ILD PRESERVATION IN THE MULTICHANNEL WIENER FILTER FOR BINAURAL HEARING AID APPLICATIONS

Márcio H. Costa[1], *Patrick A. Naylor*[2]

[1]Universidade Federal de Santa Catarina
Florianópolis, Brazil
costa@eel.ufsc.br

[2]Imperial College
London, UK
p.naylor@imperial.ac.uk

ABSTRACT

This work presents a new method for noise reduction in binaural hearing aid applications that preserves the interaural level difference. A bounded symmetrical approximation of the logarithm is employed to estimate the interaural level difference, resulting in identical values for symmetrical (left/right) frontal angles. It proposes a new cost function to be used in association with the multichannel Wiener filter technique to provide a trade-off between noise reduction and distortion of the localization cues. Simulations of a binaural setup and comparisons with a previously developed technique show that the new method gives a signal to noise ratio improvement of up to 9.6 dB better than the baseline technique, for the same maximum-tolerated binaural-cue distortion.

Index Terms— Hearing-aids, noise reduction, binaural, wiener filter, speech processing

1. INTRODUCTION

Human ability to localize, separate and track sound sources is a complex perceptual process involving integration of multiple variables, in which interaural time (ITD) and level (ILD) differences play a major role [1].

Spatialization for hearing impaired people can be severely affected by noise reduction processing in hearing aid devices that do not preserve the original acoustic localization cues of the target speech and interfering sources. For example, in [2] it has been shown that bilateral hearing aids do not preserve the listener's sense of auditory space, distorting the original direction of arrival of both target sound and interference. As a result, localization of sounds is generally best achieved by turning off the noise reduction processing. This represents a major disadvantage to the user since the immediate localization of the sources of interest is highly important for social situations (the 'Cocktail Party Effect'), for safety (traffic, safety warnings), and for lip-reading.

Despite many advances in hearing aid technology, noise reduction techniques still constitute a significant source of distortion in the subjective perception of acoustic source direction of arrival. A common approach for noise reduction is the use of spatial filtering techniques. Among them, the most popular are the linearly constrained minimum variance beamformer and the generalized side-lobe canceller [3]. Both methods rely on prior knowledge of the source localization and/or head related transfer functions, presenting significant performance degradation when the assumed conditions are a less than ideal match for reality.

Multichannel Wiener Filter (MWF) based techniques have been extensively explored in the scientific literature [4]-[7]. Their attractiveness comes in part from the fact that they do not need *a priori* spatial information and permit deep theoretical insights about design and performance [6]-[7]. Although it was theoretically demonstrated that the conventional MWF naturally preserves speech localization cues, the processed noise at the output inherits the input speech localization cues [6] and therefore significantly degrades the overall usefulness of speech and noise spatialization for the user after noise reduction has been applied. In [4]-[5] approximated analytical functions for ITD, ILD and interaural transfer function (ITF) were introduced as additional terms in the MWF cost function. Even though there are still no closed-form solutions to their optimal coefficients (and time consuming optimization techniques have to be used to compute them), such an approach is extremely valuable to indicate achievable performance bounds.

Although ITD is an important binaural cue, amplitude stereo panning techniques have demonstrated that ILD carries enough information for creating complex artificial auditory scenes even in headphones [8].

This work presents a new MWF-ILD based technique for noise reduction with preservation of speech and noise localization cues for use in hearing aids. The proposed approach employs a bounded approximation of the logarithm to estimate the interaural level difference.

This work was supported by CAPES, under grant BEX 10184/12-0 and CNPq under grant 304832/2012-0.

2. SIGNAL AND SYSTEM MODELS

Our scenario assumes binaural hearing aids communicating in full-duplex mode. Frequency domain decomposition is applied to the incoming signals through an N-bin Short-Time Fourier Transform. The input signal from the m^{th} left (L) microphone (of a total of M_L), for each frame λ and frequency k, is defined as

$$y_{L,m}(\lambda,k) = x_{L,m}(\lambda,k) + v_{L,m}(\lambda,k)\,, \tag{1}$$

in which x is the desired speech and v is additive noise. The collection of input signals is expressed as $\mathbf{y}(\lambda,k) = \mathbf{x}(\lambda,k) + \mathbf{v}(\lambda,k) = [\; y_{L,1}(\lambda,k) \;\ldots\; y_{L,M_L}(\lambda,k) \;\; y_{R,1}(\lambda,k)\ldots\; y_{R,M_R}(\lambda,k) \;]^{\mathrm{T}}$ (where R means right), with dimension (M×1) for $M = M_L+M_R$. Considering the deterministic vector $\mathbf{q}_L$, containing 1 in the position corresponding to a defined left reference microphone and zeros in all other positions, the left reference vector of the hearing aids is given by

$$y_{L,ref}(\lambda,k) = x_{L,ref}(\lambda,k) + v_{L,ref}(\lambda,k) = \mathbf{q}_L^{\mathrm{T}}\mathbf{y}(\lambda,k)\,. \tag{2}$$

As illustrated in Fig. 1 the left output signal of the hearing aids after processing is given by

$$z_L(\lambda,k) = z_{Lx}(\lambda,k) + z_{Lv}(\lambda,k) = \mathbf{w}_L^{\mathrm{H}}(\lambda,k)\mathbf{y}(\lambda,k) \tag{3}$$

where $\mathbf{w}_L(\lambda,k)$ is the left coefficient vector of a noise reduction linear filter to be determined.

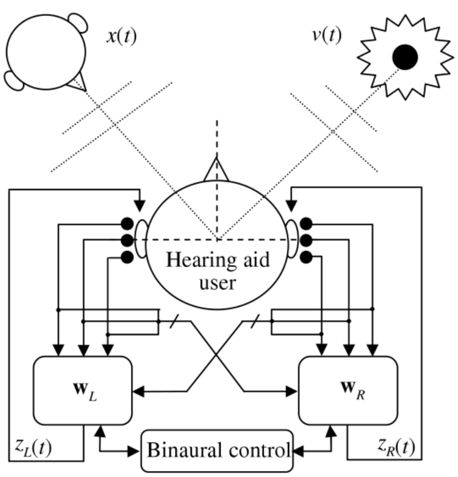

Fig. 1. Binaural system setup.

3. MULTICHANNEL WIENER FILTER

The MWF cost function is given by

$$J_{MWF}(\lambda,k) = E\left\{\left\|\begin{matrix} x_{L,ref}(\lambda,k) - \mathbf{w}_L^{\mathrm{H}}(\lambda,k)\mathbf{y}(\lambda,k) \\ x_{R,ref}(\lambda,k) - \mathbf{w}_R^{\mathrm{H}}(\lambda,k)\mathbf{y}(\lambda,k)\end{matrix}\right\|^2\right\}, \tag{4}$$

where $E\{\cdot\}$ indicates expectation, and $\|\cdot\|^2$ is the squared Euclidean norm. Manipulating (4) leads to

$$\begin{aligned} J_{MWF} = {} & \mathbf{q}_L^{\mathrm{T}}\mathbf{R}_x\mathbf{q}_L + \mathbf{q}_R^{\mathrm{T}}\mathbf{R}_x\mathbf{q}_R - \mathbf{q}_L^{\mathrm{T}}\mathbf{R}_x\mathbf{w}_L - \mathbf{q}_R^{\mathrm{T}}\mathbf{R}_x\mathbf{w}_R \\ & - \mathbf{w}_L^{\mathrm{H}}\mathbf{R}_x\mathbf{q}_L - \mathbf{w}_R^{\mathrm{H}}\mathbf{R}_x\mathbf{q}_R + \mathbf{w}_L^{\mathrm{H}}\mathbf{R}_y\mathbf{w}_L + \mathbf{w}_R^{\mathrm{H}}\mathbf{R}_y\mathbf{w}_R \end{aligned} \tag{5}$$

in which $\mathbf{R}_x(k) = E\{\mathbf{x}(\lambda,k)\mathbf{x}^{\mathrm{H}}(\lambda,k)\}$ and $\mathbf{R}_y(k) = E\{\mathbf{y}(\lambda,k)\,\mathbf{y}^{\mathrm{H}}(\lambda,k)\}$. The minimum of J_{MWF} is found by equating the partial derivatives w.r.t. the coefficients to zero, leading to:

$$\mathbf{w}_L^{opt} = \mathbf{R}_y^{-1}\mathbf{R}_x\mathbf{q}_L\;,\; \mathbf{w}_R^{opt} = \mathbf{R}_y^{-1}\mathbf{R}_x\mathbf{q}_R \tag{6}$$

It was shown that preservation of noise localization cues can also be taken into account by including auxiliary cost function terms [4]-[5], resulting in:

$$J = J_{MWF} + \alpha_A \cdot J_A \tag{7}$$

in which J_A is an additional cost function term related to the preservation of the binaural cues, and α_A is a weighting factor. Depending on J_A, optimization techniques must be used to find the point of minimum. It is clear to see that additional cost function terms like this should be designed to be as efficient as possible so as to minimize computational requirements and to avoid complex performance surfaces.

4. INTERAURAL LEVEL DIFFERENCE VARIATION

The noise ILD variation, for each frame λ and frequency k, is defined as $\Delta ILD_v = oILD_v - iILD_v$ where $oILD_v$ and $iILD_v$ are the output and input noise ILDs. It results in [9]:

$$\Delta ILD_v = 10\log_{10}\left(\frac{\mathbf{w}_L^{\mathrm{H}}\mathbf{v}\mathbf{v}^{\mathrm{H}}\mathbf{w}_L}{\mathbf{w}_R^{\mathrm{H}}\mathbf{v}\mathbf{v}^{\mathrm{H}}\mathbf{w}_R}\right) - 10\log_{10}\left(\frac{\mathbf{q}_L^{\mathrm{T}}\mathbf{v}\mathbf{v}^{\mathrm{H}}\mathbf{q}_L}{\mathbf{q}_R^{\mathrm{T}}\mathbf{v}\mathbf{v}^{\mathrm{H}}\mathbf{q}_R}\right). \tag{8}$$

Preservation of the original noise ILD requires that $\mathbf{w}_L$ and $\mathbf{w}_R$ result in small values of ΔILD_v (≅0), for all λ and k. An ILD cost function was firstly proposed in [4] to provide a trade-off between noise reduction and preservation of source lateralization. It was defined as:

$$J_1 = \left(\frac{\mathbf{w}_L^{\mathrm{H}}\mathbf{R}_v\mathbf{w}_L\mathbf{q}_R^{\mathrm{T}}\mathbf{R}_v\mathbf{q}_R - \mathbf{w}_R^{\mathrm{H}}\mathbf{R}_v\mathbf{w}_R\mathbf{q}_L^{\mathrm{T}}\mathbf{R}_v\mathbf{q}_L}{\mathbf{w}_R^{\mathrm{H}}\mathbf{R}_v\mathbf{w}_R\mathbf{q}_R^{\mathrm{T}}\mathbf{R}_v\mathbf{q}_R}\right)^2. \tag{9}$$

4. PROPOSED TECHNIQUE

Despite not being stated in [4], equation (9) can be obtained as an approximation of the mean square value of (8), assuming two conditions: (a) each ratio is approximated by a first order Taylor expansion of its mean value; (b) a first order Taylor series approximation of the logarithm ($\log_{10}(x) \cong (x\text{-}1)/\log_e(10)$). The first order Taylor series approximation of the logarithm in (8) results in a simple rational fraction and, consequently, minimizes the complexity of the resulting cost function. However, the polynomial Taylor expansion is only valid for arguments less than 2, and its first order truncated form only provides accurate estimates in the vicinity of 1.

A more accurate approximation of the logarithm can be obtained by the inverse hyperbolic tangent function. This converges more quickly than the polynomial series and provides bounded results:

$$\log_{10}(x) = \frac{2}{\log_e(10)}\sum_{i=1}^{\infty}\frac{1}{2i-1}\left(\frac{x-1}{x+1}\right)^{2i-1} \tag{10}$$

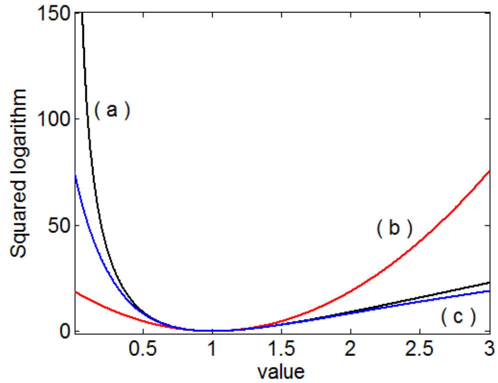

Fig. 2. Squared logarithm (a) and approximation by first order Taylor series (b), and first order hyperbolic tangent approximation (c).

Figure 2 shows a comparison between squared values of the true logarithm, its first order Taylor series, and the first order approximation using the inverse hyperbolic tangent. The first order Taylor approximation overestimates the true logarithm for arguments bigger than 1 and underestimates it for arguments smaller than 1. These are undesired characteristics, since acoustic sources in symmetrical angles may result in significantly different squared ILDs. The hyperbolic approximation underestimates the true value of the logarithm, but results in more accurate estimates. Reciprocal arguments result in the same value.

Based on these characteristics, a new cost function for preservation of the original ILD is derived from (8) through the following steps: (a) each ratio in (8) is approximated by a first order Taylor expansion (about the mean of both numerator and denominator) of its mean value [10]; (b) applying the logarithm addition property; (c) using the first order approximation of (10); (d) squaring so as to give only positive values. After some manipulation it results in:

$$J_2 = \left(\frac{\mathbf{w}_L^{\mathrm{H}} \mathbf{R}_v \mathbf{w}_L \mathbf{q}_R^{\mathrm{T}} \mathbf{R}_v \mathbf{q}_R - \mathbf{w}_R^{\mathrm{H}} \mathbf{R}_v \mathbf{w}_R \mathbf{q}_L^{\mathrm{T}} \mathbf{R}_v \mathbf{q}_L}{\mathbf{w}_L^{\mathrm{H}} \mathbf{R}_v \mathbf{w}_L \mathbf{q}_R^{\mathrm{T}} \mathbf{R}_v \mathbf{q}_R + \mathbf{w}_R^{\mathrm{H}} \mathbf{R}_v \mathbf{w}_R \mathbf{q}_L^{\mathrm{T}} \mathbf{R}_v \mathbf{q}_L} \right)^2 . \quad (11)$$

5. SIMULATION RESULTS

Simulation results are presented to demonstrate the validity and accuracy of the proposed technique to improve the trade-off between noise reduction and preservation of ILD.

5.1. ILD Estimation

Simulations were performed convolving random white Gaussian noise bandlimited to 7.5 kHz with head related impulse responses sampled at 16 kHz. The impulse responses were obtained in an anechoic chamber with 3 m of distance between the acoustic sources and a head-and-torso simulator, in which two behind-the-ear hearing aids were placed (left and right ears) [11]. Each device has three microphones ($M_L = M_R = 3$). The front microphones were assumed to be the reference microphones. Elevation angle was set to 0° and azimuth varied from -90° (left) to 90° (right) in steps of 5°. The input signal was transformed to the frequency domain by an $N = 256$ bin short-time Fourier Transform, using the weighted overlap-and-add method with an analysis window of 128 samples and 50% of overlap. Three ILD performance surfaces, denoted ILD_0, ILD_1, and ILD_2, are next considered. The 'true' ILD performance surface ILD_0, as a function of the frequency bin, was calculated as (T is the number of available frames):

$$ILD_0^2(k) = \left[\frac{10}{T} \sum_{\lambda=1}^{T} \log_{10} \left(\frac{\mathbf{q}_L^{\mathrm{T}} \mathbf{v}\mathbf{v}^{\mathrm{H}} \mathbf{q}_L}{\mathbf{q}_R^{\mathrm{T}} \mathbf{v}\mathbf{v}^{\mathrm{H}} \mathbf{q}_R} \right) \right]^2 . \quad (12)$$

Approximations of (12) were calculated using obtained correlation matrices $\mathbf{R}_v(k) = E\{\mathbf{v}(k)\mathbf{v}^{\mathrm{H}}(k)\}$ and Taylor and hyperbolic first order approximations, respectively, as:

$$ILD_1^2(k) = \left[\log_e(10)^{-1} \left(\frac{\mathbf{q}_L^{\mathrm{T}} \mathbf{R}_v \mathbf{q}_L - \mathbf{q}_R^{\mathrm{T}} \mathbf{R}_v \mathbf{q}_R}{\mathbf{q}_R^{\mathrm{T}} \mathbf{R}_v \mathbf{q}_R} \right) \right]^2 , \quad (13)$$

$$ILD_2^2(k) = \left[2\log_e(10)^{-1} \left(\frac{\mathbf{q}_L^{\mathrm{T}} \mathbf{R}_v \mathbf{q}_L - \mathbf{q}_R^{\mathrm{T}} \mathbf{R}_v \mathbf{q}_R}{\mathbf{q}_L^{\mathrm{T}} \mathbf{R}_v \mathbf{q}_L + \mathbf{q}_R^{\mathrm{T}} \mathbf{R}_v \mathbf{q}_R} \right) \right]^2 . \quad (14)$$

Comparative results are presented in Figs. 3 to 5.

5.2. Objective measures

We now present objective performance experiments using the same binaural setup of the previous section. A male speech signal and a noise signal recorded in a fully occupied cafeteria environment [11], containing babble noise and other noises, were used. Both the speech and noise elevations were set to 0°. The speech source azimuth was set to 0° and noise to -60° and 60°, resulting in the experiments S_0N_{-60} and S_0N_{60}. In each experiment, the signal to noise ratio was set to 0 dB in the ear nearest to the noise source position. Ideal voice activity detection was applied by hand-labelling of the clean speech signal. Noisy signal and noise correlation matrices, $\mathbf{R}_y(k)$ and $\mathbf{R}_v(k)$ were computed *a-priori*, directly from the original signals to avoid reinforcement of the binaural cues due to estimation errors [7] and to obtain an upper bound performance. Speech correlation matrices were calculated as $\mathbf{R}_x(k) = \mathbf{R}_y(k)\text{-}\mathbf{R}_v(k)$. The left and right coefficient vectors, $\mathbf{w}_L$ and $\mathbf{w}_R$, were obtained from minimization of (7) using J_1 from (9) and J_2 from (11) as additional cost functions terms. The performance of both methods was analysed using the wideband perceptual evaluation of speech quality (PESQ) [12], signal to noise ratio (SNR), improvement in the intelligibility weighted signal to noise ratio (IIWSNR) [13], and noise interaural level difference error, defined as:

$$\mathrm{ILDe_{No}} = \frac{10}{K_F - K_I + 1} \sum_{k=K_I}^{K_F} \log_{10} \left(\frac{\sum_{\lambda} \mathbf{w}_L^{\mathrm{H}} \mathbf{v}\mathbf{v}^{\mathrm{H}} \mathbf{w}_L \sum_{\lambda} \mathbf{q}_R^{\mathrm{T}} \mathbf{v}\mathbf{v}^{\mathrm{H}} \mathbf{q}_R}{\sum_{\lambda} \mathbf{w}_R^{\mathrm{H}} \mathbf{v}\mathbf{v}^{\mathrm{H}} \mathbf{w}_R \sum_{\lambda} \mathbf{q}_L^{\mathrm{T}} \mathbf{v}\mathbf{v}^{\mathrm{H}} \mathbf{q}_L} \right) \quad (15)$$

from 1.5 kHz (K_I) to 8 kHz (K_F). This is a key measure of the preservation of the ILD binaural cue. Subscripts are used to refer to speech (Sp), noise (No), and left (L) and right (R) ears. Results are presented in Figs. 6 to 8. Table 1 presents comparisons between both techniques assuming a weighting factor that results in $ILDe_{No} \leq 4$ dB in both S_0N_{-60} and S_0N_{60} cases. Informal listening using headphones indicated that this condition was sufficient to preserve the noise source lateralization. Table 2 shows results for $ILDe_{No} \leq 2$ dB, where approximately symmetric $ILDe_{No}$ are obtained for J_2.

Table 1. Objective measures for $ILDe_{NO} \leq 4$ dB. Weighting parameters: $\alpha_1 = 5 \cdot 10^{-2}$ (J_1), $\alpha_2 = 4 \cdot 10^{-3}$ (J_2).

	S_0N_{-60}		S_0N_{60}	
	Proposed	[4]	Proposed	[4]
$ILDe_{No}$ [dB]	3.8	1.9	2.9	3.9
$PESQ_L$	3.1	2.6	4.1	4.1
$PESQ_R$	3.7	3.7	3.3	3.5
SNR_L [dB]	26.3	23.1	34.9	35.3
SNR_R [dB]	29.8	30.1	29.4	30.7
$IIWSNR_L$ [dB]	25.4	23.5	26.5	25.9
$IIWSNR_R$ [dB]	22.1	22.4	28.6	29.8
$ILDe_{Sp}$ [dB]	0.6	0.7	0.6	0.6

Table 2. Objective measures for $ILDe_{NO} \leq 2$ dB. Weighting parameters: $\alpha_1 = 600$ (J_1), $\alpha_2 = 60$ (J_2).

	S_0N_{-60}		S_0N_{60}	
	Proposed	[4]	Proposed	[4]
$ILDe_{No}$ [dB]	2.0	1.2	1.9	2
$PESQ_L$	2.7	1.7	4.0	4
$PESQ_R$	3.7	3.2	3.1	3.1
SNR_L [dB]	23.7	14.1	34.5	34.2
SNR_R [dB]	30.1	28.8	26.7	28.2
$IIWSNR_L$ [dB]	22.9	19.3	25.6	26.4
$IIWSNR_R$ [dB]	22.3	19.1	26.3	27.3
$ILDe_{Sp}$ [dB]	0.8	1.3	0.6	0.6

6. DISCUSSION

Figure 3 agrees with the ILD characteristic in the 1 kHz band presented in [14], for a single undisturbed speaker in the free-field. Comparison of Figs. 3 and 4 indicates that the Taylor approximation results in large errors in squared ILD estimations, presenting a left side bias. This finding can be explained by the characteristics of the first order Taylor approximation described in Section 4. Comparison of Figs. 3 and 5 indicates the hyperbolic approximation provides accurate estimates of the ILD morphology in the frequency domain, as a function of the incidence angle.

Figure 6 shows that cost functions based on both approximations under study can be used to preserve ILD. Incidence angles of ±60° were chosen since they represent the most difficult situation for J_1 according to [14]. Figures 6, 7 and 8 show the trade-off between noise reduction and ILD preservation (S_0N_{-60}: J_1 in red and J_2 in blue; S_0N_{60}: J_1 in magenta and J_2 in cyan). Tables 1 and 2 show that assuming a maximum $ILDe_{NO}$ design requirement, the proposed technique results in a better quality signal in the S_0N_{-60} case. For $ILDe_{NO} \leq 2$, an increase of 1 point in PESQ and 9.6 dB in SNR_L were obtained for the left ear. In the S_0N_{60} case, the same weight factor results in a decrease of only 1.5 dB in SNR_R, maintaining the same left and right PESQ. The proposed cost function presents an increase in computation cost, compared to [4], of only one additional sum per sample.

7. CONCLUSIONS

This work presented a new multichannel Wiener filter based technique with interaural level difference preservation for binaural hearing aid applications. A bounded symmetrical approximation of the logarithm was proposed for improving interaural level difference estimations, resulting in identical estimates for symmetrical (left/right) frontal incidence angles. As a result, it was shown that improvements up to 9.6 dB in signal to noise ratio and 1 PESQ can be obtained when assuming a maximum tolerated interaural level difference distortion for both sides as design requirement.

8. REFERENCES

[1] J. Blauer, Spatial Hearing: The Psychoacoustics of Human Sound Localization, MIT Press, 1997.

[2] T. Bogaert, T.J. Klasen, M. Moonen, L. Deun, J. Wouters, "Horizontal localization with bilateral hearing aids: without is better than with," *Journal of the Acoustical Society of America*, vol. 119, no. 1, pp. 515-526, 2006.

[3] S. Haykin, Adaptive Filter Theory, Prentice Hall, 2002.

[4] S. Doclo, R. Dong, T.J. Klasen, J. Wouters, S. Haykin, M. Moonen, "Extension of the multi-channel Wiener filter with localization cues for noise reduction in binaural hearing aids," in *Proc. IEEE Intl. Work. on Acoustic Echo and Noise Control* (*IWAENC*), Sept. 2005, pp. 221-224.

[5] T. Bogaert, J. Wouters, S. Doclo, M. Moonen, "Binaural cue preservation for hearing aids using an interaural transfer function multichannel Wiener filter," in *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing* (*ICASSP*), Apr. 2007, pp. 565-568.

[6] B. Cornelis, S. Doclo, T. Bogaert, M. Moonen, J. Wouters, "Theoretical analysis of binaural multimicrophone noise reduction techniques," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 2, pp. 342-355, 2010.

[7] B. Cornelis, M. Moonen, J. Wouters, "Performance analysis of multichannel Wiener filter-based noise reduction in hearing aids under second order statistics estimation errors," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 5, pp. 1368-1381, 2011.

[8] K. Blum, G.-J. Rooyen, H.A. Engelbrecht, "Spatial audio to assist speaker identification in telephony," in *Proc. International Conference on Systems, Signals and Image Processing* (*IWSSIP*), pp. 1-4, 2010.

[9] M. Raspaud, H. Viste, G. Evangelista, "Binaural source localization by joint estimation of ILD and ITD," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 18, no. 1, pp. 68-77, 2010.

[10] G.M.P. Kempen, L.J. Vliet, "Mean and variance of ratio estimators used in fluorescence ratio imaging," *Cytometry*, vol. 39, no. 4, pp. 300-305, 2000.

[11] H. Kayser, S.D. Ewert, J. Anemüller, T. Rohdenburg, V. Hohmann, B. Kollmeier, "Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, pp. 1-10, 2009.

[12] ITU-T, P.862: Revised Annex A - Reference implementations and conformance testing for ITU-T Recs P.862, P.862.1 and P.862.2, 2005.

[13] J.E. Greenberg, P.M. Peterson, P.M. Zurek, "Intelligibility-weighted measures of speech-to-interference ratio and speech system performance," *Journal of the Acoustical Society of America*, vol. 94, no.5, pp. 3009-3010, 1993.

[14] M. Dietz, S.D. Ewert, V. Hohmann, "Auditory model based direction estimation of concurrent speakers from binaural signals," *Speech Communication*, vol. 53, pp.592-605, 2011.

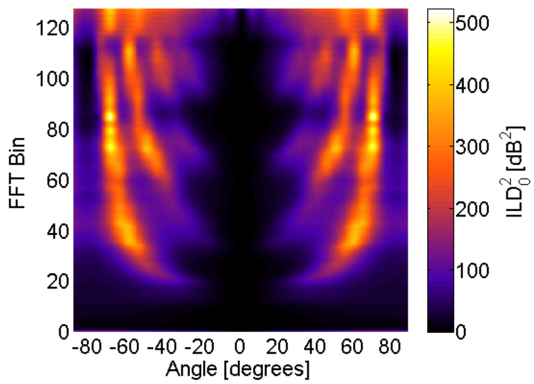

Fig. 3. ILD estimated by (12).

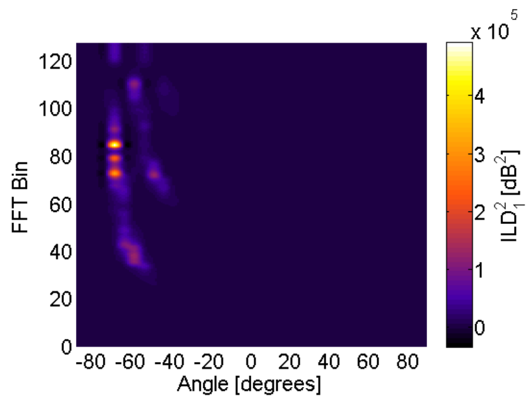

Fig. 4. ILD estimated by (13).

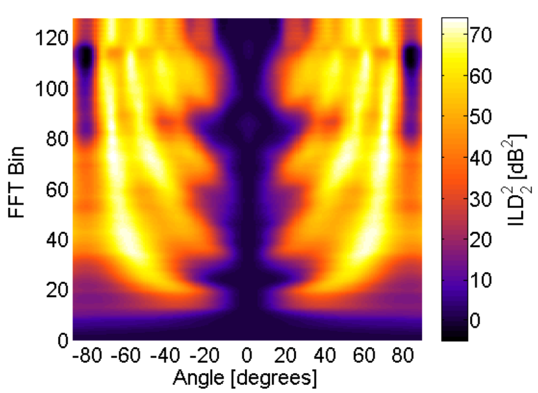

Fig. 5. ILD estimated by (14).

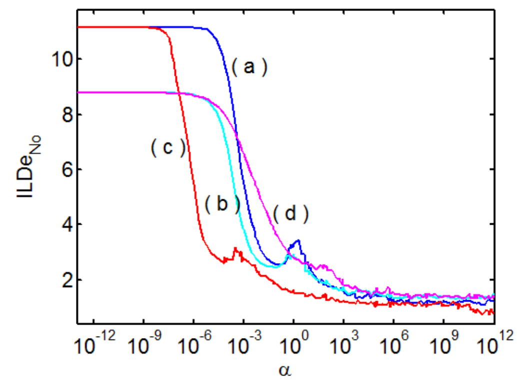

Fig. 6. Noise ILD error. Proposed technique: (a) S_0N_{-60}, (b) S_0N_{60}. Technique proposed in [4]: (c) S_0N_{-60}, (d) S_0N_{60}.

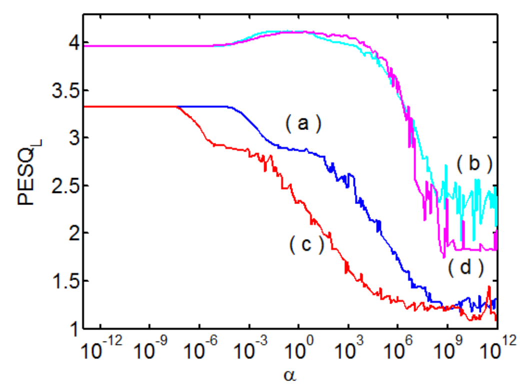

Fig. 7. PESQ in the left ear. Proposed technique: (a) S_0N_{-60}, (b) S_0N_{60}. Technique proposed in [4]: (c) S_0N_{-60}, (d) S_0N_{60}.

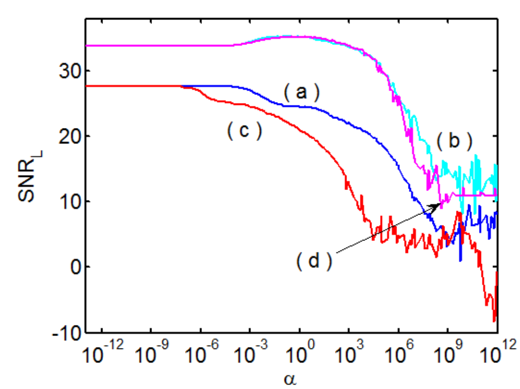

Fig. 8. SNR in the left ear. Proposed technique: (a) S_0N_{-60}, (b) S_0N_{60}. Technique proposed in [4]: (c) S_0N_{-60}, (d) S_0N_{60}.