ISSN 2229-5518

Preparation of Papers for International Journal of Scientific & Engineering Research

Malika Singh, Santosh, Ameeta Khanna

Abstract--- In this paper we examine the audio source separation problem using the general framework of Independent Component Analysis (ICA). For the greatest part of the analysis, it has been assumed that equal number of sensors and sound objects. Firstly, it explores the case that the auditory scene is modeled as instantaneous mixtures of the auditory objects, to establish the basic tools for the analysis. The case of real room recordings, modeled as convolutive mixtures of the auditory objects, is then introduced. A novel Fast ICA framework is introduced, using two possible implementations. A great number of audio source separation problems can be addressed successfully using Independent Component Analysis. And concludes by highlighting some of the as yet unsolved problems to tackle the actual audio source separation problem in full.

Index Terms— Principal component analysis, Independent component analysis, projection pursuit, blind signal separation, source separation, factor analysis representation.

s————————— ◆ —————————

1 INTRODUCTION

THIS document shows the whole dimension of communications has been changed by the rapid growth of technology. Today people are more interested in hands-free communication. The main advantage of wireless system is that more than two persons can participate in conversation while freely moving in the room, on roads etc i.e. conferencing. The presence of large acoustic coupling between speaker and microphone would produce a loud acoustic echo making the conversation difficult in this case. The term "ACOUSTIC ECHO CANCELLER" means audio source separation. Number.Click the forward arrow in the pop-up tool bar to modify the header or footer on subsequent pages. Audio source separation is the problem of automated separation of audio sources present in a room, using a set of differently placed microphones, capturing the auditory scene. The whole problem resembles the task a human can solve in a cocktail party situation, where using two sensors (ears),the brain can focus on a specific source of interest, suppressing all other sources present (cocktail party problem).

The source separation is an inductive inference problem. There is not enough information to deduce the solution, so one must use any available information to infer the most probable solution. The aim is to process these observations in such a way that the original source signals are extracted by the adaptive system. The problem of separating and estimating the original source waveforms from the sensor array, without knowing the transmission channel characteristics and the source can be briefly expressed as problems related to BSS (Blind signal/source separation). In BSS the word blind refers to the fact that I do not know how the signals were mixed or how they were generated. As such, the separation is in principle impossible. Allowing some relatively indirect and general constrains, I however still hold the term BSS valid, and separate under these conditions. There appears to be something magical about blind source separation; I am estimating the original source signals without knowing the parameters of mixing and/or filtering processes. It is difficult to imagine that one can estimate this at all. In fact, without some a priori knowledge, it is not possible to uniquely estimate the original source signals[2].

Independent component analysis (ICA) is one of the most widely used BSS techniques for revealing hidden factors that underlie sets of random variables, measurements, or signals. ICA is essentially a method for extracting individual signals from mixtures. Its power resides in the physical assumptions that the different physical processes generate unrelated signals[3] The simple and generic nature of this assumption allows ICA to be successfully applied in diverse range of research fields. In this paper, I first set the scene of the blind source separation problem. Then, Independent Component Analysis is introduced as a widely used technique for solving the blind source separation problem. A general description of the approach to achieving separation via ICA and the underlying assumptions of the ICA framework and important ambiguities that is inherent to ICA.

2. Audio Source Sepration

Humans exhibit a remarkable ability to extract a sound object of interest from an auditory scene. The human brain can perform this everyday task in real time using only the information acquired from a pair of sensors, i.e. our ears. Imagine the situation of walking down a busy street with a friend. Our ears capture a huge variety of sound sources: car noise, other people speaking, a friend speaking, mobile phones ringing. However, I can focus and isolate a specific source that is of interest at this point. For example, I may listen to what our friend is saying. Getting bored, I can over hear some body else's conversation, pay attention to an annoying mobile ringtone or even listen to a passing car's engine, only to understand it is a Porsche. The human brain can automatically focus on and separate a specific source of interest. Audio source separation can be defined as the problem of decomposing a real world sound mixture (auditory scene) into individual audio objects. The automated analysis using a computer that captures an auditory scene through a number of sensors is the main objective of this thesis. Although this is a relatively simple task for the human auditory system, the automated audio source separation can be considered one of the most challenging topics in current research. A number of different methods were proposed to solve the problem.

3. Computational Auditory Scene Analysis (CASA)

A possible approach to address the problem will be to analyze and finally emulate the way humans perform audio source separation using a computer. Psychoacoustics is a special area of research studying how people perceive, process and deduce information from sounds. Such studies construct experimental stimuli consisting of a few simple sounds such as sine tones or noise bursts, and then record human subjects' interpretation/perception of these test sounds. Computational Auditory Scene Analysis (CASA) was one of the first methods that tried to "decrypt" the human auditory system in order to per-form an automatic audio source separation system. Conceptually, CASA may be divided into two stages .In the first stage, the acoustic mixture is decomposed into sensory elements ("segments"). CASA employs either computer vision techniques or complete ear models (outer and middle ear, cochlear filtering etc) in order to segment the auditory scene into several audio elements.
The second stage ("grouping") then combines segments that are likely to have originated from the same sound source. Psychological and psychoacoustic research of this kind has uncovered a number of cues of grouping rules which may describe how to group different parts of an audio signal into a single source, such as i) common spatial origin, ii) common on set characteristics, i.e., energy appearing at different frequencies at the same time, iii) amplitude or frequency modulations in the harmonics of a musical tone or periodicity, iv) proximity in time and frequency, v) continuity (i.e. temporal coherence). Usually, CASA employs one or two sensor signals, as the main goal is to emulate human way of performing auditory scene analysis.

4. Beam Forming

Array signal processing is a research topic that developed during the late70s and 80s mainly for telecommunications, radar, sonar and seismic applications. The general array processing problem consists of obtaining and processing the information about a signal environment from the waveforms received at the sensor array (a known constellation of sensors) . The use of an array allows for a directional beam pattern. The beam pattern can be adapted to null out signals arriving from directions other than the specified look direction. This technique is known as spatial filtering or adaptive beam forming .The reception of sound in large rooms, such as conference rooms and auditoria, is typically contaminated by interfering noise sources and reverberation. One can set up an array of microphones and apply the techniques of adaptive beam forming in the same way as in telecommunications to per-form several audio processing tasks. I can enhance the received amplitude of a desired sound source, while reducing the effects of the interfering signals and reverberation. Moreover, I can estimate the direction or even the position of the sound sources in the near field present in the room (source localization). Most importantly, if the auditory scene contains more than one source, I can isolate one source of interest, whilst suppressing the others, i.e. perform source separation.Beam forming assumes some prior knowledge on the geometry of the array, i.e. the distance between the sensors and the way they are distributed in the auditory scene. Usually, linear arrays are used to simplify the computational complexity.

5. Blind Source Separation

In contrast to CASA and Beam forming, BSS is a technique in which I have a number of sources emitting signals which are interfering with one another. Familiar situations in which this occurs are a crowded room with many people speaking at the same time, interfering electromagnetic waves from mobile phones or crosstalk from brain waves originating from different areas of the brain. In each of these situations the mixed signals are often incomprehensible and it is of interest to separate the individual signals. This is the goal of Blind Source Separation. A classic problem in BSS is the cocktail party problem. The objective is to sample a mixture of spoken voices, with a given number of microphones - the observations, and then separate each voice into a separate speaker channel - the sources. The BSS is unsupervised and thought of as a black box method. In this I encounter many problems, e.g. time delay between microphones, echo, amplitude difference, voice order in speaker and underdetermined mixture signal.

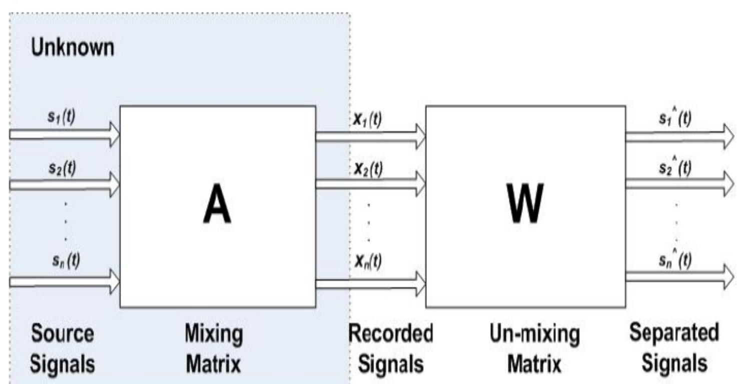

Figure: 1.1 Block Diagram of Blind Source Separation

In this block diagram s_1 ….s_n Are the source signal and A is a mixing matrix which is unknown.x_1 ……… x_n are mixed recorded signal. W is an un-mixing matrix and $\hat{s_1}$…..$\hat{s_n}$ are the separated or estimated original source signal. G.R. NAIK and D.K.KUMAR [2] proposed that, in a artificial neural network like architecture the separation could be done by reducing redundancy between signals.This approach initially lead to what is known as independent component analysis today.In ICA the general idea is to separate the signals, assuming that the original underlying source signals are mutually independently distributed. Due to the field's relatively young age, the distinction between BSS and ICA is not fully clear. When regarding ICA, the basic framework for most researchers has been to assume that the mixing is instantaneous and linear, as in infomax. ICA is often described as an extension to PCA, that un correlates the signals for higher order moments and produces a non orthogonal basis.

Result

We compared result of this technique with famous audio source separation technique Beam forming and Computational Auditory Scene Analysis (CASA) using different algorithm. We have analyzed that the result of this technique and algorithm is much better than other technique. As shown in table 1 that algorithm is achieving higher accuracy. These results are taken on our self-created database on different sound signal at different condition.

Table1 . Blind source Separation result using Independent Com ponent Analysis (ICA).

S.NO.	Method	Original Signal Rate (%)
1.	Beam Forming	85%
2.	CASA	90%
3.	Proposed BSS using (ICA)	95%

The sound signals are using in this paper is from self- created database using Matlab. But we can take any signal.

Conclusion

ICA is a very general-purpose statistical technique that is used to find underlying factors by analyzing a set of observed random data. These observed random data are linearly transformed into components that are maximally independent of each other. ICA was originally developed to deal with sound source separation for audio processing, but now has been widely used in many different areas such as biomedical signal processing, image processing, telecommunications, and econometrics. In addition, ICA can be estimated as a latent variable model. There are many approaches that can be used to estimate ICA: optimization of the maximum of non-gaussianity can be used for the estimation of the ICA model; alternatively, maximum likelihood estimation, kourtis, negntropy or minimization of mutual information can also be used to estimate ICA.An example of a real-world communications application where blind separation techniques are useful is the separation of the user's own signal from the interfering other users' signals in CDMA (Code-Division Multiple Access) mobile communications. This problem is semi-blind in the sense that certain additional prior information is available on the CDMA.

Future Scope

Some possible extensions for this project over future years could be:
1) More advanced lighting normalization algorithms before applying ICA method
2) Use of ICA to increase robustness to change in expression
3) Use of multiple subspaces to allow for detection from gen eral perspectives, i.e. Separation of signals.

.References

[1] J.S. Bridle, "Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition," *Neurocomputing—Algorithms, Architectures and Applications*, F. Fogelman-Soulie and J. Herault, eds., NATO ASI Series F68, Berlin: Springer-Verlag, pp. 227-236, 1989. (Book style with paper title and editor)

[2] W.-K. Chen, *Linear Networks and Systems*. Belmont, Calif.: Wadsworth, pp. 123-135, 1993. (Book style)

[3] H. Poor, "A Hypertext History of Multiuser Dimensions," *MUD History*, http://www.ccs.neu.edu/home/pb/mud-history.html. 1986. (URL link *include year)

[4] K. Elissa, "An Overview of Decision Theory," unpublished. (Unplublished manuscript)

[5] R. Nicole, "The Last Word on Decision Theory," *J. Computer Vision*, submitted for publication. (Pending publication)

[6] C. J. Kaufman, Rocky Mountain Research Laboratories, Boulder, Colo., personal communication, 1992. (Personal communication)

[7] D.S. Coming and O.G. Staadt, "Velocity-Aligned Discrete Oriented Polytopes for Dynamic Collision Detection," *IEEE Trans. Visualization and Computer Graphics*, vol. 14, no. 1, pp. 1-12, Jan/Feb 2008, doi:10.1109/TVCG.2007.70405. (IEEE Transactions)

[8] S.P. Bingulac, "On the Compatibility of Adaptive Controllers," *Proc. Fourth Ann. Allerton Conf. Circuits and Systems Theory*, pp. 8-16, 1994. (Conference proceedings)

[9] H. Goto, Y. Hasegawa, and M. Tanaka, "Efficient Scheduling Focusing on the Duality of MPL Representation," *Proc. IEEE Symp. Computational Intelligence in Scheduling (SCIS '07)*, pp. 57-64, Apr. 2007, doi:10.1109/SCIS.2007.367670. (Conference proceedings)

[10] J. Williams, "Narrow-Band Analyzer," PhD dissertation, Dept. of Electrical Eng., Harvard Univ., Cambridge, Mass., 1993. (Thesis or dissertation)

[11] L. Hubert and P. Arabie, "Comparing Partitions," *J. Classification*, vol. 2, no. 4, pp. 193-218, Apr. 1985. (Journal or magazine citation)

[12] R.J. Vidmar, "On the Use of Atmospheric Plasmas as Electromagnetic Reflectors," *IEEE Trans. Plasma Science*, vol. 21, no. 3, pp. 876-880, available at http://www.halcyon.com/pub/journals/21ps03-vidmar, Aug. 1992. (URL for Transaction, journal, or magzine)

[13] J.M.P. Martinez, R.B. Llavori, M.J.A. Cabo, and T.B. Pedersen, "Integrating Data Warehouses with Web Data: A Survey," *IEEE Trans. Knowledge and Data Eng.*, preprint, 21 Dec. 2007, doi:10.1109/TKDE.2007.190746.(PrePrint)