ADAPTIVE BINAURAL NOISE REDUCTION BASED ON MATCHED-FILTER EQUALIZATION AND POST-FILTERING

Masoumeh Azarpour, Gerald Enzner, and Rainer Martin

Institute of Communication Acoustics
Ruhr-Universität Bochum, Germany
{masoumeh.azarpour, gerald.enzner, rainer.martin}@rub.de

ABSTRACT

In this paper a binaural noise reduction system based on the adaptive matched filter array (MFA) and post-filtering is presented. The binaural MFA filter is formed using the interaural impulse response between left and right microphone signal which is estimated by means of the NLMS algorithm. The residual noise reduction by three post-filter types is then compared and evaluated according to objective and subjective measures. Furthermore, the performance of the algorithm for preserving binaural cues is discussed.

Index Terms— Binaural noise reduction, Matched filter array, Binaural cue preservation

1. INTRODUCTION AND RELATION TO PRIOR WORKS

Acoustic noise is ubiquitous and affects the quality and the intelligibility of speech signals. Especially for hearing aid users, the intelligibility of speech is more vulnerable to noise and a major problem for them is the degradation of their speech perception in noisy environment. A number of monaural and bilateral algorithms has been proposed to enhance the speech signal for hearing aids [1–3]. The signal processing algorithms utilized in most of hearing aids are the bilateral ones. It means the individual apparatus in each ear works independently. One of the main drawback of these algorithms is the distortion of binaural cues, particularly interaural time difference (ITD) and interaural level difference (ILD). It has been shown that the hearing impaired subjects localize sounds better when noise reduction option is switched off in their hearing aid if the binaural cues are not preserved in the noise reduction algorithms [4]. Furthermore, in the literature many more advantages of binaural signal processing have been underlined [5–9]. Motivated by this fact that the spatial perception can be used to increase the speech intelligibility [10, 11], many binaural noise reduction algorithms have been developed [12–14], [9].

The authors in [15] have proposed an extended version of the binaural speech distortion weighted multichannel Wiener filter (SDW-MWF) which was first introduced in [16]. It has been shown that the localization cues can be simultaneously preserved for the target as well as interfering signals. Furthermore, a new tradeoff parameter has been introduced to control the amount of noise reduction and binaural cues preservation. The performance of the algorithm has been demonstrated through subjective test [17]. In [18] a new binaural noise reduction with a stereo setup has been proposed. By forming a complex signal using the left microphone as real part and the right microphone as the imaginary part, the binaural noise reduction has been converted to a single-channel noise reduction. The optimal filter was derived based on the widely linear estimation theory taking advantages of the specific statistical characteristic of complex random variables, so-called noncircularity. It has been shown that it can be possible to recover the spatial information of the speech source while reducing noise by forming a time-domain MVDR noise reduction filter under the new proposed framework. In the study carried out in [19], the method which was proposed before in [20] has been extended to a Wiener filter framework. At the first stage of this binaural speech enhancement with Wiener filter (TS-BASE/WF), the interference signals are estimated by equalizing and canceling the target signal. In the second stage, a time-variant Wiener filter is applied to enhance the noisy signal given the noisy mixture signals. A binaural noise reduction algorithm based on the minimum variance distortionless response (MVDR) concept has been presented in [21]. It utilized the interaural impulse response between left and right ear to form the MVDR beamformer. It has been discussed that the accurate estimation of the interaural impulse response would affect the performance of the algorithm.

In this paper, a binaural noise reduction based on MFA using the coherence between two microphones is proposed. First the interaural impulse response is estimated using the adaptive NLMS algorithm. We greatly improve the performance of interaural impulse response estimation in [21] by using voice activity detection (VAD). Then the matched filter array is formed using the interaural impulse response in order to remove the noise with the constraints to keep the speech undistorted. The subsequent post-filter-based stage reduces the residual noise. We discuss the preservation of binaural cues, specifically ILD and ITD, in different speech source configuration.

The remainder of this paper is organized as follows. In Sec. 2 the binaural signal model is described. Sec. 3, investigates the interaural impulse response estimation. In Sec. 4 the proposed binaural noise reduction algorithm is explained. Finally, the experimental results and conclusions are presented in Secs. 6 and 7, respectively.

2. BINAURAL SIGNAL MODEL

Figure 1 depicts the schematic diagram of the signal model in conjunction with proposed binaural noise reduction.

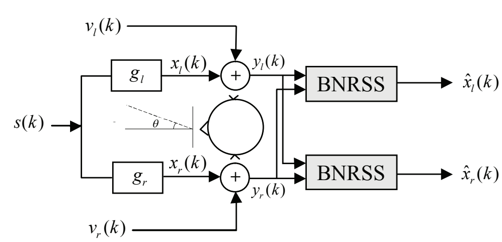


Fig. 1. Signal model and binaural noise reduction system.

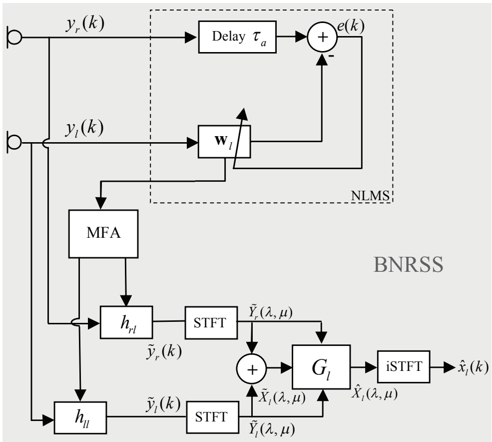


Fig. 2. The binaural noise reduction subsystem (left ear).

The source signal $s(k)$ at sampling time k convolved with acoustic room impulse responses $g_{r,l}$ is captured with two microphones one on each side of the head, i.e., $x_l(k)$ and $x_r(k)$. The subscripts $(.)_l$ and $(.)_r$ denote the signals corresponding to the left and right microphone respectively. Considering the ambient noise, the signal model can be expressed as follows,

$$y_i(k) = x_i(k) + v_i(k) = s(k) * g_i + v_i(k), \qquad i \in \{l, r\}, \quad (1)$$

where $*$ represents linear convolution, $v_i(k)$ and $y_i(k)$ are the ambient noise and the noisy signal received at the left and right microphones, respectively. It is assumed that signals $v_i(k)$ and $x_i(k)$ are zero mean and uncorrelated. The acoustic room impulse responses include room acoustics, microphone characteristics, and head shadow effect. By stacking the left and the right signal, the signal model given in (1) is cast into vector form,

$$\mathbf{y}(k) = \mathbf{x}(k) + \mathbf{v}(k) \quad (2)$$

where

$$\mathbf{y}(k) = \begin{bmatrix}\mathbf{y}_l^T(k) & \mathbf{y}_r^T(k)\end{bmatrix}^T, \quad (3)$$

is a stacked vector of length $2L$ and $\mathbf{y}_i$ is accumulation of L successive samples for the left and right microphones,

$$\mathbf{y}_i(k) = \begin{bmatrix}y_i(k) & y_i(k-1) & \ldots & y_i(k-L+1)\end{bmatrix}^T, i \in \{r, l\}, \quad (4)$$

where the superscript $(.)^T$ denotes matrix transposition. Furthermore, $\mathbf{x}(k)$ and $\mathbf{v}(k)$ are defined in the same way as $\mathbf{y}(k)$. The received noisy signals at the left and right ears are passed through the binaural noise reduction subsystem (BNRSS) which is illustrated in Fig. 2 for the left ear. In the first stage, the interaural impulse response is estimated by means of the adaptive normalized least-mean-square algorithm (NLMS). The matched filter array (MFA) filters h_i are then formed using the estimated interaural impulse response. Due to the relationship with the MVDR, the MFA filters are designed in such a way to reduce the noise while keeping the speech distortion at an acceptable level. The output of the MFA is further processed by different post-filters G_i to achieve an improved noise suppression.

3. ADAPTIVE INTERAURAL-IMPULSE-RESPONSE IDENTIFICATION

The normalized least mean square (NLMS) algorithm [22] which has been widely used in system identification problems is one of the most popular algorithms available for adaptive filtering. In our study, the NLMS algorithm is utilized to estimate the two interaural impulse responses, $\mathbf{w}_l$ and $\mathbf{w}_r$ between right and left microphones,

$$\begin{bmatrix}x_r(k-\tau_a)\\ x_l(k-\tau_a)\end{bmatrix} = \begin{bmatrix}\mathbf{w}_l^T & 0\\ 0 & \mathbf{w}_r^T\end{bmatrix}\mathbf{x}(k). \quad (5)$$

As shown in Fig. 1 and discussed in [21], to ensure causality in the system-identification problem, a delay τ_a, called *algorithmic delay*, should be introduced. The coefficient vector of the adaptive filter $\hat{\mathbf{w}}_i$ is updated according to the NLMS algorithm using noisy signals,

$$\hat{\mathbf{w}}_i(k+1) = \hat{\mathbf{w}}_i(k) + \frac{\mu(k)}{\varepsilon + \|\mathbf{y}_i(k)\|^2}\mathbf{y}_i(k)e_i(k), i \in \{r, l\}, \quad (6)$$

where ε is a small constant value to avoid division by zero. The step size μ, which governs the rate of convergence was set to a fixed value in our previous work. Since the performance of the adaptive NLMS algorithm highly depends on the step-size, a variable step size based on ideal voice activity detection (VAD) is utilized, i.e., $\mu(k) = 0.2$ in speech presence and zero otherwise. In the current work the coefficient vector is updated just in speech presence frames using the error signal

$$\begin{bmatrix}e_l(k)\\ e_r(k)\end{bmatrix} = \begin{bmatrix}y_r(k-\tau_a)\\ y_l(k-\tau_a)\end{bmatrix} - \begin{bmatrix}\hat{\mathbf{w}}_l & \mathbf{0}\\ \mathbf{0} & \hat{\mathbf{w}}_r\end{bmatrix}\mathbf{y}(k). \quad (7)$$

In [21] the performance of the NLMS algorithm in estimating the interaural room impulse response for different measured binaural room impulse responses (BRIRs) has been evaluated. It has been shown that the determination of the filter length L in reverberant environments is not easy as it depends on the reverberation condition. Furthermore, by increasing the algorithmic delay, the performance of the NLMS algorithm will be increased at the expense of longer latency. Thus, for selecting the preferred value of τ_a and L some trade-off should be taken into account.

4. INTERAURAL-IMPULSE-RESPONSE MFA

The matched filter array (MFA) processing as introduced in [23] has a distinct advantage over delay-and-sum beamforming in terms of SNR improvment for noisy and reverberant environments [21]. The aim is to recover one of the microphone signals using the noisy signals from both microphones while keeping the speech signal undistorted. It can be done so, by applying two temporal filters,

$$\tilde{\mathbf{x}}_i(k) = \sum_{j=l,r}\mathbf{H}_{ji}\mathbf{y}_j(k) = \mathbf{H}_i\mathbf{y}(k), \quad i \in \{l, r\}, \quad (8)$$

where $\mathbf{H}_i = \begin{bmatrix}\mathbf{H}_{li} & \mathbf{H}_{ri}\end{bmatrix}$. The symbol $\mathbf{H}_{il}, i \in \{l, r\}$, represents filtering matrices of size $L \times L$ which are derived from the interaural impulse response. In [24], the MFA was put into perspective with mean-square error and MVDR estimation. It has been demonstrated that the computationally inexpensive MFA can be considered as an approximation of the MVDR beamformer. In the following this approximation will be considered in detail. To this end, it was shown in Sec. 3 that we can estimate the interaural impulse responses $\hat{\mathbf{w}}_i$ between microphones and thus we assume a given linear convolutive relationship between the two microphone approximately as

$$\mathbf{x}(k) = \hat{\mathbf{W}}_i^T\check{\mathbf{x}}_i(k), \quad i \in \{l, r\}, \quad (9)$$

where $\check{\mathbf{x}}_i$ is a suitable extension of $\mathbf{x}_i$ to length $2L-1$ and $\hat{\mathbf{W}}_i = \begin{bmatrix}\hat{\mathbf{W}}_{li}^T & \hat{\mathbf{W}}_{ri}^T\end{bmatrix}, i \in \{l, r\}$. Obviously, $\hat{\mathbf{W}}_{rr} = \hat{\mathbf{W}}_{ll} = \mathbf{I}_a$, where $\mathbf{I}_a$

implements a delay of τ_a samples that is also contained in $\hat{\mathbf{W}}_{rl}$ and $\hat{\mathbf{W}}_{rl}$ according to (5). In contrast to [25], the interaural convolution matrices $\hat{\mathbf{W}}_{rl}$ and $\hat{\mathbf{W}}_{lr}$ are strictly linear convolution matrices of size $L \times (2L-1)$ based on the interaural impulse responses $\hat{\mathbf{w}}_l$ and $\hat{\mathbf{w}}_r$ respectively. Finally, the optimal filter matrix $\mathbf{H}_o$ in the MFA sense is obtained as

$$\mathbf{H}_{i,o} = \mathbf{I}_a \hat{\mathbf{W}}_i. \tag{10}$$

An optimal filter vector $\mathbf{h}_{i,o}$ can be one of the rows of $\mathbf{H}_{i,o}$, i.e.,

$$\mathbf{h}_{i,o} = \mathbf{H}_{i,o}^T \mathbf{u}_d \tag{11}$$

where

$$\mathbf{u}_d = \begin{bmatrix} \underbrace{0 \dots 0}_{\tau_d} & 1 & 0 & \dots & 0 \end{bmatrix}^T$$

is a length L unit vector representing the *decision delay* τ_d. In [21] it has been shown that the decision delay affects the performance of the algorithm. However, there is wide range of decision delays τ_d for which an improvement in noise reduction has been reported.

4.1. Binaural MFA and its Consequence on Binaural Cues

The proposed algorithmic structure is designed to keep the reference microphone signals unaffected. Using (8) and (10) with respect to the target signal components, i.e., $\mathbf{y}(k) = \mathbf{x}(k)$, the estimated speech signals at the left and right microphone can be rewritten as

$$\begin{aligned} \tilde{\mathbf{x}}_l(k) &= \mathbf{I}_a \hat{\mathbf{W}}_l \mathbf{x}(k) \\ \tilde{\mathbf{x}}_r(k) &= \mathbf{I}_a \hat{\mathbf{W}}_r \mathbf{x}(k). \end{aligned} \tag{12}$$

According to (9), there is a convolutive relationship between the left and right signals. By substituting the microphone signals $\mathbf{x}_l(k)$ and $\mathbf{x}_r(k)$ via (9), the recovered speech signals can be expressed as

$$\begin{aligned} \tilde{\mathbf{x}}_l(k) &= \mathbf{I}_a \hat{\mathbf{W}}_l \hat{\mathbf{W}}_r^T \mathbf{x}_r(k) \\ \tilde{\mathbf{x}}_r(k) &= \mathbf{I}_a \hat{\mathbf{W}}_r \hat{\mathbf{W}}_l^T \mathbf{x}_l(k), \end{aligned} \tag{13}$$

where $\hat{\mathbf{W}}_l \hat{\mathbf{W}}_r^T$ and $\hat{\mathbf{W}}_r \hat{\mathbf{W}}_l^T$ are not equal in general. This is due to the time reversal of optimal filter vectors $\mathbf{h}_{i,o}$ which is one of the properties of MFA. We observe that binaural phase is well preserved in $\tilde{\mathbf{x}}_l(k)$ and $\tilde{\mathbf{x}}_r(k)$, but the amplitude of the left and right signals are swapped as suggested by (13). According to the above equations, for keeping the speech unaffected, the output of (13) needs to be further equalized by the inverse of $\hat{\mathbf{W}}_l \hat{\mathbf{W}}_r^T$ and $\hat{\mathbf{W}}_r \hat{\mathbf{W}}_l^T$. By equalizing, the binaural cues would be significantly preserved. In other words, by equalization of the MFA approximation, we would restore the MVDR solution, but at the expense of sacrificing noise reduction as observed [21]. Therefore, a tradeoff between noise reduction and preservation of the binaural cues seems inevitable.

4.2. Estimation of Binaural Cues

The interaural time difference can be easily defined as a time delay of arrival of sound to both ears. One of the most straightforward ways to estimate ITD which is used in this contribution, is to calculate the cross correlation of the impulse responses of left and right ears with each other and measure the time from the center to its maximum. However, there are several techniques to estimate the time difference between left and right microphone signals [26]. Assuming the head is simply modeled by a sphere of radius a and if the source is assumed to be in the far field, the ITD can be computed by [27]

$$ITD = \frac{a}{c} \left(\sin\theta + \theta \right) \tag{14}$$

where $c = 340 \frac{m}{s}$ is the speed of sound and θ is the azimuth angle. Considering $a = 8.5cm$, the maximum time difference between two ears would be $643\mu s$. The interaural level difference in dB is given by the energy difference of the signals at the left and right ear,

$$ILD = 10 \log_{10} \left(\frac{E_l}{E_r} \right), \tag{15}$$

where $E_i = \sum_{k=0}^{M} x_i^2(k), \quad i \in \{l, r\}$.

5. POST-FILTERING STAGE

In the second stage, the output of the MFA is further processed by post-filters in order to increase the degree of noise reduction. The estimated components of microphone signals are obtained as

$$\hat{X}_i(\omega, \lambda) = G_i(\omega, \lambda) \tilde{X}(\omega, \lambda), i \in \{l, r\}. \tag{16}$$

Throughout the considerations, the signals are segmented into overlapping frames of length K. The windowed frames (i.e., square Hanning window) are then transformed into the frequency domain via discrete short-time Fourier transform (STFT) of length M. We obtain $\tilde{X}_i(\omega, \lambda)$ where λ and ω are time indices and frequency bins, respectively. The enhanced time-domain signals are obtained by using inverse short-time Fourier transform and an overlap-add method for the final speech signal reconstruction.

Three post-filtering techniques are presented to reduce residual noise.These post-filters are based on the correlation between the two MFA output signals $\tilde{Y}_l(\omega, \lambda)$ and $\tilde{Y}_r(\omega, \lambda)$ which are already aligned by the MFA. The resulting correlation controls the gain of each frequency band. As these filters are real-valued, the phase of signal and noise are kept, therefore no distortion will be introduced on ITD.

5.1. Zelinski Post-filter

In [28] a post-filter G_z was proposed which has been developed in further studies [29]. The filter gain is estimated from the cross and auto-power spectral densities of two channels,

$$G_z(\omega, \lambda) = \frac{Re\{\hat{\Phi}_{\tilde{y_l}\tilde{y_r}}(\omega, \lambda)\}}{\frac{1}{2}(\hat{\Phi}_{\tilde{y_l}\tilde{y_l}}(\omega, \lambda) + \hat{\Phi}_{\tilde{y_r}\tilde{y_r}}(\omega, \lambda))} \tag{17}$$

where $\hat{\Phi}_{\tilde{y_l}\tilde{y_r}}$ denotes the cross-power spectral density defined as $\hat{\Phi}_{\tilde{y_l}\tilde{y_r}}(\omega, \lambda) = E\{\tilde{Y}_l(\omega, \lambda)\tilde{Y}_r^*(\omega, \lambda)\}$, and $\hat{\Phi}_{\tilde{y_i}\tilde{y_i}}$ denotes the auto-power spectral density defined as $\hat{\Phi}_{\tilde{y_i}\tilde{y_i}}(\omega, \lambda) = E\{|\tilde{Y}_i(\omega, \lambda)|^2\}$.

5.2. Doerbecker Post-filter

In [30] an adaptive Wiener post-filter G_d based on what was proposed in [28] has been introduced. It is exactly the square of another gain function which has been proposed by Allen et al. [31]:

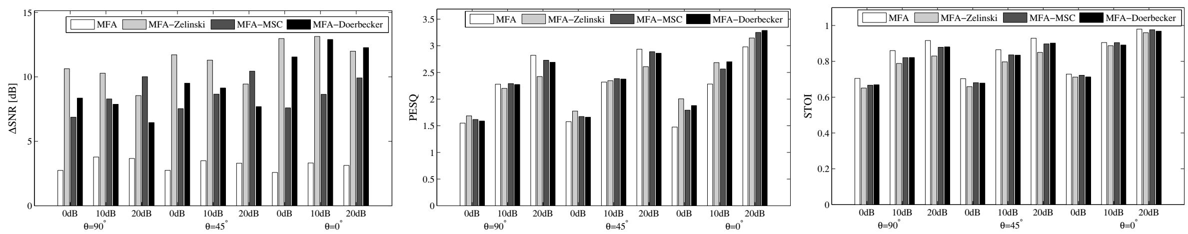
$$G_d(\omega, \lambda) = 4. \frac{\left| \hat{\Phi}_{\tilde{y_l}\tilde{y_r}}(\omega, \lambda) \right|^2}{(\hat{\Phi}_{\tilde{y_l}\tilde{y_l}}(\omega, \lambda) + \hat{\Phi}_{\tilde{y_r}\tilde{y_r}}(\omega, \lambda))^2} \tag{18}$$

5.3. Magnitude Squared Coherence Post-filter

The idea of applying the coherence function as a gain function for speech enhancement was first proposed in [31]. The magnitude squared coherence (MSC) can be expressed as

$$G_{MSC}(\omega, \lambda) = \frac{|\Phi_{\tilde{y_l}\tilde{y_r}}(\omega, \lambda)|^2}{\hat{\Phi}_{\tilde{y_l}\tilde{y_l}}(\omega, \lambda)\hat{\Phi}_{\tilde{y_r}\tilde{y_r}}(\omega, \lambda)} \tag{19}$$

where G_{MSC} is always bounded between zero and one.



(a) (b) (c)

Fig. 3. Comparison in terms of (a) ΔSNR, (b) PESQ, and (c) STOI for different algorithms in different speech configurations ($\tau_d = 400$ samples and $T_{60} = 0.69$s)

6. EXPERIMENTAL RESULTS

In order to evaluate the performance of the discussed algorithms, experiments have been carried out with the *measured* "stairway" impulse responses ($T_{60} = 690$ms, and $0° \leq \theta \leq 90°$) from the Aachen room impulse responses database [32]. The distance from the loud-speaker to the dummy head is $d = 1$m. The left and the right microphone outputs are generated by convolving the source signal (female speaker) with the binaural impulse responses. In our simulation, the additive observation noise signals picked up by the microphones are independent white Gaussian noises.

In this contribution we are just relying on (8) and (10) and presenting the result of the swapped outputs. The noise signals are added to the convolved signals. These noisy mixtures are also used as the excitation signal for the NLMS algorithm as shown in Fig.2. To evaluate the performance of the algorithm we used three performance measures, i.e., the segmental SNR which will be defined in the following, the perceptual evaluation of speech quality (PESQ) [33], and short-time objective intelligibility (STOI) [34]. It should not be overlooked that all the results are obtained in highly reverberant room in which the target speech signal is affected by reverb as well as ambient noise. The input and output signal-to-noise ratios, SNR_i and SNR_o, respectively, are evaluated as

$$\text{SNR}_i = 10\text{log}\{\frac{E[x_l^2(k)] + E[x_r^2(k)]}{E[v_l^2(k)] + E[v_r^2(k)]}\}, \tag{20}$$

$$\text{SNR}_o = 10\text{log}\{\frac{E[x_{l,f}^2(k)] + E[x_{r,f}^2(k)]}{E[v_{l,f}^2(k)] + E[v_{r,f}^2(k)]}\}, \tag{21}$$

where $x_{i,f}(k) = G\circ(\mathbf{h}_{i,o}^T\mathbf{x}(k))$ and $v_{i,f}(k) = G\circ(\mathbf{h}_{i,o}^T\mathbf{v}(k)),\ \ i \in \{l, r\}$, are the filtered speech and the filtered noise, respectively, and $\circ$ denotes concatenation of the MFA and the post-filtering.

The result of SNR improvement for different input SNR and different azimuth angles is illustrated in Fig. 3.a. As it is shown, the Zelinski post-filter outperforms all other algorithms. Moreover, the Doerbecker and Zelinski post-filter obtain fairly the same output SNR since the Doerbecker post-filter has been designed based on the Zelinski post-filter. Contrary to MSC post-filter and plain MFA, the output SNR in Doerbecker and Zelinski post-filters drop by increasing the input SNR. Figs. 3.b and 3.c demonstrate the impact of post-filtering on speech quality and intelligibility. As it can be seen, the plain MFA algorithm significantly achieves a better speech intelligibility than in combination with the post-filter algorithms according to STOI measure. It should be also taken into account that the MFA does not really significantly improve PESQ scores in comparison to the other algorithm in low SNR (i.e., 0 dB). This is due to the high level of the residual noise which outweighs speech quality.

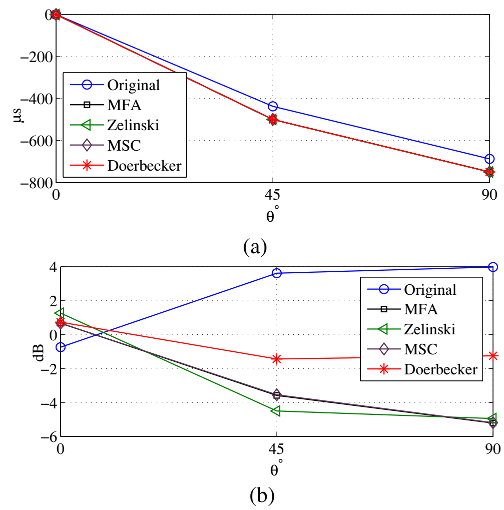


Fig. 4. Comparison of (a) ITD and (b) ILD for different azimuth.

Furthermore, it turns out that post-filtering slightly reduces speech quality and intelligibility measures.

The influence of the algorithms on binaural cues will be discussed as follows. Fig. 4.a shows the ITD as a function of azimuth angle. It clearly can be seen that not much distortion is introduced in ITD. The ILD for different speech configuration is depicted in Fig. 4.b which is affected by all tested algorithms cf. Sec.4.1. The lowest influence can be found for the frontal direction (0°).

7. CONCLUSION

This paper introduced a binaural noise reduction system. First the interaural room impulse responses between the left and right ear were estimated by means of the adaptive NLMS algorithm. The estimated interaural impulse responses were then utilized to form MFA filters. On MVDR basis these filters are formed to keep one of the microphone signals, i.e., the reference microphone, undistorted. It was shown that the MFA as a rough approximation of the MVDR beamformer results in amplitude errors which does not lead to consistent binaural cues. By equalization, the binaural cues could be preserved better at the expense of reducing the amount of noise reduction. Furthermore, three post-filters were considered to further suppress noise in the output signal. It turned out that the post-filtering boosts the output SNR at the price of introducing slightly more speech distortion, e.g., in terms of STOI.

8. REFERENCES

[1] C. Ludvigsen, C. Elberling, and G. Keidser, “Evaluation of a noise reduction method: Comparison between observed scores and scores predicted from STI,” in *Scand. Audiol. Suppl.*, 1993, vol. 38, pp. 50 – 55.

[2] J. M. Kates, “Digital hearing aids,” San Diego, CA: Plural Publishing, 2008.

[3] H. Dillon, “Hearing aids,” 2001, Boomerang Press.

[4] T. Van den Bogaert, T. J. Klasen, M. Moonen, L. Van Deun, and J. Wouters, “Horizontal localization with bilateral hearing aids: Without is better than with,” in *J. Acoust. Soc. Am.*, 2006, vol. 119, pp. 515–526.

[5] P.B. Balfour and D.B. Hawkins, “A comparison of sound quality judgments for monaural and binaural hearing aid processed stimuli,” in *Ear Hear*, 1992, vol. 13, pp. 331 – 339.

[6] S.A. Erdman and R.K. Sedge, “Subjective comparisons of binaural versus monaural amplification,” in *Ear Hear*, 1981, vol. 2, pp. 225 – 229.

[7] V. Hamacher, “Comparison of advanced monaural and binaural noise reduction algorithms for hearing aids,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2002, vol. 4, pp. 4008–4011.

[8] J. Blauert, “Spatial hearing, the psychophysics of human sound localization,” 1997, The MIT Press, Cambridge, MA.

[9] H. Huang and C. Kyriakakis, “Binaural model based adaptive binaural noise reduction,” in *Asilomar Conf. on Signals, Systems & Computers (ACSSC)*, 2006, pp. 1110 –1113.

[10] A. W. Bronkhorst and R. Plomp, “The effect of head-induced interaural time and level differences on speech intelligibility in noise,” in *J. Acoust. Soc. Am.*, 1988, vol. 83, pp. 1508–1516.

[11] T. Van den Bogaert, T. J. Klasen, M. Moonen, and J. Wouters, “Distortion of interaural time cues by directional noise reduction systems in modern digital hearing,” in *Proc. IEEE Workshop Applicat. Signal Process. Audio Acoust. (WASPAA)*, 2005, pp. 57 – 60.

[12] S. Markovich Golan, S. Gannot, and I. Cohen, “A reduced bandwidth binaural MVDR beamformer,” in *Proc. Int. Workshop on Acoustic Echo and Noise Control (IWAENC)*, 2010.

[13] S. Doclo, M. Moonen, T. Van den Bogaert, and J. Wouters, “Reduced-bandwidth and distributed MWF-based noise reduction algorithms for binaural hearing aids,” in *IEEE Trans. Audio, Speech, Lang. Process.*, 2009, vol. 17, pp. 38 – 51.

[14] K. Reindl, Y. Zheng, and W. Kellermann, “Analysis of two generic Wiener filtering concepts for binaural speech enhancement in hearing aids,” in *Proc. European Signal Processing Conf. (EUSIPCO)*, 2010, pp. 989–993.

[15] T. Klasen, T. Van den Bogaert, M. Moonen, and J. Wouters, “Binaural noise reduction algorithms for hearing aids that preserve interaural time delay cues,” in *IEEE Trans. Signal Process.*, 2007, vol. 55, p. 15791585.

[16] T. Klasen, T. Van den Bogaert, M. Moonen, and J. Wouters, “Preservation of interaural time delay for binaural hearing aids through multichannel wiener filtering based noise reduction,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2005, vol. 3, p. 29 32.

[17] T. Van den Bogaert, S. Doclo, J. Wouters, and M. Moonen, “The effect of multimicrophone noise reduction systems on sound source localization by users of binaural hearing aids,” in *J. Acoust. Soc. Amer.*, 2008, vol. 124, p. 484 497.

[18] J. Benesty, J. Chen, and Y. Huang, “Binaural noise reduction in the time domain with a stereo setup,” in *IEEE Trans. Audio, Speech, Lang. Process.*, 2011, vol. 19, pp. 2260 –2272.

[19] J. Li, S. Sakamoto, S. Hongo, M. Akagi, and Y. Suzuki, “Two-stage binaural speech enhancement with Wiener filter for high-quality speech communication,” in *Speech Commun.*, 2011, vol. 53, pp. 677 – 689.

[20] H. Nakashima, Y. Chisaki, T. Usagawa, and M. Ebata, “Frequency domain binaural model based on interaural phase and level differences,” in *Acoustic. Sci. Tech.*, 2003, vol. 24, p. 172178.

[21] M. Azarpour, G. Enzner, and R. Martin, “Distortionless-response vs. matched-filter-array processing for adaptive binaural noise reduction,” in *Proc. Int. Workshop on Acoustic Echo and Noise Control (IWAENC)*, 2012, pp. 1 –4.

[22] S. Haykin, “Least-mean-square algorithm,” in *Adaptive Filter Theory*. Prentice Hall, Upper Saddle River, New Jersey, 2001, pp. 367–438.

[23] R. J. Renomeron, D. V. Rabinkin, J. C. French, and J. L. Flanagan, “Small-scale matched filter array processing for spatially selective sound capture,” in *J. Acoust. Soc. Am.*, 1997, vol. 102, pp. 3208–3217.

[24] D. Schmid and G. Enzner, “A parametric least-squares approximation for multichannel equalization of room acoustics,” in *Proc. Int. Workshop on Acoustic Echo and Noise Control (IWAENC)*, 2010.

[25] J. Chen, J. Benesty, and Y. Huang, “A minimum distortion noise reduction algorithm with multiple microphones,” in *IEEE Trans. Audio, Speech, Lang. Process.*, 2008, vol. 16, pp. 481–493.

[26] J. Estrella, “On the extraction of interaural time differences from binaural room impulse responses,” in *Master thesis at the Technische Universität Berlin*, 2010.

[27] R. S. Woodworth and G. Schlosberg, “Experimental psychology,” Holt, Rinehard and Winston, 1962, pp. 349–361.

[28] R. Zelinski, “A microphone array with adaptive post-filtering for noise reduction in reverberant rooms,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 1988, vol. 5, pp. 2578 –2581.

[29] I.A. McCowan and H. Bourlard, “Microphone array post-filter based on noise field coherence,” in *IEEE Trans. Audio, Speech, Lang. Process.*, 2003, number 6, pp. 709 – 716.

[30] M. Dörbecker and S. Ernst, “Combination of two-channel spectral subtraction and adaptive wiener post-filtering for noise reduction and dereverberation,” in *Proc. European Signal Processing Conf. (EUSIPCO)*, 1996, p. 995 998.

[31] J. B. Allen, D. A. Berkley, and J. Blauert, “Multi-microphone signal processing technique to remove room reverberation from speech signals,” in *J. Acoust. Soc. Amer.*, 1977, vol. 62, p. 912915.

[32] M. Jeub, M. Schaefer, and P. Vary, “A binaural room impulse response database for the evaluation of dereverberation algorithms,” in *Int. Conf. on Digital Signal Processing (DSP)*, 2009, pp. 1 –5.

[33] A.W. Rix, J.G. Beerends, M.P. Hollier, and A.P. Hekstra, “Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2001, vol. 2, pp. 749–752.

[34] C.H. Taal, R.C. Hendriks, R. Heusdens, and J. Jensen, “A short-time objective intelligibility measure for time-frequency weighted noisy speech,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2010, pp. 4214 –4217.