

Avaliação da Variação Total de Espectrogramas para a Extração Cega de Fontes

Giulio G. R. Suzumura, Ricardo Suyama

Resumo—Abordagens clássicas de separação de fontes como as baseadas em Análise por Componentes Independentes e Análise por Componentes Esparsas são amplamente difundidas e, dependendo da aplicação, retornam bons resultados. Entretanto, no estudo de separação de sinais de áudio, explorar características em outros domínios pode trazer melhores resultados. Transformar sinais de áudio em espectrogramas faz com que estes possam ser interpretados como imagens e, portanto, utilizar técnicas de processamento de imagens pode trazer novas perspectivas ao problema. Nesse trabalho, baseando-se em métricas associadas à nitidez de imagens, avaliamos o uso da Variação Total do espectrograma no problema de separação cega de fontes, e os resultados preliminares indicam que a ferramenta pode ser útil para a construção de uma nova abordagem para recuperação dos sinais.

Palavras-Chave—Variação Total, Índice de Nitidez, Nitidez de Espectrogramas, Extração Cega de Fontes.

Abstract—Classical approaches to source separation such as those based on Independent Component Analysis and Sparse Component Analysis are widespread and, depending on the application, yield good results. However, in audio signal separation study, exploring features in other domains may yield better results. Turning audio signals into spectrograms causes them to be interpreted as images, therefore using imaging techniques can bring new perspectives to the problem. In this work, based on metrics associated with image sharpness, we evaluated the use of the Total Variation of the spectrogram in the problem of Blind Source Separation, and the preliminary results indicate that the tool can be useful for the construction of a new approach for sigever recovery.

Keywords—Total Variation, audio Sharpness Index, Spectrogram Sharpness, Blind Source Extration.

I. INTRODUÇÃO

O problema de separar sinais de vozes misturadas ocorre em diversas situações do nosso cotidiano e, além disso, este problema passa despercebido por grande parte das pessoas. Almoçar sozinho em uma praça de alimentação com grande quantidade de pessoas é um ambiente ideal para exemplificar esta situação. Nesta e diversas outras situações a concentração de um indivíduo deve ser capaz de separar e interpretar cada falante específico em meio aos diversos sons interferentes, conceito conhecido como audição seletiva [1].

A modelagem computacional deste problema está diretamente ligada ao estudo do problema de Separação Cega de

Giulio Guiyti Rossignolo Suzumura, UFABC, CECS, Engenharia da Informação, giulio.suzumura@ufabc.edu.br; Ricardo Suyama, UFABC, CECS, Engenharia da Informação, ricardo.suyama@ufabc.edu.br. O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, e da Universidade Federal do ABC – UFABC.

Fontes (*Blind Source Separation* – BSS) que, por se basear em uma formulação abrangente, pode ser aplicada em diferentes cenários para a recuperação dos objetos de interesse. [2], [3], [4].

Uma das principais técnicas associadas à separação de fontes é a Análise por Componentes Independentes (*Independent Component Analysis* – ICA), que apresenta bons resultados quando as fontes satisfazem duas suposições: os sinais são independentes entre si e não há mais do que uma fonte com distribuição Gaussiana [5].

Há, entretanto, outras abordagens que exploram outras características particulares dos sinais de interesse. Por exemplo, no caso de sinais de áudio, sabe-se que os sinais de interesse são esparsos no domínio tempo-frequencial, o que possibilita o uso da Análise por Componentes Esparsas (*Sparse Component Analysis* – SCA) para identificação do processo de mistura e separação dos sinais [2], [6]. Nesse contexto, uma das ferramentas usuais para a análise dos sinais é o espectrograma, que fornece uma representação visual do conteúdo espectral dos sinais na medida em que estes variam no tempo. Dessa forma, pode-se considerar que essa imagem representa a informação contida nos sinais, o que abre uma nova perspectiva para métodos de processamento dos sinais baseados em técnicas usualmente associadas ao processamento digital de imagens.

Uma possível abordagem, apontada em [7], relaciona a nitidez de uma imagem à inteligibilidade de sinais de voz. Assim como a qualidade de uma imagem está diretamente relacionada a sua nitidez, a qualidade de sinais de áudio tem relação direta com a inteligibilidade do sinal, e em ambos os casos, a qualidade está diretamente relacionada à distribuição dos pixels e coeficientes da representação tempo-frequencial, respectivamente.

A proposta em [7], denominada *audio Sharpness Index* (aSI), analisa justamente este comportamento, i. e., a alteração da “nitidez” de um sinal de áudio. Sinais de áudio limpos são conhecidos por terem um espectrograma esparsos, composto por segmentos bem definidos – harmônicos no tempo e impulsivos na frequência. De forma contrária, áudios reverberantes ou ruidosos tem os segmentos manchados, borrados, no domínio tempo-frequencial.

Assim, filtrar um sinal de áudio limpo por um ruído Gaussiano branco faz com que a esparsidade deste seja alterada de forma drástica. Por outro lado, a aplicação deste procedimento em um áudio reverberante ou ruidoso não deverá resultar em grandes alterações na esparsidade. Com o estudo da variação deste resultado sobre diversos tipos de sinais de áudio, é possível descobrir em quais momentos o sinal filtrado se torna mais, ou menos, suscetível ao ruído aplicado. Dessa forma,

comparado a outros métodos de separação de sinais, a proposta de separação baseada no aSI traz vantagens em relação a técnicas que se aproveitam da ICA: não partem do pressuposto de independência clássica e da não-Gaussianidade das fontes [7].

A proposta do aSI, de fato, se baseia diretamente em uma métrica de nitidez em imagens [8]. No processamento digital de imagens, um dos fatores que qualificam a nitidez de uma fotografia se dá a partir da análise de bordas bem definidas de sua composição. Nessa situação, o gradiente de uma imagem nítida deve conter picos indicando a presença de bordas bem definidas, apresentado assim uma estrutura esparsa - fato que é explorado na definição do Índice de Nitidez (*Sharpness Index* – SI) .

Assim, a SI fornece uma medida relacionada à sensibilidade do gradiente de uma imagem à filtragem por um ruído branco. Se uma imagem nítida é convoluída com um ruído branco, o gradiente da imagem produzida deve apresentar picos menos pronunciados, alterando significativamente a esparsidade. Por outro lado, uma imagem borrada ou ruidosa, que possui originalmente gradiente não-esparsa, continuará sendo não-esparsa quando filtrada pelo mesmo ruído.

Computacionalmente simples, a SI tem grande mérito no estudo de recuperação cega de foco em imagens (*blind image deblurring*)[9]. Nesse sentido, sabendo que o processamento de imagens e áudio compartilham abordagens semelhantes, é possível supor que a aplicação do conceito de Índice de Nitidez pode trazer uma nova perspectiva aos estudos de separação de sinais de áudio.

Dessa forma, o objetivo do presente trabalho consiste em explorar o SI e, em particular, o conceito de Variação Total (*Total Variation* – TV) do espectrograma que está presente em sua definição – para o problema de separação de fontes. Para isso, o trabalho está dividido como segue. Primeiramente apresentamos os princípios do *audio Sharpness Index* na Seção 2. Estabelecemos a aplicação do conceito de nitidez em sinais de áudio na Seção 3. Finalmente desenvolvemos comparações experimentais na Seção 4.

II. Audio Sharpness Index

O *Audio Sharpness Index*, proposto em [7], busca quantificar a esparsidade do sinal de áudio a partir de medidas dos sinais envolvidos.

Como a esparsidade dos sinais de áudio não se revela de maneira tão pronunciada no domínio temporal, em geral considera-se a representação dos sinais em outros domínios (e.g., tempo-frequencial, wavelet etc) [10], obtidos através de uma transformação linear tal que

$$S(f, t) = \mathcal{F}\{s(t)\}, \quad (1)$$

onde \mathcal{F} representa algum tipo de transformação aplicada a uma janela de amostras, como a DFT *discrete fourier transform* ou DCT *discrete cosine transform* [11], e (f, t) os índices do domínio transformado. Ao considerar a transformada de Fourier, obtemos o espectrograma do sinal, obtendo assim uma representação tempo-frequencial do mesmo.

Para se quantificar o grau da esparsidade de um sinal é possível utilizar a norma ℓ_0 , cujo cálculo se dá pela

quantificação dos valores não nulos de uma sequência. De fato esta norma diferencia sinais mais esparsos de menos esparsos, porém sinais de áudio dificilmente serão compostos por valores nulos [12]. Nesse sentido, uma métrica alternativa normalmente utilizada é a norma “*entrywise*” ℓ_1 da matriz, i.e., considerando que uma matriz $S_{m \times n}$ pode ser tratada como um vetor $S_{mn \times 1}$, de maneira que a norma pode ser calculada a partir da soma do módulo dos elementos da matriz

$$\|S\|_1 = \sum_{f,t} |S(f, t)|, \quad (2)$$

i.e., a soma do módulo de todos os coeficientes do espectrograma.

Através da Eq. 2 é possível identificar que sinais mais esparsos possuem normas ℓ_1 relativamente menores quando comparados a sinais menos esparsos¹. No restante deste trabalho quando houver a indicação da métrica norma ℓ_1 esta representará a norma “*entrywise*” ℓ_1 , indicada pela Eq. 2, além disso, os índices de tempo (t) e tempo-frequência (t, f) foram suprimidos.

A hipótese explorada em [7] é de que sinais esparsos, quando convoluídos com um ruído gaussiano, resultam em sinais com esparsidade menor (aumentam a norma ℓ_1), enquanto que sinais não-esparsos filtrados com o mesmo tipo de ruído não apresentam grande variação em sua esparsidade. Portanto, seja s' o resultado da convolução de um sinal s com um ruído Gaussiano branco w de média zero e variância conhecida, i.e., $s' = s*w$, e S', S as respectivas representações no domínio tempo-frequencial, espera-se que

$$P(\|S'\|_1 \leq \|S\|_1) \quad (3)$$

seja pequena para sinais esparsos (pois a convolução com o ruído tende a aumentar a norma ℓ_1), e não tão pequena para sinais ruidosos.

Considerando adicionalmente que $\|S'\|_1$ apresenta uma distribuição Gaussiana com média e variância definidas por

$$\|S'\|_1 \sim \mathcal{N}(\mathbb{E}[\|S'\|_1], \text{Var}[\|S'\|_1]), \quad (4)$$

[7] define o *audio Sharpness Index* (aSI) de um dado sinal s como

$$\begin{aligned} aSI(s) &= -\log(P(\|S'\|_1 \leq \|S\|_1)) \\ &= -\log Q\left(\frac{\|S\|_1 - \mathbb{E}[\|S'\|_1]}{\text{Var}[\|S'\|_1]}\right) \end{aligned} \quad (5)$$

onde $Q(\cdot)$ é função de distribuição acumulada complementar de uma distribuição Gaussiana padrão.

No artigo original [7], os autores mostram que o aSI é um bom preditor do *Speech Transmission Index* (STI), portanto pode ser utilizado como uma medida de inteligibilidade pois o STI é uma métrica objetiva bem estabelecida. Além disso, o aSI não precisa de uma referência para ser calculado, uma grande vantagem em relação a outras métricas.

¹A quantificação de esparsidade é inversamente proporcional a linguística utilizada sobre o valor calculado da norma ℓ_1 , o que pode causar certa confusão em comparações. Quanto mais esparsa um sinal menor a norma ℓ_1 e, inversamente, quanto menos esparsa um sinal maior a norma ℓ_1 .

O aSI se apresenta eficaz para classificação de inteligibilidade de áudios e como guia para algoritmos de Separação Cega de Fontes. Entretanto, o processo utilizado por esta solução apresenta um alto custo computacional, uma vez que é necessário estimar a esperança e variância de S' presentes em (5), fato que motivou a busca por uma abordagem que utilize simplificações nos princípios apresentados.

III. VARIAÇÃO TOTAL DO ESPECTROGRAMA

Analisar a nitidez do espectrograma do sinal através da Variação Total (TV) ao invés do cálculo de aSI pode trazer um novo horizonte ao estudo de separação de fontes e, possivelmente, a simplificação do processo de estimação de mistura.

Assim como na aSI, o SI verifica qual o efeito da convolução da imagem com um ruído gaussiano branco. A diferença reside na métrica utilizada, que no caso do SI corresponde a Variação Total da imagem filtrada[8].

Considerando um espectrograma S , a TV do sinal s é definida por

$$TV(s) := \|\partial_f S\|_1 + \|\partial_t S\|_1, \quad (6)$$

onde $\partial_t S = \dot{S}(f, t+1) - S(f, t)$, $\partial_f S = \dot{S}(f+1, t) - S(f, t)$ e \dot{S} representa a periodização do espectrograma S .

A tentativa de utilizar a TV ao invés do aSI se deve ao fato de que a TV explora os contornos de uma imagem, portanto os contornos dos espectrogramas.

A fim de comparar o comportamento das métricas do aSI e TV aplicadas a sinais de áudio, consideramos um vetor $\mathbf{n} = [n_1, n_2, \dots, n_N]^T$ composto por N ruídos Gaussianos brancos com média 0 e variância 0.1. Criando diferentes versões ruidosas de um sinal

$$s_i(t) = s_{i-1}(t) + n_i(t),$$

onde $i = 1, \dots, N$ e s_0 um sinal de voz limpo, foi possível verificar a variação da TV e do aSI sobre os s_i sinais distorcidos.

Os resultados apresentados pela Figura 1 indicam que as curvas associadas à média do logaritmo da Variação Total² e à obtida pela média do *audio Sharpness Index* são, mesmo que em escalas diferentes, próximas, porém o $\log(TV)$ é mais sensível do que o aSI.

É imprescindível apontar, entretanto, que os valores obtidos são resultados da média entre blocos do sinal filtrado, portanto é possível que cada bloco não apresente o mesmo valor de TV, ou seja, o ruído pode interferir de forma diferente em cada segmento do sinal de áudio original. Neste sentido é intuitivo dizer que alguns blocos podem ser mais suscetíveis ao ruído do que outros.

IV. EXTRAÇÃO DE FONTES COM A TV

Seja um sistema determinado estereofônico instantâneo, i.e., duas misturas (x_1, x_2) produzidas a partir de duas fontes (s_1, s_2) , criado através de um processo de mistura $\mathbf{x} = \mathbf{A}\mathbf{s}$,

²A função $\log(\cdot)$ foi utilizada sobre métrica TV para fins de comparação com o parâmetro aSI, o qual utiliza escala logarítmica apresentada na Eq. (5).

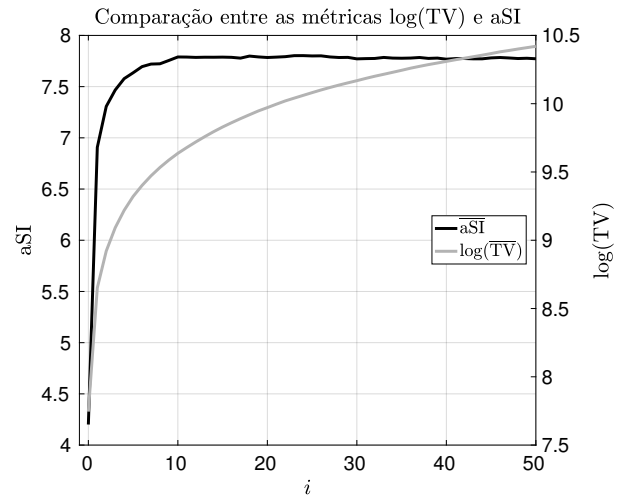


Figura 1. Valores de $\log(\overline{TV})$ e \overline{aSI} de cada sinal s_i para $i = 0, \dots, 50$, com s_0 sendo um sinal de voz masculina amostrada a 16KHz. Os traços sobre as métricas representam a média obtida a partir de 17 blocos (com sobreposição de 50%) parametrizados com tamanho de 8192 amostras cada, i.e, 512ms.

onde $A(\cdot, i) = [\cos \theta_i \sin \theta_i]^T$ e $i \in \{1, 2\}$, é possível obter as fontes estimadas \hat{s} de forma ótima através da utilização de uma matriz inversa W tal que $W.A = I$, onde $\hat{s} = W.x$.

A melhor estimativa dos sinais \hat{s}_i se dá quando os valores estimados $\hat{\theta}_i$ forem os mesmos utilizados no processo de mistura.

Desconhecendo os valores de θ_i mas sabendo que o processo de mistura foi realizado de forma ortogonal, uma boa proposta é estimar diferentes sinais em função de θ a partir de uma função que leve em consideração cada mistura em uma posição ortogonal a outra, e.g.

$$y_\theta = x_1 \sin \theta - x_2 \cos \theta. \quad (7)$$

Assim, a partir da teoria desenvolvida sobre a Variação Total, é possível percorrer todo o espaço linear encontrando diferentes sinais y_θ . Sabendo que as estimativas das fontes limpas tendem a ter menores valores de TV, minimizar esta métrica sobre o sinal y_θ pode trazer informações sobre o processo de mistura.

O valor da TV calculada sobre o espectrograma do sinal inteiro pode trazer resultados imprecisos, portanto se faz necessário dividir o sinal em blocos. Desta forma, uma análise mais eficaz pode ser realizada a partir da média entre os valores da TV de cada bloco.

Como existe a possibilidade de haver mudanças repentinas do sinal analisado nos interstícios destes segmentos, é interessante utilizar blocos sobrepostos, a fim de obter melhores estimativas.

A. Análise da TV

A partir do modelo de mistura apresentado e utilizando dois sinais de música (guitarra e piano) misturados com $\theta_1 = \pi/10$ e $\theta_2 = 4\pi/5$, foram calculados os valores de $TV(y_{\theta_{bl}}) \forall \theta \in [0, \pi]$, sendo bl diferentes blocos, segmentos, de y_θ .

Os resultados apresentados pelos valores dos blocos – Figura 2 – não apresentam vales expressivos nas posições que

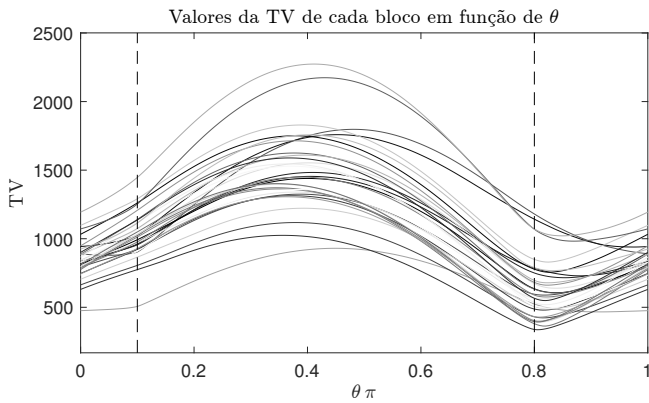


Figura 2. Variação Total de 31 segmentos do sinal y_θ com 8192 amostras cada e sobreposição de 50%. Sinal y_θ criado a partir de duas misturas de duas fontes musicais (guitarra e piano) amostradas a 32KHz.

corresponderiam à recuperação das fontes – indicadas pelas linhas tracejadas. De fato, cada bloco obtém um valor de TV distinto um do outro. Em alguns blocos é possível que uma das fontes se sobressaia sobre a outra, fazendo com que o mínimo da curva se desloque para os diferentes blocos. Dessa forma, a curva média $\overline{TV}(y_\theta)$ – Figura 3 – também não fornece informações precisas.

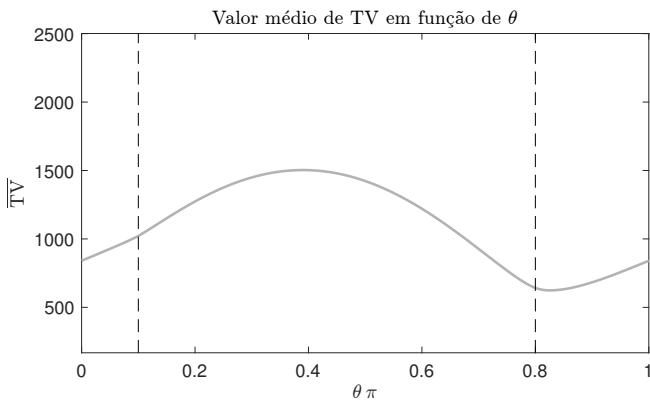


Figura 3. Curva da TV média obtida a partir das $M = 31$ curvas apresentadas na Figura 2, i.e. $\overline{TV} = M^{-1} \cdot \sum_{bl=1}^M TV(y_{\theta_{bl}})$.

Entretanto, mesmo que a média entre os blocos não traga resultados precisos sobre o processo de mistura, ela fornece ao menos uma ideia da localização dos mínimos, conforme observa-se na Figura 3. Nela, nota-se que a métrica, em geral, varia de maneira suave em grande parte do intervalo de θ , mas em algumas regiões nota-se uma variação mais aguda, que pode ser melhor avaliada através da derivada de segunda ordem da curva média - conforme pode ser visualizado na Figura 4.

As maiores variações desta medida apresenta uma ótima estatística sobre a estimativa do processo de mistura (θ_1 e θ_2). Porém, mesmo interessante de se apresentar a taxa de variação, utilizar um método de busca sobre esta curva é inviável de se reproduzir no contexto de separação de fontes. Isto se deve ao fato de que para se extrair esta informação foi necessário realizar uma varredura de θ em todo o intervalo considerado.

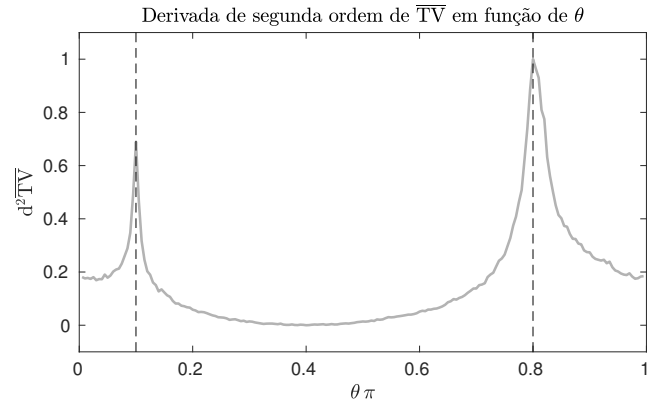


Figura 4. Taxa de variação da curva \overline{TV} . Valores obtidos a partir da derivada de segunda ordem da curva média apresentada pela Figura 3.

B. Análise da TV entre blocos

Como os sinais de áudio são esparsos no tempo e na frequência é possível que diversos blocos tenham melhores resultados do que outros para uma mesma posição no espaço, assim uma forma mais interessante de se dimensionar corretamente os resultados obtidos se dá pela normalização de cada bloco, como apresentado na Figura 5.

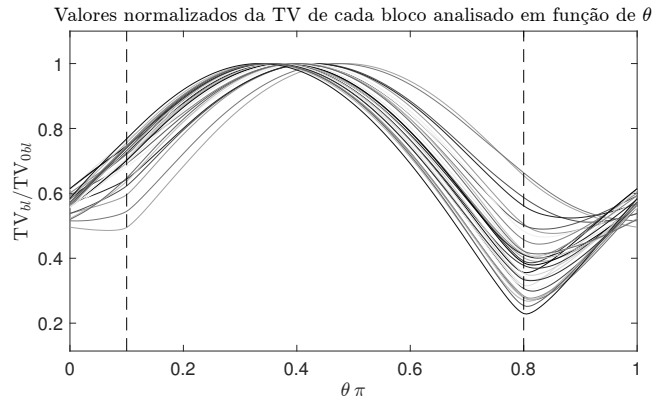


Figura 5. Valores normalizados de TV de cada bloco (bl) apresentado na Figura 2.

Mesmo que simples, esta manipulação consegue extrair uma característica interessante: as posições onde efetivamente foram realizadas as misturas apresentam as maiores variações de TV entre os blocos. Isto se deve ao fato de que em um bloco pode haver uma fonte dominante, portanto um valor de TV reduzido. Já em outro bloco esta mesma posição não sofreu dominância nenhuma.

Tendo isso em mente, a partir do conhecimento de que o processo de mistura não foi alterado entre os blocos, é possível analisar a variação entre as curvas para estimar o processo de mistura original, conforme apresentado na Figura 6.

Comparando os resultados obtidos da TV e do aSI, observa-se que ambos apresentam picos locais nas posições reais de mistura. De fato ambas métricas apresentam resultados similares relacionados com a inteligibilidade. Porém, é possível dizer que a variância da TV tem maiores distâncias entre os

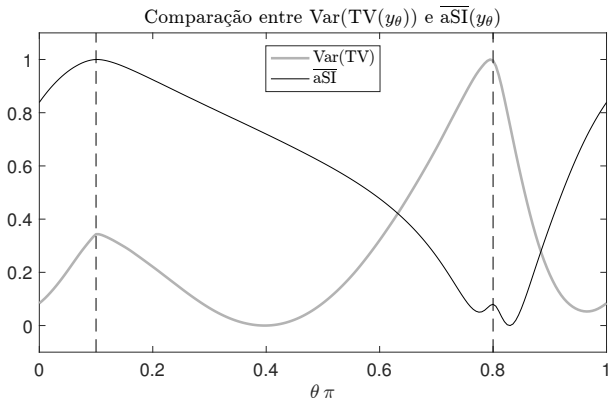


Figura 6. Variância de TV e valores de aSI normalizados entre os blocos do sinal y_θ . Sistema determinado com duas fontes musicais (guitarra e piano).

extremos máximos e mínimos, algo que poderia facilitar a otimização do problema.

Os resultados preliminares apresentados impõem a necessidade de avaliar em maior profundidade o critério, considerando tipos de áudio diferentes; tamanho e quantidade de blocos analisados; e quantidade de fontes.

Entretanto, já é possível dizer que o estudo da variância da Variação Total traz uma vantagem significativa ao estudo de separação de fontes: o custo computacional – o tempo de cálculo do aSI foi, em média, 340x superior ao cálculo relacionado a TV³.

Uma comparação entre TV e aSI utilizando um sistema subparametrizado, com três fontes de voz em duas misturas, indica – através da Figura 7 – que a Variação Total também é mais robusta em relação ao aSI quando se trata de misturas subparametrizadas.

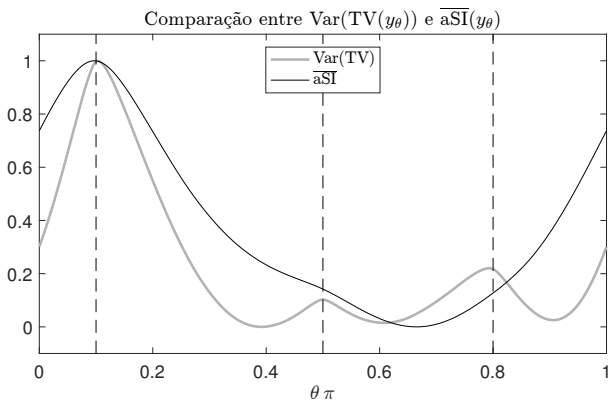


Figura 7. Variância de TV e valores de aSI normalizados entre os blocos do sinal y_θ . Sistema subparametrizado com três sinais de voz. Valores obtidos com blocos de 512ms.

A fim de maximizar localmente a função custo associada a variância, é possível utilizar um algoritmo multimodal como por exemplo o *Multiple Objective Particle Swarm Optimization* (MOPSO) [13]. Com este algoritmo é possível avaliar o custo de um mesmo espaço θ em diferentes blocos buscando

³Foi utilizado o software Matlab R2016b sobre um processador AMD FX(tm)-8350 Eight-Core.

maximizar a distância entre eles, i.e., maximizar a variância entre as partículas.

V. CONCLUSÕES

A aplicação do conceito de nitidez em espectrograma de sinais de áudio trazidos pelo *audio Sharpness Index*, idealizado para medir inteligibilidade, traz uma nova perspectiva ao conceito de separação cega de fontes de áudio.

A esparsidade pode ser medida de diferentes maneiras, mas a utilização do conceito da Variação Total se fez pertinente. Além da avaliação de sinais limpos e ruidosos, o conceito aplicado ao problema de separação de sinais de áudio pode trazer boas estimativas quando realizado a partir de subdivisões do sinal misturado.

Diferente de abordagens que utilizam grande poder computacional, avaliar a variância da Variação Total entre os blocos parece ser uma alternativa interessante, o que confere alta eficiência no processo de separação de fontes.

A partir deste trabalho é possível realizar estudos mais avançados com a utilização da Variação Total em algoritmos de separação de sinais de áudio. Além disso, é de interesse dos autores aperfeiçoar o método de análise da Variação Total para diferentes cenários e utilização de um otimizador na busca de obter as melhores estimativas dos sinais separados.

REFERÊNCIAS

- [1] N. Mesgarani and E. F. Chang, "Selective cortical representation of attended speaker in multi-talker speech perception," *Nature*, vol. 485, no. 7397, p. 233, 2012.
- [2] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press, 2010.
- [3] P. Comon, "Independent component analysis, a new concept?" *Signal processing*, vol. 36, no. 3, pp. 287–314, 1994.
- [4] T.-W. Lee, M. S. Lewicki, M. Girolami, and T. J. Sejnowski, "Blind source separation of more sources than mixtures using overcomplete representations," *Signal Processing Letters, IEEE*, vol. 6, no. 4, pp. 87–90, 1999.
- [5] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent component analysis*. John Wiley & Sons, 2004, vol. 46.
- [6] G. G. R. Suzumura and R. Suyama, "Estimação de sinais de voz esparsificados em misturas subparametrizadas," In: *XXXV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais - Anais do XXXV SBRT*, 2017.
- [7] G. Mahé, L. Moisan, and M. Mitrea, "An image-inspired audio sharpness index," in *2017 25th European Signal Processing Conference (EUSIPCO)*. IEEE, 2017, pp. 683–687.
- [8] G. Blanchet and L. Moisan, "An explicit sharpness index related to global phase coherence," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2012, pp. 1065–1068.
- [9] A. Leclaire and L. Moisan, "Blind deblurring using a simplified sharpness index," in *International conference on scale space and variational methods in computer vision*. Springer, 2013, pp. 86–97.
- [10] P. Bofill and M. Zibulevsky, "Underdetermined blind source separation using sparse representations," *Signal processing*, vol. 81, no. 11, pp. 2353–2362, 2001.
- [11] R. R. Mergu and S. K. Dixit, "A new paradigm for plotting spectrogram," *Journal of Information Systems and Communication*, vol. 3, no. 1, p. 158, 2012.
- [12] E. Z. Nadalin, R. Suyama, and R. Attux, "Um breve estudo sobre análise de componentes esparsos," In: *III Encontro de Alunos e Docentes do Departamento de Engenharia de Computação e Automação Industrial - Anais do III EADCA*, 2010.
- [13] K. E. Parsopoulos and M. N. Vrahatis, "Multi-objective particles swarm optimization approaches," in *Multi-objective optimization in computational intelligence: Theory and practice*. IGI global, 2008, pp. 20–42.