

Método cepstral para estimação da resposta ao impulso do efeito de oclusão em sistemas de auxílio à audição

Bruno C. Bispo, Renata C. Borges e Márcio H. Costa

Resumo— O efeito de oclusão é uma queixa comum de usuários de aparelhos auditivos que utilizam moldes sem ventilação ou com ventilação estreita, os quais promovem uma sensação de abafamento da própria voz. Na literatura, controladores fixos e adaptativos têm sido propostos para a redução desse efeito, o qual possui uma solução ótima não-causal. Este trabalho propõe um método cepstral para identificar a resposta ao impulso da oclusão e um sistema de controle fixo que utiliza essa estimativa para redução do efeito de oclusão em sistemas de auxílio à audição. Simulações indicam que a eficiência do método cepstral depende do tamanho da janela de análise, resultando em estimativas com desalinhamento média de $-28,77$ dB para janelas de análise de 2 s. E que o controlador aproxima o sinal a ser ouvido pelo usuário do seu valor ideal, diminuindo em $0,34$ dB a potência do erro e aumentando em $0,12$ MOS-LQO a qualidade perceptual do sinal.

Palavras-Chave— Efeito de oclusão, aparelho auditivo, análise cepstral, controle ativo.

Abstract— The occlusion effect is a common complaint of hearing aids users who use narrow- or unvented ear molds. This phenomenon is characterized by a muffled version of the user's own voice. In the scientific literature, fixed and adaptive controllers have been proposed for occlusion effect reduction. This work proposes the use of cepstral estimation to identify the impulse response of the controller to reduce the occlusion effect in hearing aid systems. Simulations have shown that the efficiency of the cepstral method depends on the size of the analysis window, resulting in estimates with a mean normalized misalignment of -28.77 dB for 2 s-long analysis window. And the controller makes the signal to be heard by the user closer to its ideal value, decreasing by 0.34 dB the error power and increasing by 0.12 MOS-LQO its perceptual quality.

Keywords— Occlusion effect, hearing aid, cepstral analysis, active control.

I. INTRODUÇÃO

O efeito de oclusão em aparelhos auditivos ocorre quando o duto de ventilação possui diâmetro insuficiente para promover a necessária dissipação de energia sonora conduzida ao canal auditivo através do crânio e da mandíbula quando o usuário da prótese auditiva fala [1], [2]. Esse efeito é modelado pela resposta ao impulso $w_o(n)$ como ilustrado na Fig. 1 na configuração chaves 1 e 2 abertas.

A voz do usuário do aparelho auditivo $u(n)$ é captada pelo microfone externo do aparelho, gerando o sinal $y(n)$, após percorrer o canal acústico representado pela resposta ao

Bruno C. Bispo e Márcio H. Costa, Departamento de Engenharia Elétrica e Eletrônica, Universidade Federal de Santa Catarina, Florianópolis-SC, e-mail: bruno.bispo@ufsc.br, costa@eel.ufsc.br. Renata C. Borges, Departamento de Engenharia Elétrica, Universidade Tecnológica Federal do Paraná, Cornélio Procópio-PR, e-mail: renatacoelho@utfpr.edu.br.

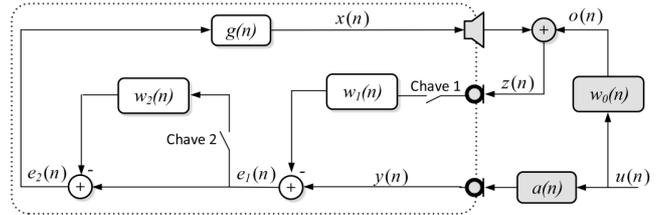


Fig. 1: Diagrama de blocos combinado: oclusão sem cancelamento se chaves abertas; cancelamento de oclusão em malha aberta se chave 1 aberta e chave 2 fechada; cancelamento de oclusão em malha fechada se chave 1 fechada e chave 2 aberta.

impulso $a(n)$. O sinal do microfone interno $y(n)$ é amplificado pelo sistema de compensação representado pela resposta ao impulso $g(n)$, gerando o sinal $x(n)$ a ser reproduzido pelo alto-falante do aparelho auditivo. O sinal $z(n)$, a ser captado por um possível microfone interno do aparelho, é de fato o sinal a ser ouvido pelo usuário, sendo definido como

$$\begin{aligned} z(n) &= x(n) + u(n) * w_o(n) \\ &= [g(n) * a(n) + w_o(n)] * u(n), \end{aligned} \quad (1)$$

onde o símbolo $*$ denota a operação de convolução. Portanto, o efeito de oclusão é caracterizado pelo acréscimo de $w_o(n) * u(n)$ no valor desejado de $z(n)$ e, assim, pode ser interpretado do ponto de vista de reverberação.

Desde 2008, estratégias fixas e adaptativas têm sido propostas para, pelo menos, atenuar o efeito da oclusão. As soluções fixas, denominadas sistemas de controle, asseguram a estabilidade do sistema mas não lidam com as mudanças dinâmicas do sistema acústico, podendo sofrer perdas de desempenho devido às variações do canal auditivo ou deslocamentos do molde auricular [1], [3]–[5]. Por outro lado, as soluções adaptativas, denominadas sistemas de cancelamento, propostas até hoje em [3], [4] apresentam convergência lenta dos coeficientes do filtro adaptativo e requerem adaptação constante em busca de uma solução ótima uma vez que o efeito de oclusão ocorre em períodos curtos de tempo.

Ambas as soluções podem ser implementadas em malha aberta ou fechada. A Fig. 1 apresenta a solução em malha aberta para a configuração chave 1 aberta e chave 2 fechada, e a solução em malha fechada para a configuração chave 1 fechada e chave 2 aberta. Note que, em malha fechada, faz-se uso de um microfone interno ao aparelho auditivo para captação e uso do sinal $z(n)$.

O objetivo desse trabalho é duplo: primeiro, propor um método cepstral para estimar a resposta ao impulso $w_o(n)$ do

sistema de oclusão utilizando a estrutura em malha fechada; segundo, propor um controlador fixo em malha aberta que utilize a estimativa de $w_o(n)$ obtida pelo método cepstral para atenuar o efeito da oclusão. Este artigo está organizado da seguinte maneira: a Seção II demonstra a solução ótima para o cancelamento não-realimentado da oclusão, justificando a necessidade de se obter uma estimativa precisa de $w_o(n)$; na Seção III, o método cepstral proposto para estimar $w_o(n)$ é detalhado; na Seção IV, as configurações das simulações realizadas são apresentadas; a Seção V apresenta e discute os resultados obtidos; e, por fim, a Seção VI conclui o trabalho.

II. SISTEMA NÃO-REALIMENTADO DE CANCELAMENTO DA OCLUSÃO

No sistema não-realimentado de cancelamento de oclusão, exemplificado pela Fig. 1 com chave 1 aberta e chave 2 fechada, pode-se demonstrar que

$$Z(e^{j\omega}) = [G(e^{j\omega})A(e^{j\omega}) - G(e^{j\omega})A(e^{j\omega})W_2(e^{j\omega}) + W_o(e^{j\omega})] U(e^{j\omega}), \quad (2)$$

Para que $Z(e^{j\omega}) = A(e^{j\omega})G(e^{j\omega})U(e^{j\omega})$, como desejado, é necessário que

$$W_2(e^{j\omega}) = \frac{W_o(e^{j\omega})}{A(e^{j\omega})G(e^{j\omega})}, \quad (3)$$

ou seja,

$$w_2(n) = w_o(n) * a^{-1}(n) * g^{-1}(n), \quad (4)$$

onde $a^{-1}(n)$ e $g^{-1}(n)$ são as respostas ao impulso dos sistemas inversos a $a(n)$ e $g(n)$, respectivamente.

Assumindo que $a(n) = \delta(n - \tau)$ e $g(n) = g$, isto é, o canal acústico entre a boca do usuário e o microfone externo do aparelho é apenas um atraso e o sistema de compensação do aparelho é apenas um ganho, como normalmente feito na literatura, a equação (4) se torna

$$w_2(n) = \frac{w_o(n + \tau)}{g}. \quad (5)$$

A partir de (5), nota-se que a solução ótima é não-causal e dependente diretamente da resposta ao impulso $w_o(n)$ do sistema de oclusão. É fácil compreender que a solução causal ótima também dependerá de $w_o(n)$ uma vez que alterações em $w_o(n)$ modificam $z(n)$ em (1) e a solução ótima para $w_2(n)$ em (4). Portanto, é de suma importância ter o conhecimento de $w_o(n)$, seja para usar num projeto de um controlador fixo ou como inicialização dos coeficientes de um filtro adaptativo.

Na próxima seção, um método para estimar $w_o(n)$ utilizando a estrutura em malha fechada e um controlador fixo em malha aberta que utilize essa estimativa de $w_o(n)$ para atenuar o efeito da oclusão serão propostos.

III. METODOLOGIA

A. Análise cepstral do sistema de cancelamento realimentado

O sistema realimentado de cancelamento de oclusão descrito na Figura 1 (chave 1 fechada e chave 2 aberta) é governado

pelos seguintes equações no domínio da frequência

$$\begin{cases} Y(e^{j\omega}) = A(e^{j\omega})U(e^{j\omega}) \\ E_1(e^{j\omega}) = Y(e^{j\omega}) - W_1(e^{j\omega})Z(e^{j\omega}) \\ X(e^{j\omega}) = G(e^{j\omega})E(e^{j\omega}) \\ Z(e^{j\omega}) = X(e^{j\omega}) + W_o(e^{j\omega})U(e^{j\omega}). \end{cases} \quad (6)$$

Assim, a relação no domínio da frequência entre o sinal de entrada do sistema $u(n)$ e o sinal de erro $e_1(n)$ é dada por

$$E_1(e^{j\omega}) = \frac{1 - W_o(e^{j\omega})W_1(e^{j\omega})/A(e^{j\omega})}{1 + G(e^{j\omega})W_1(e^{j\omega})} U(e^{j\omega}), \quad (7)$$

a qual aplicando o logaritmo natural se torna

$$\begin{aligned} \ln [E_1(e^{j\omega})] &= \ln [U(e^{j\omega})] \\ &+ \ln [1 - W_o(e^{j\omega})W_1(e^{j\omega})/A(e^{j\omega})] \\ &- \ln [1 + G(e^{j\omega})W_1(e^{j\omega})]. \end{aligned} \quad (8)$$

Se $|W_o(e^{j\omega})W_1(e^{j\omega})/A(e^{j\omega})| < 1$, então o termo do meio do lado direito de (8) pode ser expandido em série de Taylor como

$$\begin{aligned} \ln [1 - W_o(e^{j\omega})W_1(e^{j\omega})/A(e^{j\omega})] &= \\ &- \sum_{k=1}^{\infty} \frac{[W_o(e^{j\omega})W_1(e^{j\omega})/A(e^{j\omega})]^k}{k}. \end{aligned} \quad (9)$$

E se $|1 + G(e^{j\omega})W_1(e^{j\omega})| < 1$, então o termo mais à direita de (8) pode ser expandido em série de Taylor como

$$\begin{aligned} \ln [1 + G(e^{j\omega})W_1(e^{j\omega})] &= \\ &\sum_{k=1}^{\infty} (-1)^{k+1} \frac{[G(e^{j\omega})W_1(e^{j\omega})]^k}{k}. \end{aligned} \quad (10)$$

Substituindo (9) e (10) em (8), e aplicando a transformada inversa de Fourier, a relação no domínio cepstral entre o sinal de entrada do sistema $u(n)$ e o sinal de erro $e(n)$ é dada por

$$\begin{aligned} c_{e_1}(n) &= c_u(n) - \sum_{k=1}^{\infty} \frac{\{w_o(n) * w_1(n) * a^{-1}(n)\}^{*k}}{k} \\ &- \sum_{k=1}^{\infty} (-1)^{k+1} \frac{\{g(n) * w_1(n)\}^{*k}}{k}, \end{aligned} \quad (11)$$

onde $\{\cdot\}^{*k}$ denota a k -ésima potência de convolução.

Portanto, no sistema realimentado de cancelamento de oclusão, o cepstro $c_{e_1}(n)$ do sinal de erro é o cepstro $c_u(n)$ do sinal de entrada do sistema adicionado a duas séries temporais em função de $w_o(n)$, $w_1(n)$, $g(n)$ e $a^{-1}(n)$. A análise cepstral modificou a representação dos componentes do sistema em relação ao sinal de entrada do sistema $u(n)$. Em (7), o sinal $u(n)$ e os componentes do sistema estão representados no domínio da frequência. Mas em (11), o sinal $u(n)$ está representado no domínio cepstral enquanto os componentes do sistema estão representados no domínio do tempo.

Porém, deve ser ressaltado que $c_{e_1}(n)$ é definido de acordo com (11) se as condições para as expansões em séries de Taylor em (9) e (10) forem satisfeitas. Caso contrário, nada pode ser inferido sobre a definição matemática de $c_{e_1}(n)$ em função de $w_o(n)$, $w_1(n)$, $g(n)$ e $a^{-1}(n)$. E na prática onde

$c_{e_1}(n)$ é calculado por meio da transformada rápida de Fourier de uma janela de análise de $e_1(n)$, a existência das respostas ao impulso em (11) depende se a janela de análise é grande o suficiente para incluir os seus efeitos [6].

B. Método cepstral para estimação de $w_o(n)$

O objetivo do método proposto é utilizar (11) para estimar $w_o(n)$ de forma similar ao feito em [6], [7] para estimar a realimentação acústica. Pelo fato de ter-se conhecimento dos sinais $e_1(n)$ e $u(n)$, seus respectivos cepstros podem ser calculados utilizando a FFT (*Fast Fourier Transform*). A subtração de $c_{e_1}(n)$ por $c_u(n)$ resultaria então nos dois somatórios em (11).

Em geral $|w_o(n)| \ll 1$ (Figura 2) e, portanto, $w_1(n)$ pode ser inicializado de maneira que as condições para as expansões em série de Taylor sejam satisfeitas. Combinando isso com o fato de que o fator $1/k$ em (11) é o responsável por penalizar o aumento da potência da convolução, pode-se concluir que quanto maior o valor de k menor serão os valores das respostas ao impulso em (11). Devido a isto, o método objetiva primeiramente estimar a resposta ao impulso presente em (11) para $k = 1$, isto é, estimar $w_1(n) * [w_o(n) * a^{-1}(n) + g(n)]$.

Para que isso seja possível é necessário que sejam estabelecidos alguns critérios de inicialização do vetor $w_1(n)$, de forma que as respostas ao impulso em (11) para diferentes valores de k não estejam sobrepostas. Como $g(n)$ é conhecido, pode-se definir $w_1(n)$ como

$$w_1(n) = 0,1g(n - \tau). \quad (12)$$

Dadas essas condições, o método proposto calcula inicialmente $\{w_1(n) * [w_o(n) * a^{-1}(n) + g(n)]\}^{\wedge}$, uma estimativa de $w_1(n) * [w_o(n) * a^{-1}(n) + g(n)]$, ao selecionar as primeiras $2L$ amostras de $c_{e_1}(n) - c_u(n)$, onde L é o tamanho de $w_o(n)$. Em seguida, obtém-se $\{w_o(n) * a^{-1}(n) + g(n)\}^{\wedge}$, uma estimativa de $w_o(n) * a^{-1}(n) + g(n)$, da seguinte maneira

$$\{w_o(n) * a^{-1}(n) + g(n)\}^{\wedge} = \{w_1(n) * [w_o(n) * a^{-1}(n) + g(n)]\}^{\wedge} * w_1^{-1}(n), \quad (13)$$

onde $w_1^{-1}(n)$ é a resposta ao impulso do sistema inverso a $w_1(n)$. Para $w_1(n)$ definido em (12), a convolução acima pode ser entendida como um avanço temporal e uma multiplicação.

Utilizando o conhecimento de $g(n)$, calcula-se $\{w_o(n) * a^{-1}(n)\}^{\wedge}$, uma estimativa de $w_o(n) * a^{-1}(n)$, como se segue

$$\{w_o(n) * a^{-1}(n)\}^{\wedge} = \{w_o(n) * a^{-1}(n) + g(n)\}^{\wedge} * g^{-1}(n). \quad (14)$$

Note que se $g(n) = g$, então a convolução acima é apenas uma multiplicação.

Por fim, com o conhecimento de $a(n)$, obtém-se $\hat{w}_o(n)$, uma estimativa de $w_o(n)$, da seguinte maneira

$$\hat{w}_o(n) = \{w_o(n) * a^{-1}(n)\}^{\wedge} * a(n). \quad (15)$$

Nesse caso, se $a(n) = \delta(n - \tau)$, a convolução acima é apenas um atraso temporal.

C. Proposta do controlador fixo em malha aberta

O controlador proposto para cancelamento do efeito de oclusão utiliza a estrutura não-realimentada apresentada em [3], representado na Figura 1 com chave 1 aberta e chave 2 fechada, e a estimativa $\hat{w}_o(n)$ da resposta ao impulso da oclusão obtida pelo método cepstral proposto na Seção III-B.

O controlador é definido pela seguinte resposta ao impulso

$$w_2(n) = \frac{\hat{w}_o(n + \tau)}{g}v(n), \quad (16)$$

onde $v(n)$ é o degrau unitário. Portanto, a resposta ao impulso do controlador é a estimativa da solução ótima, que é não-causal, truncada de forma a se tornar causal.

IV. CONFIGURAÇÃO DAS SIMULAÇÕES

A. Base de dados

A base de dados consiste em 20 gravações (12 masculinas e 8 femininas) da vogal sustentada */a/*. A vogal */a/* foi escolhida devido à sua ampla utilização na análise acústica da voz.

A base de dados foi fornecido pelo Grupo de Pesquisa em Engenharia Médica do Conselho Nacional de Desenvolvimento Científico e Tecnológico (GPEM/CNPq). As gravações foram realizadas no Ambulatório de Voz do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HC-FMUSP) utilizando uma frequência de amostragem de 22050 Hz. Mas, neste trabalho, os sinais foram reamostrados para 16000 Hz.

Além disso, os sinais foram truncados de forma a terem 2 s de duração e os níveis de potência ativa foram normalizados para -26 dBov de acordo com o algoritmo da Recomendação ITU-T P.56 [8].

B. Configuração do sistema

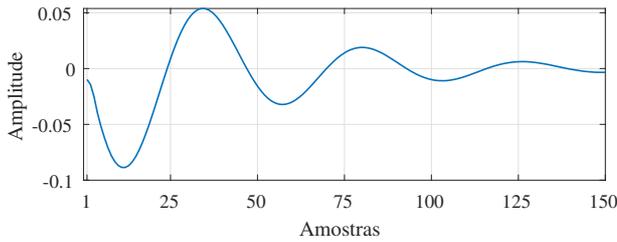
O atraso de propagação foi definido como

$$a(n) = \delta(n - \tau), \quad (17)$$

onde $\tau = 14$. Este atraso refere-se a um comprimento de propagação de 30,29 cm, no qual é considerado um comprimento médio de 15 cm entre a glote e os lábios [9], e 15,29 cm de distância entre os lábios e o microfone externo dos aparelhos auditivos [3]. Assumiu-se nesse momento que não há falantes externos nem ruído, somente o sinal de fala do usuário do aparelho auditivo. O processamento intrínseco do aparelho, $g(n)$, foi modelado como ganho de banda larga e um atraso de propagação. Os processos de filtragem associados à cadeia de aquisição (microfones, pré-amplificadores, filtros anti-aliasing, conversão analógico-digital) são considerados desprezíveis em toda a faixa de frequência [3].

C. Sistema de Oclusão

O efeito de oclusão foi simulado por um filtro linear com resposta ao impulso finita com $L = 150$ coeficientes (Figura 2), que foram obtidos em um processo de identificação real para um voluntário com um molde auricular personalizado e sem ventilação [3].


 Fig. 2: Resposta ao impulso $w_o(n)$ do sistema de oclusão.

D. Métricas de Avaliação

1) *Desalinhamento*: A precisão da estimativa da resposta ao impulso do sistema de oclusão obtida pelo método cepstral proposto foi medida por meio do desalinhamento normalizado (*Misalignment - MIS*), o qual é definido como

$$\text{MIS} = \frac{\left\{ \sum_{n=0}^{L-1} [w_o(n) - \hat{w}_o(n)]^2 \right\}^{1/2}}{\left\{ \sum_{n=0}^{L-1} w_o^2(n) \right\}^{1/2}}. \quad (18)$$

2) *Potência do erro*: A eficiência do controlado fixo proposto foi medida pela potência da diferença entre o sinal $z(n)$ e o seu valor ideal, isto é,

$$P = \frac{\sum_{n=0}^{N-1} [z(n) - g(n) * a(n) * u(n)]^2}{N}, \quad (19)$$

onde $N = 32000$ é o comprimento dos sinais de voz.

3) *W-PESQ*: O W-PESQ (*Wideband Perceptual Evaluation of Speech Quality*) é um algoritmo para avaliação objetiva da qualidade de sinais de fala de banda larga (amostrados a 16 kHz) [10]–[12]. Ele utiliza as versões de referência (original) e degradada (processada) de um sinal de fala para avaliar objetivamente a degradação perceptual da última, que é quantificada na escala 1-5 de opinião média (*Mean Opinion Score - MOS*) de acordo com a classificação da categoria de degradação (*Degradation Category Rating - DCR*), resultando na pontuação MOS-LQO (*MOS-Listening Quality Objective*). No entanto, quando os sinais de referência e degradados são idênticos, a pontuação do W-PESQ é de 4,644 MOS-LQO.

O W-PESQ possui uma correlação de 80% com MOS ao avaliar o comprometimento da fala por reverberação embora não tenha sido projetado para esse propósito [13], [14]. Não há estudos que comprovem que a oclusão, que é caracterizada pela amplificação de faixas estreitas em baixas frequências, é adequadamente quantificada por esse algoritmo. No entanto, como explicado na Seção I, o efeito da oclusão pode ser interpretado do ponto de vista de reverberação. Assim, neste trabalho, o algoritmo W-PESQ foi utilizado para avaliar perceptivamente a distorção inserida no sinal $z(n)$ em comparação com seu valor ideal $g(n) * a(n) * u(n)$.

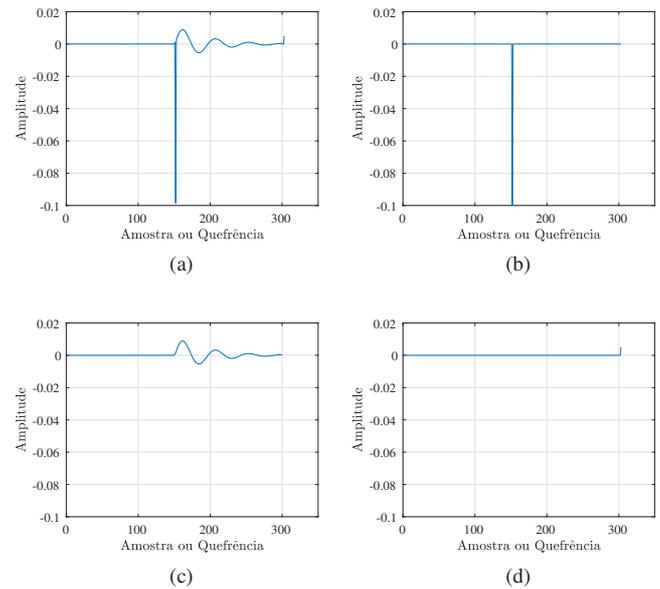
Mas deve-se ressaltar que o W-PESQ foi validado para sinais com duração entre 8 e 12 s e pode ser utilizado em sinais com, pelo menos, 3,2 s de fala [15]. Como neste trabalho os sinais têm apenas 2 s de duração, os resultados do W-PESQ precisam ser tratados cautelosamente.

V. RESULTADOS DAS SIMULAÇÕES

A. Resultados do método cepstral para estimação de $w_o(n)$

Como explicado na Seção III, o método cepstral utiliza a estrutura realimentada de cancelamento de oclusão com $w_1(n)$ definido em (12) para estimar $w_o(n)$.

Para exemplificar a equação (11), a Figura 3 mostra $c_{e_1}(n) - c_u(n)$ calculado utilizando uma janela de análise de 2 s e as respostas ao impulso $-g(n) * w_1(n)$, $-w_o(n) * w_1(n)$ e $0,5 [g(n) * w_1(n)]^{*2}$. Pode-se concluir que, de fato, essas respostas ao impulso aparecem em $c_{e_1}(n) - c_u(n)$ de acordo com (11) e podem ser estimadas desse sinal.


 Fig. 3: Visualização das respostas ao impulso presentes em $c_{e_1}(n)$ de acordo com (11): (a) $c_{e_1}(n) - c_u(n)$; (b) $-g(n) * w_1(n)$; (c) $-w_o(n) * w_1(n)$; (d) $0,5 [g(n) * w_1(n)]^{*2}$.

O desalinhamento das estimativas $\hat{w}_o(n)$ obtidas pelo método cepstral utilizando vários tamanhos da janela de análise são apresentados na Tabela 1. Pode-se notar que quanto maior a janela de análise utilizada para calcular os cepstros, melhor será a estimativa de $w_o(n)$ obtida pelo método. E, além disso, para o método obter uma estimativa com alguma precisão é necessário uma janela de análise igual ou maior que 200 ms. Os resultados obtidos utilizando janelas de análise com tamanho de 0,5 e 2 s são ilustrados na Fig. 4.

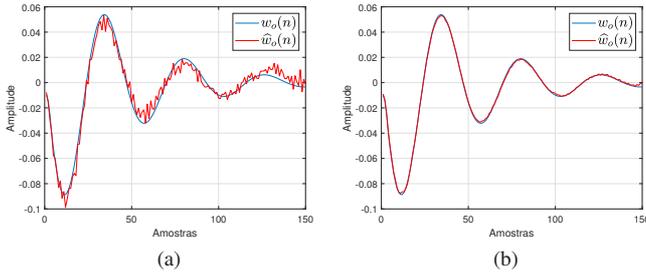
Na prática, a utilização do método cepstral para identificar $w_o(n)$ deve ser realizado quando o aparelho auditivo é ligado, num sistema de calibração, e para que a resposta ao impulso seja devidamente estimada o usuário deve fazer a locução de uma vogal por um determinado período de tempo.

B. Resultados do controlador fixo

Como explicado na Seção III-C, o controlador proposto para atenuar o efeito de oclusão utiliza a estrutura não-realimentada, representada na Figura 1 com chave 1 aberta e chave 2 fechada, e possui a resposta ao impulso definida em (16), baseada na estimativa $\hat{w}_o(n)$ obtida pelo método cepstral.

TABELA I: Média e desvio padrão do desalinhamento normalizado para vários tamanhos da janela de análise.

| Tamanho da janela (ms) | Desalinhamento médio (dB) | Desvio padrão (dB) |
|------------------------|---------------------------|--------------------|
| 20 | 10,27 | 1,34 |
| 50 | 5,37 | 1,13 |
| 75 | 2,17 | 1,67 |
| 100 | 0,16 | 2,33 |
| 300 | -7,31 | 2,62 |
| 500 | -10,60 | 3,66 |
| 750 | -12,91 | 4,96 |
| 1000 | -14,85 | 4,79 |
| 1500 | -22,81 | 5,82 |
| 2000 | -28,77 | 4,24 |

Fig. 4: Estimativas de $w_o(n)$ obtidas pelo método cepstral para alguns tamanhos da janela de análise: (a) 0,5 s; (c) 2 s.

Os resultados obtidos pelo controlador são apresentados na Tabela 2. Pode-se notar que a solução definida em (5) é realmente ótima uma vez que o erro é nulo e o MOS-LQO é máximo. Além disso, em comparação ao caso sem controlador, pode-se concluir que a utilização do controlador proposto apresenta uma redução do efeito de oclusão nas duas métricas utilizadas, diminuindo em 0,34 dB a potência do erro e aumentando em 0,12 MOS-LQO a qualidade perceptual.

Além disso, o controlador proposto resulta numa menor potência do erro em comparação com a solução proposta em [3]. Por outro lado, a degradação perceptual gerada pelo controlador é superior a da solução proposta em [3].

TABELA II: Resultados dos critérios objetivos de qualidade para diferentes controladores.

| Solução | Potência do Erro (dB) | MOS-LQO |
|---------------------------|-----------------------|-----------------|
| Solução ótima (5) | $-348,37 \pm 2,11$ | $4,64 \pm 0$ |
| Controlador proposto (16) | $-31,01 \pm 1,60$ | $4,34 \pm 0,63$ |
| Solução de [3] | $-26,08 \pm 1,47$ | $4,55 \pm 0,1$ |
| Sem controle | $-30,67 \pm 3,13$ | $4,26 \pm 0,46$ |

VI. CONCLUSÕES

Este trabalho propôs um método cepstral para identificar a resposta ao impulso do efeito de oclusão e um sistema de controle fixo que utiliza essa estimativa para reduzir o efeito de oclusão em sistemas de auxílio à audição.

Simulações demonstraram que a eficiência do método cepstral depende do tamanho da janela de análise, resultando em estimativas com desalinhamento normalizado médio igual a -28,77 dB para janelas de 2 s. E que o controlador proposto aproxima o sinal a ser ouvido pelo usuário do seu valor ideal,

diminuindo em 0,34 dB a potência do erro e aumentando em 0,12 MOS-LQO a qualidade perceptual do sinal.

Esses resultados iniciais indicam que o método cepstral para estimação da resposta ao impulso do efeito de oclusão pode continuar sendo explorado, seja para produzir um controlador fixo ainda mais eficiente ou para utilizar a sua estimativa como inicialização dos coeficientes de um filtro adaptativo.

AGRADECIMENTOS

Os autores gostariam de agradecer o Grupo de Pesquisa em Engenharia Médica do Conselho Nacional de Desenvolvimento Científico e Tecnológico (GPEM/CNPq), em particular a Prof. María E. Dajer, por fornecer a base de dados de voz.

REFERÊNCIAS

- [1] J. Mejia, H. Dillon, and M. Fisher, "Active cancellation of occlusion: An electronic vent for hearing aids and hearing protectors," *The Journal of the Acoustical Society of America*, vol. 124, no. 1, pp. 235–240, 2008.
- [2] K. Chung, "Challenges and recent developments in hearing aids: Part ii. feedback and occlusion effect reduction strategies, laser shell manufacturing processes, and other signal processing technologies," *Trends in Amplification*, vol. 8, no. 4, pp. 125–164, 2004.
- [3] R. C. Borges and M. H. Costa, "A feed forward adaptive canceller to reduce the occlusion effect in hearing aids," *Computers in biology and medicine*, vol. 79, pp. 266–275, 2016.
- [4] R. C. Borges, M. H. Costa, J. A. Cordioli, and L. F. Assuiti, "An adaptive occlusion canceller for hearing aids," in *21st European Signal Processing Conference (EUSIPCO 2013)*, pp. 1–5, IEEE, 2013.
- [5] M. Sunohara, M. Osawa, T. Hashiura, and M. Tateno, "Occlusion reduction system for hearing aids with an improved transducer and an associated algorithm," in *2015 23rd European Signal Processing Conference (EUSIPCO)*, pp. 285–289, IEEE, 2015.
- [6] B. C. Bispo and D. R. S. Freitas, "On the use of cepstral analysis in acoustic feedback cancellation," *Digital Signal Processing*, vol. 44, pp. 88–101, September 2015.
- [7] B. C. Bispo, P. M. L. Rodrigues, and D. R. S. Freitas, "Acoustic feedback cancellation based on cepstral analysis," in *Proceedings of 17th IEEE Conference on Signal Processing Algorithms, Architectures, Arrangements and Applications*, (Poznan, Poland), pp. 205–209, September 2013.
- [8] ITU-T P.56, "P. 56: Objective measurement of active speech level," 2011.
- [9] G. Fant, "A note on vocal tract size factors and non-uniform f-pattern scalings," *Speech Transmission Laboratory Quarterly Progress and Status Report*, vol. 1, pp. 22–30, 1966.
- [10] ITU-T P.862, "Perceptual evaluation of speech quality (PESQ): objective method for end-to-end speech quality assessment of narrow band telephone networks and speech codecs." International Telecommunications Union, Geneva, Switzerland 2001.
- [11] ITU-T P.862.2, "Wideband extension to recommendation P.862 for the assessment of wideband telephone networks and speech codecs." International Telecommunications Union, Geneva, Switzerland 2005.
- [12] A. A. de Lima, F. P. Freeland, R. A. de Jesus, B. C. Bispo, L. W. P. Biscainho, S. L. Netto, A. Said, T. Kalker, R. W. Schafer, B. Lee, and M. Jam, "On the quality assessment of sound signals," in *Proceedings of the IEEE International Symposium on Circuits and Systems*, (Seattle, USA), pp. 416–419, May 2008.
- [13] A. A. de Lima, F. P. Freeland, P. A. A. Esquef, L. W. P. Biscainho, B. C. Bispo, R. A. de Jesus, S. L. Netto, R. W. Schafer, A. Said, B. Lee, and T. Kalker, "Reverberation assessment in audioband speech signals for telepresence systems," in *Proceedings of 11th International Conference on Signal Processing and Multimedia Applications*, (Porto, Portugal), pp. 257–262, July 2008.
- [14] A. A. de Lima, S. L. Netto, L. W. P. Biscainho, F. P. Freeland, B. C. Bispo, R. A. de Jesus, R. W. Schafer, A. Said, B. Lee, and T. Kalker, "Quality evaluation of reverberation in audioband speech signals," in *e-Business and Telecommunications - Communications in Computer and Information Science* (J. Filipe and M. S. Obaidat, eds.), vol. 48, pp. 384–396, Springer, 2009.
- [15] ITU-T P.862.3, "Application guide for objective quality measurement based on recommendations P.862, P.862.1 and P.862.2." International Telecommunications Union, Geneva, Switzerland 2007.