring talams. Our system has to be improved with a Hidden markov model based approach where we combine the low level features and HMM for better and accurate identification. This system's basic disadvantage is the assumption of fundamental frequency and hence determination of fundamental frequency is our next task. The system has to be modified to use the identified raga to index songs and use raga as a query for music information retrieval

REFERENCES

[1] Prof. Sambamurthy, *South Indian Music*, vol. 4, The Indian Music Publishing House, Madras, India, 1982
[2] Rajeswari Sridhar and T.V. Geetha, "Swara identification of Carnatic music", IEEE Computer Society press, proceeding of ICIT 2006
[3] Rajeswari Sridhar and T.V. Geetha, "Music Information Retrieval of Carnatic Songs Based on Carnatic Music Singer Identification", IEEE conference on Computer Science, ICCEE 2008
[4] Gaurav Pandey, et al "Tansen: A system for automatic Raga identification" IICAI, 2003
[5] Ram Avtar Vir, "Theory of Indian Music" Pankaj Publications, New Delhi, 1999.
[6] Preeti Rao and Anand Raju, "Building a melody retrieval system", NCC 2002
[7] Maarten Grachten et al, "Melody retrieval based on the Implication Realization model", MIREX 2005
[8] Yun-Gang Zhang, "Separation of Voice and Music by Harmonic Structure Stability Analysis", IEEE Proceedings, 2005.
[9] Mark R Every & John E Szymanski, "A Spectral Filtering approach to music signal separation", Proceedings of the 7th International Conference on Digital Audio Effects (DAFx '04), Naples, Italy, October, 2004, pp 197 – 200
[10] Rajeswari Sridhar and T.V. Geetha, "Comparison of signal separation algorithms as applied to Carnatic music", FRSM 2007
[11] Bello, J.P., Daudet, L., Abdallah, S., Duxbury, C., Davies, M., Sandler, M.B., "A Tutorial on Onset Detection in Music Signals", IEEE Transactions on Speech and Audio Processing, 2005
[12] Kyogu Lee, "Automatic chord recognition from audio using enhanced pitch profile", ICMC 2006