Proposta de uma função custo para a preservação espacial de fontes sonoras em aparelhos auditivos biauriculares

Diego M. do Carmo, Márcio H. Costa e Ricardo Borsoi

***Resumo*—Aparelhos auditivos biauriculares possibilitam não apenas uma redução mais efetiva de ruído aditivo em comparação ao uso bilateral, como também permitem a preservação das características espaciais originais do ambiente acústico. Neste trabalho, é realizada uma análise sobre as três principais técnicas de preservação de pistas biauriculares para aparelhos auditivos. A partir das observações realizadas, é apresentada uma proposta para a preservação das pistas biauriculares, tanto para ruído proveniente de uma fonte sonora pontual quanto para aquele decorrente de um campo acústico difuso. Simulações computacionais indicam a eficácia do método proposto, possibilitando um adequado compromisso entre redução de ruído e preservação do cenário acústico original.**

Palavras-Chave— Aparelhos auditivos biauriculares, MWF, ITF, IC.

***Abstract*—Binaural hearing aids allow not only a more effective noise reduction, in comparison with the bilateral use, but also preservation of the original spatial characteristics of the acoustic scenario. In this work, an analysis about the three main techniques for binaural cue preservation in hearing aids is provided. From the obtained insight, a penalty function was proposed for binaural cue preservation of either diffuse sound field or punctual noise source. Computational simulations indicate the efficacy of the proposed method in providing an adequate trade-off between noise reduction and spatial preservation of the acoustic scenario.**

Keywords— Binaural hearing aids, MWF, ITF, IC.

I. INTRODUÇÃO

Aparelhos auditivos são dispositivos utilizados para compensar limitações na capacidade auditiva. Nos casos em que há perda em ambas as orelhas, os aparelhos biauriculares com canal de transmissão de dados são a opção de auxílio mais avançada disponível atualmente, pois permitem o compartilhamento de sinais e parâmetros entre ambos os dispositivos [1].

Métodos de redução de ruído são parte fundamental de aparelhos auditivos modernos, uma vez que a presença de ruído pode diminuir significativamente o conforto acústico e a inteligibilidade da fala. Técnicas biauriculares permitem não apenas uma maior eficiência na redução de ruído em relação às suas contrapartidas mono-auriculares, como também permitem preservar a impressão espacial original do cenário acústico no qual o usuário se insere [2]. Isso é realizado preservando as pistas biauriculares dos sinais acústicos captados.

Diego M. do Carmo, Márcio H. Costa e Ricardo Borsoi, Universidade Federal de Santa Catarina. E-mails: diego.carmoh@gmail.com, raborsoi@gmail.com, costa@eel.ufsc.br. Este trabalho foi parcialmente financiado por CAPES e CNPq, projeto 315020/2018-0.

As pistas biauriculares são definidas a partir de diferenças entre os sinais sonoros em ambas as orelhas e fornecem informação sobre as características da fonte e do ambiente acústico. A localização de fontes acústicas pontuais no plano horizontal é realizada pelo sistema auditivo humano a partir da diferença de tempo interaural (ITD – *interaural time difference*), relevante principalmente nas baixas frequências (<1,5 kHz), e da diferença de nível interaural (ILD – *interaural level difference*), relevante nas altas frequências (>1,5 kHz) [3]. No caso de um campo acústico difuso, em que não há uma localização espacial definida para a fonte sonora, a correta impressão auditiva do ambiente advém da coerência interaural (IC – *interaural coherence*), a qual mede a correlação espacial entre os sinais em ambas as orelhas [4]. A IC também está relacionada à extensão percebida da fonte sonora: quanto maior a magnitude da IC, menor é essa extensão e vice-versa [4]. Sinais sonoros que proporcionam magnitude máxima da IC são identificados como provenientes de uma fonte pontual, enquanto que sinais com magnitude mínima proporcionam uma impressão auditiva de um campo acústico difuso.

O filtro multicanal de Wiener (MWF – *multichannel Wiener filter*) é uma das abordagens mais estudadas para a redução biauricular de ruído. Ele reduz significativamente a potência dos sinais indesejados, ao mesmo tempo que limita a distorção e preserva as pistas biauriculares da fala. No entanto, o MWF altera significativamente a ILD e a ITD de sinais provenientes de uma fonte interferente pontual [5], assim como a IC, quando o campo sonoro é difuso [6].

Para evitar a distorção da percepção espacial do ambiente acústico, foram propostas modificações no MWF a partir da inclusão de funções custo auxiliares. Nessa composição, a função custo do MWF mantém seu objetivo original de reduzir ruído e limitar a distorção da fala, enquanto que as funções custo adicionais são introduzidas para penalizar distorções nas pistas biauriculares do sinal interferente. Dessa forma, tenta-se preservar a impressão auditiva original do ambiente após o processo de filtragem.

Apesar da proposição de diversas funções custo para preservação espacial, até o momento não há um método eficaz para preservação conjunta da ITD, ILD e IC. Dessa forma, há a necessidade de identificação prévia das características da fonte sonora e/ou do campo acústico para a escolha da função custo a ser utilizada.

Neste trabalho, é realizada uma investigação sobre as propriedades das três funções custo de preservação espacial mais conhecidas. A partir da análise realizada, é proposta uma função custo para penalização de erros em pistas biauriculares, com o objetivo de preservar, simultaneamente, ITD, ILD e IC.

Dessa forma, a função proposta pode ser usada, indistintamente, tanto para ruído decorrente de um campo acústico difuso quanto aquele proveniente de uma fonte acústica pontual, sem a necessidade da utilização de um classificador de ruído (difuso/pontual).

Ao longo desse texto, letras maiúsculas e minúsculas em negrito representam, respectivamente, matrizes e vetores. Variáveis escalares são apresentadas em itálico, enquanto que constantes escalares são apresentadas sem formatação.

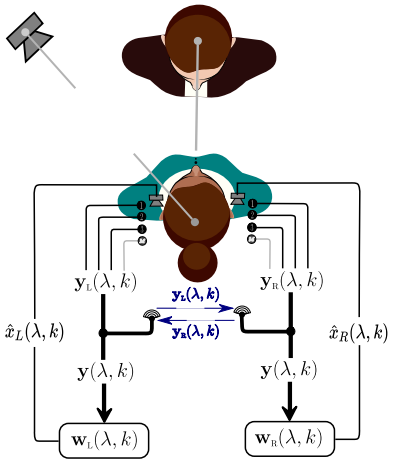

Fig. 1. Diagrama de um sistema binaural *full duplex*.

II. Aparelho Auditivo Biauricular

A Fig. 1 ilustra os estágios de aquisição, transmissão e processamento de sinais de um sistema biauricular de auxílio à audição. O sistema é composto por dois aparelhos auditivos, um na orelha esquerda (L) e outro na orelha direita (R), em que cada um contém um arranjo com M microfones, totalizando $2M$ microfones no sistema completo. A representação tempo-frequência dos sinais contaminados captados pelos microfones, na k-ésima frequência do λ-ésimo conjunto de amostras, definido como $\mathbf{y}(k,\lambda) \in \mathbb{C}^{2M}$, é dada por:

$$\mathbf{y}(k,\lambda) = \mathbf{x}(k,\lambda) + \mathbf{v}(k,\lambda) \ , \tag{1}$$

em que $\mathbf{x}(k,\lambda)$ denota o sinal da fonte de interesse e $\mathbf{v}(k,\lambda)$ o sinal interferente. Considera-se que o vetor em (1) é conhecido em ambos os aparelhos auditivos através de um canal de comunicação *full duplex* e satisfaz $\mathbf{y}(k,\lambda) = [\ \mathbf{y}_L^T(k,\lambda)\ \mathbf{y}_R^T(k,\lambda)\ ]^T$, em que $(\cdot)^T$ é a operação de transposição e $\mathbf{y}_l \in \mathbb{C}^M$ é um vetor que contém os sinais recebidos nos M microfones do lado $l \in \{\ L, R\ \}$. Em geral, os microfones frontais de ambos os lados são definidos como microfones de referência (*REF*). O sinal contaminado nesses microfones é dado por:

$$y_{REF_l}(k,\lambda) = x_{REF_l}(k,\lambda) + v_{REF_l}(k,\lambda) = \mathbf{q}_l^T \mathbf{y}(k,\lambda) \ , \tag{2}$$

em que $\mathbf{q}_L = [\ 1\ 0\ \ldots\ 0\ 0\ 0\ \ldots\ 0\ ]^T$ e $\mathbf{q}_R = [\ 0\ 0\ \ldots\ 0\ 1\ 0\ \ldots\ 0\ ]^T$.

Considerando o sistema biauricular descrito, o problema de redução de ruído consiste em estimar, para cada frequência, a componente de fala recebida nos microfones de referência utilizando as amostras observadas pelos $2M$ microfones. As componentes de fala estimadas, no domínio da frequência, enviadas aos alto-falantes dos aparelhos auditivos, são definidas por:

$$\hat{x}_L(k,\lambda) = \mathbf{w}_L^H(k)\mathbf{y}(k,\lambda), \ \hat{x}_R(k,\lambda) = \mathbf{w}_R^H(k)\mathbf{y}(k,\lambda), \tag{3}$$

em que $(\cdot)^H$ representa o transposto hermitiano e $\mathbf{w}_L(k)$ e $\mathbf{w}_R(k) \in \mathbb{C}^{2M}$ representam o par de filtros de redução de ruído na esquerda e na direita, respectivamente.

As matrizes de correlação do sinal contaminado, fala e ruído são definidas, respectivamente, por: $\mathbf{\Phi}_y(k,\lambda) = E\{\mathbf{y}(k,\lambda)\mathbf{y}^H(k,\lambda)\}$, $\mathbf{\Phi}_x(k,\lambda) = E\{\mathbf{x}(k,\lambda)\mathbf{x}^H(k,\lambda)\}$ e $\mathbf{\Phi}_v(k,\lambda) = E\{\mathbf{v}(k,\lambda)\mathbf{v}^H(k,\lambda)\}$, cujo operador $E\{\cdot\}$ representa a esperança estatística em relação à λ. Assumindo que fala e ruído são descorrelacionados, tem-se que: $\mathbf{\Phi}_y(k,\lambda) = \mathbf{\Phi}_x(k,\lambda) + \mathbf{\Phi}_v(k,\lambda)$. Por simplicidade de notação, os índices k e λ serão omitidos no restante deste artigo.

III. Redução de Ruído e Preservação Espacial

No MWF, os coeficientes $\mathbf{w}_L$ e $\mathbf{w}_R$ são obtidos a partir da minimização da seguinte função custo:

$$J_{\text{MWF}} = E\{|\, x_L - \mathbf{w}_L^H \mathbf{y}\,|^2\} + E\{|\, x_R - \mathbf{w}_R^H \mathbf{y}\,|^2\}\,, \tag{4}$$

em que $|\cdot|$ é a operação de módulo.

A solução que minimiza (4) resulta naturalmente na preservação da ILD, ITD e IC da fala de interesse [5]. Contudo, os coeficientes obtidos distorcem as pistas biauriculares do ruído interferente fazendo com que seu azimute percebido seja o mesmo associado ao sinal da fonte de fala [5] [9].

Com o intuito de solucionar esse problema, funções custo que penalizam distorções nas pistas biauriculares têm sido adicionadas a (4), resultando na seguinte função custo genérica:

$$J_T = J_{\text{MWF}} + \alpha J_{\text{BPC}}^{\text{v}} \ , \tag{5}$$

em que a $J^{\text{v}}_{\text{BPC}}$ é uma função custo relacionada à preservação das pistas biauriculares do sinal interferente; e o fator α pondera a importância de $J^{\text{v}}_{\text{BPC}}$ em relação à J_{MWF} [6]. As três principais funções de preservação utilizadas na literatura, para cenários acústicos constituídos por uma única fonte pontual de interesse cujo sinal é contaminado por ruído, são apresentadas a seguir.

A. *Função custo associada à ILD-ITD*

Em [7], foi proposta a primeira função custo para preservação espacial das pistas biauriculares de uma fonte pontual interferente. A ideia principal é penalizar distorções na ILD e na ITD do ruído interferente, a partir da minimização da seguinte soma de funções custo:

$$J_{BPC}^{\text{v}} = J_{\text{ILD-ITD}}^{\text{v}} = J_{\text{ILD}}^{\text{v}} + \beta J_{\text{ITD}}^{\text{v}} \ , \tag{6}$$

em que

$$J_{\text{ILD}}^{\text{v}} = \left| \overline{ILD}_{\text{ou}}^{\text{v}} - \overline{ILD}_{\text{in}}^{\text{v}} \right|^2 \ , \tag{7}$$

$$\overline{ILD}_{\text{ou}}^{\text{v}} = \frac{\mathbf{w}_L^H \mathbf{\Phi}_v \mathbf{w}_L}{\mathbf{w}_R^H \mathbf{\Phi}_v \mathbf{w}_R}, \ \overline{ILD}_{\text{in}}^{\text{v}} = \frac{\mathbf{q}_L^T \mathbf{\Phi}_v \mathbf{q}_L}{\mathbf{q}_R^T \mathbf{\Phi}_v \mathbf{q}_R} \ , \tag{8}$$

e

$$J_{\text{ITD}}^{\text{v}} = 1 - \cos(\overline{ITD}_{\text{ou}}^{\text{v}} - \overline{ITD}_{\text{in}}^{\text{v}}) \ , \tag{9}$$

$$\overline{ITD}_{\text{ou}}^{\text{v}} \cong \angle \mathbf{w}_L^H \mathbf{\Phi}_v \mathbf{w}_R, \quad \overline{ITD}_{\text{in}}^{\text{v}} = \angle \mathbf{q}_L^T \mathbf{\Phi}_v \mathbf{q}_R \ , \tag{10}$$

em que $\angle\{\cdot\}$ é a fase do argumento e β é um coeficiente de ponderação.

B. *Função custo associada à ITF*

Considerando ainda o caso de uma fonte interferente pontual, em [8] foi proposta uma função custo baseada na função de transferência interaural (ITF – *interaural transfer*

function), que contém informação intrínseca sobre a ITD e a ILD [9]. Esse enfoque faz uso de apenas uma função custo, baseada no erro quadrático médio entre a ITF média do ruído de entrada ($\overline{ITF}^{\mathrm{v}}_{\mathrm{in}}$) e a ITF instantânea do ruído processado ($ITF^{\mathrm{v}}_{\mathrm{ou}}$), dada por:

$$J^{\mathrm{v}}_{\mathrm{BPC}} = J^{\mathrm{v}}_{\mathrm{ITF}} = E\left\{\left|ITF^{\mathrm{v}}_{\mathrm{ou}} - \overline{ITF}^{\mathrm{v}}_{\mathrm{in}}\right|^2\right\} , \tag{11}$$

em que:

$$\overline{ITF}^{\mathrm{v}}_{\mathrm{in}} \cong \mathbf{q}_{\mathrm{L}}^{\mathrm{T}}\mathbf{\Phi}_{\mathrm{v}}\mathbf{q}_{\mathrm{R}} / \mathbf{q}_{\mathrm{R}}^{\mathrm{T}}\mathbf{\Phi}_{\mathrm{v}}\mathbf{q}_{\mathrm{R}} \quad \text{e} \quad ITF^{\mathrm{v}}_{\mathrm{ou}} = \mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\mathbf{v} / \mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\mathbf{v} . \tag{12}$$

Substituindo (12) em (11) e aproximando-se o valor esperado da razão pela razão dos valores esperados, obtém-se [8]:

$$J^{\mathrm{v}}_{\mathrm{ITF}} = E\{|\,\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\mathbf{v} - \overline{ITF}^{\mathrm{v}}_{\mathrm{in}}\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\mathbf{v}\,|^2\} / E\{|\,\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\mathbf{v}\,|^2\} . \tag{13}$$

Até o momento a $J^{\mathrm{v}}_{\mathrm{ITF}}$ é a função custo mais efetiva para preservação das pistas biauriculares de uma fonte pontual.

C. *Função custo associada à IC*

Em [6], foi proposta uma função custo para a preservação espacial da impressão auditiva de um campo acústico difuso, definida por:

$$J^{\mathrm{v}}_{\mathrm{IC}} = \left|\overline{IC}^{\mathrm{v}}_{\mathrm{ou}} - \overline{IC}^{\mathrm{v}}_{\mathrm{in}}\right|^2 , \tag{14}$$

em que $\overline{IC}^{\mathrm{v}}_{\mathrm{ou}}$ e $\overline{IC}^{\mathrm{v}}_{\mathrm{in}}$ são definidos, respectivamente, por

$$\overline{IC}^{\mathrm{v}}_{\mathrm{ou}} \triangleq (\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\mathbf{\Phi}_{\mathrm{v}}\mathbf{w}_{\mathrm{R}}) / (\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\mathbf{\Phi}_{\mathrm{v}}\mathbf{w}_{\mathrm{L}}\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\mathbf{\Phi}_{\mathrm{v}}\mathbf{w}_{\mathrm{R}})^{-1/2}, \tag{15}$$

$$\overline{IC}^{\mathrm{v}}_{\mathrm{in}} \triangleq (\mathbf{q}_{\mathrm{L}}^{\mathrm{T}}\mathbf{\Phi}_{\mathrm{v}}\mathbf{q}_{\mathrm{R}}) / (\mathbf{q}_{\mathrm{L}}^{\mathrm{T}}\mathbf{\Phi}_{\mathrm{v}}\mathbf{q}_{\mathrm{L}}\mathbf{q}_{\mathrm{R}}^{\mathrm{T}}\mathbf{\Phi}_{\mathrm{v}}\mathbf{q}_{\mathrm{R}})^{-1/2} . \tag{16}$$

As definições em (15) e (16) podem ser representadas na forma polar como:

$$\overline{IC}^{\mathrm{v}}_{\mathrm{in}} = |\,\overline{IC}^{\mathrm{v}}_{\mathrm{in}}\,|\, \mathrm{e}^{\mathrm{j}\overline{IPD}^{\mathrm{v}}_{\mathrm{in}}}, \qquad \overline{IC}^{\mathrm{v}}_{\mathrm{ou}} = |\,\overline{IC}^{\mathrm{v}}_{\mathrm{ou}}\,|\, \mathrm{e}^{\mathrm{j}\overline{IPD}^{\mathrm{v}}_{\mathrm{ou}}}, \tag{17}$$

em que $\overline{IPD}^{\mathrm{v}}_{\mathrm{in}}$ é a diferença interaural de fase média (IPD – *interaural phase difference*) do ruído de entrada e $\overline{IPD}^{\mathrm{v}}_{\mathrm{ou}}$ é a IPD média do ruído processado, definidas como:

$$\begin{aligned} \overline{IPD}^{\mathrm{v}}_{\mathrm{in}} &= -(2\pi k f_s / N)\ \overline{ITD}^{\mathrm{v}}_{\mathrm{in}}, \\ \overline{IPD}^{\mathrm{v}}_{\mathrm{ou}} &= -(2\pi k f_s / N)\ \overline{ITD}^{\mathrm{v}}_{\mathrm{ou}}, \end{aligned} \tag{18}$$

sendo f_s a frequência de amostragem e N o total de frequências discretas indexadas por k.

Em [10], foi demonstrado que a função custo em (14) também pode ser usada para preservar a ITD de fontes pontuais.

IV. Análise das Funções de Preservação Espacial

Nesta seção é realizada uma análise das características e propriedades das três funções custo apresentadas na Seção III, de forma a subsidiar a proposta apresentada na Seção V.

Utilizando-se as definições da ITF, ILD e IC, pode-se chegar em:

$$\overline{ITF}^{\mathrm{v}}_{\mathrm{in}} = (\overline{ILD}^{\mathrm{v}}_{\mathrm{in}} \cdot |\,\overline{IC}^{\mathrm{v}}_{\mathrm{in}}\,|^2)^{1/2}\ \mathrm{e}^{\mathrm{j}\overline{IPD}^{\mathrm{v}}_{\mathrm{in}}} \tag{19}$$

$$\overline{ITF}^{\mathrm{v}}_{\mathrm{ou}} = (\overline{ILD}^{\mathrm{v}}_{\mathrm{ou}} \cdot |\,\overline{IC}^{\mathrm{v}}_{\mathrm{ou}}\,|^2)^{1/2}\ \mathrm{e}^{\mathrm{j}\overline{IPD}^{\mathrm{v}}_{\mathrm{ou}}} . \tag{20}$$

Considerando (19) e (20), três são as condições necessárias para a preservação das pistas biauriculares. Para uma fonte pontual no campo distante que possui $|\overline{IC}^{\mathrm{v}}_{\mathrm{in}}|=1$, as pistas biauriculares do sinal interferente processado devem atender às seguintes condições:

$$\mathrm{A}_1\!: \overline{ILD}^{\mathrm{v}}_{\mathrm{ou}} = \overline{ILD}^{\mathrm{v}}_{\mathrm{in}};\ \mathrm{A}_2\!: \overline{IPD}^{\mathrm{v}}_{\mathrm{ou}} = \overline{IPD}^{\mathrm{v}}_{\mathrm{in}};\ \mathrm{A}_3\!: \left|\overline{IC}^{\mathrm{v}}_{\mathrm{ou}}\right| = 1 . \tag{21}$$

No caso de um campo acústico difuso, a preservação da impressão auditiva exige que:

$$\mathrm{B}_1\!: \overline{ILD}^{\mathrm{v}}_{\mathrm{ou}} = 1;\ \mathrm{B}_2\!: \overline{IPD}^{\mathrm{v}}_{\mathrm{ou}} = \overline{IPD}^{\mathrm{v}}_{\mathrm{in}};\ \mathrm{B}_3\!: \left|\overline{IC}^{\mathrm{v}}_{\mathrm{ou}}\right| = \left|\overline{IC}^{\mathrm{v}}_{\mathrm{in}}\right| . \tag{22}$$

De maneira genérica, a preservação perfeita das pistas biauriculares é dada pelas seguintes condições:

$$\mathrm{C}_1\!: \overline{ILD}^{\mathrm{v}}_{\mathrm{ou}} = \overline{ILD}^{\mathrm{v}}_{\mathrm{in}};\ \mathrm{C}_2\!: \overline{IPD}^{\mathrm{v}}_{\mathrm{ou}} = \overline{IPD}^{\mathrm{v}}_{\mathrm{in}};\ \mathrm{C}_3\!: \left|\overline{IC}^{\mathrm{v}}_{\mathrm{ou}}\right| = \left|\overline{IC}^{\mathrm{v}}_{\mathrm{in}}\right| . \tag{23}$$

A. *Função custo associada à ILD-ITD*

Apesar de sua estreita relação com os fenômenos físicos envolvidos na preservação espacial de uma fonte pontual, como também dos promissores resultados obtidos através de critérios objetivos de desempenho, experimentos psicoacústicos indicam a ineficácia de (9) na preservação da ITD. Uma investigação das possíveis causas é apresentada em [10]. Entretanto, a preservação da ILD através de (7) é suficiente para a manutenção da percepção acústica original do ruído, tendo sido apresentado um aperfeiçoamento em [11].

B. *Função custo associada à ITF*

Para investigar o impacto da minimização da $J^{\mathrm{v}}_{\mathrm{ITF}}$ nas pistas biauriculares do ruído processado, manipula-se (13) para que possa ser descrita como:

$$J^{\mathrm{v}}_{\mathrm{ITF}} = J^{\mathrm{v}}_1 + J^{\mathrm{v}}_2 , \tag{24}$$

em que

$$J^{\mathrm{v}}_1 = \left|\overline{ITF}^{\mathrm{v}}_{\mathrm{ou}}\right|^2 + \left|\overline{ITF}^{\mathrm{v}}_{\mathrm{in}}\right|^2 - 2\left|\overline{ITF}^{\mathrm{v}}_{\mathrm{ou}}\right|\left|\overline{ITF}^{\mathrm{v}}_{\mathrm{in}}\right|\cos(\theta^{\mathrm{v}}) , \tag{25}$$

$$J^{\mathrm{v}}_2 = \overline{ILD}^{\mathrm{v}}_{\mathrm{ou}}(1 - |\,\overline{IC}^{\mathrm{v}}_{\mathrm{ou}}\,|^2) = \overline{ILD}^{\mathrm{v}}_{\mathrm{ou}} - |\,\overline{ITF}^{\mathrm{v}}_{\mathrm{ou}}\,|^2 ; \tag{26}$$

em que $\overline{ITF}^{\mathrm{v}}_{\mathrm{in}}$ é definido em (12) e

$$\overline{ITF}^{\mathrm{v}}_{\mathrm{ou}} = E\left\{\frac{\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\mathbf{v}}{\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\mathbf{v}}\right\} = E\left\{\frac{\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\mathbf{v}}{\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\mathbf{v}} \cdot \frac{\mathbf{v}^{\mathrm{H}}\mathbf{w}_{\mathrm{R}}}{\mathbf{v}^{\mathrm{H}}\mathbf{w}_{\mathrm{R}}}\right\} \cong \frac{\mathbf{w}_{\mathrm{L}}^{\mathrm{H}}\mathbf{\Phi}_{\mathrm{v}}\mathbf{w}_{\mathrm{R}}}{\mathbf{w}_{\mathrm{R}}^{\mathrm{H}}\mathbf{\Phi}_{\mathrm{v}}\mathbf{w}_{\mathrm{R}}} . \tag{27}$$

para $\theta^{\mathrm{v}} = \overline{IPD}^{\mathrm{v}}_{\mathrm{ou}} - \overline{IPD}^{\mathrm{v}}_{\mathrm{in}}$, cujo valor principal é dado por $0 \le \theta^{\mathrm{v}} < 2\pi$.

Como $J^{\mathrm{v}}_1 \ge 0$ e $J^{\mathrm{v}}_2 \ge 0$, o valor mínimo de $J^{\mathrm{v}}_{\mathrm{ITF}}$ (i.e. $J^{\mathrm{v}}_{\mathrm{ITF}} = 0$), ocorre se e somente se $J^{\mathrm{v}}_1 = 0$ e $J^{\mathrm{v}}_2 = 0$. Analisando-se J^{v}_1, verifica-se que o ponto de mínimo ocorre quando $|\overline{ITF}^{\mathrm{v}}_{\mathrm{ou}}| = |\overline{ITF}^{\mathrm{v}}_{\mathrm{in}}|$ e $\theta^{\mathrm{v}} = 0$. Utilizando-se (19) e (20), isso implica que $\overline{ILD}^{\mathrm{v}}_{\mathrm{ou}} \cdot |\overline{IC}^{\mathrm{v}}_{\mathrm{ou}}|^2 = \overline{ILD}^{\mathrm{v}}_{\mathrm{in}} \cdot |\overline{IC}^{\mathrm{v}}_{\mathrm{in}}|^2$ e que $\overline{IPD}^{\mathrm{v}}_{\mathrm{ou}} = \overline{IPD}^{\mathrm{v}}_{\mathrm{in}}$. Sabendo que $\overline{ILD}^{\mathrm{v}}_{\mathrm{ou}} > 0$ (por definição), verifica-se então que o valor mínimo de J^{v}_2 ocorre quando $|\overline{IC}^{\mathrm{v}}_{\mathrm{ou}}|^2 = 1$. Portanto, o valor mínimo da $J^{\mathrm{v}}_{\mathrm{ITF}}$ ocorre quando $\overline{ILD}^{\mathrm{v}}_{\mathrm{ou}} = \overline{ILD}^{\mathrm{v}}_{\mathrm{in}} \cdot |\overline{IC}^{\mathrm{v}}_{\mathrm{in}}|^2$, $\overline{IPD}^{\mathrm{v}}_{\mathrm{ou}} = \overline{IPD}^{\mathrm{v}}_{\mathrm{in}}$ e $|\overline{IC}^{\mathrm{v}}_{\mathrm{ou}}|^2 = 1$, que, em um cenário com fonte pontual, resulta em: $\overline{ILD}^{\mathrm{v}}_{\mathrm{ou}} = \overline{ILD}^{\mathrm{v}}_{\mathrm{in}}$, $\overline{IPD}^{\mathrm{v}}_{\mathrm{ou}} = \overline{IPD}^{\mathrm{v}}_{\mathrm{in}}$ e $|\overline{IC}^{\mathrm{v}}_{\mathrm{ou}}|^2 = 1$.

Considerando as condições em (21), verifica-se que o termo $J^{\mathrm{v}}_{\mathrm{ITF}}$ permite que as condições A_1, A_2 e A_3 sejam atendidas. Contudo, pode-se mostrar que minimizar $J^{\mathrm{v}}_{\mathrm{ITF}}$ não permite o atendimento completo às condições B e C, descritas em (22) e (23). Dessa forma, $J^{\mathrm{v}}_{\mathrm{ITF}}$ não é adequada para a preservação espacial da impressão auditiva de sinais que contenham $|\overline{IC}^{\mathrm{v}}_{\mathrm{in}}|<1$. Nessa condição, a minimização de $J^{\mathrm{v}}_{\mathrm{ITF}}$ resulta em $\overline{ILD}^{\mathrm{v}}_{\mathrm{ou}} \neq \overline{ILD}^{\mathrm{v}}_{\mathrm{in}}$ e $\overline{IC}^{\mathrm{v}}_{\mathrm{ou}} \neq \overline{IC}^{\mathrm{v}}_{\mathrm{in}}$. A limitação de $J^{\mathrm{v}}_{\mathrm{ITF}}$

para sinais provenientes de um campo acústico difuso foi demonstrada a partir de simulações computacionais em [6], porém, o equacionamento apresentado neste trabalho é a primeira contribuição formal para a demonstração dessa propriedade da função custo.

C. *Função custo associada à IC*

Para analisar as propriedades de J_{IC}, expande-se sua definição em (14), resultando em:

$$J_{\mathrm{IC}}^{\mathrm{v}} = \left|\overline{IC}_{\mathrm{ou}}^{\mathrm{v}}\right|^2 + \left|\overline{IC}_{\mathrm{in}}^{\mathrm{v}}\right|^2 - 2\left|\overline{IC}_{\mathrm{ou}}^{\mathrm{v}}\right|\left|\overline{IC}_{\mathrm{in}}^{\mathrm{v}}\right|\cos(\theta^{\mathrm{v}}) . \qquad (28)$$

Considerando $|\overline{IC}^{\mathrm{v}}_{\mathrm{ou}}|$ como a variável de decisão de (28), verifica-se $J^{\mathrm{v}}_{\mathrm{IC}}$ é mínima para $|\overline{IC}^{\mathrm{v}}_{\mathrm{ou}}| = |\overline{IC}^{\mathrm{v}}_{\mathrm{in}}|$ e $\theta^{\mathrm{v}} = 0$, ou seja, $|\overline{IC}^{\mathrm{v}}_{\mathrm{ou}}| = |\overline{IC}^{\mathrm{v}}_{\mathrm{in}}|$ e $\overline{IPD}^{\mathrm{v}}_{\mathrm{ou}}=\overline{IPD}^{\mathrm{v}}_{\mathrm{in}}$. Logo, apesar de permitir a perfeita preservação tanto da IC quanto da IPD, a $J^{\mathrm{v}}_{\mathrm{IC}}$ não impõe restrição ao valor da $\overline{ILD}^{\mathrm{v}}_{\mathrm{ou}}$. Analisando-se as condições apresentadas em (21), (22) e (23) verifica-se que a $J^{\mathrm{v}}_{\mathrm{IC}}$, por si só, não garante a preservação em nenhuma das três situações analisadas.

V. Proposta de uma Função Custo para Preservação das Pistas Biauriculares

Em função da inexistência de uma função custo genérica para preservação das pista biauriculares e a partir dos resultados apresentados em [10] e [11], propõe-se a seguinte função custo para a preservação da ILD, ITD e IC do ruído interferente:

$$J_{\mathrm{BPC}}^{\mathrm{v}} = J_{\mathrm{ILD\text{-}IC}}^{\mathrm{v}} = J_{\mathrm{ILD}}^{\mathrm{v}} + \beta J_{\mathrm{IC}}^{\mathrm{v}} , \qquad (29)$$

em que $J^{\mathrm{v}}_{\mathrm{ILD}}$ é descrita em (7) e $J^{\mathrm{v}}_{\mathrm{IC}}$ em (14).

A. *Análise da função custo associada à ILD-IC*

Ambas $J^{\mathrm{v}}_{\mathrm{ILD}}$ e $J^{\mathrm{v}}_{\mathrm{IC}}$ são funções não negativas. Assim, a minimização da $J^{\mathrm{v}}_{\mathrm{ILD}}$ resulta naturalmente em $\overline{ILD}^{\mathrm{v}}_{\mathrm{ou}} = \overline{ILD}^{\mathrm{v}}_{\mathrm{in}}$, enquanto que a minimização da $J^{\mathrm{v}}_{\mathrm{IC}}$, resulta em $|\overline{IC}^{\mathrm{v}}_{\mathrm{ou}}| = |\overline{IC}^{\mathrm{v}}_{\mathrm{in}}|$ e $\overline{IPD}^{\mathrm{v}}_{\mathrm{ou}}=\overline{IPD}^{\mathrm{v}}_{\mathrm{in}}$. Dessa forma, os três conjuntos de condições apresentados em (21), (22) e (23) são atendidos por (29).

VI. Configuração dos Experimentos

O desempenho da função custo proposta foi quantificado a partir de métricas objetivas calculadas em experimentos realizados com cenários acústicos simulados.

A. *Cenário acústico simulado*

Dois cenários acústicos, ambos contendo uma fonte de fala e um sinal interferente de ruído foram simulados: no primeiro cenário, o sinal interferente está associado a uma fonte pontual de ruído; no segundo cenário, o ruído resulta de um campo acústico difuso.

Como sinal de interesse, utilizou-se um áudio com tom grave com 3 segundos de duração [12]. Como sinal interferente foi utilizado um ruído com características espectro-temporais semelhantes às da fala, desenvolvido para uso em testes clínicos com aparelhos auditivos [13].

As fontes pontuais foram modeladas a partir de respostas ao impulso (RI) obtidas em um ambiente anecoico. Elas caracterizam o caminho acústico entre a fonte e os microfones de um par de aparelhos auditivos retroauriculares, cada um contendo $M = 3$ microfones, e carregam informação sobre os efeitos físicos da cabeça e do tronco sobre as ondas sonoras [14]. A fonte de interesse foi associada ao azimute 0º, com distância de 80 cm dos microfones, enquanto que a fonte interferente refere-se ao azimute −60º (lado esquerdo), distante 300 cm dos aparelhos auditivos. Os sinais captados foram obtidos convolvendo-se os sinais de fala e ruído com as RI das suas respectivas fontes.

O campo acústico difuso foi simulado a partir do método descrito em [15], cuja matriz de coerência espacial entre os microfones foi obtida utilizando a base de dados disponível em [14], da mesma forma realizada em [9]. A função densidade de probabilidade do sinal interferente captado foi definida de acordo com o ruído ICRA tipo I [13].

O sinal contaminado foi obtido através da soma do sinal de interesse e do sinal interferente, de maneira que a razão sinal-ruído (SNR) fosse 0 dB.

A representação tempo-frequência dos sinais recebidos foi feita utilizando uma janela de 128 amostras, ponderadas por uma janela de análise, e uma transformada de Fourier de tempo curto (STFT) com $N = 256$ pontos, utilizando acréscimo de zeros. Após processadas, as amostras foram transformadas para o domínio do tempo utilizando a STFT inversa e ponderadas por uma janela de síntese, como descrito em [16]. As janelas de análise e síntese são iguais e determinadas pela raiz quadrada da janela de Hanning. A sobreposição entre as janelas de amostras adjacentes é de 50%, e a frequência de amostragem utilizada foi $f_s = 16$ kHz.

B. *Métricas de desempenho*

Quatro métricas objetivas foram utilizadas para a avaliação dos experimentos realizados. Três delas quantificam o erro determinado pelas pistas biauriculares do sinal no microfone de referência e do sinal processado: (1) ΔILD – que mede o erro entre as ILDs, calculada apenas para valores de k associados a frequências superiores a 1,5 kHz [10]; (2) ΔITD – que quantifica o erro entre as ITDs, calculada apenas para valores de k associados a frequências inferiores a 1,5 kHz [10], e (3) ΔMSC – que mede o erro entre a magnitude quadrada da coerência [6], calculada para todos os valores de k. A última métrica é a (4) ΔSNR – dada pela diferença entre a SNR do sinal recebido e do sinal processado, que mede a redução global de ruído. O sinal contaminado captado pelo microfone de referência apresenta os seguintes indicadores: $SNR_L = 0$ dB e $SNR_R = 3{,}1$ dB.

VII. Resultados e Discussão

As Figs. 2 e 3 apresentam o desempenho dos métodos MWF-ITF, MWF-IC, e da técnica proposta, denominada MWF-ILD-IC, respectivamente, para os cenários acústico associados à fonte pontual de ruído e ao campo acústico difuso. A região plana nos gráficos ($\alpha \rightarrow 0$) corresponde à região cujo desempenho das técnicas é aproximadamente igual ao desempenho do MWF.

As Figs. 2(a) e 2(b) apresentam, respectivamente, a ΔILD^{v} e ΔITD^{v} do ruído. Considerando ΔILD^{v}, o MWF-ILD-IC apresenta, consistentemente, erros na ILD da fonte interferente que são menores ou iguais aos erros apresentados pelo MWF-ITF e MWF-IC em todos os valores analisados de α. Na região $10^{-8} < \alpha < 10^{-1}$, o MWF-ILD-IC apresenta resultados erros menores aos apresentados pelas demais técnicas. Considerando a ΔITD^{v}, as três técnicas apresentam desempenho de

preservação semelhantes. A Figura 2(c) apresenta a ΔMSC^v para as três técnicas analisadas. Na região em que $10^{-8} < \alpha < 10^{-1}$, o MWF-ILD-IC apresenta desempenho inferior ao MWF-ITF e o MWF-IC. O que significa que um parâmetro α escolhido nessa região pode levar a uma maior variabilidade na localização percebida da fonte de ruído. Considerando a faixa $\alpha > 10^{-1}$, o desempenho das técnicas é aproximadamente o mesmo, pois o erro na MSC^v é menor que 0,2 [10]. Analisando a ILD e a ITD da fala, nas Figuras 2(e) e 2(f), respectivamente, o MWF-IC e o MWF-ILD-IC apresentam erros próximos de zero, o que indica preservação da localização da fala. Por outro lado, o MWF-ITF apresenta erros crescentes, especialmente na região $\alpha > 10^{-1}$. Os erros na *MSC* da fala (Figura 2(c)) indicam que as três técnicas possuem desempenho qualitativo praticamente igual. As Figs. 2(g) e 2(f) apresentam a variação de SNR global nas orelhas esquerda e direita, respectivamente.

A Fig. 3 apresenta o desempenho das técnicas analisadas para o cenário com campo difuso. Em todas as métricas consideradas, o MWF-ILD-IC e o MWF-IC apresentam, em termos práticos, o mesmo desempenho. Já o MWF-ITF distorce a ILD e a MSC do ruído, assim como a SNR global, presente nas Figs. 3(g) e 3(f).

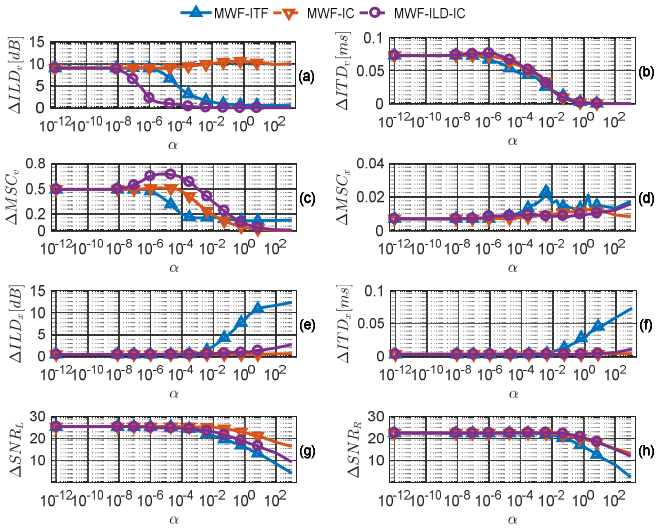

Fig. 2. Resultados para ruído proveniente de uma fonte pontual.

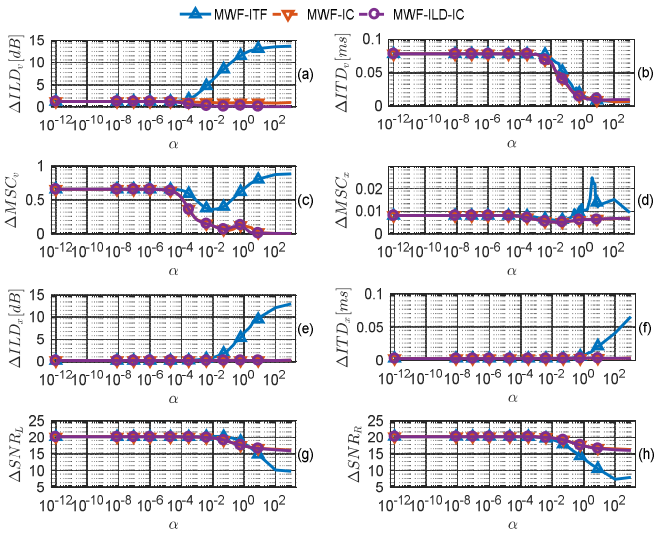

Fig. 3. Resultados para ruído proveniente de campo acústico difuso.

VIII. Conclusões

Este trabalho apresentou uma análise das limitações de funções custo utilizadas para preservação do cenário acústico espacial em aparelhos auditivos biauriculares. Como resultado foi proposta uma nova função custo capaz de preservar as pistas biauriculares de ruídos provenientes de fontes acústicas pontuais, como também de ruído proveniente de um campo acústico difuso. A capacidade de manutenção da ILD, ITD e a IC originais permite sua utilização sem a necessidade de conhecimento prévio sobre as características do campo sonoro. Os resultados obtidos em simulações corroboram o desempenho esperado, indicando vantagens na comparação com funções custo do estado da arte.

Referências

[1] V. Hamacher *et. all*, "Binaural signal processing in hearing aids: Technologies and algorithms", *Advances in digital speech transmission*, v 14, pp. 401-429, 2008.

[2] S. Doclo *et al.*, "Binaural speech processing with application to hearing devices", *Audio Source Separation and Speech Enhancement*, ch. 18, pp. 413-442, 2018.

[3] J. Blauert, Spatial hearing: the psychophysics of human sound localization. MIT press, 1997.

[4] S. Doclo *et. all*, "Binaural speech processing with application to hearing devices," in *Audio Source Separation and Speech Enhancement*, Wiley, ch. 18, pp. 413-442, 2018.

[5] B. Cornelis *et al.*, "Theoretical analysis of binaural multimicrophone noise reduction techniques", *IEEE Transactions on Audio, Speech, and Language Processing*, v. 18, n. 2, pp. 342-355, 2009.

[6] D. Marquardt, V. Hohmann e S. Doclo, "Interaural coherence preservation in multi-channel Wiener filtering-based noise reduction for binaural hearing aids", *IEEE/ACM Transactions on Audio, Speech and Language Processing*, v. 23, n. 12, pp. 2162-2176, 2015.

[7] S. Doclo *et al.*, "Extension of the multi-channel Wiener filter with localisation cues for noise reduction in binaural hearing aids", *International Workshop on Acoustic, Echo and Noise Control* (IWAENC), 2005, pp. 221-224.

[8] S. Doclo *et al.*, "Theoretical analysis of binaural cue preservation using multi-channel Wiener filtering and interaural transfer functions", *International Workshop on Acoustic, Echo and Noise Control* (IWAENC), 2006, pp. 1-4.

[9] D. Marquardt *et al.*, "Theoretical analysis of linearly constrained multi-channel Wiener filtering algorithms for combined noise reduction and binaural cue preservation in binaural hearing aids", *IEEE/ACM Transactions on Audio, Speech and Language Processing*, v. 23, n. 12, pp. 2384-2397, 2015.

[10] F. P. Itturriet e M. H. Costa, "Perceptually relevant preservation of interaural time differences in binaural hearing aids", *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, v. 27, n. 4, pp. 753-764, 2019.

[11] M. H. Costa e P. A. Naylor, "ILD preservation in the multichannel Wiener filter for binaural hearing aid applications," *in Proc. Eur. Signal Process. Conf.* (EUSIPCO), 2014, pp. 636-640.

[12] ITU-T Recommendation P.50, Appendix 1, "Test signals," 1998.

[13] W. A. Dreschler *et al.*, "ICRA noises: artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment", *Audiology*, v. 40, n. 3, pp. 148-157, 2001.

[14] H. Kayser *et al.*, "Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses", *EURASIP Journal on Advances in Signal Processing*, v. 2009, pp. 1-10, 2009.

[15] E. A. Habets, I. Cohen e S. Gannot, S., "Generating nonstationary multisensor signals under a spatial coherence constraint", *The Journal of the Acoustical Society of America*, v. 124, n. 5, pp. 2911-2917, 2008.

[16] R. Crochiere *et al.*, "A weighted overlap-add method of short-time Fourier analysis/synthesis", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, v. 28, n. 1, pp. 99-102, 1980.