

CENTRO UNIVERSITÁRIO DA FEI

ROBERTO CYRULNIK

**REDUÇÃO DE RUÍDOS LATERAIS
EXPLORANDO A DIRECIONABILIDADE DE
ARRANJO DE MICROFONES**

São Bernardo do Campo
2011

ROBERTO CYRULNIK

**REDUÇÃO DE RUÍDOS LATERAIS
EXPLORANDO A DIRECIONABILIDADE DE
ARRANJO DE MICROFONES**

Relatório de Qualificação de Mestrado apresentada ao Centro Universitário da FEI, como parte dos requisitos necessários para obtenção do título de Mestre em Engenharia Elétrica, orientada pelo Prof. Dr. Ivandro Sanches.

LISTA DE ILUSTRAÇÕES

1.1	Protetor auditivo com comunicador (PELTOR, 2011).	2
1.2	Protetores auditivos com comunicação (SENSEAR, 2011).	2
2.1	Exemplo diagrama polar de um beamforming formado por dois microfones.	5
2.2	Diagrama ilustrando um beamforming constituído por J microfones.	5
2.3	Representação do atraso τ_j no arranjo de microfones.	6
2.4	Diagrama da técnica delay and sum.	7

SUMÁRIO

1	Introdução	1
1.1	Objetivo	3
1.2	Organização do Trabalho	3
2	Revisão Bibliográfica	4
2.1	Introdução	4
2.2	Técnica Delay and Sum	6
2.3	Técnica MVDR (Minimum Variance Distortionless Response)	7
2.4	Conclusão	9
3	Trabalhos Subsequentes e Cronograma	10
3.1	Trabalhos Subsequentes	10
3.2	Cronograma	11
	REFERÊNCIAS	12

1 INTRODUÇÃO

Comunicação em ambientes ruidosos é um grande desafio para trabalhadores, militares e policiais pois a inteligibilidade fica comprometida. Os dispositivos de comunicação como rádios comunicadores ou telefones celulares utilizados nessas áreas captam a voz do locutor adicionada ao ruído, prejudicando a comunicação com o receptor.

São exemplos de ambientes extremamente ruidosos: pistas de aeroportos, estádios de futebol, shows, indústrias siderúrgicas e indústrias de mineração, ambientes onde o ruído pode alcançar níveis de pressão sonora muito elevados. Um policial durante sua guarda em um estádio de futebol recebe ordens enquanto a torcida gera ruídos extremos. Um operário em uma indústria siderúrgica deve comunicar-se com seu colega em um ambiente extremamente ruidoso. Um soldado deve transmitir informações sobre o inimigo dentro do campo de batalha imerso ao som de metralhadoras, explosões, gritos, etc. Em todos esses ambientes, o dispositivo de comunicação, que pode ser um rádio comunicador ou um aparelho celular, deve possuir um meio de suprimir o ruído ambiente para que o receptor consiga receber a informação com qualidade.

Segundo (MTE, 2011), um trabalhador que fica exposto a um nível de pressão sonora de no mínimo 85dB, contínuo ou intermitente, deve utilizar algum tipo de proteção auditiva. Porém, ao utilizar um protetor auditivo tipo plugue ou concha, a comunicação com outras pessoas é prejudicada já que a voz também é atenuada.

Existem portanto, duas questões a serem analisadas: proteger a saúde do trabalhador através de um protetor auditivo e melhorar a comunicação entre os trabalhadores.

Atualmente, existem protetores auditivos com rádio comunicação como o (PELTOR, 2011). Esse equipamento possui sistema de cancelamento de ruído, porém, a comunicação é realizada somente por rádio one way, isto é, deve-se pressionar um botão PTT (push to talk) para comunicar-se com outros trabalhadores.

Já o (SENSEAR, 2011) possibilita a comunicação face a face sem a necessidade de um rádio comunicador. O equipamento processa o sinal através de microfones externos, atenuando o ruído externo e amplificando o sinal de voz. Além de reduzir o ruído externo o dispositivo oferece proteção auditiva. Esses equipamentos podem ser visualizados na figura 1.2 e são comercializados em dois modelos: a) plugue e b) concha.

Como sou profissional da área de Segurança no Trabalho, o meu objetivo é dissertar sobre



Figura 1.1: Protetor auditivo com comunicador (PELTOR, 2011).



Figura 1.2: Protetores auditivos com comunicação (SENSEAR, 2011).

um algoritmo de alta eficiência para atenuar o ruído, garantindo que o sinal de voz recebido garanta a inteligibilidade da comunicação em ambientes ruidosos.

Atualmente existem técnicas avançadas de supressão de ruído como o CASA (Computational Auditory Scene Analysis) (WANG; BROWN, 2005), o BSS (Blind Source Separation) (WANG; BROWN, 2005) e o Beamforming. Segundo (WANG; BROWN, 2005), o CASA utiliza o conceito sensorial da audição humana para separar diferentes fontes de sinais. Esse conceito utiliza um processo de dois estágios: a segmentação e o agrupamento. Durante a segmentação, o sinal de entrada é decomposto em uma coleção de regiões tempo-frequência chamados segmentos. No segundo estágio, o agrupamento, os segmentos são agrupados de acordo com regras perceptuais, para separar as diferentes fontes que geraram o sinal de entrada. Essa é uma técnica muito recente e ainda necessita de mais estudos para que possa ser implementada em dispositivos móveis.

Alguns algoritmos de BSS, utilizam a técnica ICA (independent component analysis). Porém, segundo (WANG; BROWN, 2005) esses algoritmos tendem a perder performance drasticamente em ambientes que possuem mais fontes sonoras do que microfones.

Beamforming é um método de filtragem espacial que diferencia os sinais desejados do ruído

e interferência, através de um arranjo de microfones. Essa técnica já está sendo implementada em DSP's (Digital Signal Processing) (ELLEDDGE et al., 2000) e FPGA (VU et al., 2010).

Nas seções a seguir, algumas técnicas de Beamforming serão exploradas através de um resumo de suas características, suas vantagens e desvantagens.

1.1 Objetivo

O objetivo desse trabalho é estudar uma técnica de supressão de ruídos para ser utilizada em um dispositivo comercial na área de Segurança no Trabalho.

Como já descrito na seção anterior, o beamforming é um método de filtragem espacial que diferencia os sinais desejados do ruído e interferência, através de um arranjo de microfones. O beamforming é uma técnica atual e já foi implementada em DSP (ELLEDDGE et al., 2000) e FPGA (VU et al., 2010).

Existem alguns métodos para implementar o beamforming que serão estudados nas seções a seguir. No final do capítulo será realizada uma comparação entre os métodos e será adotado para a dissertação o que possuir maior compatibilidade com o objetivo desse trabalho.

1.2 Organização do Trabalho

Esse trabalho está organizado em 3 capítulos. O capítulo 2 apresenta uma revisão bibliográfica abordando o beamforming e duas técnicas conhecidas de beamforming que são o Delay and Sum e o MVDR (Minimum variance distortionless response). Ainda nesse capítulo, é apresentada uma conclusão da técnica de beamforming a ser utilizada na dissertação. Por fim, o capítulo 3 apresenta a proposta dos próximos passos da pesquisa juntamente com o cronograma das atividades.

2 REVISÃO BIBLIOGRÁFICA

2.1 Introdução

Beamforming é um método de filtragem espacial que diferencia os sinais desejados do ruído e interferência, através de um arranjo de microfones. Segundo (VEEN; BUCKLEY, 1998), essa técnica surgiu para realizar filtragem espacial nas seguintes aplicações:

- a) radar: controle de tráfego aéreo;
- b) sonar: localização e classificação;
- c) imagem: ultrassonografia e tomografia;
- d) exploração geofísica: mapeamento da crosta terrestre e exploração de petróleo;
- e) exploração astrofísica: imagens de alta resolução do universo;
- f) biomédica: monitoração de batimentos cardíacos de fetos, hipertermia do tecido humano para tratamento de doença e aparelhos auditivos;
- g) comunicações: transmissões e recepções direcionais, transmissão setorial em comunicação via satélite.

Como essa técnica é efetiva em filtragem espacial, ela foi adaptada para aplicações em sinais de voz. O termo “beamforming” ou, em português, “formação de feixe”, provém da formação de um feixe direcional de ganho em um diagrama polar, como pode-se verificar na figura 2.1.

A figura 2.2 ilustra o conceito de um beamforming constituído por J microfones. Frentes de onda estão irradiando os microfones com o sinal alvo $s(t)$ e ruído. As posições dos microfones são conhecidas. O sinal detectado por cada um dos microfones pode ser representado pela equação a seguir:

$$x_j(t) = a_j s(t - \tau_j) + n_j(t), \quad (2.1)$$

onde $x_j(t)$ representa o sinal detectado pelo microfone j sendo que $j = [1, 2, \dots, J]$, a_j é um parâmetro que inclui o ganho intrínseco do microfone, além da perda de propagação do sinal desejado até o microfone, $s(t - \tau_j)$ representa o sinal original desejado com o seu respectivo atraso e $n_j(t)$ representa o ruído detectado pelo microfone.

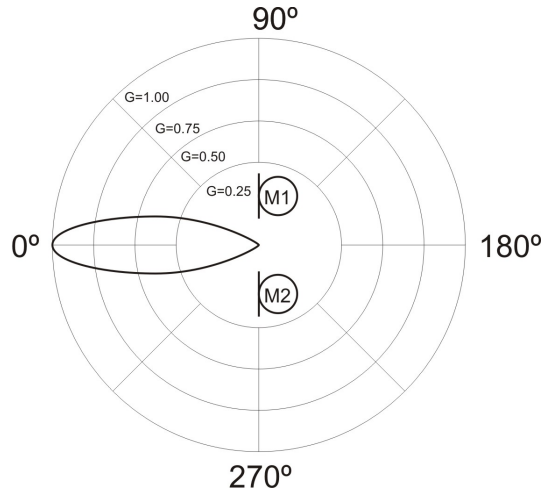


Figura 2.1: Exemplo diagrama polar de um beamforming formado por dois microfones.

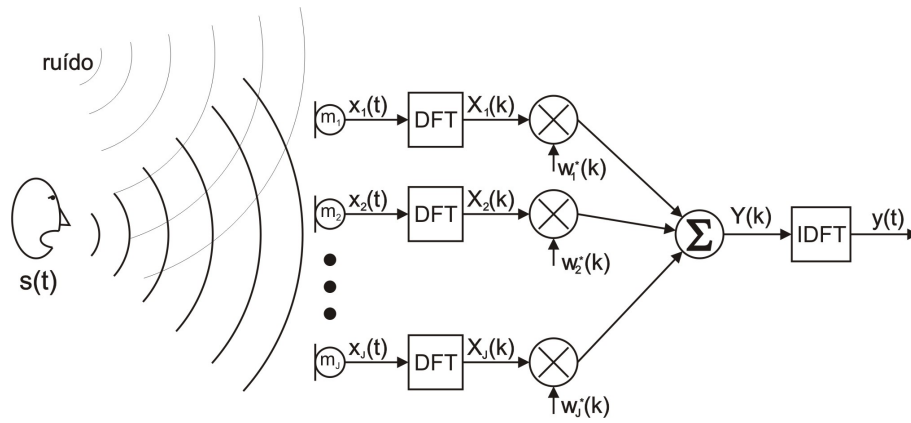


Figura 2.2: Diagrama ilustrando um beamforming constituído por J microfones.

O atraso τ_j pode ser ilustrado através da figura 2.3. Um sinal $s(t)$ incide sobre o arranjo de microfones em um ângulo θ . No microfone de referência m_1 , o atraso τ_1 é zero. No microfone m_2 , o sinal $s(t)$ incide após τ_2 segundos e no microfone m_J , o sinal $s(t)$ incide após τ_J segundos.

A seguir, a DFT (Discrete Fourier Transform) é calculada, obtendo a equação 2.2, onde k é a frequência discreta:

$$\mathbf{X}(k) = \mathbf{S}(k)\mathbf{d}(k) + \mathbf{N}(k), \quad (2.2)$$

onde:

$$\mathbf{X}(k) = [X_1(k), X_2(k), \dots, X_J(k)]^T, \quad (2.3)$$

$$\mathbf{d}(k) = [a_1 e^{-jk\tau_1}, a_2 e^{-jk\tau_2}, \dots, a_J e^{-jk\tau_J}]^T, \quad (2.4)$$

$$\mathbf{N}(k) = [N_1(k), N_2(k), \dots, N_J(k)]^T. \quad (2.5)$$

O vetor $\mathbf{d}(k)$ é chamado de vetor direção ou steering vector que indica o ângulo de incidência θ do sinal desejado. Algumas técnicas de beamforming, com por exemplo, o delay and

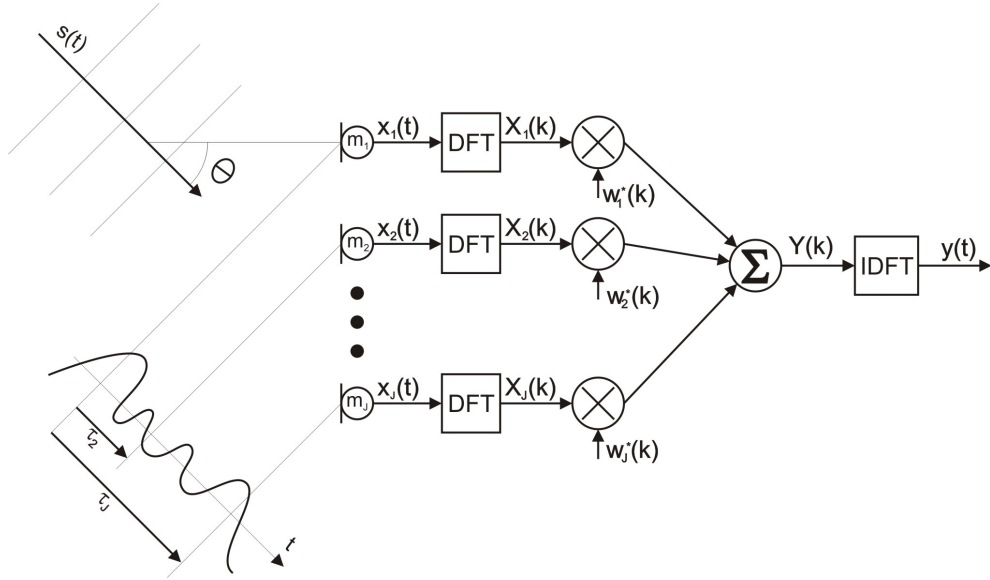


Figura 2.3: Representação do atraso τ_j no arranjo de microfones.

sum e o MVDR, utilizam esse vetor para filtrar o sinal desejado.

O algoritmo de beamforming calcula os pesos $w_j(k)^*$ para estimar $S(k)$, onde o símbolo $*$ representa o conjugado complexo de $w_j(k)$. Portanto, a saída do sistema será uma combinação linear dos sinais obtidos pelos microfones:

$$\mathbf{Y}(k) = \sum_{j=1}^J w_j(k)^* X_j(k), \quad (2.6)$$

$$= \mathbf{w}(k)^H \mathbf{X}(k), \quad (2.7)$$

onde $\mathbf{w}(k)^H$ é o transposto conjugado complexo de $\mathbf{w}(k)$. Para obter a saída $y(t)$ no domínio do tempo, calcula-se a transformada discreta inversa (IDFT) de $\mathbf{Y}(k)$.

Na próxima seção, serão demonstradas duas técnicas para calcular os valores de $\mathbf{w}(k)$.

2.2 Técnica Delay and Sum

A técnica Delay and Sum compensa o atraso dos sinais detectados por cada microfone apropriadamente antes que sejam combinados utilizando uma operação de soma. Os atrasos dos sinais são compensados de acordo com o vetor direção $\mathbf{d}(k)$, descrito na equação 2.4, que fornece a informação da direção do sinal desejado. O sinal combinado após a operação de atraso e soma apresenta uma versão reforçada do sinal desejado com diminuição do ruído detectado nos diferentes microfones. A Figura 2.4 ilustra a técnica delay and sum.

Segundo (VU et al., 2010), para o delay and sum, o vetor peso $\mathbf{w}(k)$ é calculado para satisfa-

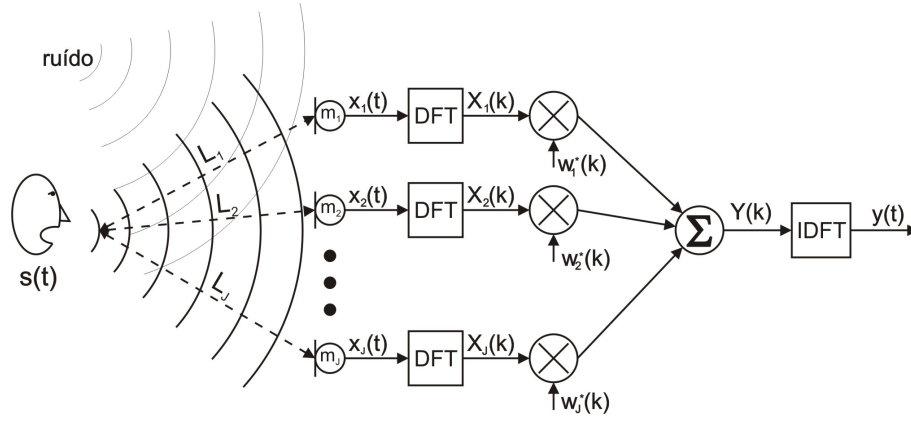


Figura 2.4: Diagrama da técnica delay and sum.

zer a seguinte equação:

$$\mathbf{w}(k)^H \mathbf{d}(k) = cte \quad (2.8)$$

Essa equação garante que o ganho na direção indicada pelo vetor $\mathbf{d}(k)$ é constante. Geralmente o valor da constante é unitário.

Segundo (VU et al., 2010), para uma fonte localizada próxima ao arranjo de microfones, a frente de onda gerada por essa fonte quando atinge os microfones, deve ser considerada de formato esférico, diferente de quando a fonte sonora localiza-se longe dos microfones, onde a frente de onda que atinge os microfones é considerada plana. Segundo (MAGRI; PEDRONI; ZUBELLI, 2001), os valores de a_j e τ_j , componentes do vetor direção $\mathbf{d}(k)$, quando a fonte sonora está próxima ao arranjo de microfones, podem ser calculados por:

$$a_j = \frac{L_{ref}}{L_j}, \tau_j = \frac{L_j - L_{ref}}{c}, \quad (2.9)$$

onde L_{ref} e L_j são, respectivamente, as distâncias entre a fonte sonora desejada $s(t)$ e o microfone de referência e a distância entre a fonte sonora e o microfone j . Se, por exemplo, o microfone de referência adotado for o m_1 , então $L_{ref} = L_1$. O valor c é a velocidade do som. O sinal de saída no domínio da frequência $\mathbf{Y}(k)$ é então calculado através da equação 2.10. O sinal $y(t)$ é calculado através da transformada inversa de $\mathbf{Y}(k)$.

2.3 Técnica MVDR (Minimum Variance Distortionless Response)

A técnica MVDR (Minimum Variance Distortionless Response), também conhecida como MOE (Minimum Output Energy), utiliza um critério diferente de otimização: minimizar a energia total na saída mantendo o ganho do arranjo de microfones direcionado no sinal desejado. Ou seja, qualquer sinal proveniente de outras direções que não seja a direção estabelecida terá o seu ganho atenuado. Segundo (ADVE, 2011) essa técnica pode ser representada matematicamente como:

$$\mathbf{w}_{MV}(k) = \arg \min_{\mathbf{w}(k)} E \{ |\mathbf{Y}(k)|^2 \} = \arg \min_{\mathbf{w}(k)} E \{ |\mathbf{w}(k)^H \mathbf{X}(k)|^2 \}, \quad (2.10)$$

restrito a:

$$\mathbf{w}(k)^H \mathbf{d}(k) = C, \quad (2.11)$$

onde C , é um número real.

A equação 2.10 significa que deseja-se obter o vetor peso $\mathbf{w}(k)$, chamado de $\mathbf{w}_{MV}(k)$ que fornece a mínima energia na saída do sistema (ou na mínima variância), porém, restrito a condição de 2.11, isto é, mantendo o sinal desejado proveniente da direção especificada pelo vetor direção $\mathbf{d}(k)$ com o ganho C . O peso $\mathbf{w}^H(k)$ é o transposto complexo do vetor peso $\mathbf{w}(k)$.

A minimização de $\mathbf{w}(k)$ segundo a equação 2.10, restrito a equação 2.11 pode ser resolvida utilizando o multiplicador de Lagrange, encontrando $\min_{\mathbf{w}(k)} [\Lambda(\mathbf{w}(k); \lambda)]$. A resolução encontra-se abaixo:

$$\begin{aligned} \Lambda(\mathbf{w}(k); \lambda) &= E \{ |\mathbf{w}(k)^H \mathbf{X}(k)|^2 \} + \lambda [\mathbf{w}(k)^H \mathbf{d}(k) - C], \\ &= E \{ \mathbf{w}(k)^H \mathbf{X}(k) \mathbf{X}(k)^H \mathbf{w}(k) \} + \lambda [\mathbf{w}(k)^H \mathbf{d}(k) - C], \\ &= \mathbf{w}(k)^H \mathbf{R}(k) \mathbf{w}(k) + \lambda (\mathbf{w}(k)^H \mathbf{d}(k) - C), \\ \frac{\partial \Lambda}{\partial \mathbf{w}(k)^H} &= \mathbf{R}(k) \mathbf{w}(k) + \lambda \mathbf{d}(k) = 0 \\ \Rightarrow \mathbf{w}(k)_{MV} &= -\lambda \mathbf{R}(k)^{-1} \mathbf{d}(k). \end{aligned} \quad (2.12)$$

$$\begin{aligned} \frac{\partial \Lambda}{\partial \lambda} &= \mathbf{w}(k)^H \mathbf{d}(k) - C = 0, \\ \Rightarrow \mathbf{w}(k)^H \mathbf{d}(k) &= C. \end{aligned} \quad (2.13)$$

Substituído 2.12 em 2.13, obtemos:

$$-\lambda = \frac{C}{\mathbf{d}(k)^H \mathbf{R}(k)^{-1} \mathbf{d}(k)}. \quad (2.14)$$

Substituindo 2.14 em 2.12:

$$\mathbf{w}_{MV}(k) = C \frac{\mathbf{R}(k)^{-1} \mathbf{d}(k)}{\mathbf{d}(k)^H \mathbf{R}(k)^{-1} \mathbf{d}(k)} \quad (2.15)$$

Considerando-se o ganho C unitário, o sinal desejado obtido na saída não será distorcido, isto é, a amplitude do sinal na saída será igual a amplitude do sinal desejado capturado na entrada. Essa é a razão do termo "distortionless response".

2.4 Conclusão

A técnica delay and sum, apresentada na seção 2.2, possui a vantagem de utilizar baixa carga computacional, porém, não atenua o ruído efetivamente. Somente realiza o atraso e soma dos sinais detectados, sem enfatizar a minimização do ruído.

(LOCKWOOD et al., 2004) realizou testes de performance entre alguns beamformers: o MVDR no domínio da frequência ou FMV (Frequency minimum variation), o algoritmo de Frost no domínio do tempo (FROST, 1972) e o GSC (generalized sidelobe canceler) de (GREENBERG, 1998). Os testes foram realizados utilizando dois microfones, sendo: 2 microfones omnidirecionais separados de 15cm; 2 microfones cardioides separados de 15cm e microfones montados nos ouvidos ouvidos de um manequim Kemar (BURKHARD; SACHS, 1978). Foram utilizados 3 quartos de dimensões diferentes para realizar as medições. A performance de cada algoritmo foi medida pela relação sinal ruído e pela distorção do sinal na saída.

Como resultado dos testes, entre os beamformers, o MVDR no domínio da frequência apresentou a melhor relação sinal ruído e a menor distorção do sinal desejado na saída do sistema.

Mesmo utilizando dois microfones, a performance do MVDR foi superior a dos outros algoritmos. Como o objetivo é estudar um algoritmo para ser utilizado em um dispositivo para proteção ao trabalhador, um número menor de microfones reduz o custo do equipamento e as suas dimensões. Além disso, esse algoritmo já foi implementado em um DSP da Texas Instruments modelo C62X (ELLEDEGE et al., 2000), o que aumenta a probabilidade de sucesso se utilizado em um equipamento móvel.

Portanto, devido a sua performance superior na relação sinal ruído, baixa distorção do sinal desejado e seu comprovado funcionamento em um DSP, o algoritmo escolhido foi o MVDR no domínio da frequência. Outra vantagem desse algoritmo é que ele possui todas essas características comprovadas com a utilização de somente dois microfones, o que viabiliza a construção de um dispositivo de comunicação e proteção do trabalhador em ambientes de altos níveis de pressão sonora.

3 TRABALHOS SUBSEQUENTES E CRONOGRAMA

Nesse capítulo são apresentados os trabalhos subsequentes e o cronograma das atividades desenvolvidas juntamente com a proposta das atividades finais.

3.1 Trabalhos Subsequentes

A próxima etapa do trabalho contempla um estudo mais detalhado do algoritmo MVDR (Minimum Variance Distortionless Response) binaural com domínio na frequência, segundo descrito por (LOCKWOOD et al., 2004). Esse algoritmo foi escolhido por critérios já comentados na seção anterior. Após o estudo detalhado dessa técnica, o algoritmo será implementado e serão realizados testes preliminares, para comprovar o seu funcionamento e eficácia. Por fim, serão realizados testes em campo, com dados reais, para sustentar a eficácia do algoritmo.

3.2 Cronograma

Segue abaixo o cronograma das atividades realizadas até o momento e também previsão das próximas atividades.

	2011											2012		
Atividade	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez	Jan	Fev	Mar
Revisão Bibliográfica preliminar	X	X	X	X	X	X								
Definição da técnica utilizada	X	X	X	X	X	X								
Redação da qualificação						X	X	X						
Exame de qualificação								X						
Estudo da técnica MVDR e obtenção de resultados experimentais								X	X	X	X			
Elaboração da dissertação										X	X	X	X	
Apresentação da dissertação														X

Tabela 3.1: Cronograma das atividades.

REFERÊNCIAS

- ADVE, R. Optimal beamforming. University of Toronto, 2011. Acesso em 21/08/2011. Disponível em: <<http://www.comm.utoronto.ca/rsadve/Notes/BeamForming.pdf>>.
- BURKHARD, M. D.; SACHS, R. M. Anthropometric manikin for acoustic research. Illinois, EUA, 1978. Acesso em 21/08/2011. Disponível em: <<http://www.gras.dk/documents/00316.pdf>>.
- ELLEDGE, M. E. et al. Real-time implementation of a frequency-domain beamformer on the ti c62x evm. *10th Annual DSP Technology Education and Research Conference*, Texas Instrument, Houston, TX, 2000.
- FROST, O. L. An algorithm for linearly constrained adaptive array processing. *Proceeding of the IEEE*, IEEE, v. 60, n. 8, p. 926–935, 1972.
- GREENBERG, J. E. Modified lms algorithms for speech processing with an adaptive noise canceller. *IEEE Transactions on speech and audio processing*, IEEE, v. 6, p. 338–351, 1998.
- LOCKWOOD, M. E. et al. Performance of time- and frequency-domain binaural beamformers based on recorded signals from real rooms. *J. Acoustic. Soc. Am.*, Acoustical Society of America, v. 115, p. 379–391, 2004.
- MAGRI, F.; PEDRONI, M.; ZUBELLI, J. P. 2001. K. U. Simmer, Superdirective microphone arrays, in *Microphone Arrays*, M. S. Brandstein and D. B. Ward, Eds., chapter 2, pp. 19-38. Springer, 2001.
- MTE, M. d. T. e. E. *Norma Regulamentadora NR-15. Anexo 1. Atividades e Operações insalubres*. 2011. Acesso em 31/08/2011. Disponível em: <<http://www.mte.gov.br>>.
- PELTOR, C. *Abafador de ruído com comunicador PowerComm*. 2011. Acesso em 31/08/2011. Disponível em: <<http://www.aearo.com/>>.
- SENSEAR, P. *Protetores auditivos com comunicação Sensear*. 2011. Acesso em 31/08/2011. Disponível em: <<http://www.sensear.com>>.
- VEEN, B. D. V.; BUCKLEY, K. M. Beamforming: A versatile approach to spatial filtering. *IEEE ASSP Magazine*, IEEE, p. 338–351, abr 1998.
- VU, N. V. et al. Small footprint implementation of dual-microphone delay-and-sum beamforming for in-car speech enhancement. 2010. Artigo apresentado na ICASSP.
- WANG, D.; BROWN, G. J. *Computational Auditory Scene Analysis. Principles, Algorithms and Applications*. 1st ed.. ed. Hoboken, New Jersey: Wiley-Interscience, 2005. 1–37 and 187–205 p.