

Separation of Singing Voice and Music

Tengli Fu (tf236@cornell.edu) Advisor: Dr. Bruce R. Land
School of Electrical and Computer Engineering, Cornell University

Overview

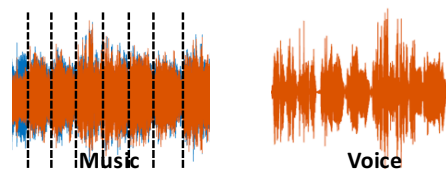
Separation of singing voice and music is an interesting research area since singing voice contains abundant information, such as melody, singer's characteristic, lyrics, emotion, etc. All of these resources in singing voice are useful for music information retrieval, singer identification, melody extraction, audio content analysis, or even karaoke gaming.

Motivation

There are many ways to do separation:

- Non-negative matrix factorization;
- Robust principal component analysis;
- Predominant pitch detection.

However, repeating pattern is a significant difference between background music and foreground voice.

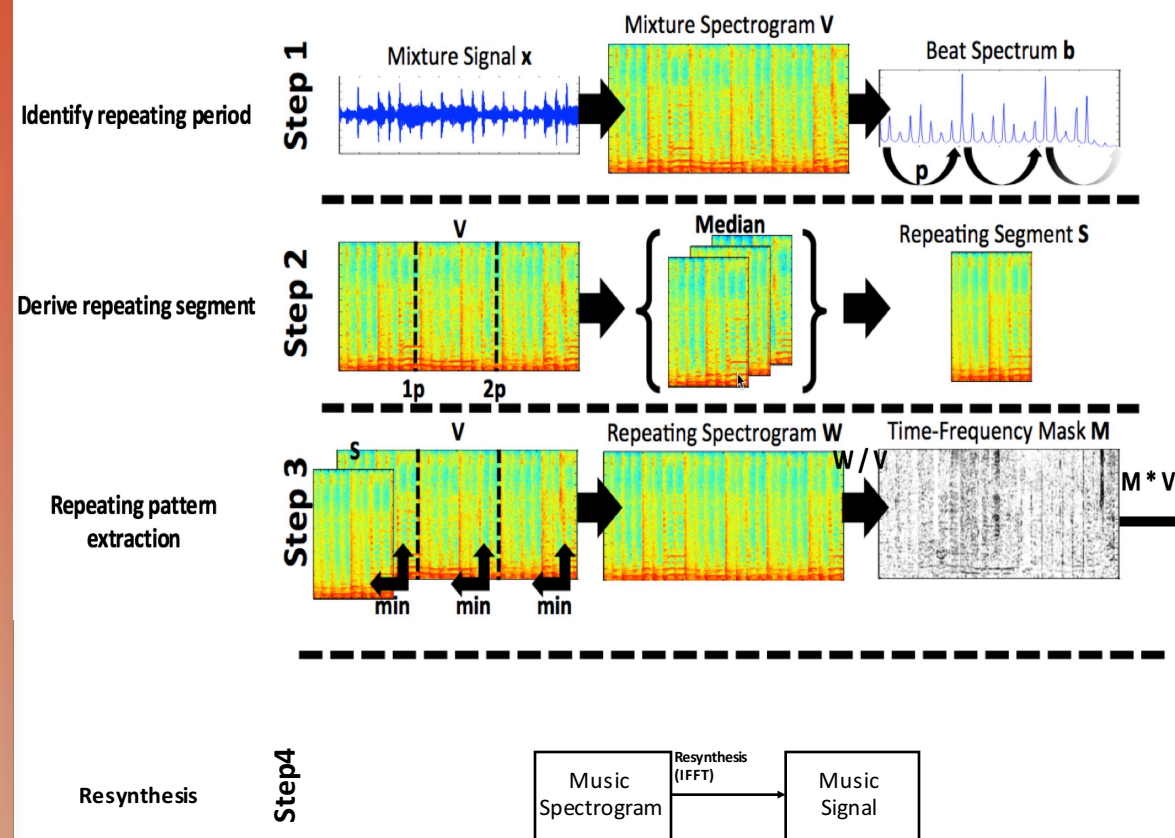


Advantage:

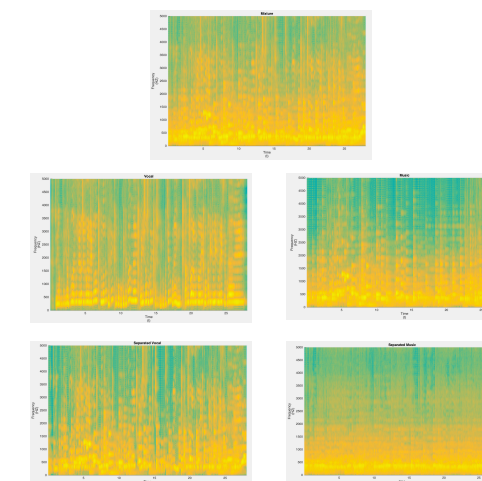
- This algorithm does not rely on complex frameworks of the audio;
- This algorithm could reply to any audio;
- This algorithm is simple, fast, blind and completely automatable.

Implementation

Identify the repeating period of mixture and use it to construct repeating spectrogram, with which we could do separation through time-frequency mask.



Result



$$E \propto \frac{1}{N} \sum_{i=1}^N V_i^2$$

$E_{mix} = 0.0106$, $E_{vocal} = 0.0032$, $E_{music} = 0.0073$
 $E_{vocal-s} = 0.0043$, $E_{music-s} = 0.0022$

Conclusion

1. This algorithm could find the repeating pattern of background music and separate singing voice.
2. From the separated vocal signal, we could clearly hear that music is filtered.
3. This algorithm still assigns some music in separated vocal.
4. Only parts that have highly repeating pattern of music get separated.
5. This algorithm is fast and automatable.