DEVELOPMENT OF SOURCE SEPARATION ALGORITHM IN AUDIO APPLICATION

NURUL AMIZA BINTI AMIR HAMZAH

This Report Is Submitted In Partial Fulfillment Of Requirements For The Bachelor Degree of Electronic Engineering (Electronic Telecommunication)

Fakulti Kejuruteraan Elektronik dan Kejuruteraan Komputer

Universiti Teknikal Malaysia Melaka

JUNE 2013

UNIVERSTI TEKNIKAL MALAYSIA MELAKA
FAKULTI KEJURUTERAAN ELEKTRONIK DAN KEJURUTERAAN
KOMPUTER

BORANG PENGESAHAN STATUS LAPORAN
PROJEK SARJANA MUDA II

Tajuk Projek : **DEVELOPMENT OF SOURCE SEPARATION ALGORITHM IN AUDIO APPLICATION**

Sesi Pengajian :

1	2	/	1	3

Saya   **NURUL AMIZA BINTI AMIR HAMZAH**

(HURUF BESAR)

mengaku membenarkan Laporan Projek Sarjana Muda ini disimpan di Perpustakaan dengan syarat-syarat kegunaan seperti berikut:

1. Laporan adalah hakmilik Universiti Teknikal Malaysia Melaka.

2. Perpustakaan dibenarkan membuat salinan untuk tujuan pengajian sahaja.

3. Perpustakaan dibenarkan membuat salinan laporan ini sebagai bahan pertukaran antara institusi pengajian tinggi.

4. Sila tandakan (√) :

☐   **SULIT***   *(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)

☑   **TERHAD****   **(Mengandungi maklumat terhad yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

☐   **TIDAK TERHAD**

_____
(TANDATANGAN PENULIS)

Disahkan oleh:

_____
(COP DAN TANDATANGAN PENYELIA)
Dr. Abdul Majid bin Darsono
Pensyarah
Fakulti Kejuruteraan Elektronik Dan Kejuruteraan Komputer
Universiti Teknikal Malaysia Melaka (UTeM)
Hang Tuah Jaya
76100 Durian Tunggal, Melaka

Tarikh: 14/6/2013   Tarikh: 18/06/2013

"I acknowledge that this report is entirely my work except summary and passage I have quoted the source"

Signature : ....

Author Name : NURUL AMIZA BINTI AMIR HAMZAH

Date : 14/6/2013

'I acknowledge that I have read this thesis and in my opinion this thesis is sufficient in terms of scope and quality for the award of Bachelor of Electronic Engineering (Telecommunication Electronics)"

Signature : ....

Supervisor Name : DR ABD MAJID BIN DARSONO

Date : 14/06/2013 ....

I dedicate this report to my father and mother, Amir Hamzah Bin Husain and Sharipah Halishah Binti Syed Baharom, my brother and sister, and my friend for their encouraging and love.

ACKNOWLEDGEMENT

First of all, I am grateful to Allah for blessing me good health in completing this project. I would like to dedicate my gratitude towards my parents for their love and supports. Without their support, I do not think that I can stand here completing this project.

I also want to say thanks towards all my friends because they give me courage and assistance throughout the whole semester for PSM 1 which have help me a lot during completing this project. Special thanks to Dr. Abd Majid Bin Darsono for his guidance and help during the period of this project.

Next, thanks to all lecturers for giving me knowledge while I am studying in UteM. Last but not least, thanks to UteM and UteM staffs for letting me study and give me so much knowledge and experience throughout whole semester.

ABSTRACT

Blind source separation (BSS) is the separation signals from a set of mixed signal without the aid information about the source signals or the mixing process. Source separation problems are occurred when a number of sources is mixed and try to separate to individual source. The problem is how to construct a signal, which is able to overcome the characteristics of various sound sources. The main objective of this project is to develop and implement the algorithm of source separation using the Nonnegative Matrix Factorization 2Deconvolution (NMF2D) based on least squares function (LS). It can estimate original signals from observed signals, which consists of mixed signal and noise. By using this method, the convolutive in time and frequency is represented where it can factorize a spectrogram of music instruments corresponding to individual instruments. The music instruments can be separated using spectrogram masking based on factorization 2-de-convolution. In this project, the basic theory and literature review of this method is applied in separating audio using Matlab software. The result shows that the music instruments are piano and trumpet can be separated effectively from mixture source to original source. The proposed algorithm has applications in wide range including computational auditory scene analysis, music information recapture and automatic music transcription.

ABSTRAK

Pemisahan sumber buta (BSS) adalah isyarat pemisahan daripada satu set isyarat bercampur tanpa maklumat bantuan tentang isyarat sumber atau proses pencampuran. Masalah pemisahan sumber yang berlaku apabila beberapa sumber yang bercampur dan cuba memisahkan sumber individu. Masalahnya ialah bagaimana untuk membina isyarat, yang mampu untuk mengatasi ciri-ciri sumber bunyi pelbagai. Tujuan utama objektif adalah memberi tumpuan untuk membangunkan dan melaksanakan algoritma pemisahan sumber menggunakan matriks pemfaktoran bukan negatif algoritma 2-*De-convolutive* berdasarkan *least square, LS*. Ia boleh mengaggarkan isyarat yang asal daripada isyarat pemerhatian yang mengandungi campuran isyarat dan hingar atau bunyi bising. Dengan menggunakan kaedah ini, *convolutive* dalam masa dan kekerapan diwakili di mana ia boleh Pendarab a spectrogram instrumen muzik yang sepadan dengan instrumen individu. Instrumen muzik boleh dipisahkan menggunakan pelekat spektrogram berdasarkan pemfaktoran 2-*De-convolutive*. Dalam projek ini, teori asas dan kajian *literature* bagi teknik ini digunakan dalam memisahkan audio menggunakan perisian Matlab. Hasil kajian menunjukkan bahawa instrumen muzik piano dan trumpet boleh dipisahkan dengan berkesan dari sumber campuran kepada sumber asal. Algoritma yang dicadangkan itu mempunyai pelbagai aplikasi termasuk pendengaran untuk di analisis, penghasilan semula segala maklumat muzik yang dikehendaki serta transkripsi muzik secara automatik.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

This chapter is introduced the blind source separation (BSS) in audio source separation. Then, objectives and problem statements of this project are explained in this chapter. Besides that, the scope of project based on this thesis is described too. Project methodology is defined based on method used in this project

1.1   Overview

Humans are surrounded by sound. We noticed that if we try to concentrate and listen carefully what happen in our environment, we must able to identify more than one source of sound. This is important of human hearing where it is ability to distinguish individual sound sources from complex mixtures of sound. It is usually relates with "cocktail party problem". The cocktail party is described as focus one is listening attention on a single talker with combine of mixture of conversations and background noises, ignoring all the other conversations and enabling us to talk in a noisy place [1].

The human brain can detect the sound in real time environment using an ear. Let us imagine that situation of walking down a busy street with a friend. Our ears heard several of sound sources such as car noise, a friend speaking, mobile phones ringing and other sound. However, we cannot focus to all sound but only the specific sound that we want to hear clearly. For example, we are listened what our friend said. Then, when we in feeling bored, we can hear another friend said or whatever sound that we want to hear. Thus, the human brain can directly focus on and make easy to separate the sources.

Here, audio source separation is one problem to compose and decompose a real world sound mixture into individual audio sources that we need. Most of musical are mixtures and playing with different instrument concurrently. In this case, mostly we want to know which one of these instruments. This is complex way to understand of sound scene to humans. However, audio source separation methods aimed at maybe to our hearing sense and brain to understand of complex sound material.

Here, the Blind Source Separation is introduced. Usually, is also known, as BSS It is most popular in the signal processing. It is a method to estimate the original signals from observed signal, which contains of mixed original sources and noise. Some of original sources have a several noises and we need to put away from original sources when we get the estimation of original source of signal. Therefore, this method is already mentioned that used to estimate the original sources that having a noise. The blind word is refers that we do not know how the signals were mixed or how they were generated [2].

Figure 1.1 shows that audio source separation scheme which used to understand of complex sound scenes. This method is used to estimate and then extract of each individual audio signal, which is from sound mixture. The mixture sounds is piano, trumpet and drum. This combination of three instruments is make the mixture of them and hard to separate them to individual instruments. Consequently, using the appropriate method in audio source separation, it can separate those instruments as respectively. It making their instruments can separate efficiently using the suiTABLE algorithm
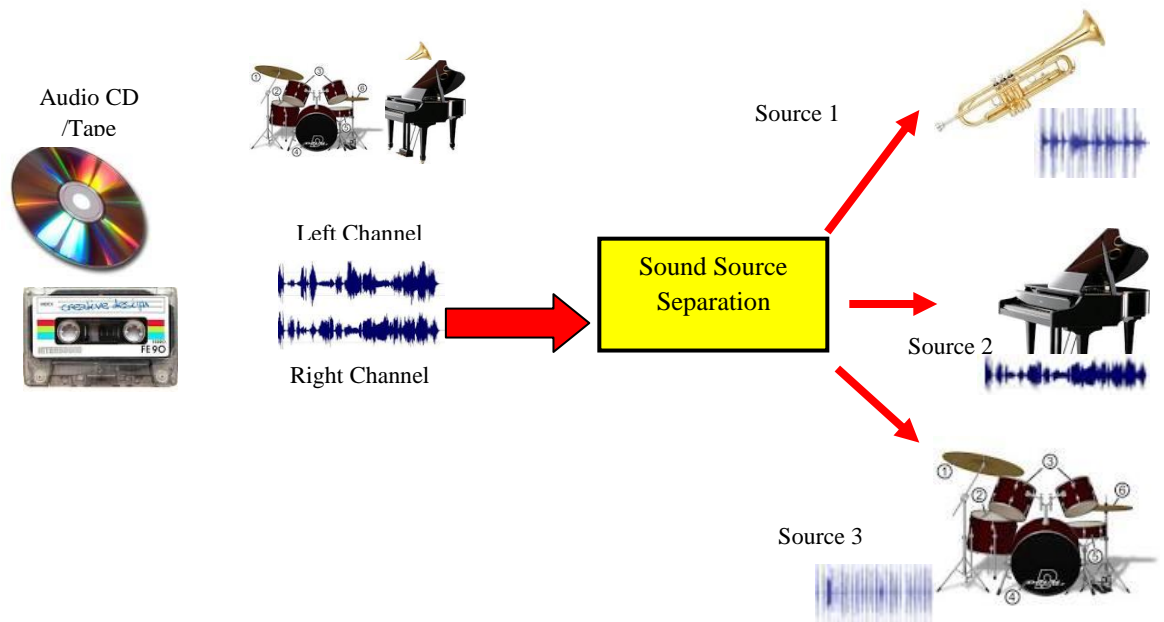
Figure 1.1: Audio Source Separation Scheme

Figure 1.2: Blind Source Separation (BSS)
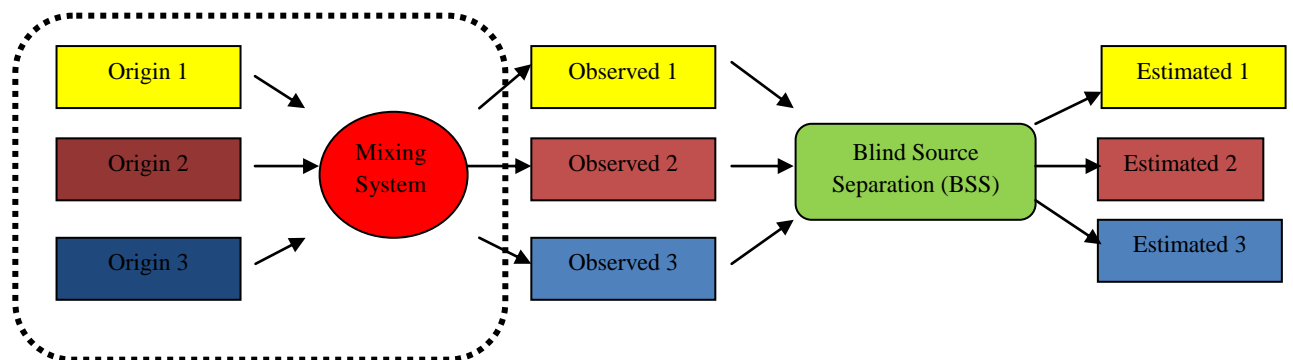
Figure 1.2 shows that the original source signals is mixed, which are, can estimate without knowing the parameters of mixing and/or filtering processes [3]. Generally, it is not possible to estimate any original source signals when we do not have any knowledge to estimate it [1]. This technique is used to extract the observed signal from mixture signal to estimation signal.

In fact, several observations of the mixture are available in many methods, but their performance is related based on the mixing environment. The separation of mixed audio signals can be applied through audio or image separation, music transcription and video conferring. The audio source separation is one of the extracting signals of each sound source from of mixture sources. This is one way of signal processing technique that used to estimate a single source where the of sources is a signal mix [2]

Thus, the separation of multiple sounds in audio instruments to a single channel recording is a difficult problem to solve. Many proposed method is observed based on the matrix decompositions of a spectrogram. It is demonstrate the sound to individual instruments. Thus, Non-negative matrix factorization (NMF) is one method that useful to separate the mixture audio to original sources. [1].

NMF gives a solution to use in separating of original sources that applied in audio application. This one proposed method used which is suiTABLE algorithm in separating audio. NMF has many application can be applied in research world because it is appropriate method to separate piano from polyphonic music using NMF. Besides that, NMF also used of separating automatic description of polyphonic music.

When polyphonic music is set by using factorizing of magnitude spectrogram with NMF, each instrument is applied by an instantaneous frequency means can vary over time. Nonnegative matrix factorization 2 de-convolution (NMF2D) algorithm has been introduced where of each instrument that we want to observe is applied by a time frequency, which is varies over time. Thus, it can symbolize components with temporal characteristics such as noise, standard deviation and likeness. The characteristics are important because through this it can be used to separate any individual instrument for example pianos from a real recording of piano sounds.

In this thesis, a new method is selected to factorize a log-frequency spectrogram by using NMF2D, which can determine both temporal characteristics and pitch change that occurs when an instrument plays different notes. Furthermore, a log frequency spectrogram is used in pitch change to a frequency axis. NMF2D is one solution in separating of various audio and effectively solves the blind source separation problem in music instruments. Not to forget, how the algorithm function is determined in separating music instruments to get the individual instruments.

1.2   Objectives

The objectives of the project are as shown below:

i.   To study mathematical of algorithm in development of audio source separation. The appropriate algorithm is finding to be used in audio source separation.

ii.   To learn an algorithm by using the Matlab software. Then, we relate the algorithm with source separation in audio application and then implement the algorithm in Matlab software.

iii.   To apply an algorithm through Matlab software. Thus, we purpose algorithm related to the processing used in audio source separation and apply it through Matlab software

1.3   Problem Statement

Audio source separation is the problem of automated separation of audio sources. For example, the people are present in a room using differently placed microphones, at the same time, also capturing the auditory scene [3]. That means the voice from the accompanied music can recognize clearly. The problems of source separation in digital processing are some of signal have been mixed together and need to find out the original signals from it

For example, known as cocktail party problem where the brain can read and handle the audio source separation problem with focus on and directly separate based on the specific sound source [3]. However, the digital signal processing cannot be recognizing. This is problem in digital signal processing

Another example, one room is crowded of people and two speaker's person that want both of them want to talk at same time. The crowded is making a little noise among of them. Mostly, some people like to hear and pay attention to two speakers speaking in front. However, when enough with background noise, the audience not giving any comment and attention which one they want to hear. The best solution is there are two microphones at different locations, recording the speaker's voices as well with noise coming from the crowded. Through this, the problem can separate the voice directly of each speaker whereas ignoring the background noise as Figure 1.3
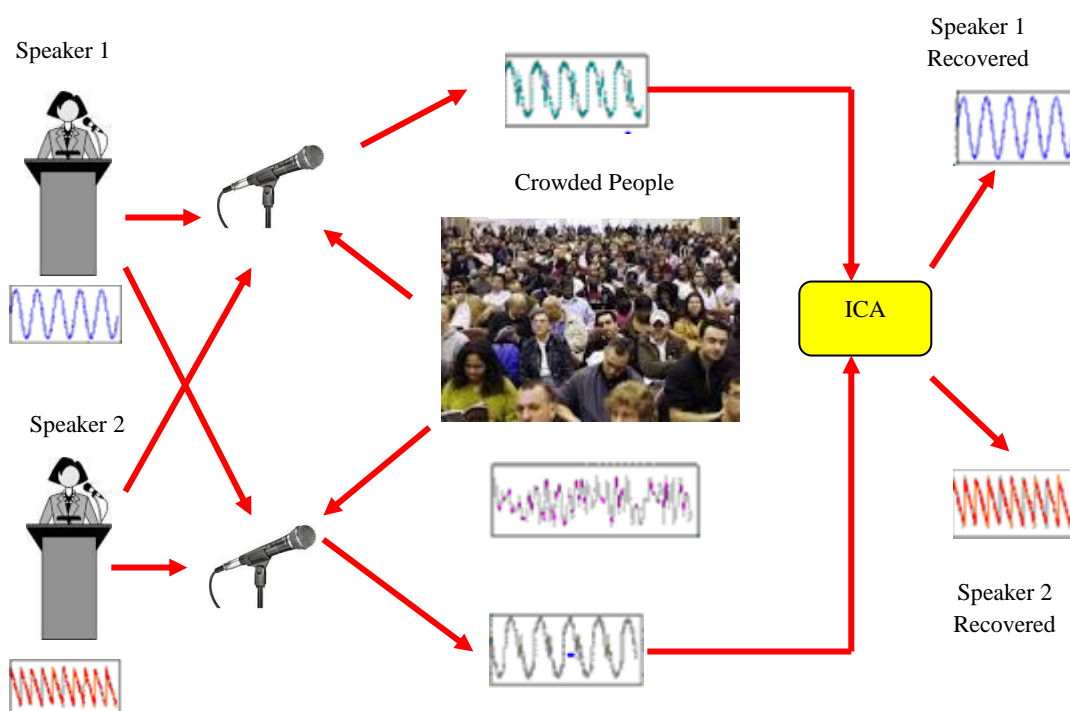

Figure 1.3: Audio source separation using ICA method

Besides that, biomedical applications are applied to separate audio in electrical recordings of brain activity using electroencephalogram (EEG). The data from EEG contains of recording of electrical potentials in many different locations on the scalp at the head [5]. These potentials are coming from mixing some components of brain activity. This situation is similar with cocktail party problem where the original sources of brain activity are needs to find from observed mixtures of the components. Through this, separating source signals can disclose this information in brain activity using both techniques.

1.4 Scope of Project

Some particular aims appear from this main scope, which are presented as follows:

i.   Study and investigate the new algorithm of source separation in audio application. We can explore the many algorithms in this application and consequently propose new solutions of separating signals.

ii. Implement the algorithm of source separation using Matlab software. We examine the relationship between objective performances with algorithm used by source separation.

iii. Develop new methods of source separation based on techniques such as Nonnegative Matrix Factorization (NMF). We can know the suitability, advantages and disadvantages of source separation techniques to use in separating signals

1.5   Project Methodology Overview

The research is to develop for separating musical audio into individual sound source such as instruments or voices. In musical audio and recording room, it is usually that many instruments are played at the same time. Moreover, some instruments play the same note at the same time. It is difficult to identify which one of each sound because they share many common frequencies and amplitude. We can say that maybe they have a same frequency in pitch. That is why, it making some not easy to separate them. Therefore, this project is related to blind source separation and familiar techniques that used to extract the single sources from mixture signals is known as non-negative matrix factorization (NMF).

1.6   Thesis Overview

The objective of thesis is to provide an understanding, propose and implement appropriate algorithms for nonnegative matrix factorization (NMF). Here, the chapters are organized and presented as follow:

- Chapter 1: **Introduction.** This chapter is give a comprehensive overview of blind source separation, audio source separation and related it. This includes the objective and problem statement of this thesis, scope of project, and project methodology overview.

- Chapter 2: **Literature Review**. This chapter presents methods for the separation of sources in audio application. It is described about the separation that achieve by using these method.

- Chapter 3: **Methodology**. In this chapter, investigations are carried out to clarify some algorithms. The first part of the chapter described how these technique that used in separating signal in audio applications. Then, why these technique is chosen in separating signal.

- Chapter 4: **Result & Discussion**. This chapter describes result that achieve in this thesis is used in separating sources and discussion about the results including result and analysis of other audio.

- Chapter 5 : **Conclusion & Recommendation**. Finally, conclusion of this thesis is obtained including the recommendation for future work in audio source separation and potential commercialization through this project.

CHAPTER 2

LITERATURE REVIEW

This chapter covers the literature review, which is related with audio source separation. The several technique/method such as Independent Component Analysis (ICA) and Nonnegative Matrix Factorization (NMF) is described in the chapter that used in audio source separation The algorithms that related with own method is explained to describe the different between the methods.

2.1   Independent Component Analysis (ICA)

Independent Component Analysis (ICA) is one of BSS techniques for extracting individual signals from mixtures. That means the original underlying source signals are mutually independently distributed [4].These are statistical techniques of decomposing a complex data into independent parts of the separated signals. It is used to exploit the non-Gaussianity of source signals and assumes statistical independence of the separated signals to perform separation [3].