

# Human Voice and Background Music Separation Python App.

Prajakta Rode<sup>1</sup>, Omkar Tanpure<sup>2</sup>, Nishigandha Panditrao<sup>3</sup>, Prof. Deepali Patil<sup>4</sup>  
<sup>1, 2, 3, 4</sup>Computer, SPPU, India

**Abstract:** Alteration of singing voice and music is an interesting research topic since singing voice contains large information, such as melody, singer's characteristic, lyrics, emotion, etc. All of these sources in singing voice are useful for music knowledge retrieval, singer identification, melody extraction, audio content analysis, or even karaoke gaming. At the same time, it is also a challenging topic because existing methods are still not so practical. Repetition is a special characteristic of music. Most songs have their own iterations. Support structures over which the singers lay varying vocals on them. This work studies the repeating structure of music and implements the algorithm based on the repeating pattern of the music background. Using Robust Principal Component Analysis, the singing voice from music has its advantage of being simple, blind and automatic.

**Keywords:** Music, Extraction, Abundant, Karaoke

## I. INTRODUCTION

Music is the path to a spirit just as otherworldly and enthusiastic singing. Individuals lose numerous buildings when they sing so anyone might hear, and regularly, this opens the boundaries among hearts and brains.

Regardless of whether a voice has no expert procedure, singing is as yet an extraordinary encounter. Karaoke has breathed life into back an old extraordinary custom of singing and in this way today we can sing, yet additionally have an awesome side interest. These days a karaoke box is the most famous kind of karaoke setting. A karaoke box is a little or medium-sized room containing karaoke hardware leased constantly or half-hour, giving an increasingly private climate. Karaoke settings of this sort are regularly committed organizations, some with numerous floors and an assortment of luxuries including sustenance administration, yet inns and business offices now and again give karaoke boxes also.

## II. PROPOSED SYSTEM

A Music source separation would be utilized to demonstrate the separation of any particular .mp3 song into 2 .mp3 files containing one of singer and other of its background music (Karaoke). Generally, they are made for Generating the Karaoke song which would be available for anyone to utilize without any paid subscription or service & internet connection. In any case, it can likewise be made for web applications, Mobile apps etc.

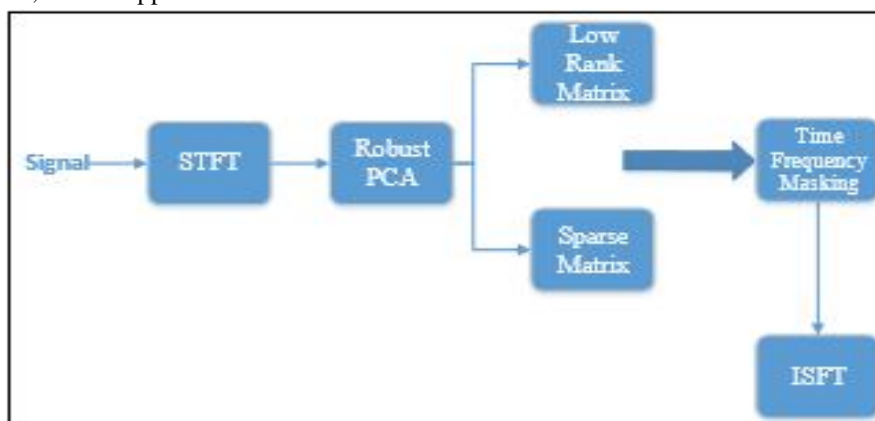


Fig:1 Block Diagram

The below given diagram flow depicts the flow of the music signal of .mp3 format which uses the STFS, RCPA and then later on STFS to obtain the processed .mp3 files containing human or singers voice and background music.

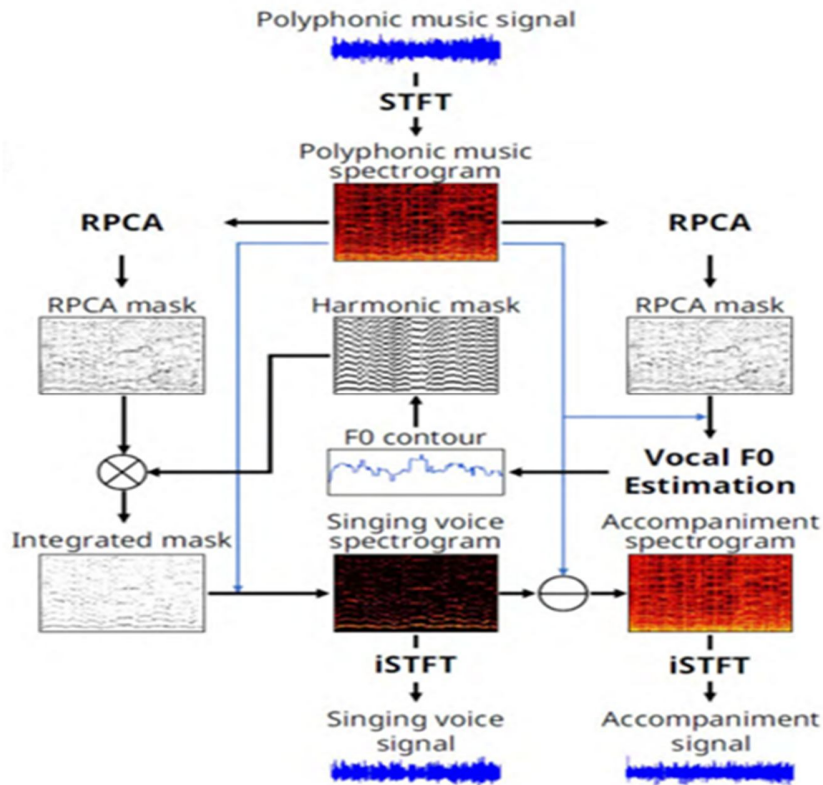


Fig:2 Diagram Process Flow

### ILLINDENTATIONS AND EQUATIONS

#### A. Robust Principal Component Analysis (RPCA)

1) Decomposes M into a low rank matrix L plus a sparse matrix S

$$\min \|L\|_* + \lambda \|S\|_1$$

$$\text{s.t. } L + S = M$$

2) L1 norm:  $\|S\|_1 = \sum_{ij} |S_{ij}|$  (Sum of absolute values; a convex relaxation of l0 norm)

3) Nuclear norm:  $\|L\|_* = \sum_i \sigma_i(L)$  (Sum of singular values; a convex relaxation of the rank of the matrix)

4) Minimize Subject to  $\|L\|_* + \lambda \|S\|_1 + \frac{1}{2\tau} \|M - L - S\|_F^2$  LaGrange.

$$\mathcal{L}(L, S; Y) = \|L\|_* + \lambda \|S\|_1 + \frac{1}{\tau} \langle Y, M - L - S \rangle + \frac{1}{2\tau} \|M - L - S\|_F^2$$

$$\arg \min_L \mathcal{L}(L, S, Y) = \mathcal{D}_\tau(M - S + Y)$$

$$\arg \min_S \mathcal{L}(L, S, Y) = \mathcal{S}_{\lambda\tau}(M - L + Y)$$

6) Easy to minimize over L and S respectively

$$\mathcal{S}_\tau[x] = \text{sgn}(x) \max(|x| - \tau, 0)$$

Scalar Shrinkage:

7) Component Wise Thresholding  $\mathcal{S}_\tau(X)$

8) Singular Value thresholding  $\mathcal{D}_\tau(X)$

$$\mathcal{D}_\tau(X) = U \mathcal{S}_\tau(\Sigma) V^* \quad X = U \Sigma V^*$$

**IV. TABLES**

In this strategy, inspectors are approached to assess the detachment nature of the performing voice and foundation music pieces acquired by REPET, RPCA and the proposed technique. Analysts tune in to the pieces by utilizing exceptional earphones in a peaceful room and they are permitted to tune in to the pieces the same number of times as they wish. They are mentioned to score each bit of music from 1 to 5 dependent on the appraisals recorded in Table I.

The test was led with 18 unique audience members to look at REPET, RPCA and the created strategy by the MOS test. Math normal of the scores given by all audience members for each bit of music is acquired as given in (13):

$$MOS = \frac{\sum R_n}{N}$$

$$n=0 \text{ N (13)}$$

where N speaks to the quantity of analysts, R<sub>n</sub> speaks to the individual MOS for the particular piece, which is either the isolated ambient melodies or the performing voice. The outcomes of the MOS test yield of the proposed technique and other strategies are appeared in the Table II.

At the point when MOS results are inspected, it is seen that the proposed strategy gives preferred outcomes over other two techniques

**TABLE I. RATING SCALE OF MEAN OPINION SCORE**

Mean Opinion Score (MOS)	
MOS	Rating
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

**TABLE II. RESULT OF THE MOS TEST**

Music Piece	RPCA		REPET		Proposed Method	
	Vocal	Music	Vocal	Music	Vocal	Music
Ani_1_07	2.12	3.06	2.41	2.53	2.88	3.24
Kenshin_2_05	2.24	3.00	2.47	2.59	2.59	3.06
Leon_8_10	2.18	2.76	2.47	2.47	2.94	3.00
Titon_2_09	2.88	3.06	2.53	2.76	3.00	3.59
Bobon_5_06	2.12	2.35	2.41	2.00	3.24	2.53

**V. CONCLUSION**

In this paper, we depict another technique for the programmed partition of vocal and music parts of melodies. The proposed strategy consolidates the Robust Principal Component Analysis (RPCA) technique so as to utilize the upsides of the one strategy to develop a partition technique with a higher division exhibition. In this technique, the detachment of performing voice and mood melodies is performed dependent on the any song of type mp3. In the proposed strategy, RPCA is utilized is intended to lessen the mood melodies in the performing voice yields of music. To test the execution of the proposed technique, a particular mp3 file is utilized, which utilizes various Libraries created for Python. Sound tracks for the tests are chosen from the local storage of Desktop. It has been appeared that the proposed strategy isolates performing voice and ambient melodies superior to RPCA techniques easily with Python on mp3 as compared to .wav with MATLAB.



#### REFERENCES

- [1] 2018 5th International Conference on Electrical and Electronics Engineering- Music/Singing Voice Separation Based on Repeating Pattern Extraction Technique and Robust Principal Component Analysis Sait Melih Doğan, Özgül Salor Department of Electrical and Electronics Engineering Gazi University, Faculty of Engineering Ankara, Turkey e-mail: [saitmelihdogan@hotmail.com](mailto:saitmelihdogan@hotmail.com), [salordurna@gazi.edu.tr](mailto:salordurna@gazi.edu.tr)
- [2] Z. Rafii and B. Pardo, "Repeating pattern extraction technique (REPET): a simple method for music/voice separation," *IEEE Transl. on audio, speech, and language processing*, vol. 21, no. 1, pp. 73-84, January 2013.
- [3] REPEATING PATTERN EXTRACTION TECHNIQUE (REPET) A SIMPLE METHOD FOR MUSIC/VOICE SEPARATION M. Joel Dixit, K. Narender
- [4] Separation of Singing Voice from Music Background: Harshada Burute Research Student, Electronics Department AISSMSs, IOIT, Pune, India, P.B. Mane, PhD Principal AISSMSs, IOIT, Pune, India
- [5] Study of Algorithms for Separation of Singing Voice from Music Madhuri A. Patil<sup>1</sup>, Harshada P. Burute<sup>2</sup>, Kirti Malini B. Chaudhari<sup>3</sup>, Dr. Pradeep B. Mane
- [6] 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS) Auditory mask estimation by RPCA for monaural speech enhancement Year: 2017, Volume: 1, Pages: 179-184 DOI Bookmark: 10.1109/ICIS.2017.7959990