SBrT 2019 1570558865

Um Estimador de Envoltória Temporal para Redução de Ruído em Implantes Cocleares

Rafael A. Chiea, Márcio H. Costa, Maria M. C. Pinheiro e Julio A. Cordioli

***Resumo*—Esse trabalho apresenta uma nova técnica de redução de ruído para implantes cocleares, baseada na estimação da envoltória temporal da fala. Simulações numéricas utilizando um critério objetivo de inteligibilidade, corroboradas por um experimento psicoacústico com um usuário de implante coclear, indicam um aumento consistente na inteligibilidade da fala em relação a sinais contaminados por ruído aditivo, como também a sinais processados pelo filtro de Wiener. Essas observações são válidas para razões sinal-ruído (SNR) menores que 10 dB e os ganhos de inteligibilidade são maiores à medida que a SNR decresce.**

Palavras-Chave — Redução de ruído, Implantes cocleares, Filtragem no domínio do tempo, Filtro de Wiener.

Abstract—**This work presents a new noise reduction technique for cochlear implants based on speech time-envelope estimates. Numerical simulations using an objective intelligibility criterion, corroborated by a psychoacoustic experiment with a cochlear implant user, indicate a consistent increase in the intelligibility in relation to raw contaminated speech signals as well as to Wiener filter processed signals. These observations are valid for signal to noise ratios (SNR) less than 10 dB. In addition, the intelligibility gain increases with the decrease of the SNR.**

Keywords — Noise reduction, Cochlear implants, Time-domain filtering, Wiener Filter.

I. INTRODUÇÃO

Implantes cocleares (IC) são as mais bem-sucedidas interfaces cérebro-máquina desenvolvidas até os dias de hoje [1]. São próteses auditivas que possibilitam a percepção de sons complexos e elevada inteligibilidade da fala humana por pessoas com perda auditiva severa a profunda ou mesmo com surdez total [2].

Diferentemente dos aparelhos auditivos, que têm como saída um sinal acústico, os implantes cocleares estimulam diretamente o nervo auditivo com corrente elétrica aplicada através de um conjunto de 16 a 24 eletrodos posicionados no interior da cóclea [1] [2].

As estratégias de estimulação da maior parte dos implantes cocleares atuais são baseadas em *vocoders* [1] [3]. Nesse tipo de estratégia, o sinal captado pelo microfone de aquisição é processado por um banco de filtros, gerando distintas saídas, associadas a diferentes bandas de frequência. As envoltórias temporais dos sinais resultantes são calculadas, sendo utilizadas para a modulação de trens de pulsos com taxa constante que acionam fontes de corrente elétrica aplicadas aos diversos eletrodos. Como consequência, a informação de fase, também chamada de informação temporal fina, não é utilizada, acarretando restrições nas informações disponibilizadas ao usuário.

Rafael Attili Chiea, Márcio Holsbach Costa, Maria Madalena Canina Pinheiro e Julio Apolinário Cordioli, Universidade Federal de Santa Catarina, e-mails: rafaelchiea@gmail.com, costa@eel.ufsc.br, madalena.pinheiro@ufsc.br, julio.cordioli@ufsc.br. Este trabalho foi parcialmente financiado por CAPES e CNPq, projetos 304867/2015-2 e 315020/2018-0.

Apesar das limitações na resolução temporal e espectral dos estímulos transmitidos ao usuário do implante, a inteligibilidade da fala se aproxima da encontrada em ouvintes normais. Essa observação pode ser verificada em uma série de testes de audição [4]. Entretanto, na presença de ruído e reverberação, esse desempenho cai significativamente. Em [5], foi verificado que embora em situações de silêncio a inteligibilidade de usuários de implantes cocleares seja de cerca de 80% do total de palavras ouvidas, para uma razão sinal-ruído de 5 dB a inteligibilidade decresce para cerca de 20%. Em recente trabalho, Bergeron e Hotton [6] mostraram que a inteligibilidade média de implantados é de 80%, 60%, 40% e 10%, respectivamente, para as condições de silêncio e SNRs de 10 dB, 5 dB, e 0 dB.

A partir do exposto, a utilização de técnicas de redução de ruído na cadeia de processamento do implante coclear pode trazer incrementos de inteligibilidade para seus usuários em ambientes ruidosos. Uma das principais formas de se realizar essa supressão de ruído é a utilização de máscaras tempo-frequência [7] [8].

Máscaras tempo-frequência são definidas no domínio da frequência, geralmente considerando sinais transformados pela transformada de Fourier de tempo curto (STFT) [9]. Entretanto, a decomposição em bandas de frequência no processamento de implantes cocleares geralmente é realizada utilizando-se bancos de filtros no domínio do tempo [1]. Embora a literatura aponte a preponderância de bancos de filtros do tipo *gammatone* baseados no sistema auditivo humano [10], outros tipos de bancos de filtros também são utilizados [11].

Em [12], foi demonstrado que não há impacto de desempenho significativo na inteligibilidade de indivíduos implantados em relação ao uso de arquiteturas baseadas em bancos de filtros ou STFT. Entretanto, em [13] verificou-se que, em comparação com a STFT, o banco de filtros *gammatone* apresenta um melhor potencial para separação de tons harmônicos complexos. Adicionalmente, bancos de filtros do tipo *gammatone* podem ser implementados com menor atraso que sistemas baseados na STFT, para uma mesma resolução em baixas frequências [13].

Neste trabalho, é proposto um novo método de redução de ruído para implantes cocleares baseado na estimação da envoltória temporal da fala. É utilizada uma estratégia de máscara tempo-frequência aplicada nos sinais provenientes de um banco filtros. Simulações obtidas por um critério objetivo de inteligibilidade e experimentos psicoacústicos com um usuário de implante coclear são apresentados para demonstração de seu desempenho.

Na seção II é apresentado o método proposto para estimação da envoltória temporal. Os métodos utilizados para realização das simulações numéricas e experimentos

psicoacústicos são apresentados na seção III e os resultados e discussão são apresentados na seção IV. A seção V finaliza o trabalho com a apresentação das conclusões.

II. ESTIMADOR DE ENVOLTÓRIA TEMPORAL

Representa-se a fala contaminada por ruído aditivo na forma $y[n] = x[n]+v[n]$, em que $x[n]$ é o sinal de fala desejado, $v[n]$ é o ruído e n é o tempo discreto. Assume-se que ambos $x[n]$ e $v[n]$ são não observáveis individualmente e descorrelacionados entre si. Considera-se que o sinal $y[n]$ é processado por um banco de $k = 1,\ldots,M$ (em que M é o número de canais) filtros, apresentando na saída de cada canal o sinal $y_k[n] = x_k[n]+v_k[n]$ limitado em banda. A envoltória temporal de cada um desses sinais é dada por [14]:

$$Y_k[n] = \left|y_{ak}[n]\right|^{1/2} = \sqrt{y_k^2[n]+\left(H\{y_k[n]\}\right)^2} \quad . \tag{1}$$

em que $H\{\cdot\}$ é o operador da transformada de Hilbert e $y_{ak}[n]$ é o sinal analítico. As envoltórias temporais $Y_k[n]$ são utilizadas para a modulação dos trens de pulsos que acionam as fontes de corrente para estimulação da cóclea.

Nesse trabalho propõe-se a utilização de um filtro temporal de um coeficiente para estimação da envoltória temporal do sinal de fala, minimizando a influência do ruído de contaminação presente em $y_k[n]$. Para tanto, propõe-se minimizar a seguinte função custo:

$$J = E\left\{|\, x_{ak}[n]\,|^2 - |\,\hat{x}_{ak}[n]\,|^2\right\} \ , \tag{2}$$

em que

$$\hat{x}_{ak}[n] = g_k[n]y_{ak}[n] \ ; \tag{3}$$

$g_k[n]$ é o coeficiente do filtro de estimação; e $\hat{x}_{ak}[n]$ é a estimativa do sinal analítico de $x_k[n]$.

Substituindo-se (3) em (2), derivando-se em relação a $g_k^*[n]$ e igualando-se a zero chega-se em

$$g_k[n] = \sqrt{\frac{E\{x_{ak}^4[n]\}+E\{x_{ak}^2[n]\}E\{v_{ak}^2[n]\}}{E\{x_{ak}^4[n]\}+E\{v_{ak}^4[n]\}+4E\{x_{ak}^2[n]\}E\{v_{ak}^2[n]\}}} \ , \tag{4}$$

em que $E\{x^2_{ak}[n]\}$ é a potência do sinal analítico da fala e $E\{v^2_{ak}[n]\}$é a potência do sinal analítico do ruído. Os sinais $x_k[n]$ e $v_k[n]$ são obtidos na saída de um filtro linear. Dessa forma, pode-se assumir que tenham distribuição gaussiana, e, neste caso, a equação (4) pode ser escrita como:

$$g_k[n] = \sqrt{\frac{2\xi_k^2[n]+\xi_k[n]}{2\xi_k^2[n]+4\xi_k[n]+2}} \ , \tag{5}$$

sendo $\xi_k[n] = E\{x^2_k[n]\}/E\{v^2_k[n]\}$ a razão sinal-ruído (SNR) local (no canal índice k); e $E\{x^2_k[n]\}$ e $E\{v^2_k[n]\}$, respectivamente, as variâncias dos sinais de fala e de ruído em cada um dos k canais.

O coeficiente $g_k[n]$ é recalculado a cada quadro temporal, funcionando de maneira análoga a uma máscara tempo-frequência.

III. AVALIAÇÃO DE DESEMPENHO

Para avaliar o filtro proposto, foram usadas como fala todas as 720 frases das listas foneticamente balanceadas do banco de dados IEEE corpus [15]. Como ruído aditivo foi utilizado um ruído de balbuciação, formado artificialmente pela concatenação de quatro sequências do sinal "*International Speech Test Signal*" (ISTS) [16], aleatoriamente deslocadas no tempo. As misturas foram construídas para as seguintes razões sinal-ruído (SNR): −20 dB, −15 dB, −10 dB, −5 dB, 0 dB, 10 dB e 20 dB.

O banco de filtros utilizado é do tipo *gammatone* [10] com 22 canais. O sinal em cada saída do banco de filtros foi tratado separada e independentemente. A cada trecho de 1 ms, foram calculadas as SNRs de cada canal k e os respectivos ganhos $g_k[n]$. Em seguida, o ganho $g_k[n]$ é multiplicado ao respectivo trecho do sinal ruidoso $y_k[n]$.

Nesse experimento, em função da contaminação artificial da fala pelo ruído, tem-se acesso individual aos sinais $x_k[n]$ e $v_k[n]$, o que permite o cálculo exato de $\xi_k[n]$ e, consequentemente, do ganho $g_k[n]$ ideal. Em aplicações reais, $\xi_k[n]$ deve ser estimado, o que pode ser feito, por exemplo, através do método de decisão direta e suas variações [17] [18].

Para fins de comparação, os mesmos sinais foram processados pelo filtro de Wiener (WF) de um coeficiente [19]:

$$w_k[n] = \frac{\xi_k[n]}{\xi_k[n]+1} \ . \tag{6}$$

O WF é a solução que minimiza o erro quadrático médio entre os sinais estimado e desejado. No presente trabalho não é analisado o impacto de erros nas estimativas das SNRs sobre o desempenho dos métodos de redução de ruído abordados.

A Fig. 1 apresenta uma comparação entre as funções de supressão w_k e g_k em função de ξ_k. Pela Fig. 1 pode-se verificar que g_k é mais permissiva, ou seja, aplica ganhos maiores que w_k, sobretudo para valores de ξ_k negativos (em dB).

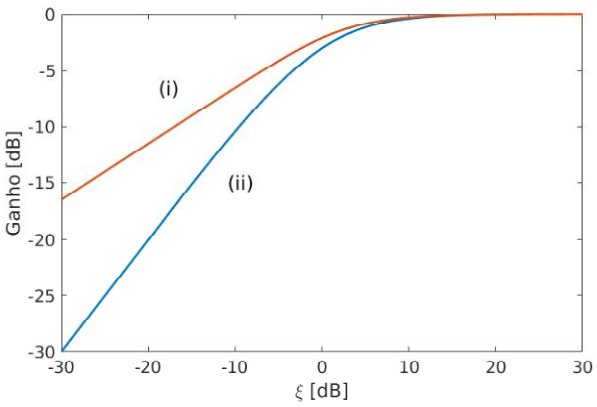

Fig. 1 – Gráficos das funções de supressão do método proposto g_k (i) e do filtro de Wiener w_k (ii), em função da SNR local (ξ_k).

A métrica *Speech-to-reverberation modutation energy ratio for cochlear implants* (SRMR-CI) normalizada [20] foi calculada para os sinais processados pela técnica proposta e pelo filtro de Wiener, de forma a prever o impacto de resultados psicoacústicos de inteligibilidade:

$$\mathrm{SRMR}_{\mathrm{norm}} = \frac{\mathrm{SRMR}_t}{\mathrm{SRMR}_x} \ , \tag{7}$$

em que SRMR_t é a métrica do sinal em teste e SRMR_x é o valor da métrica para o sinal desejado (fala). Essa métrica tem alta correlação com a inteligibilidade de usuários de IC.

A. Experimento Psicoacústico

De forma a verificar o desempenho do método proposto foi realizado um experimento psicoacústico com um indivíduo

implantado. Esse experimento seguiu os procedimentos aprovados pelo Comitê de Ética em Pesquisa com Seres Humanos através do certificado CEPSH-UFSC número 66301917.7.0000.0121.

- *Interface de pesquisa em implantes cocleares*

A interface de pesquisa em implantes cocleares é uma plataforma desenvolvida por pesquisadores da Universidade de Texas em Dallas (UTD) [21]. Essa plataforma permite o envio de estímulos ao implante do paciente, a partir de um computador, *tablet* ou *smartphone*. Desta forma, a cadeia de processamento do implante pode ser completamente definida pelo pesquisador e novas estratégias podem ser testadas.

Um exemplar da plataforma de pesquisa foi cedido pelos seus desenvolvedores ao grupo de pesquisas em implantes cocleares da Universidade Federal de Santa Catarina em 2016, estabelecendo-se um acordo de colaboração. Esta é a primeira unidade disponibilizada a um grupo de pesquisa externo aos Estados Unidos da América.

A interface suporta diversas linguagens de programação, como por exemplo: Matlab, C, C++, JAVA e LabView. A interface se comunica com a unidade processadora através de um cabo USB ou conexão *Wi-Fi*. Na versão disponível, apenas o protocolo de comunicação com implantes da fabricante *Cochlear* está implementado.

- *Cadeia de processamento do implante coclear*

A cadeia de processamento de sinais utilizada nos experimentos psicoacústicos realizados nesse trabalho é baseada na estratégia *Automatic combination encoder* (ACE) [1] [21]. O diagrama de blocos que ilustra esse processamento é apresentado na Fig. 2. A frequência de amostragem utilizada é 16 kHz.

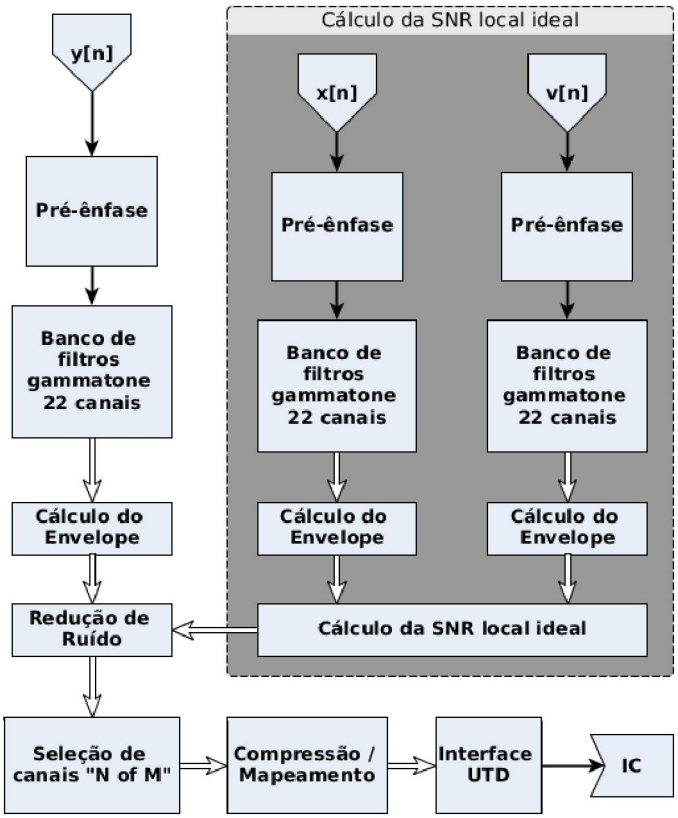

Fig. 2 – Diagrama de blocos da cadeia de processamento de sinais para os experimentos psicoacústicos. As setas negras indicam comunicação monocanal, enquanto as brancas largas indicam comunicação paralela multicanal.

Inicialmente, o sinal contaminado $y[n]$ passa por um filtro de ênfase e em seguida o sinal é separado em sub-bandas através de um filtro *gammatone* de 22 canais. Os sinais na saída do banco de filtros têm suas envoltórias temporais calculadas de acordo com a equação (1). A envoltória temporal é então dividida em blocos de tamanho e sobreposição determinados de acordo com o mapa de compensação do usuário. O processamento em cada canal ocorre de maneira independente.

Nesse trabalho a razão sinal-ruído (ξ_k) é calculada a partir dos sinais $x_k[n]$ e $v_k[n]$, previamente (*off-line*) à aplicação do sinal, sendo definida como uma constante ao longo de toda a frase. Os ganhos das máscaras são então calculados e multiplicados aos sinais contaminados que serão enviados para a estimulação da cóclea. A estimulação segue o método *N-of-M*, no qual, a cada quadro temporal, é realizada a seleção dos N canais (em que N é uma constante definida durante o processo de *fitting*) de maior magnitude no universo de $M = 22$ canais (eletrodos). Finalmente os sinais sofrem um processo de compressão e mapeamento em amplitude de corrente elétrica, de acordo com o mapa de parâmetros do voluntário. A sequência de pulsos de corrente resultante desse processamento é transmitida ao implante do paciente através da interface de pesquisa.

TABELA I – *PARÂMETROS DE CONFIGURAÇÃO DO IMPLANTE COCLEAR DO VOLUNTÁRIO.*

Canal	Freq. Inf. [Hz]	Freq. Sup. [Hz]	T-Level	C-Level
22	188	313	102	185
21	313	438	101	185
20	438	563	101	185
19	563	688	101	184
18	688	813	100	184
17	813	938	100	183
16	938	1063	100	183
15	1063	1188	101	183
14	1188	1313	102	183
13	1313	1563	103	184
12	1563	1813	104	184
11	1813	2063	106	185
10	2063	2313	102	182
9	2313	2688	98	180
8	2688	3063	95	177
7	3063	3563	91	175
6	3563	4063	88	173
5	4063	4688	89	171
4	4688	5313	90	170
3	5313	6063	91	168
2	6063	6938	92	170
1	6938	7938	93	172

- *Descrição do Experimento*

O participante desse experimento foi um indivíduo do sexo feminino com 32 anos de idade e usuário de implante coclear por 4 anos. Avaliação preliminar indica que o voluntário apresenta elevada capacidade de compreensão da fala e capacidade para desenvolver conversações em diferentes cenários acústicos. O aparelho utilizado é da marca *Cochlear*®, modelo CI422. A etiologia da surdez é desconhecida.

O conjunto de configurações individuais (Tabela 1) do implante do participante foi obtido por fonoaudiólogo capacitado e carregado na interface de pesquisa.

No experimento foi utilizado um corpus de fala em português do Brasil [22], consistindo em um conjunto de 20 listas foneticamente balanceadas, sendo 10 gravadas com voz masculina e outras 10 gravadas com voz feminina. Cada lista contém 10 frases de cerca de 3s de duração. O mesmo sinal de balbuciação descrito na seção para o cálculo do critério $SRMR_{norm}$ foi utilizado como ruído aditivo, assumindo SNRs de −25 dB a −5 dB com passos de −5 dB.

Sinais contaminados nas SNRs descritas foram processados tanto pelo método proposto quanto pelo WF. Para cada uma dessas técnicas de redução de ruído e valor de SNR global, duas listas, uma com voz masculina e outra com voz feminina, foram selecionadas aleatoriamente e sem reposição. Assim, no experimento completo foi utilizado um total de 200 frases (5 SNRs × 2 listas × 10 frases × 2 métodos de redução de ruído).

Antecipadamente ao início do experimento, o participante ouviu uma série de frases aleatórias, de forma a ajustar o volume de maneira confortável. Essas frases eram diferentes das utilizadas no experimento e o volume ajustado não foi alterado ao longo de todo o processo.

Uma interface gráfica (GUI) foi desenvolvida para a condução do experimento (Fig. 3). A cada iteração, uma frase entre as 200 disponíveis foi selecionada aleatoriamente, sem reposição. O participante pôde ouvir cada frase apenas uma vez, e foi instruído a repetir em voz alta as palavras que havia compreendido. As palavras corretas foram registradas pelo pesquisador através da interface gráfica. O processo foi repetido até que todas as frases fossem esgotadas. Durante a apresentação das frases, o microfone conectado à interface permaneceu desligado e, assim, nenhum som ambiente pôde ser ouvido pelo participante. O botão "*Push to talk*" na GUI acionava o microfone para a comunicação entre o pesquisador e o participante quando necessário.

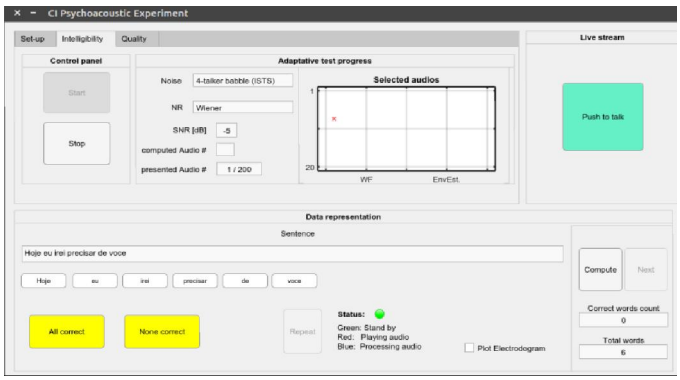

Fig. 3 – Interface gráfica para realização dos experimentos psicoacústicos.

IV. RESULTADOS E DISCUSSÃO

A Fig. 4 apresenta o resultado das simulações numéricas associadas ao critério de inteligibilidade $SRMR_{norm}$, indicando que, em relação ao sinal não processado, tanto o filtro de Wiener como o método proposto aumentam a inteligibilidade de fala, principalmente para valores de SNR menores que 10 dB. As médias da $SRMR_{norm}$ para o método proposto são maiores que as obtidas com o WF. Em todos os experimentos, quanto menor a SNR maiores são as diferenças de desempenho entre o os dois métodos de redução de ruído analisados.

Na Tabela II, são encontrados os percentuais médios de palavras corretas obtidos no experimento psicoacústico com o voluntário, em função da SNR, considerando-se conjuntamente sinais de locutores masculinos e femininos. O maior valor para cada SNR é apresentado em negrito. Em todas as situações o método proposto apresentou maior inteligibilidade que o método de Wiener. Ainda, na Tabela II, verifica-se que as diferenças de inteligibilidade a favor do método proposto são significativamente maiores para valores de SNR iguais ou menores que −10 dB.

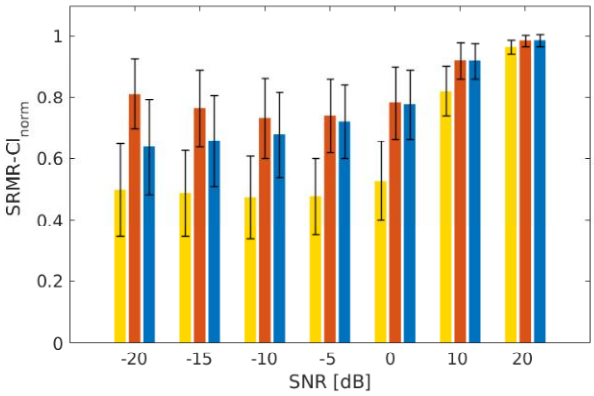

Fig. 4 – $SRMR_{norm}$ do sinal não processado (amarelo, à esquerda) e resultante do processamento com o método proposto (laranja, ao centro) e filtro de Wiener (azul, à direita). As barras indicam a média e os bigodes, a variância.

TABELA II – *VALORES PERCENTUAIS DE PALAVRAS CORRETAS OBTIDAS COM O MÉTODO PROPOSTO E O FILTRO DE WIENER NO DOMÍNIO TEMPO PARA CADA VALOR DE SNR.*

Método	SNR [dB]				
	−25	−20	−15	−10	−5
Proposto	**82,0%**	**94,2%**	**93,5%**	**99,28%**	**97,8%**
WF	13,1%	59,7%	46,9%	73,2%	95,8%

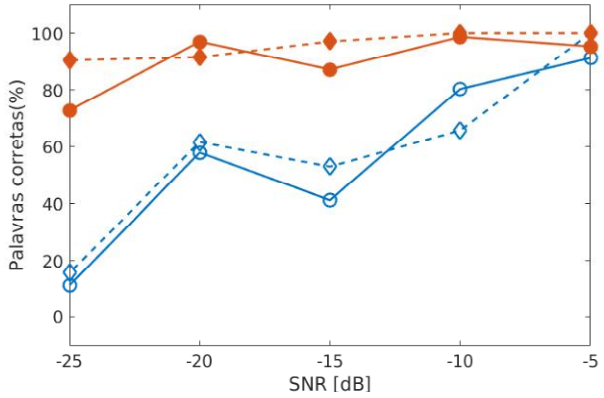

Fig. 5 – Percentuais de palavras corretas para fala masculina (losangos, linha pontilhada) e feminina (círculos, linha contínua) em função da SNR. Método proposto em laranja com marcadores sólidos e filtro de Wiener em azul com marcadores vazados.

Na Fig. 5 são apresentados separadamente os desempenhos obtidos para as falas femininas e masculinas em ambos os métodos. Os resultados concordam com relato do voluntário, de que não houve diferença perceptual com relação ao sexo do locutor.

Os resultados obtidos através de simulações numéricas com o critério objetivo de inteligibilidade $SRMR_{norm}$ foram corroborados pelo experimento psicoacústico, indicando melhor desempenho, principalmente em pequenas SNRs, do método proposto em relação ao filtro de Wiener. Entretanto, na literatura, há indicação de grande variabilidade na percepção de diferentes usuários de IC [8]. Dessa forma, apesar dos

resultados consistentes, um maior número de voluntários deve ser avaliado para que se possam produzir evidências mais conclusivas sobre o desempenho do método proposto.

V. CONCLUSÕES

Esse trabalho apresentou a proposta de uma nova técnica de redução de ruído para implantes cocleares baseada na estimação da envoltória temporal da fala. Seu potencial em aplicações práticas foi avaliado em caráter preliminar através de simulações numéricas utilizando um critério objetivo de inteligibilidade, como também através de um experimento psicoacústico com um usuário de implante coclear. Os resultados obtidos concordam em apontar um aumento da inteligibilidade em relação a sinais contaminados e processados pelo filtro de Wiener. Para o caso de uma SNR de −25 dB, o método proposto resultou em 82% de inteligibilidade, apresentando 68,9% a mais que o obtido com o filtro de Wiener.

AGRADECIMENTOS

Os autores gostariam de agradecer à CAPES, CNPq, Wavetech-ST e aos pesquisadores Hussnain Ali e John Hansen da UTD.

REFERÊNCIAS

[1] J. Wouters, H. J. McDermott, T. Francart, “Sound coding in cochlear implants”, *IEEE Signal Processing Magazine*, v. 32, n. February, pp. 67-80, 2015.

[2] J. Wouters et al., “Sound processing for better coding of monaural and binaural cues in auditory prostheses”, *Proceeding of the IEEE*, v. 101, n. 9, pp. 1986-1997, 2013.

[3] P. C. Loizou, “Speech processing in vocoder-centric cochlear implants” *Cochlear and Brainstem Implants*, v. 64, pp. 109-143, 2006.

[4] B. S. Wilson, M. F. Dorman, “The surprising performance of present-day cochlear implants”, *IEEE Transactions on Biomedical Engineering,* v. 54, n. 6, pp. 969-972, 2007.

[5] A. Hast et al., “Speech perception of elderly cochlear implant users under different noise conditions”, *Otology and Neurology*, v. 36, n. 10, pp. 1638-1643, 2015.

[6] F. Bergeron, M. Hotton, “Perception in noise with the Digisonic SP cochlear implant: Clinical trial of Saphyr processor's upgraded signal processing”, *European Annals of Otorhinolaryngology, Head and Neck Diseases*, pp. 3-5, 2016.

[7] R. Koning, N. Madhu, J. Wouters, “Ideal time-frequency masking algorithms lead to different speech intelligibility and quality in normal-hearing and cochlear implant listeners”, *IEEE Transactions on Biomedical Engineering,* v. 62, n. 1, pp. 331-341, 2015.

[8] S. J. Mauger, P. W. Dawson, A. A. Hersbach, “Perceptually optimized gain function for cochlear implant signal-to-noise ratio based noise reduction”, *The Journal of the Acoustical Society of America*, v. 131, n. 1, pp. 327-336, 2012.

[9] C. Loizou, *Speech enhancement - Theory and practice* 2nd. CRC Press, 2013.

[10] P. I. M. Johannesma, “The pre-response stimulus ensemble of neurons in the cochlear nucleus”, *Symposium on Hearing Theory*, pp. 58-69, 1972.

[11] S. Cosentino et al., “Cochlear implant filterbank design and optimization: A simulation study”, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, v. 22, n. 2, pp. 347-353, 2014.

[12] M. A. M. Jong et al., “Take-home trial comparing fast Fourier transformation-based and filter bank-based cochlear implant speech coding strategies”, *BioMed Research International*, pp. 1-7, 2017.

[13] S. Tabibi et al., “Investigating the use of a gammatone filterbank for a cochlear implant coding strategy”, *Journal of Neuroscience Methods*, v. 277, pp. 63-74, 2016.

[14] J. Justice, “Analytic signal processing in music computation”, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, v. 27, n. 6, pp. 670-684, 1979.

[15] IEEE recommended practice for speech quality measurements, *IEEE Transactions on Audio and Electroacoustics*, v.17 n. 3, pp. 225-246, 1969.

[16] I. Holube et al., “Development and analysis of an international speech test signal (ISTS)”, *International Journal of Audiology*, v. 49, n. 12, pp. 891-903, 2010.

[17] Y. Hu et al., “Use of a sigmoidal-shaped function for noise attenuation in cochlear implants”, *The Journal of the Acoustical Society of America*, v. 122, n. 4, pp. 128-134, 2007.

[18] Y. Ephraim, D. Malah, “Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator”, *IEEE Transactions on Acoustics, Speech and Signal Processing*, n. 6, pp. 1109-1121, 1984.

[19] J. Chen et al., “New insights into the noise reduction Wiener filter”, *IEEE Transactions on Audio, Speech, and Language Processing*, v. 14, n. 4, pp. 1218-1234, 2006.

[20] J. F. Santos, T. H. Falk, “Updating the SRMR-CI metric for improved intelligibility prediction for cochlear implant users”, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, v. 22, n. 12, pp. 2197-2206, 2014.

[21] H. Ali et al., “Design and evaluation of a personal digital assistant-based research platform for cochlear implants”, *IEEE Transactions on Biomedical Engineering*, v. 60, n. 11, pp. 3060-3073, 2013.

[22] A. Alcaim, J.A. Solewicz, J.A. Moraes, “Freqüência de ocorrência dos fones e listas de frases foneticamente balanceadas no português falado no Rio de Janeiro”, *Revista da Sociedade Brasileira de Telecomunicações*, v. 7, n. 1, pp. 23-41, 2015.