Blind Source Separation of audio files using Independent Component Analysis

Aldrin C. Racusa
Department of Physical Sciences and Mathematics
College of Arts and Sciences
University of the Philippines Manila
Manila, Philippines
acracusa@up.edu.ph

Lorenz Timothy Barco Ranera
Department of Physical Sciences and Mathematics
College of Arts and Sciences
University of the Philippines Manila
Manila, Philippines
lbranera@up.edu.ph

***Index Terms*—Independent Component Analysis, Audio file, WAV, Blind Source Separation, Machine Learning**

I. Introduction

Blind signal separation, also known as blind source separation, is the classification of a set of source signals from a set of one big mixed signals, without the help of information about the source signals or the mixing process. [1]

A common application of Blind Source Separation is the *Cocktail Party problem* where given a situation that suppose you're in a cocktail party, there is a a lot of people talking at the same time. You have multiple microphones picking up mixed signals, but you want to isolate the speech of a specific person. Blind Source Separation can be used to separate the individual sources by using mixed signals. And when there is a presence of noise, optimization criteria need to be used in order to remove it. [1]

One method to implement the Blind source separation is by using the Independent Component Analysis. ICA works by separating multivariate signals into additive subcomponents. [2]

The aim of this study is to implement Blind Source Separation using Independent Component Analysis in two audio files (WAV) mixed together in order to produce the two different audio source files.

II. Dataset

The dataset used in this paper is two audio file (WAV) with the title: *Kathang Isip by Ben&Ben* and *Perfect by Ed-Sheeran*.

III. Independent Component Analysis

Independent Component Analysis is a statistical and computational technique for uncovering hidden factors that underlie sets of random variables, measurements or signals.

ICA defines a generative model for the observed multivariate data, which is typically given as a large database of samples. In the model, the data variables are assumed to be linear mixtures of some unknown latent variables, and the mixing system is also unknown. The latent variables are assumed non-gaussian and mutually independent, and they are called the independent components of the observed data. These independent components, also called sources or factors, can be found by ICA. [3]

The information investigated by ICA could start from a wide range of sorts of utilization fields, including computerized pictures, report databases, financial pointers and psychometric estimations. Much of the time, the estimations are given as a lot of parallel flags or time arrangement; the term blind source separation is used to characterize this problem. Common examples are mixtures of simultaneous speech signals that have been picked up by several microphones, mind waves recorded by different sensors, meddling radio signs touching base at a cell phone, or parallel time series acquired from some industrial procedure. [3]

IV. Methodology

To read the WAV files in python, the package *wave* was used.

In order to avoid Memory Error in python, the audio files have been cut into segments with 1min / 1min and 30 seconds long. And to combine the two audio files, A function zip (formerly known as *izip* from the *itertools* package) was used. *zip* returns an iterative tuple of given set of data. We then store the returned tuple in a list.

And to implement the Independent Component Analysis, we used *FastICA* function from the decomposition module of *sklearn* package with the parameter of *n_components* = 2 which then returns 2 arrays. And that two arrays returned are the independent components.

To visualize and to save the returned values as a WAV file, we used the *matplotlib* package for visualization, and *scipy.io* for saving the returned arrays into a WAV file.

The following figures shows the visualizations of the initial independent audio files

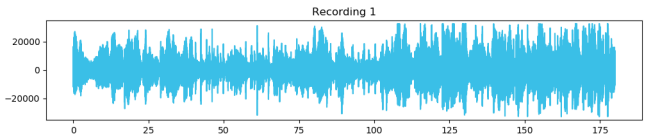

Fig. 1. 0 - 1:30min of audio 1 which is Kathang Isip by Ben&Ben

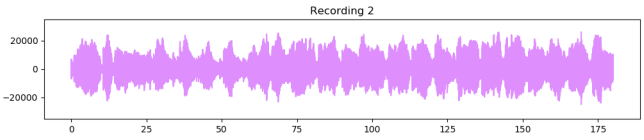

Fig. 2. 0 - 1:30min of audio 2 which is Perfect by Ed-Sheeran

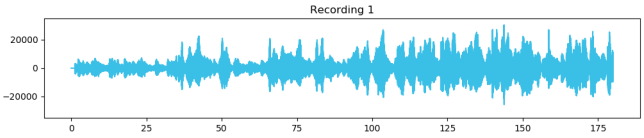

Fig. 3. 1:30 - 3min of audio 1 which is Kathang Isip by Ben&Ben

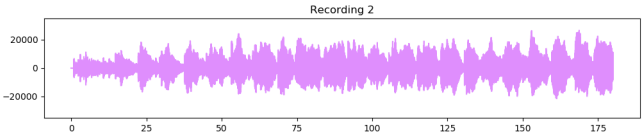

Fig. 4. 1:30 - 3min of audio 2 which is Perfect by Ed-Sheeran

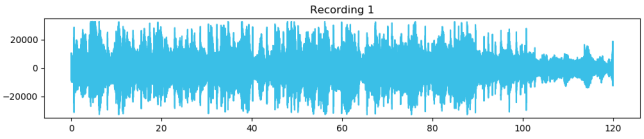

Fig. 5. 3 - 4min of audio 1 which is Kathang Isip by Ben&Ben

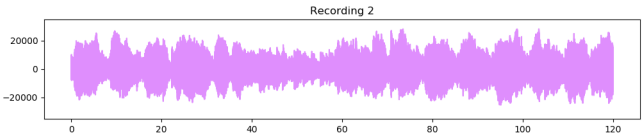

Fig. 6. 3 - 4min of audio 2 which is Perfect by Ed-Sheeran

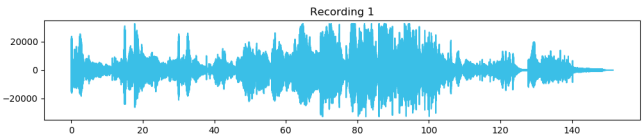

Fig. 7. 4min - end of audio 1 which is Kathang Isip by Ben&Ben

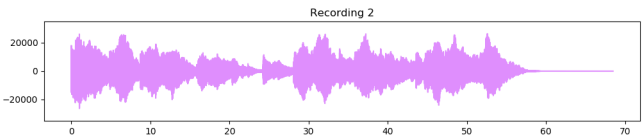

Fig. 8. 4min - end of audio 2 which is Perfect by Ed-Sheeran

V. Results and Discussion

After performing Independent Component Analysis on each tuples of audio files, The following figures are the results of *FastICA* which are then saved as an audio file.

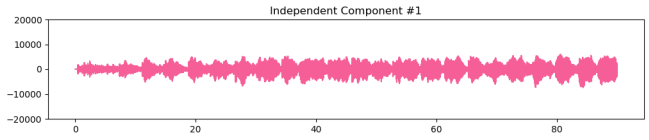

Fig. 9. 0 - 1:30min of Independent Component 1 which is Kathang Isip by Ben&Ben

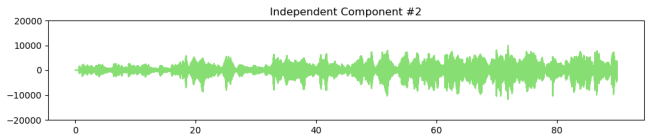

Fig. 10. 0 - 1:30min of Independent Component 2 which is Perfect by Ed-Sheeran

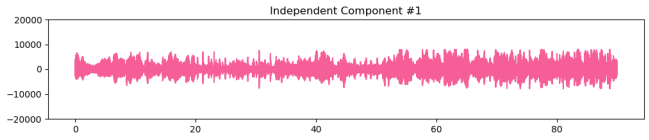

Fig. 11. 1:30 - 3min of Independent Component 1 which is Kathang Isip by Ben&Ben

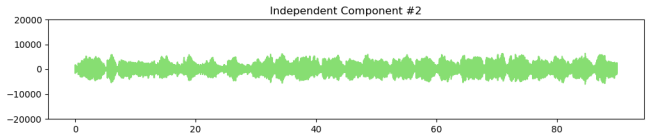

Fig. 12. 1:30 - 3min of Independent Component 2 which is Perfect by Ed-Sheeran

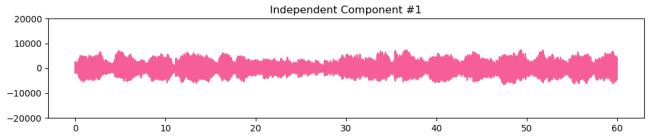

Fig. 13. 3 - 4min of Independent Component 1 which is Kathang Isip by Ben&Ben

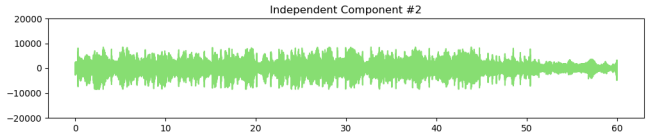

Fig. 14. 3 - 4min of Independent Component 2 which is Perfect by Ed-Sheeran

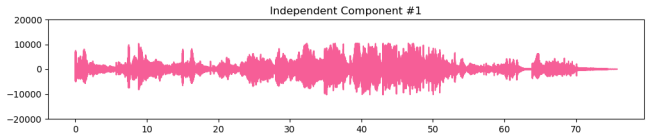

Fig. 15. 4min - end of Independent Component 1 which is Kathang Isip by Ben&Ben

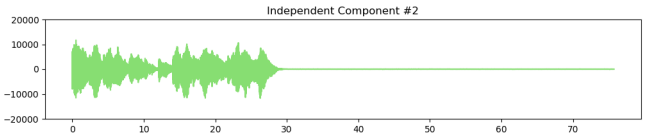

Fig. 16. 4min - end of Independent Component 2 which is Perfect by Ed-Sheeran

As you can see from the results of performing Independent Component Analysis, the Independent components have been separated but the catch is that it has been scaled down which lowers the volume and the frame rate has been cut into half, making it 0.5x slower than the original WAV file, so to bring back to the original frame rate, we multiply the original frame rate by 2 when saving the WAV file.

VI. Conclusion

It is possible to perform Blind source separation on audio files by performing Independent Component Analysis. As seen from the results, the the independent components are quite similar to the initial audio source files but only scaled down which lowers the volume of the audio file, and slowing down the frame rate by 0.5x.

References

[1] “Blind Signal Separation ”. Retrieved in https://en.wikipedia.org/wiki/Blind_signal_separation
[2] “Independent component analysis ”. Retrieved in https://en.wikipedia.org/wiki/Independent_component_analysis
[3] “What is Independent Component Analysis”. Retrieved in https://www.cs.helsinki.fi/u/ahyvarin/whatisica.shtml