MUSIC/VOICE SEPARATION FROM INDIAN FOLK MUSIC USING ESTABLISHED TIME FREQUENCY MASKING BASED TECHNIQUE

[1]RAHUL YADAV, [2]SHUBHAM MISHRA, [3]NITIN MUKESH, [4]NEELESH KUMAR

[1,2,3,4] Department of Computer Science and Engineering Technology, National Institute of Technology, Agartala
Email: [1]ram.2010.rahul@gmail.com, [2]shubhammishra5252@gmail.com

Abstract - This paper concentrates on music voice separation from mixture of audio specifically from Indian folk music. It had been seen that recent music/voice separation techniques specifically works efficiently for a particular type of music such as "REpeating Pattern Extraction Technique" is well suited for "pop music" because the music piece consist of stable repeating accompaniment. There are some latest algorithm such as Blind source separation, repeat pattern extraction and more. So this paper investigates the effectiveness of established music voice separation technique on different type of music, the case of Indian folk music was considered as a specific example and it was found that Human-Assisted Time-Frequency Masking method is more efficient for Indian folk music.

Keywords - REPET, STFT, TFM, DFT, IFFT,IPD

I. INTRODUCTION

Music consist of two components: vocal and instrumental. Separation of vocal and instrumental is a point of interest for researchersas well as for commercial users such as DJ's. It is also useful for creating karaoke and extracting melody. There are several approaches for voice and music separation from a mixture of an audio of specific type. So the algorithm used by them comes up with satisfactory results. Here we have proposed an investigation on extraction of vocals and instruments from a various type of music such as Indian classical music, pop music and specially from Indian folk Music using "REpeating pattern extraction [1] and Human-Assisted Time-Frequency Masking method[2]and on comparing this method,it was found that time frequency masking is more accurate and robust on Indian folk music dataset.

II. LITERATURE REVIEW

There are various methods for vocal separation from an audio which are based on pitch extraction when the background is non-repeating. Pitch extraction based approach gives better results but when the background (instruments) in non-repeating it fails to give the accuracy [5]. There is another approach which is suitable for repeating patterns. Our work is fundamentally laid on following research work.In 2013 Rafi and Prado [1]proposed a simple music voice separation method called "REpeating Pattern Extraction Technique" which separates the repeating background from the non –repeating foreground in an audio.This method basically find the repetition in the mixture and then extracts the repeating instruments . The proposed method is very simple,fast but not efficient on variety of music.

In REPET the very first step is to identify the periodically repeating segments, to do this STFT has been applied using half overlapping hamming window then they have derived the spectrogram then autocorrelation is computed. Then they have computed repeating segment model and finally using binary time frequency masking repeating segment has been extracted.
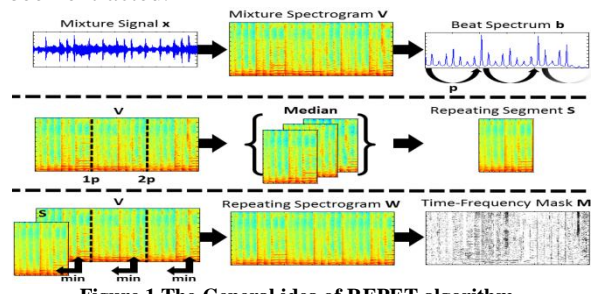

Figure 1 The General idea of REPET algorithm

MarC Vinyes, Jordi Bonada, Alex Loscos[2] proposed a method for singing voice separation and instrument separation from commercial tracks.Here Time-Frequency Masking de-mixing technique is used in which human assisted selection of DFTcoefficient grouped in pan and phase difference. This method does not works efficiently for all tracks. If vocals and instruments are separated in different audio tracks the method gives solution but it's not fully generic and the problem is still not solved.

TFM follows following steps:

(i)   Signal split into overlapped frames of fixed size in time.
(ii)   FFT
(iii)   Binary mask applied
(iv)   IFFT
(v)   Overlap-and-add process.

III. MOTIVATION

REPET algorithm works efficiently, fast but not works much efficiently for non -repeating background. Algorithm is well suited for specific type of music. TFM method proposed by [1] is also restrictedto specific to the track have which have certain mathematical characteristic. Melody

extraction from music is also a scope of music/voice separation as discussed[3].So we were motivated to work with varies type of music sample's such as Indian classical music ,Indian folk music, pop song and were interested about the finding of the results so that we can come up with an algorithm that is efficient irrespective of the characteristic of music type.
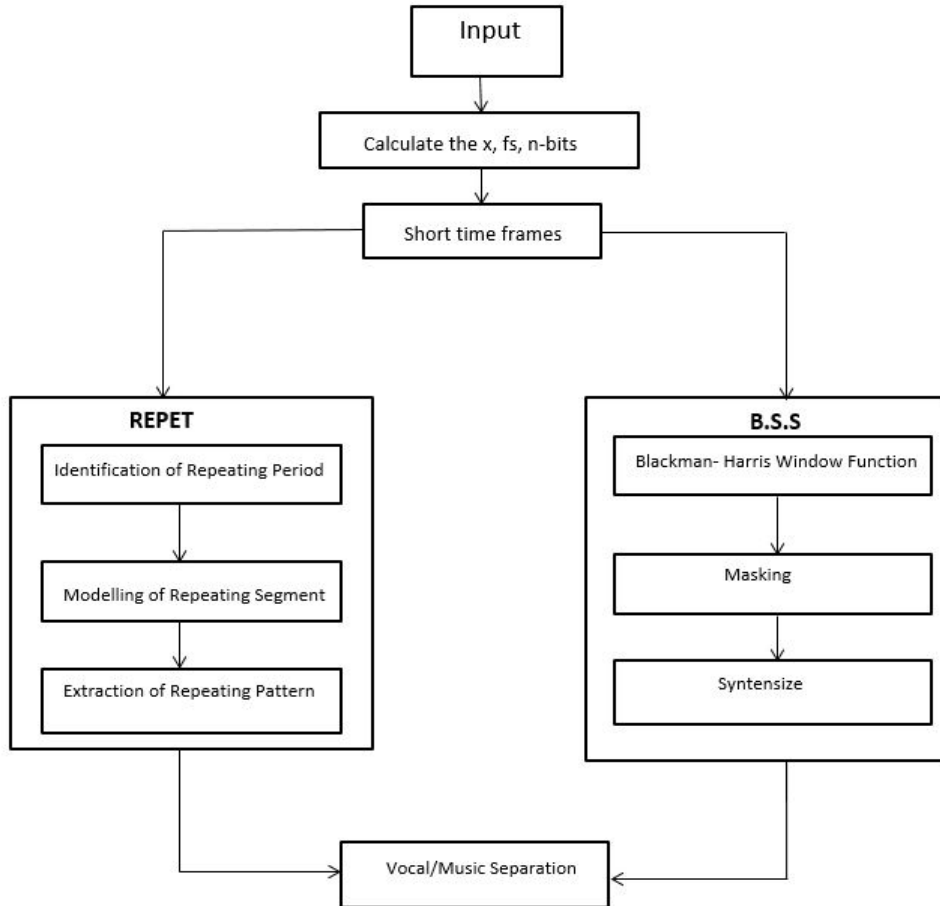
IV. ARCHITETURE
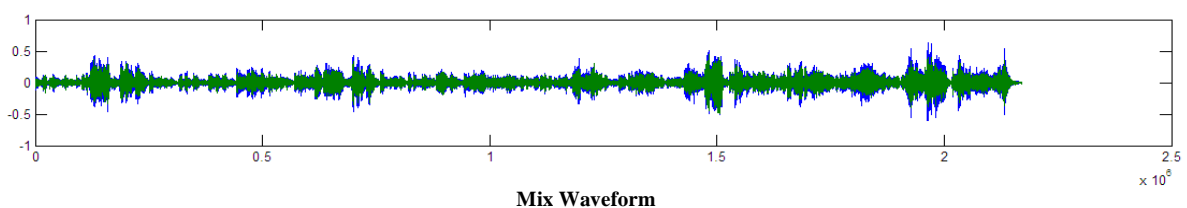

Figure 2 System Architecture
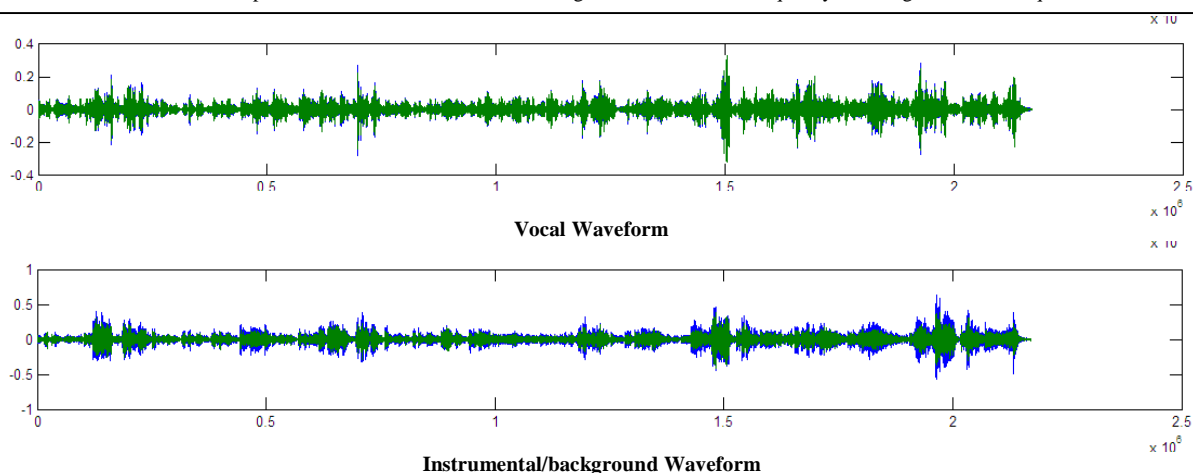
V. DESCRIPTION

We have taken our data set as Indian folk music song, Indian classical song and pop songs. At first according to our system diagram we have computed sampling frequency, signalvector, number of bits sampled. Then as our diagram show we have divided the signal into short time frames. Now we have applied two methods which fundamentally based on time frequencymasking. As per REPET algorithm, given a mixture signal x, we have calculated Short-Time Fourier Transform X, using half-overlapping Hamming windows of N samples. Now we have estimated the repeating time period then we have computed repeated segment. At last we have applied binary time frequency masking for extracting repeating pattern.
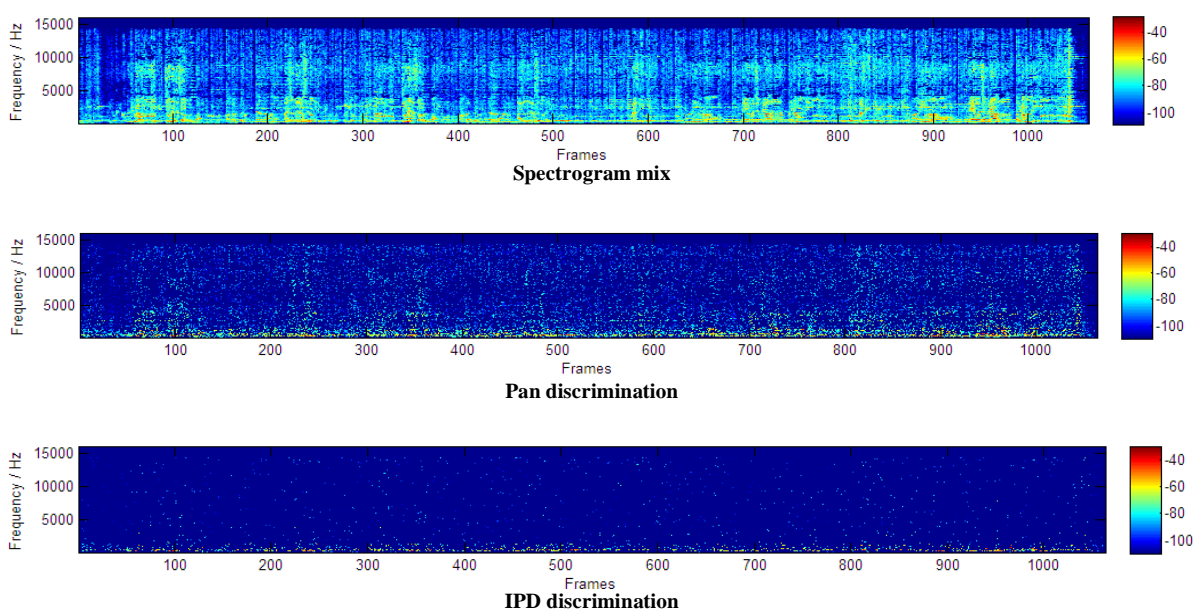
In our second method we have calculated FFT by using Blackman-Harris window function.After that we have calculated FFT of the signal. Then we have performed binary masking.
Post to that we have applied IFFT and finally we have overlapped frame using triangular window.

VI. EXPERIMENTAL RESULT

Mix Waveform

Vocal Waveform

Instrumental/background Waveform

a) Waveform

Spectrogram mix

Pan discrimination

IPD discrimination

b) Spectrogram

Figure (a) is the result of the Indian folk music sample waveform recovered by time frequency masking method and figure (b) is the spectrogram of the recovered vocal andinstrumental/background. When we listen to the recovered vocals and background using human assisted time frequency masking and on the basis of graphical representation we perceive them highly accurate in case of Indian folk music butusing REPET accuracy is very less.

CONCLUSION

This paper introduced the implementation of TFM based approach of music voice separation specifically from Indian folk music. In an order to find a method which is irrespective of the music type, we have applied time frequency based masking on data set of Indian folk music, Indian classical music and pop songs. It has been seen that TFM based approach is working efficiently for Indian folk music samples and extracted vocal and instrumental are highly accurate, even though the background is not stable.It has also seen that in case of repeat method which is also TFM based approach results are not efficient at all because this method requires stable repetition of background. So TFM based approach can be more useful for producing more commercial de -mixing and music information retrieval purpose.

ACKNOWLEDGEMENT

REFERENCES

[1] Zafar Rafii, Student Member, IEEE, and Bryan Pardo, Member, IEEE (2013) " REpeating Pattern Extraction Technique (REPET): A Simple Method for Music/Voice Separation"IEEE transactions on audio, speech, and language processing, vol. 21, no.1, pp. 71-82.

[2] MarC Vinyes, Jordi Bonada, Alex LoscosPompeu Fabra University, Audiovisual Institute, Music Technology Group, Barcelona, 08003, Spain." Demixing Commercial Music Productions via Human-Assisted Time-Frequency Masking" Convention Paper Presented at the 120th Convention 2006 May 20–23 Paris, France.

[3] Vishweshwara Rao and Preeti Rao"Vocal Melody Detection In The Presence Of Pitched Accompaniment Using Harmonic Matching Methods"Proc. of the 11th Int. Conference on Digital Audio Effects (DAFx-08), Espoo, Finland, September 1-4, 2008.

[4] J. Salamon and E. Gómez, "Melody Extraction from Polyphonic Music Signals using Pitch Contour Characteristics",IEEE Transactions on Audio, Speech and Language Processing, 20(6):1759-1770, Aug. 2012.

[5] Pooja Gautam1and B S Kaushik ,"Voice & Music Pattern Extraction: A Review" 2nd International Seminar On "Utilization of Non-Conventional Energy Sources for Sustainable Development of Rural Areas ISNCESR'16 17th & 18th March 2016.

★ ★ ★