

# Métodos PHAT, FLOS-PHAT e ML-Energia para Localização de Fontes de Sinais Acústicos com Características Impulsivas em Ambiente Não-Reverberante

E. Dranka Junior e R. Coelho

**Resumo**— Este trabalho avalia o desempenho de três métodos para a localização de fontes de sinais acústicos com características impulsivas. Diversas simulações foram realizadas considerando quatro sinais acústicos impulsivos reais e valores de razão sinal ruído (SNR - *Signal to Noise Ratio*) entre 0 dB e 25 dB. Além disto, os métodos de localização foram investigados para casos em que o ruído aditivo é impulsivo e representado por distribuições alfa-estáveis. Os resultados mostram que a impulsividade atua diretamente no desempenho da localização.

**Palavras-Chave**— Localização de Fonte, PHAT, FLOS-PHAT, Impulsividade.

**Abstract**— This paper evaluates the performance of three methods for acoustic signals sources localization with impulsive characteristics. Several simulations were performed considering four impulsive acoustic signals and values of signal to noise ratio (SNR) between 0 dB and 25 dB. In addition, the methods have been investigated for the location where the impulsive noise is additive and is represented by alpha-stable distributions. The results show that impulsiveness acts directly on the final performance of the location.

**Keywords**— Source Localization, PHAT, FLOS-PHAT, Impulsiveness.

## I. INTRODUÇÃO

A localização de fontes acústicas é de grande interesse em diversas áreas de pesquisa. Dentre algumas aplicações pode-se citar o uso em sonares, onde um conjunto (*array*) de hidrofones é utilizado para localizar objetos em ambiente subaquático [1]. Em teleconferências, pode ser empregada para localizar o sinal de voz e direcionar a câmera [2]. Em campo aberto, permite o rastreamento da posição de um veículo em movimento [3]. Além destas, aplica-se também às áreas sísmicas, geofísicas, e para a localização de alvos. Os principais desafios da estimação da posição de uma fonte acústica se referem a presença de ruído aditivo, a ocorrência de reverberação em ambientes fechados, e a largura de banda do sinal da fonte. Sabe-se que as fontes acústicas e os ruídos reais podem ser impulsivos e não-Gaussianos [4], [5].

Os métodos para a estimação da localização de fonte acústica propostos na literatura são basicamente três: estimação do retardo temporal (TDE - *Time Delay Estimation*), determinação do ângulo de chegada (DOA - *Direction of*

*Arrival*) e medições da energia ou intensidade do sinal (Energia). O método DOA utiliza-se da informação da diferença de fase dos sinais recebidos em distintos sensores. Ele é utilizado quando os sinais são coerentes e de banda estreita [6]. A técnica TDE é aplicável para sinais de banda larga (por exemplo, o áudio) e consiste em determinar a diferença de tempo (atraso) que o sinal gerado pela fonte leva para atingir os sensores (microfones) que se encontram em posições diversas. Conhecendo-se a geometria espacial dos sensores, a velocidade do som e o atraso estimado, é possível determinar a posição geográfica da fonte acústica [7]. Um dos métodos mais utilizados de TDE é denominado de transformação de fase (PHAT - *Phase Transform*) que usa a informação da fase do sinal para melhorar a estimação do atraso. A terceira técnica baseia-se em medições de energia acústica e no conhecimento de que esta apresenta decaimento de sua intensidade de forma proporcional à distância entre o sensor e a fonte. Por meio de medições de energia, em posições diferentes, é possível estimar a localização da fonte do sinal acústico [8]. Um dos algoritmos baseados nesta técnica é denominado de ML-Energia, por estimar a posição por meio da função de máxima verossimilhança [6].

Este trabalho analisa a influência da característica da impulsividade das fontes acústicas no desempenho dos métodos de localização. As técnicas escolhidas são as baseadas no TDE e na intensidade acústica, pois estes são mais adequados para localização de fontes com banda larga. Também é avaliado o método baseado em TDE denominado de transformação de fase com estatísticas fracionárias de ordem inferior (FLOS-PHAT - *Fractionary Lower Order Statistics Phase Transform*), proposto pelos autores em [4] com finalidade de melhorar a estimação quando o ruído aditivo possui comportamento impulsivo. Diversas simulações foram realizadas empregando quatro fontes de sinais acústicos impulsivos reais (Metralhadora, Avião, Voz e *Ringtone*) e diferentes valores de SNR entre 0 dB e 25 dB. Além disso, foram considerados ruídos com características impulsivas distintas representadas por distribuições alfa-estáveis. A localização é em ambiente aberto, ou seja, a reverberação é desprezada.

O restante do trabalho está organizado em quatro seções. A seção II define a impulsividade e apresenta os métodos utilizados para sua estimação. A seção III aborda a fundamentação teórica com relação aos métodos de localização de fonte propostos na literatura. Os resultados são apresentados e debatidos

na seção IV.

## II. IMPULSIVIDADE

O grau de impulsividade ( $0 < \alpha \leq 2$ ) de uma determinada distribuição estável pode ser definido pelo decaimento de sua função cauda. A distribuição Gaussiana ( $\alpha = 2$ ) possui decaimento de cauda exponencial. Quanto menor o valor de  $\alpha$  mais impulsiva é a distribuição e mais se distancia de uma Gaussiana. Alguns estudos [5], [9] demonstram que os sinais acústicos reais são não-Gaussianos e com diferentes graus de impulsividade. O grau de impulsividade de um sinal ou ruído pode ser representado por uma distribuição alfa-estável ( $\alpha$ -estável) [9].

A propriedade da estabilidade diz que uma variável aleatória  $X$  é estável se, e somente se, para quaisquer variáveis aleatórias independentes  $X_1$  e  $X_2$  com mesma distribuição de  $X$  e para constantes arbitrárias  $a_1$  e  $a_2$  existem constantes  $a$  e  $b$  tais que  $a_1 X_1 + a_2 X_2 \stackrel{d}{=} aX + b$  onde a notação  $X \stackrel{d}{=} Y$ , significa que  $X$  e  $Y$  possuem semelhantes distribuições de probabilidade.

O Teorema do Limite Central Generalizado (TLCG) [10] afirma que a soma de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) tenderá para uma variável aleatória  $X$  com distribuição  $\alpha$ -estável. Este teorema permite mostrar que o TLC clássico é um caso especial do TLCG.

Não existe uma forma analítica fechada para representar as distribuições  $\alpha$ -estáveis, com exceção dos casos limites, tais como  $\alpha = 2$  em que a distribuição é Gaussiana ou  $\alpha = 1$ , quando tem-se uma distribuição de Cauchy. Deste modo, é comum utilizar-se da sua função característica  $\phi(t)$  dada por

$$\phi(t) = \exp(j\lambda t - \gamma|t|^\alpha [1 + j\beta \text{sign}(t)\omega(t, \alpha)]) \quad (1)$$

onde,

$$\omega(t, \alpha) = \begin{cases} \text{tg} \frac{\alpha\pi}{2}, & \alpha \neq 1 \\ \frac{2}{\pi} \log |t|, & \alpha = 1 \end{cases} \quad (2)$$

$$\text{sign}(t) = \begin{cases} 1, & t > 0 \\ 0, & t = 0 \\ -1, & t < 0 \end{cases} \quad (3)$$

$\alpha$  é o expoente da função característica ou grau de impulsividade.

$\lambda$  é o parâmetro de localização, equivalente a média da distribuição se  $1 < \alpha \leq 2$  ou à mediana se  $0 < \alpha \leq 1$ .

$\gamma$  é o parâmetro de dispersão ( $\gamma > 0$ ) e determina o espalhamento das amplitudes em torno de  $\lambda$ .

$\beta$  é o índice de simetria ( $-1 < \beta < 1$ ). A distribuição será simétrica em relação a  $\lambda$  se  $\beta = 0$ .

Uma característica importante das distribuições  $\alpha$ -estáveis é que os seus momentos de ordem  $p$  são finitos apenas quando  $p < \alpha$ . Define-se a classe das distribuições  $\alpha$ -estáveis simétricas (S $\alpha$ S), quando a mesma apresenta simetria ( $\beta = 0$ ) e momento de ordem 1 (média) finito ( $1 < \alpha \leq 2$ ).

### A. Estimadores do grau de impulsividade ( $\alpha$ )

Diversos algoritmos foram propostos para estimar os parâmetros de uma distribuição  $\alpha$ -estável, dentre eles o grau de

impulsividade. Fama introduziu um método baseado nos valores dos *quantis* de uma certa distribuição [11]. Este estimador foi aperfeiçoado por McCulloch [12] e permite a estimação de uma maior faixa de valores de  $\alpha$  (entre 0,6 e 2). O método baseado nos momentos fracionários foi proposto por Nikias [9] e modificado por Georgiou [4], mas apesar de ser rápido, é apropriado apenas para a classe de distribuições S $\alpha$ S ( $\beta = 0$ ). Por sua maior faixa de estimação e boa precisão obtida nos estudos de [13], o estimador escolhido neste trabalho foi o de McCulloch. Sejam  $L$  amostras independentes de um processo aleatório ordenadas em ordem crescente,  $x_1, x_2, \dots, x_L$  e,  $\hat{x}_p$ , a estimação da amostra correspondente ao valor do *quantil*  $p$ . Define-se

$$\begin{aligned} \hat{v}_\alpha &= \frac{\hat{x}_{0.95} - \hat{x}_{0.05}}{\hat{x}_{0.75} - \hat{x}_{0.25}} \\ \hat{v}_\beta &= \frac{\hat{x}_{0.95} + \hat{x}_{0.05} - 2\hat{x}_{0.5}}{\hat{x}_{0.95} - \hat{x}_{0.05}} \end{aligned} \quad (4)$$

Os valores de  $\hat{v}_\alpha$  e  $\hat{v}_\beta$  dependem apenas dos parâmetros  $\alpha$  e  $\beta$  da distribuição. Com ajuda de tabelas previamente determinadas, é possível, por interpolação, se estimar os parâmetros  $\alpha$  e  $\beta$ .

## III. MÉTODOS DE LOCALIZAÇÃO

Esta seção descreve os métodos de localização de fonte acústica baseados no TDE (PHAT e o FLOS-PHAT) e nas medições da intensidade acústica (ML-Energia) avaliados neste artigo.

### A. PHAT

O princípio do PHAT [14] baseia-se na correlação entre os sinais de uma fonte acústica recebidos por dois microfones que se encontram em posições distintas. Considere dois sinais acústicos  $r_1(t)$  e  $r_2(t)$ ,

$$\begin{aligned} r_1(t) &= x(t) + w_1(t) \\ r_2(t) &= x(t - \tau) + w_2(t) \end{aligned} \quad (5)$$

onde  $x(t)$  é o sinal acústico gerado pela fonte,  $w_1(t)$  e  $w_2(t)$  são ruídos assumidos com média zero e descorrelatados entre si e com o sinal  $x(t)$ . O atraso temporal a ser determinado é dado por  $\tau$ . Passando os sinais para o domínio da frequência, tem-se,

$$\begin{aligned} R_1(k) &= X(k) + W_1(k) \\ R_2(k) &= X(k)e^{-j\omega_k\tau} + W_2(k). \end{aligned} \quad (6)$$

A função de correlação cruzada é determinada no domínio da frequência,

$$\begin{aligned} C_{R_1 R_2}(k) &= \mathbb{E}[R_1(k)R_2(k)^*] \\ C_{R_1 R_2}(k) &= \mathbb{E}[|X(k)|^2 e^{j\omega_k\tau}] + \mathbb{E}[W_1(k)W_2^*(k)] + \\ &\quad \mathbb{E}[X(k)W_2^*(k)] + \mathbb{E}[X^*(k)W_1(k)e^{j\omega_k\tau}], \end{aligned} \quad (7)$$

onde o símbolo  $*$  representa o complexo conjugado. Como assumiu-se que os ruídos são descorrelatados entre si e o sinal de interesse, as três últimas parcelas de (7) são iguais a zero.

De acordo com o método PHAT, a função de correlação cruzada é suavizada normalizando-se todos os seus valores.

Aplicando-se a transformada inversa de Fourier nesta função suavizada obtêm-se um pico agudo (idealmente um impulso) no valor de tempo correspondente ao atraso temporal entre os sinais.

$$C_{R_1 R_2}^{wv}(k) = \frac{C_{R_1 R_2}(k)}{|C_{R_1 R_2}(k)|} = e^{j\omega_k \tau}. \quad (8)$$

### B. FLOS-PHAT

O método FLOS-PHAT é uma generalização do método PHAT com o objetivo de melhorar o desempenho da estimação quando o ruído aditivo possui comportamento impulsivo [4]. Para isso, emprega a função de correlação de ordem fracionária, definida por

$$A_{R_1 R_2}(k) = \mathbb{E}[R_1^*(k)^{\langle p \rangle} R_2(k)^{\langle q \rangle}] \quad (9)$$

onde  $R_1(k)^{\langle p \rangle} = |R_1(k)|^{p-1} R_1(k)^*$  e  $p$  e  $q$  são valores reais ( $p + q < \alpha$ ), sendo  $\alpha$  o índice de impulsividade do sinal. Da mesma forma que no método PHAT, a função de correlação é suavizada, resultando em

$$A_{R_1 R_2}^w(k) = \frac{A_{R_1 R_2}(k)}{|A_{R_1 R_2}(k)|} = e^{j\omega_k \tau}. \quad (10)$$

A transformada de Fourier inversa apresenta o ponto de máximo no valor de tempo correspondente ao atraso temporal estimado. Convém notar que, para os valores de  $p = q = 1$ , o método se reduz ao PHAT.

### C. ML-Energia

Sejam  $J$  sensores posicionados em um campo de medição, no interior do qual  $K$  fontes acústicas omnidirecionais emitem sinais sonoros. A energia dos sinais recebida em cada um dos  $J$  sensores é dada por

$$y_i(t) = g_i \cdot \sum_{k=1}^K \frac{S(t - t_{ki})}{|\rho_k(t - t_{ki}) - \mathbf{r}_i|^a} + \epsilon_i(t) \quad (11)$$

onde,  $S_k(t)$  é a energia ou intensidade do sinal acústico gerado pela fonte  $k$ ,  $g_i$  é o ganho do  $i$ -ésimo sensor e  $t_{ki}$  é o tempo que o sinal da fonte  $k$  leva para chegar da fonte ao sensor  $i$ .  $\rho_k(t)$  é o vetor de coordenadas cartesianas da  $k$ -ésima fonte acústica e  $\mathbf{r}_i(t)$  é o vetor que indica a posição espacial de cada um dos  $J$  sensores. De acordo com [8], a constante de decaimento  $a$  é aproximadamente igual a 2 e  $\epsilon_i(t)$  representa o erro de modelagem dos parâmetros  $a$ ,  $g_i$  e  $\mathbf{r}_i$  além do ruído aditivo. Considerando o teorema do limite central,  $\epsilon_i(t)$  pode ser representado por uma distribuição Gaussiana, com média  $\mu_i = \nu_i^2$  e variância  $\sigma_i^2 = 2\nu_i^4/M$ , onde  $\nu_i^2$  é a variância do ruído aditivo Gaussiano Branco presente em cada sensor e  $M$  é a quantidade de amostras (quadro) do sinal utilizada para estimar a energia. Neste trabalho, adota-se que a fonte acústica é monitorada em um campo de tamanho 100m x 100m, o que faz com que o atraso máximo de propagação seja de 0,41 segundos (correspondente à maior distância possível, que seria a diagonal do quadrado). Para fins de simplificação, considera-se que neste intervalo de tempo a energia do sinal e a posição da fonte não variam de forma significativa, resultando em um modelo mais simples,

$$y_i(t) = g_i \cdot \sum_{k=1}^K \frac{S_k(t)}{d_{ik}^2(t)} + \epsilon_i(t) \quad (12)$$

onde  $d_{ik}(t) = \|\rho_k(t) - \mathbf{r}_i(t)\|$  é a distância Euclidiana entre o sensor  $i$  e a fonte  $k$ . Omitindo-se o índice do tempo  $t$  e definindo-se as seguintes matrizes

$$\begin{aligned} \mathbf{Z} &= \begin{bmatrix} \frac{y_1 - \mu_1}{\sigma_1} & \dots & \frac{y_J - \mu_J}{\sigma_J} \end{bmatrix}^T \\ \mathbf{G} &= \text{diag} \left[ \frac{g_1}{\sigma_1} \quad \frac{g_2}{\sigma_2} \quad \dots \quad \frac{g_J}{\sigma_J} \right] \\ \mathbf{D} &= \begin{bmatrix} \frac{1}{d_{11}^2} & \frac{1}{d_{12}^2} & \dots & \frac{1}{d_{1K}^2} \\ \frac{1}{d_{21}^2} & \frac{1}{d_{22}^2} & \dots & \frac{1}{d_{2K}^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{d_{J1}^2} & \frac{1}{d_{J2}^2} & \dots & \frac{1}{d_{JK}^2} \end{bmatrix} \\ \mathbf{S} &= \begin{bmatrix} S_1 & S_2 & \dots & S_K \end{bmatrix}^T \\ \mathbf{H} &= \mathbf{GD} \\ \epsilon &= \begin{bmatrix} \epsilon_1 & \epsilon_2 & \dots & \epsilon_J \end{bmatrix}^T \end{aligned} \quad (13)$$

Pode-se reescrever (12) como

$$\mathbf{Z} = \mathbf{GDS} + \epsilon \quad (14)$$

A função densidade de probabilidade conjunta de  $\mathbf{Z}$  pode ser expressa por

$$f(\mathbf{Z}|\theta) = (2\pi)^{-N/2} \cdot \exp\left\{-\frac{1}{2}(\mathbf{Z} - \mathbf{HS})^T(\mathbf{Z} - \mathbf{HS})\right\} \quad (15)$$

onde,

$$\theta = \begin{bmatrix} \rho_1^T & \rho_2^T & \dots & \rho_K^T & S_1 & S_2 & \dots & S_K \end{bmatrix}^T \quad (16)$$

é o vetor que contém a posição  $\rho_k$  de cada fonte e sua respectiva energia  $S_k$ .

A função de máxima verossimilhança a ser minimizada será dada por  $L(\theta) = \|\mathbf{Z} - \mathbf{GDS}\|^2$ . Por abordar o problema por este método, o mesmo passará a ser chamado de ML-Energia.

O algoritmo proposto por Sheng [6] para minimizar a função de máxima verossimilhança é o da busca de multiresolução (MR). Este método consiste em efetuar a varredura do campo dos sensores em uma resolução grosseira inicialmente e após encontrar o ponto de mínimo, efetuar buscas mais precisas em torno do ponto encontrado.

## IV. RESULTADOS EXPERIMENTAIS E DISCUSSÃO

Os experimentos realizados neste trabalho são apresentados para quatro sinais acústicos reais, sendo dois deles obtidos da base NOISEX-92 [15] (Avião e Metralhadora), um sinal de voz da base TIMIT [16] e um ruído (*Ringtone*) disponível em <http://www.findmidis.com>. Os sinais da NOISEX apresentam taxa de amostragem de 19,98kHz e os sinais de Voz e *Ringtone*, 16kHz.

### A. Estimação do Grau de Impulsividade

Para a estimação do grau de impulsividade dos sinais acústicos foi aplicado o estimador de McCulloch no sinal completo. Também foi feita a estimação por quadros (4096 amostras) e ao final, calculada a média da estimação obtida em cada um dos quadros. Os valores obtidos nos dois casos são apresentados na Tab. I. Como pode-se observar, o sinal Metralhadora apresenta elevado grau de impulsividade ( $\alpha = 0,57$ ). O sinal de Voz também possui impulsividade elevada

TABELA I  
VALORES DE IMPULSIVIDADE ESTIMADOS

Fonte	$\alpha$ sinal completo	$\alpha$ estimação por quadros
Avião	2,00	1,98
Ringtone	1,48	1,58
Voz	0,98	1,08
Metralhadora	0,57	1,42

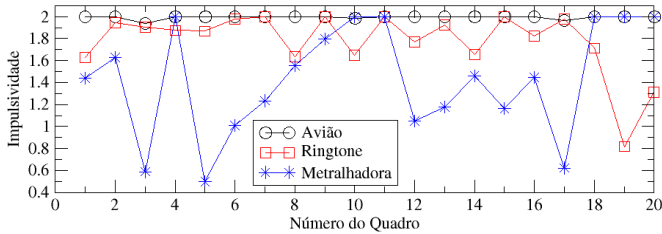


Fig. 1. Estimação do grau de impulsividade ( $\alpha$ ) para as fontes acústicas, feita em blocos de 4096 amostras.

( $\alpha = 0,98$ ). Para a fonte Avião obteve-se o valor de  $\alpha$  igual a 2, sendo, portanto, o que possui comportamento mais próximo da distribuição Gaussiana. A fonte denominada por *Ringtone* tem grau moderado de impulsividade.

Na estimação dos valores de  $\alpha$  por quadros, pode-se verificar que este varia em função do tempo. A Fig. 1 ilustra os valores obtidos em função do número do quadro. Nota-se a fonte Metralhadora apresenta uma maior variabilidade de seu grau de impulsividade que as outras fontes. A fonte Avião apresenta uma menor variabilidade.

### B. Resultados da localização usando PHAT e FLOS-PHAT

Neste artigo, utiliza-se o conceito de razão sinal ruído generalizado (GSNR - *Generalized Signal to Noise Ratio*), por ser mais apropriada para ruídos com distribuição  $\alpha$ -estável [17]. Nos experimentos realizados para avaliar o desempenho dos métodos PHAT e FLOS-PHAT foram realizadas estimativas de localização com as fontes acústicas. Duas situações foram analisadas. Para a primeira (situação 1), um retardo temporal correspondente ao valor de 100 amostras é aplicado juntamente com um ruído aditivo Gaussiano ( $\alpha = 2$ ), em diversos valores de SNR. Na segunda situação (situação 2), utilizou-se do mesmo valor de retardo porém é adicionado um ruído impulsivo gerado pelo método de Chambers-Mallows-Stuck [18] com parâmetro  $\alpha = 1,4$ . Os algoritmos de localização foram aplicados para tamanhos de blocos ou quadros de 4096, 2048 e 1024 amostras. Um limiar de tolerância de 0 (zero) intervalo de tempo foi adotado para considerar a estimação como um acerto. Para o método FLOS-PHAT foi definido  $p = q = 0,2$ , sendo tais valores considerados ótimos para este método [4].

A Fig. 2 apresenta o desempenho dos métodos de localização das fontes acústicas em função da impulsividade do ruído aditivo para as diversos valores de GSNR. Veja que a fonte mais impulsiva (Metralhadora, com  $\alpha = 0,57$ ) gerou mais erros de estimação para ambos os métodos. Nota-se que a localização pelo método FLOS-PHAT apresenta menos erros em relação ao PHAT quando existe a presença de ruído

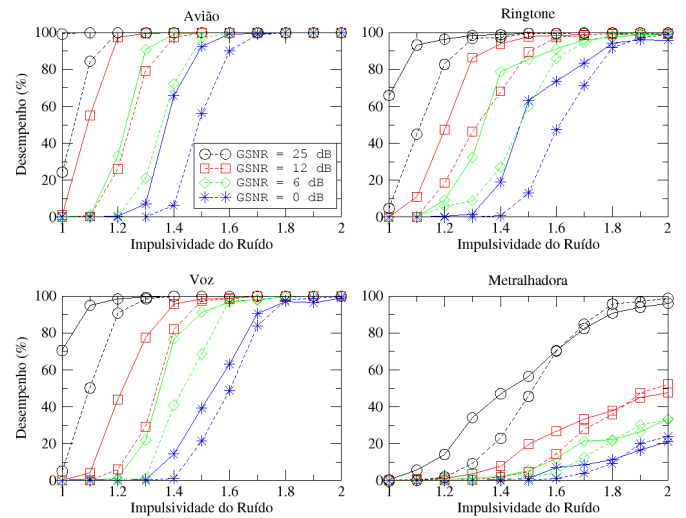


Fig. 2. Resultados do desempenho dos algoritmos PHAT e FLOS-PHAT para as diversas fontes. As linhas contínuas apresentam o desempenho médio obtido pelo método PHAT e as linhas tracejadas pelo método FLOS-PHAT.

TABELA II  
MÉDIA DE ACERTOS PARA DIVERSOS SNR COM RUÍDO  $\alpha = 2,0$

Fonte	M = 4096		M = 2048		M = 1024	
	PHAT	FLOS	PHAT	FLOS	PHAT	FLOS
Metralhadora	61,86	56,25	59,03	55,95	52,13	49,56
Voz	100,00	100,00	99,98	99,98	99,88	99,71
Avião	100,00	99,99	100,00	99,99	100,00	99,99
Ringtone	99,54	98,38	99,68	97,45	99,63	98,49

impulsivo. A média de acertos de cada fonte para as duas situações é mostrado nas Tab. II e Tab. III, respectivamente.

Com relação a situação 1, pode-se verificar que a fonte Avião (a menos impulsiva) apresenta melhor resultado dentre as demais para os métodos PHAT e FLOS-PHAT, com média de acertos de 100%. A fonte Metralhadora (a mais impulsiva) teve o pior resultado, com média de erro de mais de 50% utilizando o método FLOS-PHAT e  $M = 1024$ . Observa-se que o tamanho do quadro contribui para melhoria da média de acertos, sendo que quanto mais amostras empregadas no algoritmo melhor é localização.

Convém ressaltar que comparando-se o desempenho entre o PHAT e o FLOS-PHAT, o primeiro é sempre superior para a situação 1, tendo em vista que o ruído adicionado é o Gaussiano ( $\alpha = 2$ ), e o método PHAT é ótimo para este caso.

Com relação a situação 2, pode-se perceber que a fonte Metralhadora mantém uma média de erros maior dentre as fontes, seguida pelo sinal de Voz. Verifica-se também um menor erro de localização para o método FLOS-PHAT se comparado com o PHAT. Para o caso da fonte Avião, com  $M = 1024$ , o método FLOS-PHAT obteve média de acerto de aproximadamente 73% enquanto que o PHAT apresentou média de 49%.

### C. Resultados da localização pelo ML-Energia

Para avaliar o desempenho do método ML-Energia para fontes acústicas impulsivas, é simulado um campo de localização de tamanho de 100m x 100m com 4 sensores posicionados aleatoriamente. Um ruído branco é adicionado nos valores de

TABELA III

MÉDIA DE ACERTOS PARA DIVERSOS SNR COM RUÍDO  $\alpha = 1, 4$ 

Fonte	M = 4096		M = 2048		M = 1024	
	PHAT	FLOS	PHAT	FLOS	PHAT	FLOS
Metralhadora	16,07	21,79	8,47	16,52	6,28	14,42
Voz	59,09	63,80	56,64	67,55	56,03	71,75
Avião	80,41	93,04	74,33	92,09	69,07	91,23
Ringtone	43,44	56,04	51,32	64,26	48,39	72,65

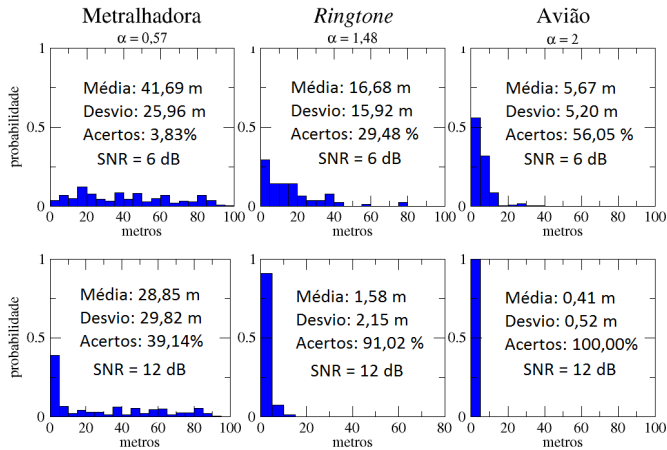


Fig. 3. Estimação da localização pelo ML-Energia. O erro absoluto é mostrado para faixas de 5 em 5 metros para cada quadro do sinal (normalizado). A primeira linha de gráficos mostra os resultados para SNR de 6dB, a segunda 12 dB

SNR de 6 dB e 12 dB. O tamanho do bloco escolhido foi o de 4096 amostras para facilitar a comparação com os métodos TDE. O método MR é empregado com dois níveis de busca: 5m x 5m e 1m x 1m. Após o término das estimações, o valor do erro absoluto, definido como a distância entre o ponto estimado e a posição real, é determinado, bem como a sua média e desvio padrão. Além disto, se para um erro absoluto inferior a 5 metros, considera-se uma estimativa correta, determina-se o desempenho médio pela razão entre o número de acertos e o total de quadros do sinal. Os histogramas referentes aos valores de erro absoluto são apresentados na Fig. 3, juntamente com os valores de média, desvio padrão e desempenho médio (acertos).

Pode-se perceber que a fonte Metralhadora obteve o pior resultado, com um desvio padrão elevado. O aumento do valor de SNR possibilitou uma sensível melhora com relação à média do erro e a taxa de acertos. A fonte menos impulsiva (Avião) obteve o melhor resultado para ambos os casos. O sinal Ringtone apresentou resultados intermediários em relação às duas outras fontes, refletindo o fato de ter sua impulsividade também de valor intermediário.

De uma análise comparativa entre os métodos ML-Energia e TDE, para o valor de SNR 12 dB e  $\alpha = 2$ , nota-se que a fonte Metralhadora obteve média de acertos de 52% para o método PHAT e apenas 39% com o ML-Energia. Já para a fonte Avião, o desempenho de ambos os métodos foi praticamente o mesmo. Para o Ringtone o PHAT obteve desempenho próximo a 100% enquanto o ML-Energia atingiu 91%. Observa-se que o método ML-Energia apresenta menor desempenho, demonstrando sensibilidade, quando as fontes são impulsivas.

## V. CONCLUSÃO

Este trabalho apresentou uma análise dos principais métodos de localização de fontes de sinais acústicos com diferentes graus de impulsividade. O estudo mostrou que a impulsividade é um fator determinante no desempenho dos métodos de localização.

A fonte acústica mais impulsiva (Metralhadora) obteve uma maior média de erro dentre todas as outras. A fonte de maior grau de impulsividade (Avião) foi a que resultou em localizações mais precisas por todos os métodos. Quando o ruído aditivo também exibe comportamento impulsivo, o método FLOS-PHAT resulta em melhores estimativas que o PHAT.

Quando a fonte é muito impulsiva, os métodos PHAT e FLOS-PHAT realizaram uma localização mais precisa que o ML-Energia. O desempenho foi semelhante quando a fonte apresenta maior grau de impulsividade.

O conhecimento da impulsividade a priori, pode ser útil para busca de melhorias nos métodos de localização de fontes de sinais acústicos.

## REFERÊNCIAS

- [1] A. Tolstoy, "Mached field processing for underwater acoustics," *World Scientific Publications*, 1993.
- [2] H. Wang and P. Chu, "Voice source localization for automatic camera pointing system in videoconferencing," in *Proc. ICASSP*, vol. 1, pp. 187–190, 1997.
- [3] C. W. Reed, R. E. Hudson, and K. Yao, "Direct joint source localization and propagation speed estimation," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process.*, Phoenix, Ariz, USA, May 1999.
- [4] P. G. Georgiou, P. Tsakalides, and C. Kyriakakis, "Alpha-stable modeling of noise and robust time-delay estimation in the presence of impulsive noise," in *IEEE Trans. on multimedia*, vol. 1, pp. 291–301, Sep 1999.
- [5] R. Webster, "Ambient noise statistics," in *IEEE Trans. on Signal Process.*, vol. 41, pp. 2249–2253, Jun 1993.
- [6] X. Sheng and Y. H. Hu, "Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks," in *IEEE Trans. on Signal Process.*, vol. 53, pp. 44–53, Jan 2005.
- [7] G. C. Carter, "Coherence and time delay estimation," *IEEE Press*, 1993.
- [8] D. Li and Y. H. Hu, "Energy based collaborative source localization using acoustic micro-sensor array," vol. 4, pp. 321–337, 2003.
- [9] C. L. Nikias and M. Shao, "Signal processing with fractional lower order moments: stable process and their applications," in *Proc. of IEEE*, vol. 81, pp. 986–1010, July 1993.
- [10] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. II. New York: Wiley, 1971.
- [11] E. F. Fama and R. Roll, "Some properties of symmetric stable distributions," *J. Amer. Statist. Assoc.*, vol. 63, pp. 817–836, 1968.
- [12] J. H. McCulloch, "Simple consistent estimators of stable distribution parameters," *Commun. Statist. Simul.*, vol. 15, pp. 1109–1136, 1986.
- [13] S. Bates and S. McLaughlin, "The estimation of stable distribution parameters," in *Proc. of IEEE SP Workshop on Higher-Order Statistics*, pp. 390–394, 1997.
- [14] C. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 24, no. 4, pp. 320–327, 1976.
- [15] A. Varga and H. Steeneken, "Assessment for automatic speech recognition ii: Noisex-92: a database and an experiment to study the effect of additive noise on speech recognition systems," *Speech Commun.*, vol. 12, no. 3, pp. 247–251, 1993.
- [16] J. S. Garofolo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, D. S. Pallett, and N. L. Dahlgren, "DARPA TIMIT acoustic phonetic continuous speech corpus CDROM," 1993.
- [17] C. L. Nikias and M. Shao, "Signal processing with alpha-stable distribution and applications," 1995.
- [18] B. W. S. J. M. Chambers, C. L. Mallows, "A method for simulating stable random variables," *Journal of the American Statistical Association*, vol. 71, no. 354, pp. 340–344, 1976.