Decomposition of music signal into multiple instruments

Deep Pujara, Trivedi Jay
Nirma University, Electronics and Communication Department, Ahmedabad, India

Abstract. In the present project report the procedure, results and conclusion of the simple two step algorithm to extract the instruments is given. This project tries to extract five commonly used instruments for production of pop western as well as Indian Bollywood songs. Decomposition of Music signal into multiple instruments is still in the research phase and is considered to be very complex challenge.

Keywords: Fourier transform, bandpass filter,

1. Introduction

Over the past few years music signal processing is considered to be a mandatory aspect in the field the of music production. From recording songs in studio to streaming it on different online platforms, from exploring new genres, artists to protecting the rights of music creators, with everything regarding music and its production being now completely dependent on the computers. Source Separation is one such challenge in this domain. The applications of extracting the various instruments from the music can be found in karaoke systems, beat matching, mix taping, mashups etc. for DJs, and many more. In this project the problem is approached to be solve by first removing the vocals of the song and then filtering out the portion of the song which contains the frequency range of given the instrument.

2. Vocal Removal Using Stereo Method

For the first step to extract out the content of a given instrument we tried to remove the vocals. As vocals contain a huge and almost the coinciding frequency spectrum with almost every instrument, especially string-based instrument it is necessary to remove that in the beginning. For the same we have applied the stereo method. Vocal removal by stereo method is very simple yet an effective method. This method includes subtracting the right channel audio from the left channel audio. Figure 2.1 shows the dual audio input which is a song by Green Day called 21 Guns. Separate left and right audio signals are represented in Figure 2.2 The reason behind removal of majority of vocal content is that the vocal content is present equally on both the channels unlike the instruments. These removed majority of vocal content from audio, and somewhere it added an echo effect as well. That can be further removed by using echo removal techniques. The output after using this method is mono channel which again was converted into dual channel by copying it in the second channel. Figure 2.3 shows the unvoiced signal which is mono channel signal

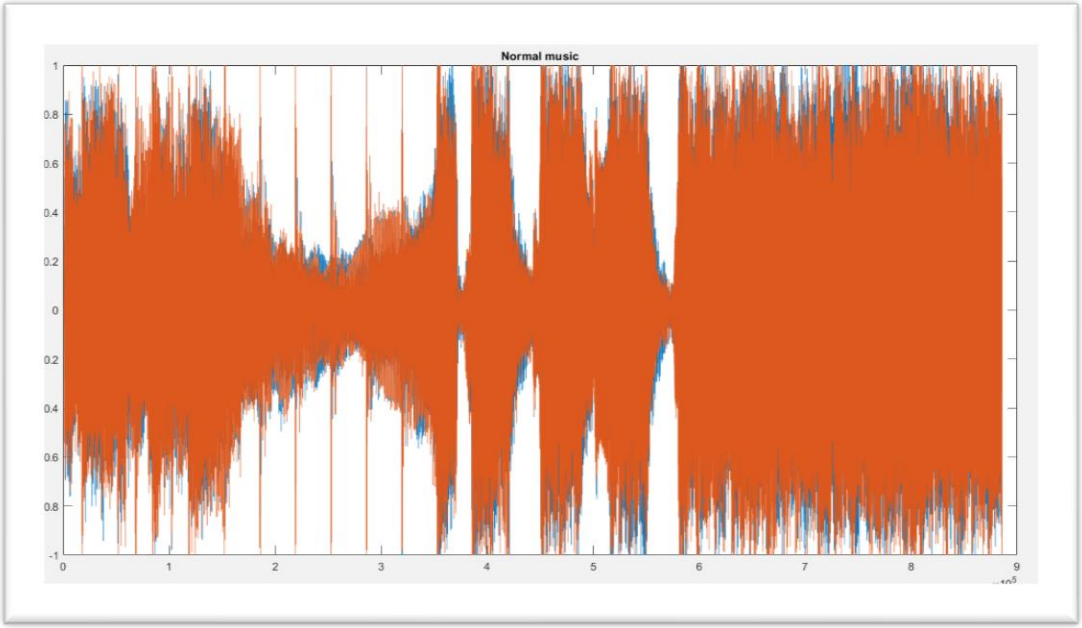
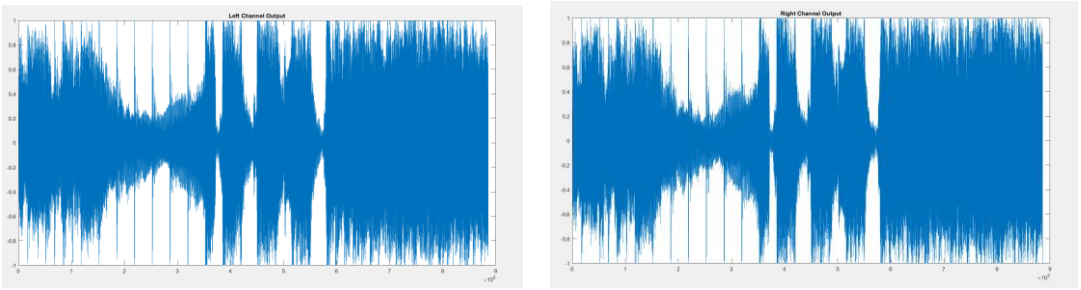


Figure 2.1 ‘21 Guns’ Audio Signal



(a) (b)

Figure 2.2 (a) Left Channel , (b) Right Channel

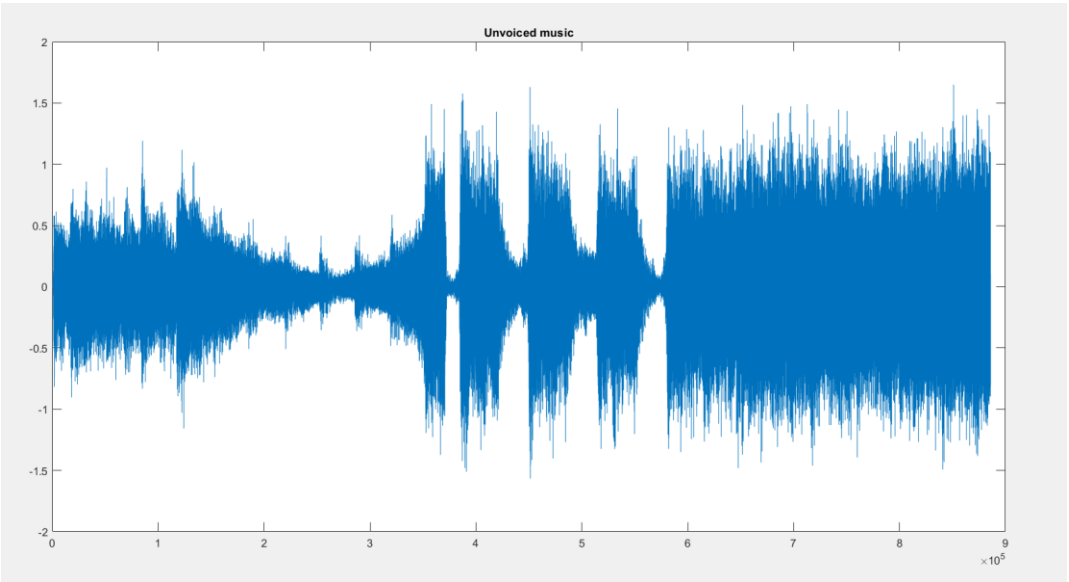


Figure 2.3 Unvoiced Signal

3. Extracting Instruments

The audio with the vocal content removed is then passed through a simple bandpass filters. The cut off frequencies of these bandpass filters are according to the instruments frequency range. For this, conversion of the time domain signal to frequency domain signal was done which is shown in the Figure 3.1. Fourier Transform of the Output signal when passed through the bandpass filter corresponding to violin is shown Figure 3.2. After filtering out the unwanted frequencies again a conversion of frequency domain signal to time domain signal is carried out to get the audio signal back. Figure 3.3 shows the output audio signal.

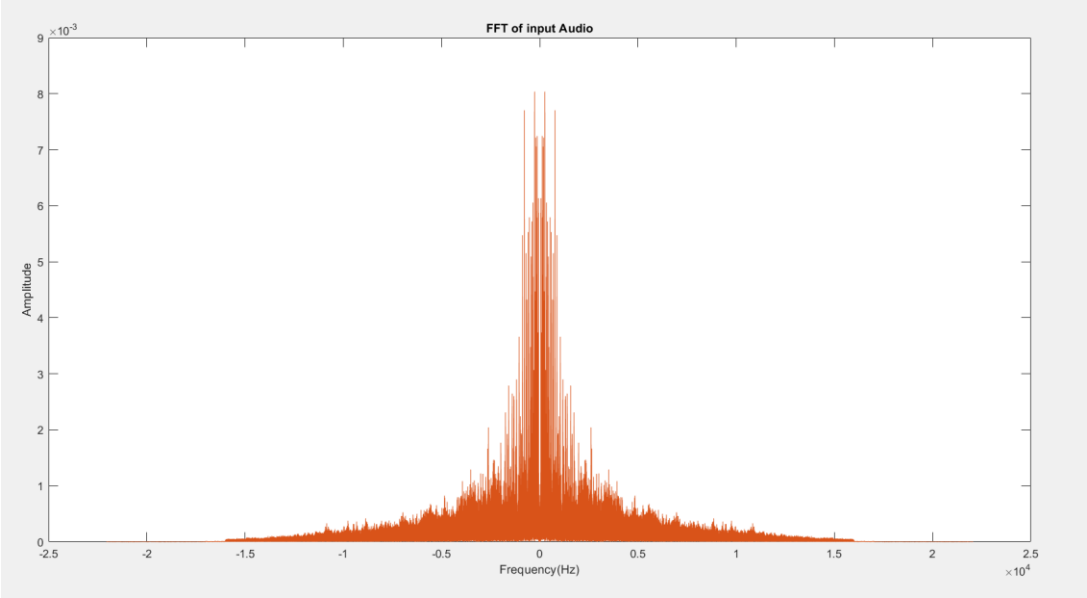


Figure 3.1 FFT of Unvoiced Signal

The instruments we tried to extract and their frequency response are given below.

Sr. No	Instrument Name	Lower Frequency (Hz)	Higher Frequency (Hz)
1	Human Voice	350	1500
2	Drum	0	210
3	Guitar	400	1200
4	Piano	1200	2000
5	Violin	350	660
6	Tabla	3000	4000

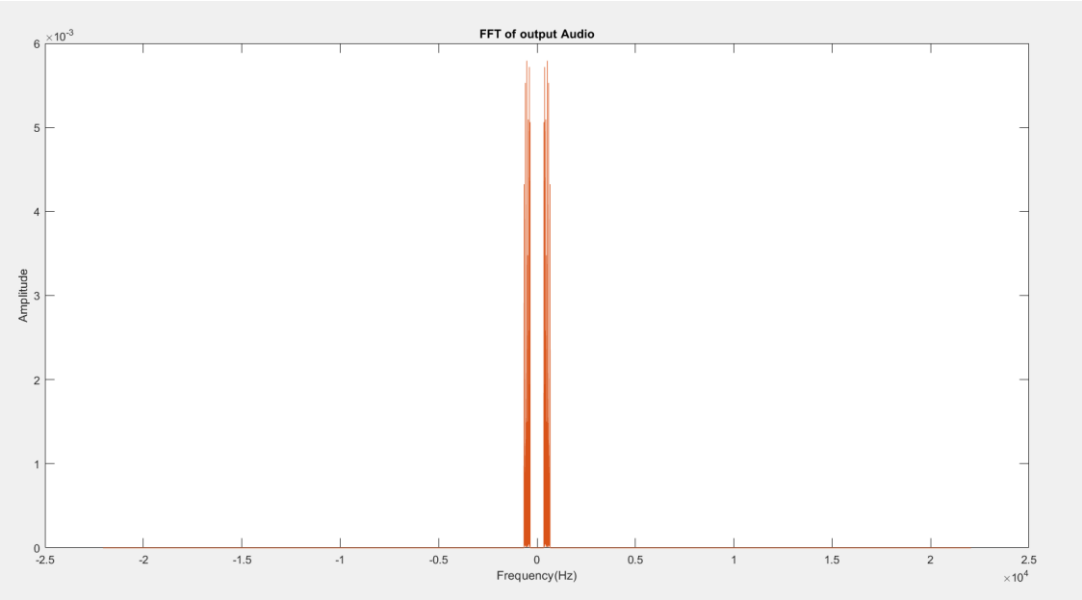


Figure 3.2 FFT of Output Signal

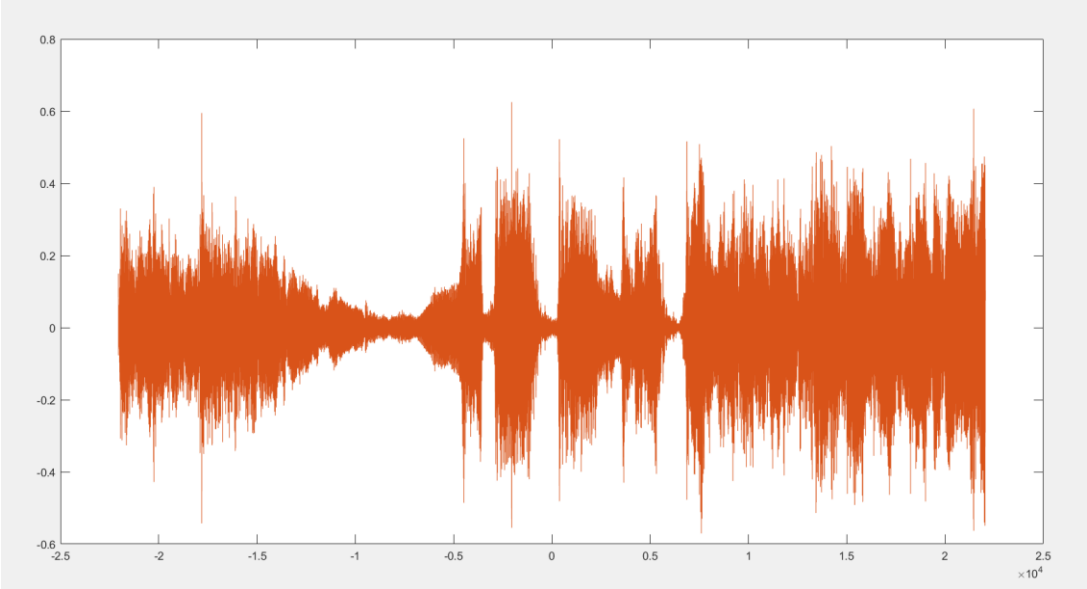


Figure 3.3 Output Signal

4. Acknowledgement

We would like to acknowledge the efforts of Professor Rutul A Patel and Dr. Bhupendra Fataniya, Nirma University for their invaluable guidance and motivation behind completion of the Project. Majority of the project execution was made possible on the basis of their domain knowledge and constant supervision.

Conclusion

Source Separation in itself is major field of research and as the present work is being done with the logic of the first intuition the result we found are slightly unsatisfactory. Though for extracting instruments it is found that vocal removal improves the results because of broad frequency range of it. Another thing to observe is that an instrument is never found to have an independent frequency range i.e. guitar and violin have their contents overlapping leading to improper removal of specific instrument. The present work is capable enough to turn a song into its instrumental version which can be used for karaoke competitions or entertainment.

References

1. Bryan Pardo, A. L. (2018). *Applying source separation to music,Audio source separation and speech Enhancement.* Wiley.
2. Ced. (19, 10 1). *Youtube.* Retrieved from Youtube: https://www.youtube.com/watch?v=Vitu0up63Qw&t=8s
3. *Mathworks*. (n.d.). Retrieved from Matlab Help: https://in.mathworks.com/help/matlab/ref/fftshift.html
4. painkra, D. (2019, 2 19). *Youtube.* Retrieved 10 1, 19, from Youtube.
5. V. Valimaki, M. I. (2004). Decomposition and modification of musical instrument sounds using a fractional delay allpass filter. *Proceedings of the 6th Nordic Signal Processing Symposium*, 208-211.

Caption List