

# Alguns Aspectos do Problema de Localização de Fontes Acústicas

Sergio Vieira Bueno<sup>1</sup>, Rodolfo José Leite Netto<sup>1</sup>, Everton Zaccaria Nadalin<sup>1</sup>, Wesley Pavan<sup>2</sup>, Romis Attux<sup>1</sup>, João Marcos Travassos Romano<sup>2</sup>

1 - Departamento de Engenharia de Computação e Automação Industrial (DCA)

2 – Departamento de Microonda e Óptica (DMO)

Faculdade de Engenharia Elétrica e de Computação (FEEC)

Universidade Estadual de Campinas (Unicamp)

Caixa Postal 6101, CEP 13083-970 – Campinas, SP, Brasil

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

**Abstract** – In this work, we present elements of the problem of source localization, focusing on key aspects like tracking, distance between sources and sensors and distance between sensors, as well as on the main characteristics of a number of well-established approaches.

**Keywords** – source localization, beamforming, acoustic signal processing.

## 1. Introdução

A estimação da posição de uma ou várias fontes de som é uma tarefa fundamental em várias aplicações, como localização de locutor [1][2] e mesmo filtragem espacial, caso em que a estimação da direção da fonte do sinal viabiliza um sistema que direciona o diagrama de irradiação de um conjunto de antenas, o que permite obter uma melhor relação sinal-ruído [3][4][5]. Também poderiam ser citados exemplos de aplicação em navegação de robôs, próteses auditivas, sistemas de reconhecimento e gravação automática [6].

Neste trabalho, buscaremos discutir alguns fundamentos do problema de localização de áudio, com ênfase em uma série de aspectos-chave que, futuramente, serão explorados em nossos trabalhos de pesquisa.

## 2. O Problema de Localização de Áudio

As duas técnicas mais clássicas para a localização de fontes acústicas são:

(i) *Conformação de feixes (Beamforming - BF)* [4], abordagem que se caracteriza por uma estimativa da distribuição de potência do espectro espacial dos sinais e, a partir daí, da direção de chegada como aquela em que há a maior potência espectral. [7].

Dentre os algoritmos para BF destacam-se o DS [8] e o FBS [2], que possuem a vantagem de serem simples, exigindo um baixo custo computacional e permitindo inferir o número de sinais presentes pela quantidade de picos encontrados. No entanto, ambos oferecem uma resolução relativamente baixa. Já o método proposto por Capon [9] oferece resolução melhor que o DS, mas possui um custo computacional adicional devido à necessidade de determinadas manipulações matriciais.

Vale ressaltar que um dos limitantes dos métodos acima citados é que estes foram propostos para trabalhar com sinais de banda estreita, sendo, no entanto, o sinal de interesse de banda larga.

(ii) *Estimação de atraso temporal (Time Delay Estimation - TDE)*, que estima a localização com base no tempo de atraso de chegada de sinais em pares de receptores.

Destacam-se os seguintes algoritmos: *Generalized Cross Correlation* (GCC) [10], que investiga a correlação cruzada entre as saídas dos elementos do arranjo, de forma a determinar para que atraso  $\tau$  a função de correlação cruzada atinge seu máximo; *Interaural Time Difference* (ITD) [1], que visa a simular de forma simples a audição - usa-se uma janela de som e é procurado o máximo da função de correlação; *Eigenvalue Decomposition Algorithm* (EDA) [11], que baseia-se em uma decomposição de autovalores. O autovetor correspondente ao autovalor mínimo da matriz de covariância dos sinais do microfone contém a resposta ao impulso entre a fonte e os sinais do microfone, e, portanto, todas as informações necessárias para a estimativa do tempo de atraso.

Por fim, deve-se sempre lembrar que métodos de BF não exigem o cálculo de diferenças de tempo exigido pelos métodos de TDE, mas uma necessária pesquisa multidimensional sobre locais de origem das potências percebidas pode ser computacionalmente muito exigente. Assim, é comum o uso de uma combinação entre alguns dos métodos citados.

Os métodos TDE clássicos transformam o frame de dados recebidos em uma função localização (FL) que exibe um pico na localização correspondente à fonte. A desvantagem prática destas abordagens é que a captação de áudio é tipicamente convolutiva devido à reverberação, o que causa picos que podem ter maior amplitude do que o pico devido à verdadeira origem, de modo que simplesmente escolher o pico máximo para estimar a localização da fonte pode não gerar resultados precisos.

### **3. Problemas de Rastreamento de Fontes Sonoras Não-Estacionárias**

Técnicas baseadas puramente em estatísticas fornecidas durante um “quadro atual” não levam em conta informações sobre os padrões de movimento da fonte. Para a localização de uma fonte sonora móvel,

algumas técnicas exploram a expressiva dependência em curtas durações do sinal de áudio. [12,13].

De forma a ser possível rastrear de maneira seletiva o emissor, um cálculo em tempo real da direção de chegada (*DOA tracking*) para sinais banda-larga provenientes de múltiplas fontes móveis foi feito por Yan e Fan [14].

Os autores supõem que os sinais captados possuem ciclo-estacionariedade de forma a ser possível perseguir de maneira seletiva o emissor. O princípio utilizado é o de que interferências e ruídos que não compartilham da mesma frequência de ciclo dos sinais desejados, ou que não exibem ciclo-estacionariedade, podem ser suprimidos, de forma a melhorar a estimativa DOA [24].

Uma abordagem bayesiana para a filtragem espacial por meio de estimativa DOA quando estão presentes incertezas nas medidas devido ao ruído, é apresentada por Song [15]. Em boas condições da relação sinal-ruído (*Signal-to-Noise Ratio* - SNR) inferida, dá-se maior ênfase nas estimativas DOA. Para SNR baixas, as estimativas DOA têm menor importância e a estimativa final considera com maior peso a informação *a priori* da fonte de sinal. Considera-se que é conhecida a função de densidade de probabilidade da estimativa DOA e, assim como em [14], que ambos, sinal e ruído, seguem a distribuição Gaussiana estacionária descorrelacionada de média zero.

Já segundo Vincent [16], um modelo de fonte gaussiana não-estacionária seria mais apropriado para fontes de áudio do que um modelo estacionário não-gaussiano, indicando que ainda não há um consenso sobre o melhor modelamento estatístico para as fontes de áudio.

Por fim, um conjunto de métodos de detecção que são baseados na teoria de informação e a integração de um sistema de detecção de voz ativa, considerando as lacunas da fala humana, é uma proposta dada por Talanzis [13].

## 4. Sensores

### 4.1 Sensores - Número e Geometria

O número de sensores utilizado em um arranjo determina o número máximo de sinais cuja direção de chegada é passível de determinação. Acontece que, em ambientes reais onde ocorrem ondas sonoras, além da ocorrência da reverberação, existem múltiplas fontes, levando tipicamente à presença de mais fontes que sensores. Já o impacto da geometria da matriz de sensores para a exatidão a localização não é totalmente compreendido ainda, o que encoraja o estudo de vários formatos.

### 4.2 Distância entre os sensores e a fonte

Os emissores de onda devem estar suficientemente distantes dos sensores, de forma que as frentes de onda incidentes possam ser consideradas planas, mas não tão distantes que as ondas não possam ser percebidas pelos receptores.

Para estimação de localização de sinais sonoros, em especial, considera-se que, para um erro menor que  $10$  graus, o afastamento de  $15$  vezes a distância entre dois elementos adjacentes dos sensores é suficiente para uma aproximação de frente de onda plana [4].

### 4.3 Distância entre os sensores

Tem-se, pela versão espacial do teorema de Nyquist, que a distância entre os sensores deverá ser tal que  $d \leq \lambda/2$ , ou seja,  $d \leq v/2f$ , onde  $v = 340m/s$ . Assim, tomando-se o valor de  $100Hz$  para  $f$ , pois, os valores onde estão os componentes espectrais de maior potência para um sinal sonoro ou de voz estão entre  $100Hz$  e  $5kHz$ , a distância  $d$  deve então ficar abaixo de  $1,7 cm$ .

### 4.4 Seletividade do Arranjo em Função da Distância entre os Sensores

Considerar-se a distância entre sensores é também importante, pois, quanto maior for tal espaçamento, maior será a defasagem entre os sinais percebidos por elementos distintos do arranjo [4], conforme se pode ver na Fig. 1, na qual foi tomada uma distancia de  $17 cm$ , privilegiando assim a frequência de  $1kHz$ .

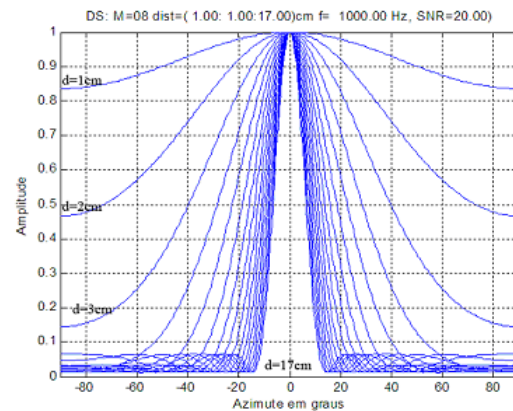


Figura 1: Largura do pico de potência em função da separação dos sensores

## 5. Conclusões e perspectivas

Neste trabalho, apresentamos os fundamentos do problema de localização de fontes acústicas, bem como o *modus operandi* de algumas abordagens clássicas. Foram discutidos de maneira sistemática pontos como rastreamento, número de sensores, distância entre fontes e sensores, distância entre sensores e seletividade do arranjo, o que permite traçar um panorama que, nos próximos passos deste trabalho, será devidamente explorado.

### Agradecimentos

Os autores agradecem à CAPES e ao CNPq o apoio financeiro.

### Referências

- [1] Vieira, José M. N. and Almeida, Luís and Santos, Paulo and Reis, João,

- "Goodears - Robot Orientado por Farol Acústico", em 7as Jornadas Hispano-Lusas de Ingeniería Eléctrica, Madrid, Espanha, Julho 2001.
- [2] Y. Tamai, Y. Sasaki, S. Kagami, and H. Mizoguchi, "Three ring microphone array for 3d sound localization and separation for mobile robot audition", em Proc. of 2005 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS2005), Edmonton, Canada, pp. 903-908, Agosto 2005.
- [3] L. C. Godara, "Application of antenna arrays to mobile communications—part II: beam-forming and direction-of-arrival considerations", em Proceedings of the IEEE, vol. 85, no. 8, pp. 1195–1245, 1997.
- [4] Gontijo, A. T., "Estimador de direção de chegada em tempo real com arranjo de microfones", Dissertação (Mestrado em Engenharia Elétrica) Universidade de Brasília, Brasília, 2010.
- [5] Reindl, K., Zheng, Y., Kellermann, W., "Speech enhancement for binaural hearing aids based on blind source separation" em Proc. 4th Int. Symp. Commun., Control, Signal Process. (ISCCSP), Limassol, Chipre, 2010.
- [6] Talantzis, F., Constantinides, A. G., "Using Information Theory To Detect Voice Activity", em Proceeding ICASSP '09 Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing.
- [7] A. Brutti, M. Omologo and P. Svaizer, "Comparison between different sound source localization techniques based on a real data collection", Hands-Free Speech Communication and Microphone Arrays (HSCMA), Trento, Itália, Maio 2008.
- [8] Veen, B. D. V., and Buckley, K. M., "Beamforming A Versatile Approach To Spatial Filtering", IEEE Signal Processing Magazine, vol. 5, no. 2. pp. 4-24, Abril 1988.
- [9] Capon, J., "High-Resolution Frequency-Wavenumber Spectrum Analysis", em Proceedings of the IEEE, 57, no. 8, pp. 1408-1419, 1969.
- [10] Knapp, C. H., Carter, G. C., "The Generalized Correlation Method for Estimation of Time delay", Acoustics, Speech and Signal Processing, IEEE Transactions In Acoustics, Speech, and Signal Processing [see also IEEE Transactions on Signal Processing], IEEE Transactions on, vol. 24, no. 4, pp. 320-327, 29 de Janeiro 2003.
- [11] Benesty, J., "Adaptive eigenvalue decomposition algorithm for passive acoustic source localization", J. Acoust. Soc. Am., vol. 107, no. 1, pp. 384-391, Janeiro 2000.
- [12] Ward, D. B., Lehmann, E. A., and Williamson, R. C., "Particle Filtering Algorithms for Tracking an Acoustic Source in a Reverberant Environment", IEEE Transactions on Speech and Audio Processing, pp. 826~836, 2003.
- [13] Talantzis, F., "An Acoustic Source Localization and Tracking Framework Using Particle Filtering and Information Theory", IEEE Transactions on Audio, Speech, and Language Processing - Special issue on processing reverberant speech: methodologies and applications archive, vol. 18, no. 7, Setembro 2010.
- [14] Yan, H., Fan, H.H., "A Wideband Kalman DOA Tracking Algorithm", em Sensor Array and Multichannel Processing. Fourth IEEE Workshop, pp. 45-49, Julho 2006.
- [15] Song, X., Wang, J., Han, Y., Meng, Y., "Robust Adaptive Beamforming Using a Bayesian Approach", em TENCON 2006. IEEE Region 10 Conference, pp. 1-4, Novembro 2006.
- [16] Vincent E., Deville Y.: "Audio Applications", capítulo 19, p.788, Comon P. Jutten, C. (Eds.): *Handbook of Blind Source Separation*. Academic Press, 2010. .