

ESTIMATIVA DA DIREÇÃO DE CHEGADA DE MÚLTIPLOS SONS EM AMBIENTES INTERNOS

ROGÉRIO RÉUS SANTIN*, JUGURTA MONTALVÃO*

* *Universidade Federal de Sergipe (UFS)*
São Cristóvão, Sergipe, Brasil

Emails: rrsantin@yahoo.com.br, jmontalvao@ufs.br

Abstract— Due to the worldwide people ageing phenomenon, the health care at home (*home care*), a more humanized and cost saving method, for elderly people and long term care, is nowadays more and more usual. The problem is how to assure a prompt help when it is necessary. A surveillance system can be the solution to improve caregivers awareness and patients safeness. This paper investigates the possibility of selection and determination of the direction of arrival of indoor sounds, corresponding to monitoring signals from a patient fall or collapse. Simulating a microphone array, several speech sounds suffered phase displacement and distortion, trying to emulate a typical effect of indoor environments, the reverberation, and white noise was then added in a controlled amount too. After that, they were processed by *Direction-Of-Arrival* (DOA) estimating algorithms, focusing in the case of dual sound sources, one of them been considered the signal of interest, and the other an interference. Performances were measured by taking the home care perspective into account.

Keywords— Direction of Arrival, Generalized Cross Correlation, Phase Transform, DUET.

Resumo— Devido ao fenômeno global do envelhecimento populacional, o tratamento domiciliar (*home care*) está se tornando cada vez mais usual, especialmente para os casos de tratamentos longos e de idosos, por tratar-se de um método mais humano e econômico. O problema decorrente desta prática é como assegurar uma rápida intervenção caso necessário. Um sistema de tele vigilância médica pode ser a solução para manter os cuidadores cientes e aumentar a segurança dos pacientes. Neste trabalho é investigada a possibilidade de se selecionar e determinar a direção de chegada de sons em ambientes internos, tentando diagnosticar uma situação de emergência, tal como uma queda do paciente. Através da simulação de um arranjo de microfones, diversos sons de fala sofreram deslocamentos de fase e distorções, objetivando emular um efeito típico de ambientes fechados, que é a reverberação, e tiveram níveis controlados de ruído branco acrescentados a eles também. Em seguida, eles foram processados por algoritmos de estimação de ângulo de chegada (*Direction-of-Arrival* – DOA), focando os trabalhos no caso de duas fontes sonoras, sendo uma delas considerada o sinal de interesse e a outra uma interferência. No final do trabalho, através da perspectiva da tele vigilância médica, avaliou-se o desempenho dos diversos algoritmos.

Palavras-chave— Direção de chegada, Correlação Cruzada Generalizada, Transformação de Fase, DUET.

1 Introdução

O rápido envelhecimento populacional que ocorre há décadas nos países desenvolvidos tem, progressivamente, se tornado um fenômeno global neste século, afetando também os países em desenvolvimento. Os idosos hoje constituem 19% da população nos países desenvolvidos, e serão 33% (37% na Europa) em 2050 (Doukas and Maglogianis, 2011). No Brasil, a população de idosos representa um contingente de quase 15 milhões de pessoas com mais de 59 anos de idade (8,6% da população) e, segundo o Instituto Brasileiro de Geografia e Estatística – IBGE, nos próximos 20 anos, a população idosa do país poderá ultrapassar os 30 milhões de pessoas e representar quase 13% da população. Baseado nisto, o governo brasileiro criou um programa chamado *Melhor em Casa*, o qual prevê a assistência médica domiciliar para pacientes do Sistema Único de Saúde (SUS). A iniciativa, que já é encontrada em alguns pontos do País, permite que pessoas com doenças crônicas, em recuperação de cirurgias e idosas sejam atendidas em casa, reduzindo assim a necessidade e o tempo de internação. A meta é implantar 1000 equipes até 2014, além de 400 equipes multidisci-

plinares de apoio (ASCOM/MS, 2011).

Neste contexto, o grande desafio para as políticas de saúde é prover cada vez mais serviços de qualidade para todos, com racionalidade econômica, antevendo o grande crescimento do número de idosos. Os progressos em tecnologias de vigilância e assistência médica remota prometem ser uma boa aposta para ajudar a atingir estas metas. O cuidado domiciliar (*home care*) com monitoramento remoto é uma solução muito mais acolhedora e econômica que a internação hospitalar (Vacher et al., 2006).

Visando dar uma maior segurança para as pessoas envolvidas neste processo, o monitoramento diuturno do ambiente em que vive o paciente é vital para a detecção de emergências, especialmente no caso de pacientes sofrendo de doenças crônicas ou pessoas idosas vivendo sozinhas. O principal objetivo deve ser a detecção de situações de emergência tais como quedas, as quais podem provocar lesões severas e a necessidade de uma rápida intervenção.

Os requisitos mandatórios para um sistema de segurança pessoal consistem na utilização de tecnologias de monitoramento pouco invasivas, que não exponham seus usuários e que permi-

tam a mobilidade. Sistemas que tenham estas características tem uma maior aceitação por parte das pessoas a serem monitoradas (Stout, 2010). Ao mesmo tempo, estes sistemas devem possuir uma taxa elevada de acertos na interpretação dos eventos. Neste estudo, como parte de um conjunto de esforços paralelos no âmbito do projeto AtenSom (MCT/CNPq \ N° 14/2010), é avaliada uma solução para os problemas de invasibilidade, mobilidade, privacidade e assertividade na detecção de eventos anômalos, através da estimativa da direção de chegada de sons em ambientes internos com a presença de fontes interferentes. Esta etapa poderá servir de base para um sistema de vigilância para detecção de emergências, tais como quedas de idosos ou desmaios de pacientes.

Desde (Knapp and Carter, 1976), a Correlação Cruzada Generalizada (*Generalized Cross Correlation* - GCC) e sua evolução, a Correlação Cruzada Generalizada com Transformação de Fase (*Generalized Cross Correlation - PHase Transform* - GCC-PHAT), têm sido os algoritmos preferidos para a estimativa de direção de chegada de sons. Em (Azaria and Hertz, 1984) tem-se uma comparação entre a Máxima Verossimilhança (*Maximum Likelihood* - ML) e o método de erro quadrático mínimo (*Least Squares Error* - LSE) para estes algoritmos. Em (Varma, 2002) tem-se uma comparação entre as técnicas de Apontamento Direcionado (*Steered Beamformer* - SB), Decomposição em Autovalores (*Eigen Value Decomposition* - EVD) e GCC-PHAT. Já em (Talantzis et al., 2005) tem-se a comparação entre GCC-PHAT e um método de Estimativa de Atraso por Teoria da Informação (*Information Theory Delay Estimation*), levando-se em conta a presença de reverberações no sinal recebido.

Este artigo está organizado como segue: na seção 2, é apresentado o conjunto de sons reais digitalizados utilizados nos experimentos, enquanto que na seção 3 são demonstradas as equações e explicado o princípio de funcionamento de cada um dos três algoritmos utilizados. Já na seção 4 são apresentados os resultados obtidos nas diversas situações experimentais investigadas neste trabalho.

2 Sons Utilizados nos Testes e Avaliações

Para poder avaliar o desempenho dos algoritmos, foram utilizados cinco sons monofônicos de fala da língua inglesa, obtidos em um repositório público (Open_Speech_Repository, 2012). Eles foram digitalizados com taxa de amostragem de 8000 amostras por segundo, e codificados em 16 bits por amostra. Foram utilizados trechos de fala mais intensa, com duração de pouco menos de três segundos para cada sinal. Os sinais foram duplicados para formarem sinais estereofônicos, sendo que as cópias dos sinais sofreram defasagens variá-

veis, de modo que os atrasos relativos entre os dois canais dos sinais estereofônicos fossem diferentes para cada sinal. Os diversos trechos de sinais estereofônicos foram adicionados aos pares, para posterior detecção dos dois atrasos relativos a cada mistura. Foram obtidas quatro misturas, sendo M1 a mistura de duas vozes femininas (timbres semelhantes), e M2 à M4 misturas de vozes feminina e masculina, apresentando timbres distintos. Na etapa seguinte foram acrescentados ruídos não correlacionados em quantidades controladas (-20 e -40 dB) aos sinais estereofônicos. Além disto, um dos sinais de voz feminina foi processado por um software utilizado em estúdios de gravações sonoras para geração de reverberações. O resultado foi a geração de distorções de fase mínima diferenciadas por canal, emulando reverberações, que são um dos grandes problemas de detecção e análise de sons em ambientes internos. Com estes procedimentos obteve-se a geração de misturas diversas, na tentativa de abranger uma ampla gama de situações.

3 Algoritmos Utilizados

A simplicidade do algoritmo chamado Correlação Cruzada (*Cross Correlation* - CC) e o fato de que se pode obter uma solução de forma fechada, ao invés de procurar em todas as direções, tem-no feito uma escolha muito popular para a estimativa da DOA utilizando arranjos de microfones (Varma, 2002). Em virtude destas características, o mesmo serve de base para os três algoritmos utilizados neste trabalho, conforme demonstrado adiante.

A estimativa da diferença de tempo (*Time Delay Estimation* - TDE) entre os sinais que chegam a um par de microfones pode ser feita através da função de correlação cruzada entre os dois sinais. A diferença de tempo que maximiza a função de correlação cruzada é o atraso (*time delay*) entre os dois sinais captados e, através deste atraso, pode-se deduzir o ângulo de chegada do sinal, de acordo com a fórmula:

$$\theta = \arcsin \frac{\tau \cdot d}{c} \quad (1)$$

onde θ é o ângulo de incidência, τ o atraso relativo, c a velocidade do som no ar e d a distância entre os microfones. Assim, conhecendo-se a geometria do sistema, podemos estimar o ângulo de chegada do som através do atraso relativo entre os microfones.

A seguir, considerando dois microfones, representados por i e j , e sendo $x_i(n)$ e $x_j(n)$ os sinais captados por eles e digitalizados, respectivamente, onde n é o índice da amostra no tempo, aplicamos janelamento (para evitar que ocorram descontinuidades ao início e final de cada quadro) e Transformada Discreta de Fourier (*Discrete Fourier Transform* - DFT) a ambos. Esse

processamento resulta em $X_i(k)$ e $X_j(k)$, onde k é o contador de frequência discreto, conforme a relação:

$$\omega_k = \frac{2 \cdot \pi \cdot k}{N} \quad (2)$$

A Densidade Espectral da Potência Cruzada (*Cross Power Spectral Density* – XPSD) entre estes dois sinais é dada pela seguinte expressão:

$$\Phi_{X_i X_j} = X_i(k) \cdot X_j^*(k) \quad (3)$$

A correlação cruzada entre estes dois sinais é dada pela transformada inversa de Fourier da XPSD, conforme a expressão:

$$R_{X_i X_j}(\tau) = \int_{-\infty}^{\infty} \Phi_{X_i X_j}(\omega) \cdot e^{\frac{j \cdot \omega \cdot \tau}{M}} d\omega \quad (4)$$

Porém, utiliza-se uma aproximação para sinais amostrados e frequências discretas:

$$\hat{R}_{X_i X_j}(l) = \frac{1}{M} \sum_{k=0}^{M-1} \hat{\Phi}_{X_i X_j}(k) \cdot e^{\frac{j \cdot 2 \cdot \pi \cdot k \cdot l}{M}} \quad (5)$$

onde M é o número de termos da XPSD e l é o índice relativo ao atraso. Como a XPSD pode ser apenas estimada para observações finitas, pode-se apenas estimar o atraso com o valor desta correlação. A correlação pode ser calculada para atrasos positivos e negativos. O índice para o qual o valor da correlação é máximo é o número de amostras correspondente ao atraso entre os dois sinais.

A expressão que determina este atraso é:

$$\tau_{i,j} = \frac{1}{f_s} \cdot \operatorname{argmax}[\hat{R}_{X_i X_j}(l)] \quad (6)$$

3.1 Correlação Cruzada Generalizada

Conforme visto na fórmula 5, a correlação cruzada pode ser definida como:

$$\hat{R}_{X_i X_j}(l) = \frac{1}{M} \sum_{k=0}^{M-1} \hat{\Phi}_{X_i X_j}(k) \cdot e^{\frac{j \cdot 2 \cdot \pi \cdot k \cdot l}{M}}$$

onde Φ é a XPSD e é definida como (Fórmula 3):

$$\Phi_{X_i X_j} = X_i(k) \cdot X_j^*(k)$$

De forma mais geral, conforme (Knapp and Carter, 1976), tem-se a generalização da correlação cruzada (*Generalized Cross Correlation* – GCC) por primeiramente filtrar-se os sinais originais:

$$y_i = h_i(n) * x_i(n) \quad y_j = h_j(n-l) * x_j(n-l) \quad (7)$$

Como consequência, no domínio frequência tem-se a Densidade Espectral da Potência Cruzada Generalizada (*Generalized Cross Power Spectral Density* – GXPSD):

$$\Phi_{Y_i Y_j}^g(k) = [H_i(k) \cdot X_i(k)] \cdot [H_j(k) \cdot X_j(k)]^* \quad (8)$$

Pode-se também combinar os dois pré-filtros:

$$\Psi_{i,j}(k) = H_i(k) \cdot H_j^*(k) \quad (9)$$

Obtendo-se assim:

$$\Phi_{Y_i Y_j}^g(k) = \Psi_{i,j}(k) \cdot X_i(k) \cdot X_j^*(k) \quad (10)$$

De onde pode-se obter finalmente a GCC, calculando-se a inversa da DFT da GXPSD:

$$R_{Y_i Y_j}^g(l) = \frac{1}{M} \sum_{k=0}^{M-1} \Phi_{Y_i Y_j}^g(k) \cdot e^{\frac{j \cdot 2 \cdot \pi \cdot k \cdot l}{M}} \quad (11)$$

A partir deste ponto, faz-se necessário modificar o algoritmo original para que o mesmo possa encontrar diferentes valores de atraso para diferentes frequências em diferentes quadros, ao invés de simplesmente calcular o valor que melhor se adequa a todos os pontos. Resulta desta modificação a determinação de mais de um atraso para o conjunto de pontos, possibilitando assim a separação posterior dos sinais através do agrupamento de seus pontos em torno de valores médios de atraso que maximizam a Correlação Cruzada Generalizada para cada ponto de frequência em cada quadro.

3.2 Correlação Cruzada Generalizada com Transformação de Fase

A Transformação de Fase (*PHase Transform* – PHAT) é a função mais utilizada como pré-filtro para a GCC. Ela garante uma maior robustez do algoritmo, atuando como um nivelador de amplitudes, tentando obter a composição do ruído branco, o qual só se correlaciona consigo mesmo, obtendo para o atraso correspondente um alto pico (a função impulso) e para os demais atrasos algo próximo a zero, já que o ruído branco é não-correlacionado com os outros sinais. Esta é sua expressão, para dois sinais idênticos defasados no tempo:

$$\begin{aligned} \Psi^g(\omega) &= \frac{1}{|\hat{\Phi}_{X_i X_j}(\omega)|} = \\ &= \frac{1}{|\hat{\Phi}_{X_i X_i}(\omega) \cdot e^{-j \cdot \omega \cdot D}|} \end{aligned} \quad (12)$$

Pode-se então calcular a estimativa da GCC–PHAT através da seguinte expressão:

$$R_{X_i X_j}^{PHAT}(\tau) = \frac{1}{2 \cdot \pi} \int_{-\pi}^{\pi} \frac{1}{|\hat{\Phi}_{X_i X_j}(\omega)| \cdot \hat{\Phi}_{X_i X_j}(\omega) \cdot e^{j \cdot \omega \cdot \tau}} d\omega \quad (13)$$

Utilizando a expressão:

$$\hat{\Phi}_{X_i X_j}(\omega) = \hat{\Phi}_{X_i X_i}(\omega) \cdot e^{-j \cdot \omega \cdot D} \quad (14)$$

Obtém-se:

$$R_{X_i X_j}^{PHAT}(\tau) = \frac{1}{2 \cdot \pi} \int_{-\pi}^{\pi} \frac{1}{|\hat{\Phi}_{X_i X_i}(\omega) \cdot e^{-j \cdot \omega \cdot D}| \cdot \hat{\Phi}_{X_i X_i}(\omega) \cdot e^{-j \cdot \omega \cdot D} \cdot e^{j \cdot \omega \cdot \tau}} d\omega \quad (15)$$

E então:

$$R_{X_i X_j}^{PHAT}(\tau) = \frac{1}{2 \cdot \pi} \int_{-\pi}^{\pi} e^{-j \cdot \omega \cdot D} \cdot e^{j \cdot \omega \cdot \tau} d\omega \quad (16)$$

Vê-se assim que a estimativa se reduz a função impulso para um atraso igual a D , e zero para os demais valores:

$$R_{X_i X_j}^{PHAT}(\tau) = \delta(\tau - D) \quad (17)$$

Para o caso de sinais amostrados, utiliza-se como pré-filtro, ou função peso, a expressão:

$$\Psi^g(k) = \frac{1}{|\hat{\Phi}_{X_i X_j}(k)|} = \frac{1}{|\hat{\Phi}_{X_i X_i}(k) \cdot e^{-j \cdot 2 \cdot \pi \cdot k \cdot D}|} \quad (18)$$

3.3 Técnica de Separação e Estimação de Sinais Degenerados

A Técnica de Separação e Estimação de Sinais Degenerados (*Degenerate Unmixing and Estimation Technique* – DUE) (Jourjine et al., 2000) é utilizada quando o número de misturas disponíveis é menor que o número de fontes, já que, neste caso, o problema não pode ser resolvido através da estimativa da inversa da matriz de misturas, advindo daí a denominação de sinais degenerados. Utiliza como premissa a existência de transformadas separadas (*W-disjoint orthogonality*) para cada sinal.

Tal técnica baseia-se no algoritmo Correlação Cruzada Generalizada para obtenção da atenuação e defasagem (*delay*) relativos dos sinais correspondentes às misturas, transformados pela Transformada Discreta de Fourier, isto para cada frequência em cada quadro (*time-frequency point*). A seguir, ao invés de achar o atraso que melhor atende a todas as frequências, a técnica utiliza um algoritmo de agrupamento (*clustering*) similar ao algoritmo Esperança - Maximização (*Expectation-Maximization* – EM), utilizado para o cálculo de

Máxima Verossimilhança (*Maximum Likelihood* – ML). Este algoritmo agrupa os pontos tempo-frequência utilizando como peso no cálculo do centro do grupo (*cluster*) o valor calculado da Correlação Cruzada Generalizada para cada ponto contido no grupo. Como observado em (Rickard and Dietrich, 2000), as equações utilizadas iterativamente para realizar o agrupamento são as seguintes:

$$\hat{\delta}_i = \frac{\sum_{(\omega, \tau) \in \omega_i} \|X_1^W(\omega, \tau)\|^2 \gamma(\frac{X_2^W(\omega, \tau)}{X_1^W(\omega, \tau)})/\omega}{\sum_{(\omega, \tau) \in \omega_i} \|X_1^W(\omega, \tau)\|^2} \quad (19)$$

$$\hat{a}_i = \frac{\sum_{(\omega, \tau) \in \omega_i} \|X_1^W(\omega, \tau)\|^2 \|\frac{X_2^W(\omega, \tau)}{X_1^W(\omega, \tau)}\|}{\sum_{(\omega, \tau) \in \omega_i} \|X_1^W(\omega, \tau)\|^2} \quad (20)$$

onde γ é o operador correspondente à parte imaginária do logaritmo de base decimal, ω_i é o conjunto de pontos do agrupamento (*cluster*) i , X_i^W é o sinal da mistura i janelado e transformado pela DFT, $\hat{\delta}_i$ é o atraso estimado do agrupamento i , e $\hat{a}_i$ é a atenuação estimada do agrupamento i .

4 Análise do Desempenho dos Algoritmos

Foram realizadas diversas simulações com as misturas elaboradas: sinais com atrasos, sinais com reverberações e sinais com ruído branco adicionado. Os resultados das simulações indicam o erro de detecção de cada atraso em número de amostras, correspondendo a um erro na estimativa da direção de chegada para cada um dos sinais da mistura. Estes erros encontram-se abaixo tabelados por tipo de sinal e por algoritmo utilizado:

Mistura	M1		M2		M3		M4	
Atrasos	1°	2°	1°	2°	1°	2°	1°	2°
GCC	-7	0	0	0	0	-16	0	-2
PHAT	-1	0	0	0	0	0	0	0
DUE	-6	0	-4	0	0	1	0	-2

Table 1: Tabela de erros de detecção para as diversas misturas

Mistura	M1		M2		M3		M4	
Atrasos	1°	2°	1°	2°	1°	2°	1°	2°
GCC	-2	0	0	0	0	-18	0	0
PHAT	-1	0	0	0	0	0	0	0
DUE	-7	0	-4	0	0	1	0	-2

Table 2: Tabela de erros de detecção para misturas com reverberações

A análise das tabelas mostra um desempenho ruim do algoritmo GCC em todas as situações, devido principalmente a sua incapacidade de diferen-

Mistura	M1		M2		M3		M4	
Atrasos	1°	2°	1°	2°	1°	2°	1°	2°
GCC	-4	0	0	0	0	-17	0	-2
PHAT	-1	1	0	0	0	0	0	0
DUET	-6	0	-4	0	0	0	0	-2

Table 3: Tabela de erros de detecção para misturas com SNR=40dB

Mistura	M1		M2		M3		M4	
Atrasos	1°	2°	1°	2°	1°	2°	1°	2°
GCC	-1	0	0	0	0	-18	0	0
PHAT	-1	1	0	0	0	0	0	0
DUET	-7	0	-4	-1	0	1	0	-2

Table 4: Tabela de erros de detecção para misturas com SNR=40dB e reverberações

Mistura	M1		M2		M3		M4	
Atrasos	1°	2°	1°	2°	1°	2°	1°	2°
GCC	-2	0	-1	-1	0	-18	-1	0
PHAT	-2	2	0	0	0	-11	0	0
DUET	-6	0	-4	0	0	0	0	-2

Table 5: Tabela de erros de detecção para misturas com SNR=20dB

ciar ângulos próximos de chegada, já que o mesmo foi criado para a detecção de apenas um sinal.

O algoritmo GCC-PHAT, embora derivado do algoritmo GCC, consegue uma discriminação muito melhor de sinais com ângulos de chegada próximos, sendo que o erro que o mesmo apresenta para uma das misturas de sinais (composta por duas vozes femininas com atrasos diferindo em apenas uma amostra) é praticamente desprezível, conforme observado em todos os gráficos. Já para a situação em que a relação sinal/ruído cai muito (SNR = 20 dB), observamos que o comportamento do algoritmo deixou a desejar para a terceira mistura, embora tenha detectado razoavelmente todas as demais.

Já o terceiro algoritmo, o DUET, mostrou-se praticamente insensível a ruídos e reverberações, devido ao seu processamento mais intenso, que inclui a informação de amplitude no agrupamento dos indícios espectrais de cada fonte sonora. Entretanto, pode-se observar desde o início um erro que não afeta muito o seu trabalho de separação dos sons na maior parte dos casos, mas que mostra-se demasiadamente grande para a estimativa da direção de chegada dos sons, que é o principal objetivo do algoritmo que estamos buscando.

Para verificar o efeito da reverberação no funcionamento de cada algoritmo foram escolhidos os sinais componentes da mistura M2, os quais apresentam uma situação de continuidade sonora, sem intervalos aparentes entre as palavras. Observa-se na sequência gráficos contendo a distribuição dos valores de atraso detectados por cada algoritmo,

com os valores de variância junto à legenda. Os dois primeiros são para os sinais originais, e os dois últimos para os sinais com reverberações:

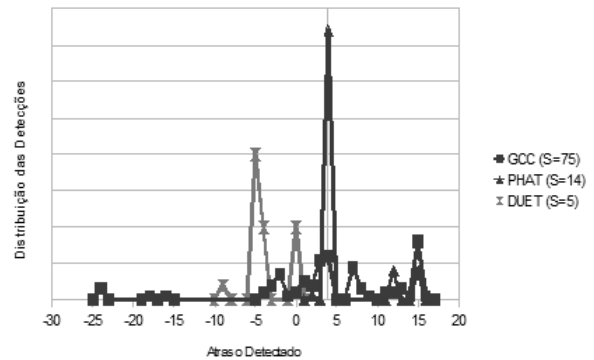


Figure 1: Histograma do atraso detectado para o 1º sinal da mistura

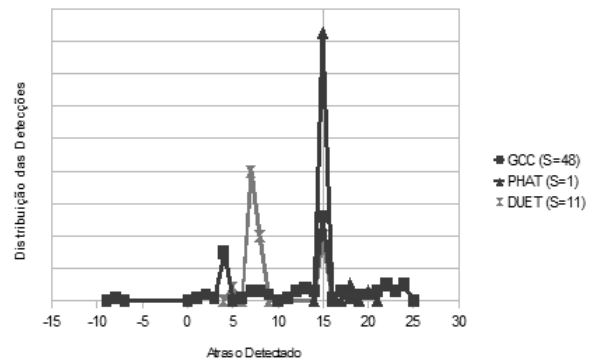


Figure 2: Histograma do atraso detectado para o 2º sinal da mistura

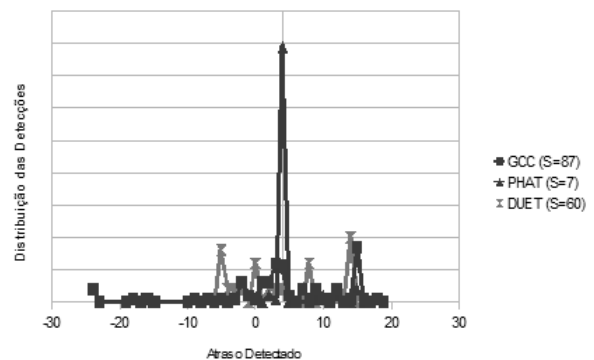


Figure 3: Histograma do atraso detectado para o 1º sinal com reverberação

Através dos gráficos percebe-se que os algoritmos GCC e, em menor escala, DUET, possuem um grande espalhamento dos atrasos detectados,

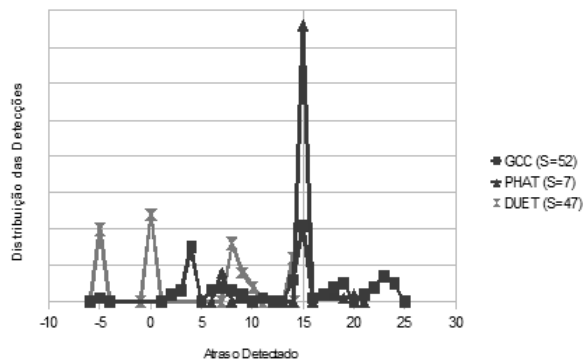


Figure 4: Histograma do atraso detectado para o 2º sinal com reverberação

sendo que as reverberações ampliam ainda mais esta tendência no caso do segundo. Já o algoritmo GCC-PHAT apresenta uma grande concentração em torno do valor de pico, com pequena variância. A mistura escolhida apresenta atrasos bem distintos (4 e 15 amostras), tendo sido estes atrasos representados nos gráficos pelo eixo vertical.

5 Conclusão

Embora as reverberações constituam o principal problema causado por ambientes fechados, com a metodologia utilizada neste trabalho não foram detectados grandes efeitos negativos, sendo que os três algoritmos apresentaram um desempenho pouco afetado pelas mesmas. Dentre os algoritmos analisados, o GCC-PHAT demonstrou possuir uma precisão muito maior que os demais, tendo, porém, o seu desempenho degradado em função do ruído do tipo aleatório (ruído branco) para uma das misturas. Devemos tratar, na prática, deste tipo de problema através da escolha de equipamentos, cabos e sensores adequados para a redução do ruído a um nível aceitável para o algoritmo. Este algoritmo possui, ainda, um tempo de processamento menor, em termos de carga computacional, que o DUET, sendo este fator essencial para aplicações que devem ser "real time", ou seja, que devem processar continuamente todos os sinais captados, sem perda de informação. Assim, recomenda-se para aplicações semelhantes ao propósito deste trabalho o emprego do GCC-PHAT pela sua simplicidade, economicidade e assertividade.

References

ASCOM/MS (2011). Melhor em casa leva atenção à casa do brasileiro, <http://portalsaude.saude.gov.br/portalsaude/noticia/3061/162/melhor-em-casa-leva-atencao-a-casa-do-brasileiro.html>.

Azaria, M. and Hertz, D. (1984). Time delay estimation by generalized cross correlation methods, *Acoustics, Speech and Signal Processing, IEEE Transactions on ASSP-32*(2): 280 – 285.

Doukas, C. N. and Maglogiannis, I. (2011). Emergency fall incidents detection in assisted living environments utilizing motion, sound, and visual perceptual components, *IEEE Transactions on Information Technology in Biomedicine* 15(2): 277–289.

Jourjine, A., Rickard, S. and Yilmaz, O. (2000). Blind separation of disjoint orthogonal signals: demixing n sources from 2 mixtures, *Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on*, Vol. 5, pp. 2985 –2988 vol.5.

Knapp, C. and Carter, G. (1976). The generalized correlation method for estimation of time delay, *Acoustics, Speech and Signal Processing, IEEE Transactions on* 24(4): 320 – 327.

Open_Speech_Repository (2012). The open speech repository, http://www.voiptroubleshooter.com/open_speech/index.html.

Rickard, S. and Dietrich, F. (2000). Doa estimation of many w-disjoint orthogonal sources from two mixtures using duet, *Statistical Signal and Array Processing, 2000. Proceedings of the Tenth IEEE Workshop on*, pp. 311 – 314.

Stout, H. (2010). Technologies help adult children monitor aging parents, <http://www.nytimes.com/2010/07/29/garden/29parents.html>.

Talantzis, F., Constantinides, A. and Polimenakos, L. (2005). The generalized correlation method for estimation of time delay, *Communications and Signal Processing, IEEE Collections on* 12(8): 561 – 564.

Vacher, M., Serignat, J.-F., Chaillol, S., Istrate, D. and Popescu, V. (2006). Speech and sound use in a remote monitoring system for health care, *TSD*, pp. 711–718.

Varma, K. (2002). Time-delay-estimate based direction-of-arrival estimation for speech in reverberant environments, <http://scholar.lib.vt.edu/theses/>.