

Separação Cega de Sinais de Áudio: Fundamentos e Perspectivas

Wesley Pavan¹, Rodolfo José Leite Netto², Sergio Vieira Bueno², Everton Zaccaria Nadalin²,
Romis Attux², João Marcos Travassos Romano¹

1 – Departamento de Microonda e Óptica (DMO)

2 – Departamento de Engenharia de Computação e Automação Industrial (DCA)

Faculdade de Engenharia Elétrica e de Computação (FEEC)

Universidade Estadual de Campinas (Unicamp)

Caixa Postal 6101, CEP 13083-852 – Campinas, SP, Brasil

{rodolfo, sbueno, nadalin, attux}@dca.fee.unicamp.br, wesleypavan@gmail.com,
romano@dmo.fee.unicamp.br

Abstract – In this work, we discuss the problem of blind source separation (BSS) having in view an application in the context of acoustic signal processing (ASP). After presenting the fundamentals of BSS, we introduce the concept of independent component analysis (ICA) and indicate some aspects that are peculiar to ASP, concluding with some research perspectives.

Keywords – Audio source separation, blind source separation, independent component analysis, sparse component analysis, convolutive mixtures.

1. Introdução

Neste trabalho, pretendemos expor os fundamentos do problema de separação cega de fontes e da estratégia mais consolidada para sua solução, aquela baseada em análise de componentes independentes, para, em seguida, traçarmos um panorama da peculiaridades desse problema no contexto de sinais de áudio. Por fim, apresentaremos algumas perspectivas associadas à pesquisa que temos buscado desenvolver.

O trabalho está estruturado da seguinte forma: na seção 2, apresentamos o problema de separação cega de fontes e a idéia de usar análise de componentes independentes para resolvê-lo; na seção 3, discutimos alguns aspectos relevantes quando se lida com sinais de áudio, e, no capítulo 4, apresentamos nossas conclusões e perspectivas.

2. Separação Cega de Fontes

Em processamento de sinais, é comum a necessidade de se trabalhar com informação que sofreu algum tipo de distorção tendo em vista recuperar ou equalizar algum atributo específico. Uma aplicação desse tipo, conhecida como separação cega de fontes (BSS, do inglês *blind source separation*) [1], pode ser ilustrada pela Fig. 1, na qual um conjunto de fontes de informação, ao qual não se tem acesso, fornece uma série de sinais resultantes de um sistema de mistura, que são os dados disponíveis. O objetivo é então projetar um sistema separador capaz de estimar os sinais originais.

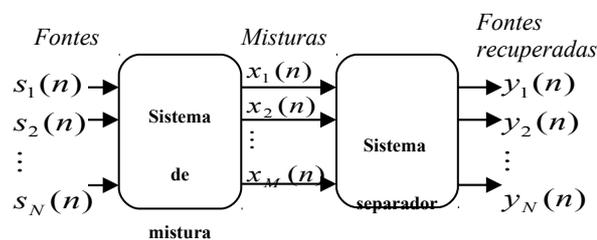


Figura 1 – Problema de BSS.

Conforme a Fig. 1, que ilustra fundamentalmente um sistema *MIMO* (*Multiple-Input/Multiple-Output*) [2], nota-se que $s_i(n)$, com $i=1, \dots, N$, representa as N fontes de informação originais, $x_j(n)$, com $j=1, \dots, M$, os sinais aos quais temos acesso (misturas) e $y_i(n)$ as N fontes recuperadas pelo sistema separador. Como se trata de um problema de processamento não-supervisionado, considera-se que não há informações específicas sobre o sistema de mistura. O modelo usualmente adotado para esse sistema é linear, instantâneo, com número de misturas igual ao número de fontes ($M=N$) e sem ruído. Para esse caso, uma representação matricial é dada por,

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (1)$$

em que \mathbf{A} é a matriz $N \times N$ que representa o sistema de mistura, \mathbf{x} é o vetor de misturas e \mathbf{s} é o vetor de fontes.

Intuitivamente, o sistema separador para esta configuração é estruturalmente similar e inverso ao sistema de mistura. Assim, a estrutura

$$\mathbf{y} = \mathbf{W}\mathbf{x} \quad (2)$$

com a matriz $N \times N$ \mathbf{W} correspondendo a um dispositivo de separação inverso à matriz \mathbf{A} , seria uma forma simples de anular o efeito de

sobreposição do canal misturador. Todavia, lembrando que o foco é obter os sinais originais, usualmente se aceita recuperar as fontes ponderadas por fatores multiplicativos e em qualquer ordem, sendo esses efeitos descritos matematicamente com o auxílio de uma matriz diagonal \mathbf{D} e de uma matriz \mathbf{P} de permutação. Assim, idealmente,

$$\mathbf{W} = \mathbf{DPA}^{-1} \quad (3)$$

Logo, para solucionar o problema de separação de fontes, é necessário encontrar uma matriz \mathbf{W} que obedeça à definição acima, o que, no caso em que a matriz \mathbf{A} e os sinais originais são desconhecidos, revela-se uma tarefa nada trivial.

Uma abordagem clássica para a obtenção da matriz de separação \mathbf{W} , neste contexto, é supor que as fontes são sinais estatisticamente independentes. Dessa forma, um critério que busque recuperar esta condição de independência na saída do separador é uma possibilidade viável, como provou Comon [3]. Essa constatação é o que une a ideia de análise de componentes independentes (ICA, do inglês *independent component analysis*) ao problema de BSS [1].

2.1 Análise de Componentes Independentes

A hipótese de que as fontes de informação são estatisticamente independentes permite que a metodologia de ICA seja adotada para inspirar soluções para o problema de BSS. Ocorre independência estatística sempre que é válida uma expressão do tipo:

$$p(y_1, y_2, \dots, y_K) = \prod_1^K p(y_i) \quad (4)$$

i.e. quando a densidade de probabilidade conjunta associada a K variáveis aleatórias é igual ao produto das correspondentes densidades marginais.

Comon provou que ajustar a matriz \mathbf{W} de modo a aproximar tanto quanto possível a condição expressa em (3) é suficiente para garantir a separação das fontes, desde que duas condições sejam garantidas: que haja no máximo uma fonte com distribuição gaussiana e que a matriz \mathbf{A} seja inversível [3]. Nesse contexto, resta apenas encontrar um modo de quantificar a independência, de modo a definir uma função de contraste capaz de ser aplicada para fornecer estimativas apropriadas para o sistema separador.

Os principais conceitos utilizados na metodologia ICA com a finalidade de recuperar de alguma forma a independência entre as misturas são não-gaussianidade, estimação de máxima verossimilhança e informação mútua

[4]. O uso de não-gaussianidade parte do teorema central do limite, o qual afirma, em termos simples, que uma soma de variáveis aleatórias independentes aproxima-se de uma variável com distribuição de probabilidade normal conforme o número de variáveis tende a infinito. Desse fato conclui-se, intuitivamente, que, ao tornar os dados “minimamente gaussianos”, estaríamos anulando o processo de misturas, ou seja, recuperando as fontes.

Dois funções de contraste são comumente utilizadas quando se lida com não-gaussianidade: a curtose,

$$kurt(y) = E\{y^4\} - 3(E\{y^2\})^2 \quad (5)$$

que se torna nula caso a distribuição de y seja gaussiana; e a negentropia, derivada da entropia diferencial, que é sempre não-negativa e nula apenas para uma distribuição gaussiana,

$$J(y) = H(y_{\text{GAUSS}}) - H(y) \quad (6)$$

onde $H(\cdot)$ é a entropia e y_{GAUSS} denota uma variável aleatória com a mesma estrutura de covariância de y .

Para lidar com o problema de ICA via estimação de máxima verossimilhança, pode-se considerar na formulação as densidades de probabilidade dos sinais $x_j(n)$ e a matriz $B = (b_1, b_2, \dots, b_n)^T$, escolhida como a inversa da matriz \mathbf{A} . A função de verossimilhança, expressa por $L(\cdot)$, é dada por,

$$L(B) = \prod_{t=1}^T \prod_{i=1}^n p_i(b_i^T x(t)) |\det B| \quad (7)$$

sendo que trabalha-se freqüentemente com o logaritmo da função (8) dividido pelo número de observações de \mathbf{x} . Ainda é pertinente acrescentar que, a rigor, as funções de densidade de probabilidade $p_i(\cdot)$ devem ser conhecidas *a priori*, o que não ocorre usualmente em casos reais.

Por fim, a informação mútua, um dos conceitos de grande importância na teoria de informação [5], proporciona a mais direta abordagem de ICA, já que quantifica de modo natural o grau de independência entre variáveis aleatórias. Essa grandeza é descrita, matematicamente, como

$$I(y_1, y_2, \dots, y_N) = \sum_{i=1}^N H(y_i) - H(y) \quad (8)$$

O ponto crucial é que $I(\cdot)$ sempre retorna um valor não-negativo, fornecendo um valor nulo apenas no caso em que as variáveis aleatórias são mutuamente independentes.

3. Particularidades do Problema de Separação de Sinais de Áudio

O arcabouço teórico visto até agora é aplicável ao problema de BSS que surge quando se busca separar fontes acústicas, ou seja, sinais sonoros capturados, na forma de misturas, por sensores (e.g. microfones). Isso é ilustrado pelo clássico exemplo correspondente ao *cocktail party problem*: várias pessoas conversam simultaneamente numa sala, e se deseja perceber com clareza o que cada uma delas diz, o que corresponde, em certa medida, a um processo de separação [1].

Não obstante, há algumas peculiaridades do problema de separação no contexto de sinais acústicos que são importantes caso se tenha em vista a construção de metodologias de alto desempenho. Essas peculiaridades surgem, interessantemente, tanto de certas propriedades das fontes quanto de aspectos relativos ao processo de misturas. Resumiremos essas peculiaridades por meio da exposição de dois pontos:

- 1) Num ambiente acústico, o processo de mistura não é, tipicamente, representado de modo fiel pela expressão (1), já que cada microfone capta tanto uma sobreposição de fontes quanto uma sobreposição de amostras dessas fontes associadas a diferentes instantes temporais. Em outras palavras, o modelo de mistura é essencialmente *convolutivo* [5]. Nesse caso, é possível construir o sistema separador como um conjunto de filtros, numa extensão da estrutura mostrada na equação (3), e abordar o problema no domínio do tempo (e.g. via predição [6]). É possível também lidar com o problema no domínio da frequência via ICA, desde que se resolvam certas ambigüidades de escala e permutação no contexto espectral [1] [5].
- 2) Sinais de fala e sinais de áudio em geral tipicamente possuem um caráter esparso, seja no domínio do tempo, seja no domínio da frequência. Isso permite que sejam construídas técnicas de BSS baseadas na noção de *sparse component analysis* (SCA), as quais podem ser extremamente relevantes em casos subparametrizados, ou seja, em que há menos sensores que fontes [5]. Em tais casos, que são muito comuns no

tratamento de sinais de áudio, métodos de ICA possuem usualmente um desempenho insatisfatório.

4. Conclusões e Perspectivas

Neste trabalho, apresentamos os fundamentos do problema de separação cega de fontes e da clássica abordagem para sua solução via ICA. Em seguida, analisamos alguns pontos que são peculiares à tarefa de BSS para sinais de áudio, como o caráter convolutivo do processo de mistura, a existência de esparsidade e o potencial de subparametrização.

No que se refere a nossos esforços de pesquisa, pretendemos realizar uma investigação do problema de BSS em domínios acústicos por meio de ensaios envolvendo sinais práticos captados em diferentes ambientes. Isso nos permitirá analisar propostas já existentes e também validar propostas teóricas desenvolvidas junto ao grupo de pesquisa, como aquelas expostas em [6] e [7].

Agradecimentos

Os autores agradecem o financiamento proporcionado por CAPES e CNPq.

Referências

- [1] A. Hyvarinen, J. Karhunen, E. Oja, *Independent Component Analysis*, Wiley, 2001.
- [2] A. V. Oppenheim, A. S. Willsky, S. Hamid Nawab, *Signals and Systems*, Prentice Hall, 1996.
- [3] P. Comon, "Independent Component Analysis: a New Concept?", *Signal Processing*, Vol. 36, No. 3, pp. 287-314, 1994.
- [4] R. Suyama, L. T. Duarte, A. Neves, R. Ferrari, R. Attux, C. C. Cavalcante, C. C. M. Junqueira, J. M. T. Romano, "Unsupervised Signal Processing: Concept, Applications and Trends", *Telecommunications: Advances and Trends in Transmission, Networking and Applications*, UNIFOR, 2006.
- [5] P. Comon, C. Jutten (eds.), *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, Academic Press, 2010.
- [6] R. Suyama, L. T. Duarte, R. Ferrari, L. E. P. Rangel, R. Attux, C. C. Cavalcante, F. J. Von Zuben, J. M. T. Romano, "A Nonlinear

Prediction Approach to the Blind Separation of Convulsive Mixtures”, EURASIP Journal on Advances in Signal Processing, Vol. 2007, Article ID 43860, 2007.

- [7] E. Z. Nadalin, R. Suyama, R. Attux, “An ICA-Based Method for Blind Source Separation in Sparse Domains”, Lecture Notes in Computer Science, Vol. 5441, pp. 597-604, 2009.