Blind Source Separation
–
A Literature Review

Anders Røsæg Pedersen
Ulf Nore

NTNU, Fall 2012

Contents

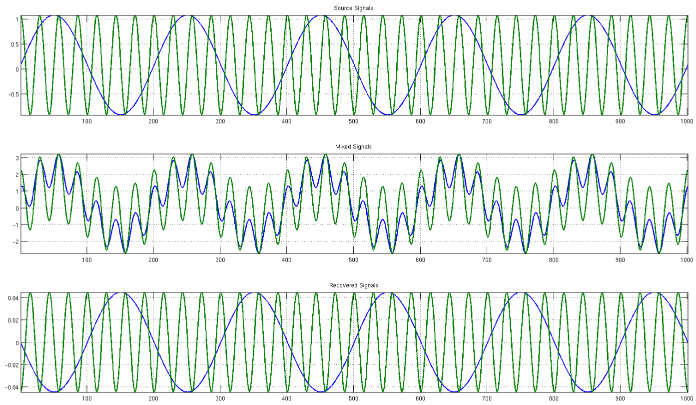

Figure 1: Stationary linear mixing process and separation.

1 Introduction

The blind source separation problem refers to the process of recovering one or more signals that have been mixed in some unknown manner and possibly also contamined by noise. Without any assumptions on the mixing process, this problem is ill-poised. In practice therefore, all BSS methods rely on some stylized fact about the nature of the signals and/or the mixing process. It is therefore useful to dichotomize BSS methods by these assumptions.

Arguably, two of the most important facts characterizing a mixing process, are its temporal dynamics and the number of degrees of freedom. The first point refers to whether the nature of the mixing process changes over time, that is if the mixing matrix at time $t+k$ is different from that at time t for $k > 0$. The number of degrees of freedom is the same concept as in linear algebra - the connection is apparent by seeing the mixing process as a system of linear equations. If m is the number of observed signals and n the number of sources, the system is said to be *underdetermined* if $m < n$ and conversely *overdetermined* if $m > n$.

We can also differentiate between method based on the nature of input data. Early BSS research often considered the case of $n = m$, which allows one to work with data in the time domain. For undetermined systems, it is commonplace to work with some transformation of the data, which in the case of audio data a time-frequency representation. Common methods include the *short-term Fourier transform* and the *wavelet transform*.

The organization of this study is as follows. Section 2 will briefly summarize the literature review process, which is further documented by the underlying research protocol given in Appendix A. Section 3 is a short description of the different techniques and methodologies found, summarised in section 4.

2 Literature Review Process

The literature review process was conducted by searching the listed databases for published articles containing the predefined search terms. The table below shows the amount of results presented to us when using the more general of our search terms.

Search term	CiteSeerX	Google Scholar	SpringerLink
blind source separation	292581	698000	13143
blind audio source separation	28612	26000	1072
single channel blind source separation	275543	78600	5024

Table 1: Magnitude of hits on the most general terms

Since the search terms we had defined gave us a large amount of results, we prioritized newer papers over older ones as per the Appendix A.

3 Literature Overview

3.1 Independent Component Analysis

Among the most common approaches to blind source separation is independent component analysis (ICA). Common definitions of ICA use either the maximization of independence or minimization of mutual information between the source signals[1]. Formally, we can state the ICA problem in terms of a generative model of the observed signals $\mathbf{x}$, and the unknown a mixing matrix $\mathbf{W}$ and source signals $\mathbf{s}$:

$$\mathbf{x} = \mathbf{W}\mathbf{s} \tag{1}$$

The AIM of the ICA process is to estimate the inverse mixing process along with the original signals.

The classical reference on ICA is [1], where the method of minimization of mutual information between sources is presented. [1] also presents an analysis of the ambiguities and limitations of ICA, hereunder the permutation of sources, scaling and non-gaussianity.

There are several equivalent statements of ICA, which yields different interpretations and computational models. [2] proposes minimizing mutual information between sources, as measured by *differential entropy*. In this implementation a feed-forward neural network structure is proposed. Other approaches include conventional maximum likelihood ([3]) and maximization of non-gausianity as measured by excess kurtosis. A popular approach is

[1]It should be noted that while this text presents ICA in terms of blind source separation, the method is applicable to a wide array of machine learning problems including dimension reduction, classification, and de-noising.

the FastICA algorithm ([6]) that minimizes mutual information expressed by *negentropy* by a fixed point method.

The classic studies on ICA focus to a large extent on developing the formal framework for ICA, and examples are largely centered on time domain analysis in systems of an equal number of sensors and sources[2]. ICA has however been extended to underdetermined systems and the extreme case of single sensor systems.

Many of these extensions are to a lesser extent changes to the previously known algorithms; rather they involve transforming the observed signals from the time domain to some other basis, the most common of which are the frequency domain (Fourier transform), the time-frequency domain (short-term Fourier transform) and the wavelet domain. Compared to the time domain, the two latter are redundant representations, but they transform the data so as to be suited for ICA. [10] surveys variations on ICA as applied to single channel recordings, hereunder single channel ICA (SCICA) and wavelet ICA (WICA), in addition to proposing an algorithm that combines ICA with empirical mode decompostion (EMD). EMD decomposes a signal into independent components in the spectral domain and can be viewed as similar to STFT.

The abovementioned approaches represent a select set of common approaches to the BSS problem. Other approaches rely to a larger extent on direct application of knowledge about the human auditory system. As an example [5] focuses on the problem on single channel speech separation in the spectral domain by means of feature maps where the features roughly corresponds to "audible" features such as common onset, pitch, timbre and so forth.

[2]For a much more thorough survey on the classical literature on ICA, see [4].

3.2 Factoral Hidden Markov Models

The Hidden Markov modelling has been around since the late 1960s. Roweis[7] proposes a technique called refiltering, where the idea is to separate sources in a mixed or corrupted recording. This is achieved through non stationary masking of the different frequency sub-bands from the target recording. Different sources may be isolated in the recording by changing the masking parameters. Using regularities in the spectrogram produced by a recording, it is possible to set the masking parameters, eg. common onset and offset.

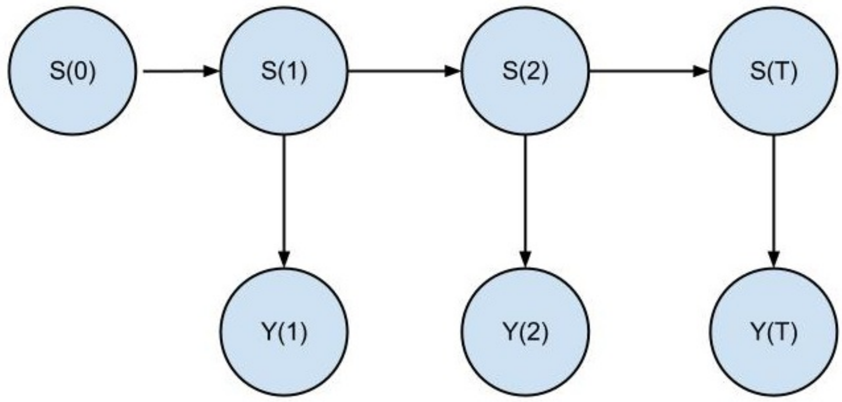

Figure 2: Hidden Markov Model

Training speaker dependent HMMs on isolated data from the sources to be separated, these models are then combined together in an architecture called factorial-max HMMs. The different HMMs evolve independently and for each observation vector produced at time t by each HMM, the elementwise maximum is chosen to create an observation. This is because the log magnitude spectrogram of a mixture of sources is very similar to the elementwise maximum of the individual spectrograms[3]. Separation is performed by setting the various masking signals to 1 or 0, depending on the observation vector at time t for frequency band i.

The full generative model is given in Equations 2 - 4.

$$p(x_t = j | x_{t-1} = i) = T_{ij} \tag{2}$$

$$p(z_t = j | z_{t-1} = i) = U_{ij} \tag{3}$$

$$p(y_t | x_t, z_t) = N(max[a_{x_t}, b_{z_t}], R) \tag{4}$$

[3]This example was performed on two speakers. a_{x_t} is the observation vector for speaker x at time t, likewise is b_{z_t} the observation vector for speaker z at time t

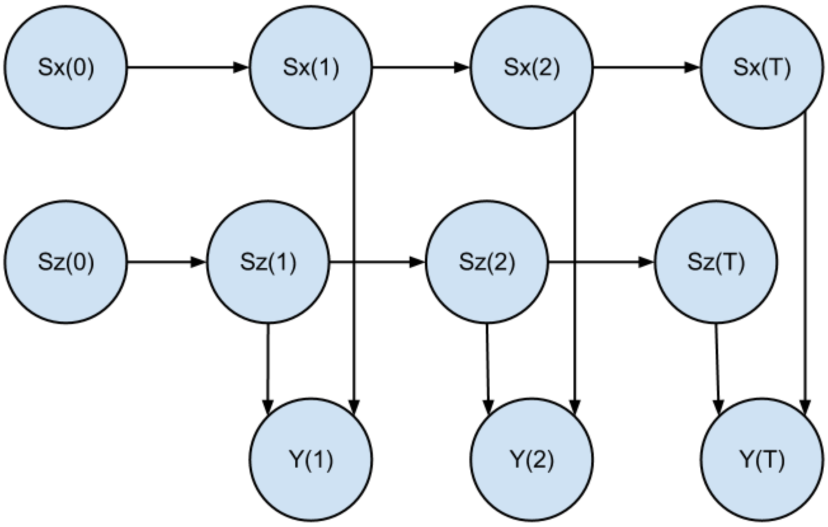

Figure 3: Factorial Hidden Markov Model

4 Conclusion

In this survey we have provided an overview over some techniques in blind source separation. Early work on blind source separation focused to a large extent on time domain ICA. In extending the BSS problem to multiple sources, the classical ICA method has been augmented by adopting different signal representations, where the time-frequency domain is particularly common. Different methods have also been introduced, some borrowing from the human auditory system attempting to hard-code domain specific knowledge. Others adopt different algorithms; one important example here being hidden markov models (HMMs). This is a very flexible approach to BSS as it allows for non-stationary mixing, and relaxes many of the stringent assumptions of classical ICA.

A Research Agenda

The aim of this study is to systematically review current technology for blind source separation (BSS), with particular emphasis on the particular subproblem of single channel blind source separation (SCBSS); that is, the recovery of several source signals from one observed signals.

A.1 Background

The blind source separation problem consists transforming a set of observed signals that has undergone some particular mixing process back to the original unobserved signals. The "blind" part of the problem refers to the fact that the nature of the mixing process is unknown. From original research on the blind source separation problem, focus has shifted from the case where with as many, or more recording channels than original sources, to the case of fewer channels than original sources. An important subproblem that we wish to focus on is where we have only one recording and attempt to recover multiple sources.

Our approach is two-fold: firstly we wish to look at studies about the performance of current single channel separation methods. Secondly, we wish to gain a broader overview over the state of research on BSS.

A.2 Research Questions

1. What are the different variations on the blind source separation problem, in particular as pertains to audio data.

2. Which methodologies and algorithms are applied to the different variations of the blind source separation problem as identified in Question 1.

3. What are the theoretical properties of the techniques identified in Question 2, and what assumptions do they make about the nature of the sources and the mixing process?

4. What empirical evidence is there to document the performance of the techniques identified in Question 2 as applied to the problems identified in Question 1?

A.3 Search Strategy

In reviewing the BSS literature we conduct a search of the below databases based on a set of keywords listed below. To filter the results we introduce a set of criteria to judge the relevance and quality of the results.

A.3.1 Databases

- SpringerLink
- CiteSeerX
- Google Scholar

A.3.2 List of Search Terms

blind source separation, single channel blind source separation, single mixture blind source separation, hidden markov blind source, single microphone blind source separation, blind source separation review, blind source separation survey, pca blind source separation, ica blind source separation, principal component analysis blind source separation, independent component analysis blind source separation.

A.3.3 Inclusion and Quality Criteria

We wish to study how various methods and/or approaches by which blind source problem is solved, which constraints are imposed by these methods, and how well a BSS system based on these ideas perform on real-life data. To filter out the most important studies to this end, we adopt the following criteria.

Inclusion Criteria

1. The main concern of the study is the BSS problem.
2. The algorithmic design decisions in the study must be justified.
3. The study describes a reproducible algorithm/method.
4. The study focuses on blind source separation of auditory signals.

Quality Criteria

1. The study presents empirical results.
2. More recent studies are preferred.
3. The described test data set is reproducible.
4. The study should present novel theoretical approaches/methodologies OR empirical results about previously known methods.
5. Literature reviews should discuss single channel blind source separation.
6. The study should describe which other algorithms/methods the proposed solution can be compared with and the performance measure used in comparison.

References

[1] Comon, P. (1994). “Independent Component Analysis: a new concept?”, Signal Processing, 36(3):287–314

[2] Bell, A.J. and Sejnowski, T.J. (1995)., “An information maximization approach to blind separation and blind deconvolution”, Neural Computation, 7, 1129-1159

[3] Pearlmutter, B. A. and Parra, L. C.(YYYY), “A Context-Sensitive Generalization of ICA”.

[4] Hyvarinen A (2001), “Independent Component Analysis”, Neural Computing Surveys, Neural Computing Surveys, vol 2.

[5] Bach, F.R. and Jordan, M.I. (2004), “Blind One-Microphone Speech Separation: A Spectral Learning Approach”.

[6] Hyvärinen, A. (1999), “Fast and Robust Fixed-Point Algorithms for Independent Component Analysis”. IEEE Transactions on Neural Networks 10(3):626-634, 1999.

[7] Sam T. Roweis(2001). “Neural Information Processing Systems 13” (NIPS’00). pp. 793-799

[8] Davies, M.E. and James, C.J. (2007). “Source separation using single channel ICA”, Signal Process., vol. 87, no. 8, pp. 1819–1832, 2007.

[9] Cardoso, J. L.(1998), “Multidimensional Independent Component Analysis”, Proceedings of ICASSP 1998.

[10] Mijovic, B. De Vos, M., Gligorijevic, I., Taelman, J. and Van Huffel, S. (2010), “Source Separation From Single-Channel Recordings by Combining Empirical-Mode Decomposition and Independent Component Analysis”, IEEE Transactions on Biomedical Engineering, vol. 57, no. 9, September 2010