WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Music Content Classification using Carnatic Music Qualities

[1]Sourav Verma, [2]Dr.Kiran Hooda

[1] Research Scholar of OPJS University, Churu, Rajasthan

[2]AssociateProfessor, OPJS University, Churu, Rajasthan

Introduction

We live in an era when the Internet is widely used for almost all types of information searches. Text, audio, and video are all examples of information available on the internet. The difficult task of audio and video extraction is critical in the entertainment industry. Speech and music extraction are two types of audio extraction. The number of people looking for music on the internet is growing every day. In the last decade, the availability of Hindustani and Carnatic music on the internet has increased dramatically. Manually providing textual metadata about a specific piece of music is one simple way of searching and retrieving music from the web. However, in order to index music automatically, the music signal must be intelligently processed and represented. Musical elements such as the Raga and the Tala, as well as nonmusical elements such as the Singer, the Instrument, the Genre, and the Emotion, are all present in Indian music. These components can be used as key values in a music information retrieval (MIR) system to perform automatic indexing and subsequent retrieval based on these key values.

The goal of this thesis is to use Carnatic music's unique characteristics in music signal processing. From the first two minutes of the song, which may contain an alaap followed by pallavi or the pallavi alone, a combination of signal processing techniques with Carnatic music specific characteristics is used to intelligently process the music signal in order to understand and determine the content of the given musical piece (Pallavi and Alaap are discussed later).

The following section examines the distinct characteristics of Carnatic music and compares them to those of other musical systems.

Indian Music and Western Music Comparison

The general characteristics of different systems of music in relation to Carnatic music are discussed before the specific characteristics of Carnatic music are discussed.

Tzanetakis and Cook (2002) classify music according to genre or whether it is classical or non-classical. One technique to categorize music is to split it into Western and Indian systems; however, additional global music systems exist, such as Chinese, Indonesian, and Persian. In this thesis, we compare and contrast the Western and Indian music systems. In Western music, musical notes are represented by the letters C, D, E, F, G, A, and B, but in Indian music, they are represented by the swaras S, R, G, M, P, D, and N. Melody and rhythm are two common qualities of Western music, referring to note patterns and beat structure, respectively. In Indian music, notes are arranged in a pre-determined sequence to form Ragas, while the rhythm is indicated by the Tala, which has different beat structures.

Furthermore, the two systems are fundamentally different in terms of rendering style, chord existence or absence, and the amount of intervals per octave. The differences are shown in Table 1.1.

Table 1.1 Comparison between Western and Indian Music

Western Music	Indian Music	Inference
Chords, Orchestration, Harmony Present – Refers to the presence of more than one frequency at a time or one melody at the same time	Chords, Orchestration, Harmony are absent	Only one frequency is present at a time in Indian music
Vertical Arrangement of Notes – overlapping of notes	Horizontal – One Note follows the other	Some Instruments are typical for Indian music to allow for this horizontal arrangement – eg. Veena
Even Tempered system of notes	Just Tempered syste of notes	Any frequency can be used as the starting note in Indian music and the frequency of a note is continuous rather than discrete
Semitones	Microtones / Semitones	Intervals of an octave are flexible
Music is first composed and then the musician plays it Little scope for improvisation	Melody is fixed and then the musician plays it – a lot of improvisation possible	Introduces the concept of Raga and the swaras of the Raga can be improvised to do ornamentation to the Raga
Improvisation in semitones is not possible	Improvisation in Raga is a key to convey Emotion	
Sound is thought of as blocks or pieces of music	Sound is not thought of as blocks of notes, but as a thin wire of flow	
Absence of Gamakas and Meends	Presence of Gamakas and Meends	Gamakas refer to pitch inflexions, where a swara is continuous rather than discrete / Meends is a slower Gamaka.

Because of the disparities between the two musical systems, it is impossible to apply algorithms established for Western music to Indian music and expect identical results. In order to design new algorithms for processing Indian music, characteristics such as the just-tempered note system, the presence of Gamakas, and the presence of microtones in addition to semitones must be taken into account.

Motivation

The closeness between speech and music signals prompted us to apply speech signal processing techniques based on temporal, spectral, and Cepstral aspects to music signal processing. Because segmentation is the initial stage in signal analysis, segmentation algorithms developed for speech or western music might be used to analyse Carnatic music. The just-tempered music system, as well as the influence of Gamakas on pitch, making mapping different frequency components to swaras a difficult undertaking. As a result, identifying a Raga is challenging. The qualities of Carnatic music must be leveraged to obtain spectral and Cepstral properties for non-music component identification. Finally, current indexing methods must be modified to fit the needs of Carnatic Music Information Retrieval.

Another goal of this thesis is to create algorithms that take use of the peculiarities of Carnatic music at each level of the signal processing system. This involves the development of segmentation and feature extraction algorithms that are based on the features of Carnatic music. These traits might lead to the identification of Raga, Tala, Tonic, and Gamakas, which are all hallmarks of Carnatic music. The other elements of music, such as the singer, genre, instrument, and emotion, might also be recognized.

REVIEW OF LITERATURE

Fundamental frequency estimation of the audio signal is a classical problem in signal processing (Camacho and Harris 2008). The estimation of fundamental frequency has been a research topic for many years both for speech and music signal processing. Fundamental frequency is the physical term for pitch (Camacho and Harris 2008). Pitch is defined as the perceptual attribute of sound which is the frequency of a sine wave that is matched to the target sound in a psychophysical experiment. Fundamental frequency is needed in speech signal processing for determining the speaker in Speaker Verification or Recognition systems. The estimation of fundamental frequency is essential in music signal processing in order to determine pitch pattern, range of pitch frequencies, music transcription, and designing music representation systems.

Fundamental frequency is defined as the lowest frequency at which a system vibrates freely. Fundamental frequency is the reciprocal of the time period between the two lowest peak points of a given signal and hence it can also be determined by looking at the time domain representation of the signal to yield the successive lowest peak points. The features that are used for the determination can be classified as Time Domain features, Spectral features, Cepstral features and features that are based on the auditory theory. Many of the algorithms that are available for fundamental frequency estimation of speech and music are based on the estimation of frequency domain features and auditory motivated features.

A fundamental frequency estimation algorithm for speech and music, developed by Doval and Rodet (1993), is based on the evolution of the signal by assigning a probabilistic value to the pseudo-periodic signal. This algorithm is based on a HMM using the estimated spectral features to identify the fundamental frequency of the signal and hence it required lot of training to determine the evolution of the signal.

Maher and Beauchamp (1994), used a two way mismatch procedure for estimating the fundamental frequency of music signals. In this algorithm, fundamental frequency is determined by computing quasiharmonic values for short-time spectra of the input signal. The same value is determined in the neighbouring spectra and then the fundamental frequency is estimated as the least value of the sample input segment considered.

An algorithm was developed by Cheveigne and Kawahara (2002) which is a generalized algorithm for speech and music. It is based on the wellknown autocorrelation method which is in turn based on the model of auditory processing. The steps involve determining the autocorrelation value, correcting the errors in the computed value by determining the difference between the autocorrelation values, normalizing the value of the difference function by estimating the mean value, and iterating this correlation value to determine the fundamental frequency of the input signal. The time taken to correct the errors is very high as it is a generic algorithm for speech and music which did not exploit the characteristics of the signal. The accuracy of the algorithm is also not very high for their sample data.

In one of the algorithms, the fundamental frequency of speech and music signal has been estimated based on spectral comparisons (Camacho and Harris 2008). The average peak to valley distance of the frequency representation of the signal is estimated at harmonic locations. This value is determined at several segments of the input and the distance is estimated between successive average peaks to valley value. From these values fundamental frequency value is estimated as the least distance of the average peak to valley values. This work for fundamental frequency estimation was generic for speech and music. The time complexity of this algorithm is very high in the worst case situation since the distance measure needs to be computed between successive segments for all possible combinations in the input signal.

Many algorithms have also been designed for estimating multiple fundamental frequencies corresponding to the Singers and Instruments (Yeh et al 2005), (Klapuri2003). In the algorithm developed by Yeh et al (2005) a quasi-harmonic model is developed to determine the components of harmonicity and spectral smoothness, after which a score value is assigned for the computed harmonicity value and spectral smoothness and based on which the fundamental frequency is estimated.

RESEARCH METHODOLOGY

MUSIC SIGNAL PROCESSING

Preprocessing, segmentation, feature extraction, model development, and a recognizer to identify the important information are all modules in any signal processing application to recognize the contents of the signal, as mentioned in Chapter 1. In the context of Indian music, efficient algorithms are needed to extract musical and non-musical components from a Carnatic music signal, such as Shruthi, Tala, Swaras, Raga, and Singer, Instrument, Emotion, and Genre.

OVERVIEW OF THE SYSTEM

Contributions were made in this thesis to all of the modules of music signal processing that were built to extract the content. Figure 3.1 depicts the overall system architecture of the different components.

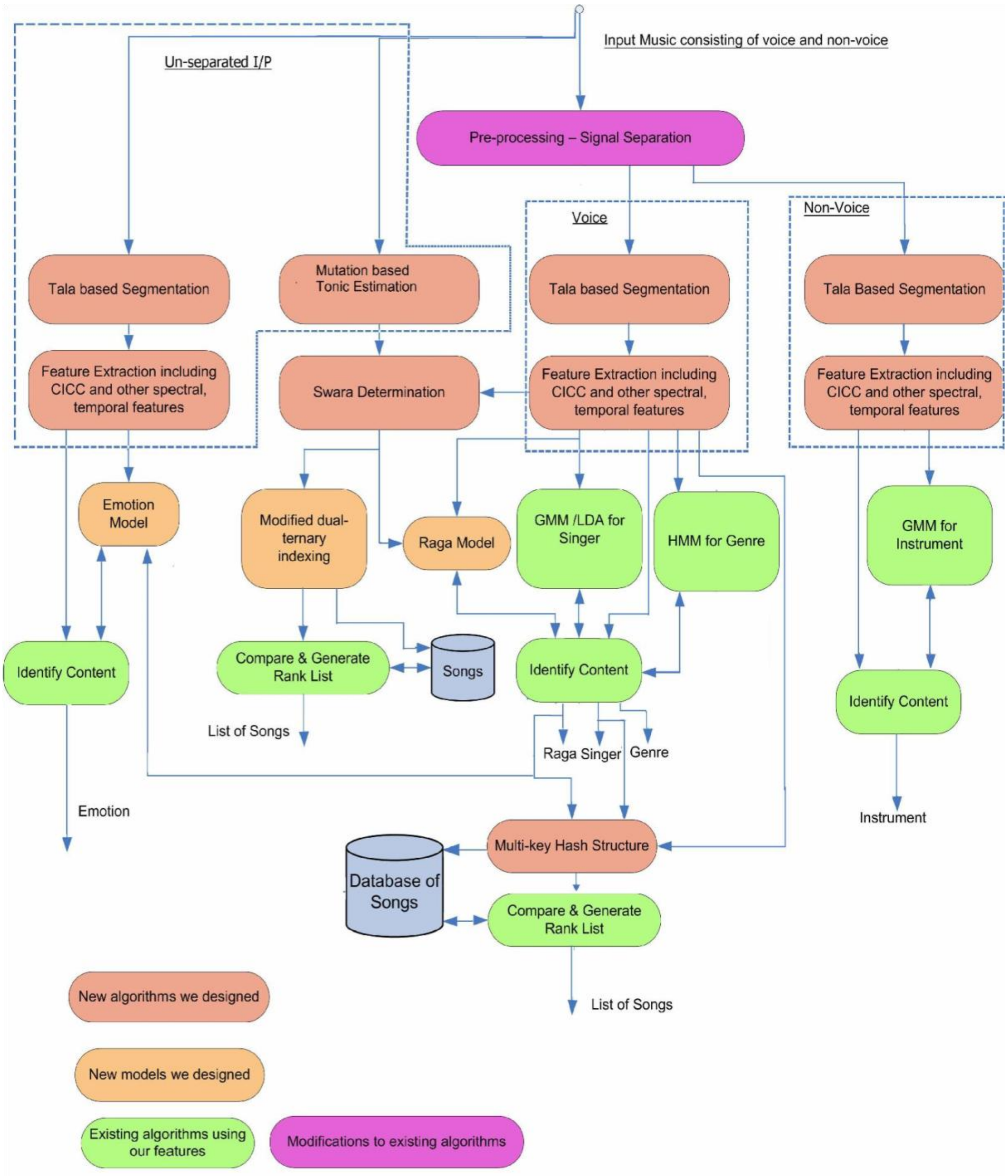

Figure 3.1 Overall System Architecture

Pre-processing, segmentation, feature extraction, model creation, and indexing for music information retrieval have all been improved, as shown in the block diagram. The music from concerts and pre-recorded audio material with a mix of vocal and non-voice (Instrument) components are utilized as input in our work. In essence, the preprocessing module takes this input and separates the voice and nonvoice components. For subsequent processing, any of the three signals, namely un-separated input, voice, and non-voice, or a mix of them, will be employed. The un-separated input, the voice, and the non-vocal signal are all segmented individually using a newly constructed Tala (unique to Carnatic music) based segmentation procedure.

The un-separated input signal is sent into the mutation based tonic estimate module, which employs a newly constructed mutation based method to determine the input signal's tonic. The newly specified Carnatic Interval Cepstral Coefficient is extracted as part of the feature extraction function (CICC). Other signal level parameters, such as spectral flux, centroid, density, energy, frequency values, and Cepstral features, such as MFCC, are computed individually from all three signals to help in the identification of the music components.

The tonic and frequency values are taken from the voice input by the swara determination module. The newly created Carnatic Raga model module determines the Raga based on the discovered swara patterns and the output of the feature extraction module.

The remaining non-music components were found using existing models, but with elements built specifically for Carnatic music. The extracted components are employed in the indexing of a Music Information Retrieval System. The Multi-key Hash structure employs the Spectral Centroid, Spectral Flux, MFCC, CICC, as well as the swara patterns, Raga, and Singer as keys, while the Modified Dual Ternary Indexing module uses the swaras as keys. The next sections go through these modules in detail.

SIGNAL SEPARATION PRIOR TO PRE-PROCESSING

The performances of two music signal separation algorithms were carried out in this study to establish their appropriateness for Carnatic music, as stated in Chapter 2. (Zhang and Zhang 2005; Every and Szymanski 2004). The one provided by Zhang and Zhang (2005) has been determined to be more suitable for Carnatic music since it basically retains the harmonic values of the frequencies. The system is fed voice and non-vocal music sounds to calculate the average harmonic structure value, and a constant value 'd' is utilized to define thresholds to discriminate between voice and non-voice signals based on the spectral peaks in each frame (Zhang and Zhang 2005). This constant value was established by the writers based on the features of Western music, and hence cannot be immediately applied to Carnatic music, which has a rich harmonic content and constantly varies the frequencies between notes (swaras). To extract the voice and non-voice components, the calculation of the threshold constant 'd' was repeatedly changed in this thesis by comparing the output signal with the input.

SEGMENTATION

The swaras must be determined in order to grasp a significant music content Raga, which demands segmentation. Every musical composition is accompanied with a pre-defined Tala, which may be used as a reference to signify the beginning and finish of a song, as detailed in Chapter 1. Tala has a pre-defined pattern, according to Sambamurthy, and each Tala should account for one, two, four, or eight swaras, resulting in different length segments (Sambamurthy 1983). Segmentation is done out in this thesis by using the Tala features of Carnatic music.

The identification of the Tala based on time is the initial stage in our Tala based segmentation. By comparing the signal's time between the onset and offset to the Tala, it may be segmented. The auto-correlation metric is used to merge segments in the case of over-segmentation. Following the segmentation and recombination process, it is assumed that each segment corresponds to a swara. The Tala found as a consequence of segmentation is also connected with the input musical composition as a music component.

EXTRACTION OF FEATURES

Another significant addition of this thesis is the creation of a novel method for determining the tonic of Carnatic music, as well as the creation and determination of a new set of Cepstral features, known as Carnatic Interval Cepstral coefficients, using Carnatic music characteristics.

LDA APPROACH

To overcome the limitations of the Arohana and Avarohana approach for Raga identification, we propose the use of a probabilistic Latent Dirichlet Allocation (LDA) model (Hu 2009), which incorporates additional Raga lakshana parameters to determine the Raga. The LDA is an unsupervised statistical model, which is being used for document classification to determine the underlying topics in a given document, under the assumption that a document contains a random mixture of topics. In this work, we have constructed the LDA for identifying the Raga(s) available in a given input music signal, based on the assumption that the musical piece is a random mixture of notes, and hence, notes map to the words in a topic, and the topics in a document map to the Raga.

Result and Findings

Complexity An examination of the Raga model algorithm

There is an O(n) comparison between the Arohana and Avarohana in terms of algorithm analysis, where 'n' is the number of Ragas in the Raga model. The next level of comparison, depending on the signal parameter, is (n), where four signal parameters must be compared for each Raga. The third level of comparison is (n), in which each of the total 'n' Ragas undergoes another (k) character comparison, resulting in (nk) comparisons, where 'k' is a constant referring to the number of character comparisons that occur during the comparison with the Raga lakshana. As a consequence, the entire running time for the comparison with the Raga model is O(n) + (n) + (nk), resulting in (nk) time, which is greater than O(n) when compared to the Arohana and Avarohana algorithms, and (n) when compared to the signal parameters-only approach. When compared to the other Raga identification techniques, the LDA approach takes longer to calculate the probability value, and hence has a larger temporal complexity.

CONCLUSION

The work we've done on content extraction from Carnatic music, which takes advantage of the genre's unique characteristics, is the first of its kind in this field. Segmentation, feature extraction, and model construction are all major modules of the music processing system that use Carnatic music characteristics in their design. Furthermore, in Carnatic music, a difficult problem in music processing – determining the tonic – was found to be absolutely necessary for Raga and Singer identification. As a result, we devised an on-the-fly tonic determination algorithm based on the biologically inspired concept of neutral mutation. We discovered that Cepstral features needed to be defined based on these two aspects rather than using either the speech specific MFCC or the Western music specific OFCC because Carnatic music components are heavily influenced by tonic and the Carnatic Octave interval. As a result, we created a new set of coefficients called the Carnatic Interval Cepstral Coefficients (CICC) to account for these two factors.

REFERENCES

1. Manoj A. and Kahar "Phenotype frequencies of blood group systems (Rh, Kell, Kidd,Duffy, MNS, P, Lewis, and Lutheran) in blood donors of south Gujarat, India" AsianJournal of Transfusion Science - Vol 8, Issue 1, January - June 2014.
2. Shamsian BS, Arzanian MT, Shamshiri AR, et al. Frequency of red cell alloimmunization in patients with β-major thalassemia in an Iranian referral hospital.Iran J Pediatr. 2008; 18:149-153.
3. Singer ST, Wu V, Mignacca R, et al. Alloimmunization and erythrocyte autoimmunization in transfusion-dependent thalassemia patients of predominantly Asian descent. Blood. 2000; 96:3369-3373.
4. Rabeya Yousuf et al. Incidence of Red Cell Alloantibody among the Transfusion Recipients of University Kebangsaan Malaysia Medical Centre. (Indian J Hematol Blood Transfus (Apr-June 2013) 29(2):65–70).
5. Salama MA, Sadek NA, Hassab HM, et al. Erythrocyte autoantibodies and expressionof CD59 on the surface of red blood cells of polytransfused patients with beta- thalassaemia major. Br J Biomed Sci. 2004; 61:88-92.
6. Hari Krishan Dhawan et al., Alloimmunization and autoimmunization in transfusiondependent thalassemia major patients in Pgimer, Chandigarh 160012 India.2014.
7. Al Joudi F et al. Prevalence and specificities of red cell alloantibodies among blood recipients in the Malaysian state of Kelantan. (Asian J Transfus Sci 2011;5:4245).
8. Spanos TH, Karageorga M, Ladis V, et al. Red cell alloantibodies in patients with thalassemia. Vox Sang. 1990; 58:50-55.
9. Ameen R, Al-Shemmari S, Al-Humood S, et al. RBC alloimmunization and autoimmunization among transfusion-dependent Arab thalassemia patients. Transfusion. 2003; 43:1604-1610.
10. Madeeha Rehan et al. Red Cell Alloantibodies in Multi - Transfused Patients of Beta Thalassaemia Major. Journal of Rawalpindi Medical College (JRMC); 2015; 19(2):152-154.
11. Davari K, Soltanpour MS. Study of alloimmunization and autoimmunization in Iranian βthalassemia major patients. (Asian J Transfus Sci 2016;10:8892).