COMPARISON OF REDUCED-BANDWIDTH MWF-BASED NOISE REDUCTION ALGORITHMS FOR BINAURAL HEARING AIDS

Simon Doclo, Tim van den Bogaert, Jan Wouters, Marc Moonen *

Katholieke Universiteit Leuven
Dept. of Electrical Engineering, SCD
Kasteelpark Arenberg 10, 3001 Leuven, Belgium
simon.doclo@esat.kuleuven.be

Katholieke Universiteit Leuven
Dept. of Neurosciences, ExpORL
Herestraat 49/721, 3000 Leuven, Belgium
tim.vandenbogaert@med.kuleuven.be

ABSTRACT

In a binaural hearing aid noise reduction system, binaural output signals are generated by sharing information between the two hearing aids. When each hearing aid has multiple microphones and all microphone signals are transmitted between the hearing aids, a significant noise reduction can be achieved using the binaural multi-channel Wiener filter (MWF). To limit the number of signals being transmitted between the hearing aids, in order to comply with bandwidth constraints of the binaural link, this paper presents reduced-bandwidth MWF-based algorithms, where each hearing aid uses only a filtered combination of the contralateral microphone signals. One algorithm uses the output of a monaural MWF on the contralateral microphone signals, whereas a second algorithm involves a distributed binaural MWF scheme. Experimental results compare the performance of the presented algorithms.

1. INTRODUCTION

Noise reduction algorithms in hearing aids are crucial to improve the speech intelligibility in background noise for hearing impaired persons. Since multi-microphone systems are able to exploit spatial information, they are typically preferred to single-microphone systems. In a dual hearing aid system, output signals for both ears are generated, either by operating both hearing aids independently (a bilateral system) or by sharing information between the hearing aids (a binaural system) [1]-[6], e.g. using a wireless link. In [3], a binaural multi-channel Wiener filter technique has been proposed that produces an estimate of the desired speech signal component in both hearing aids. It has been shown that this technique -and its extensions- achieves significant noise reduction and also partly preserves the binaural localisation cues [3, 4, 5]. Since this binaural MWF, which will be reviewed in Section 3, optimally exploits all microphone signals from both hearing aids, all microphone signals need to be transmitted over the binaural link, requiring a large bandwidth. To reduce the bandwidth requirement, alternative techniques are presented in Section 4, where each hearing aid uses only one signal transmitted from the contralateral ear. Suboptimal techniques either using the front contralateral microphone signal or the output of a monaural MWF are presented, together with an iterative distributed MWF scheme that remarkably converges to the optimal binaural MWF solution in the case of a single speech source. In Section 5 the SNR improvement and the directivity pattern of all algorithms are compared in a realistic setup, showing that the distributed binaural MWF scheme has the best performance of all reduced-bandwidth techniques and indeed approaches the optimal binaural MWF performance.

2. CONFIGURATION AND NOTATION

Consider the binaural hearing aid configuration depicted in Figure 1, where both hearing aids have a microphone array consisting of M microphones. The mth microphone signal in the left hearing aid Y_{0,m}(\omega) can be written in the frequency-domain as

\[ Y_{0,m}(\omega) = X_{0,m}(\omega) + V_{0,m}(\omega), \quad m = 0 \ldots M - 1, \] (1)

where \( X_{0,m}(\omega) \) represents the speech component and \( V_{0,m}(\omega) \) represents the noise component. Similarly, the mth microphone signal in the right hearing aid is \( Y_{1,m}(\omega) = X_{1,m}(\omega) + V_{1,m}(\omega) \).

For conciseness we will omit the frequency-domain variable in the remainder of the paper. We define the M-dimensional stacked vectors \( \mathbf{Y}_0 \) and \( \mathbf{Y}_1 \) and the 2M-dimensional signal vector \( \mathbf{Y} \) as

\[
\mathbf{Y}_0 = \begin{bmatrix} Y_{0,0} \\ \vdots \\ Y_{0,M-1} \end{bmatrix}, \quad \mathbf{Y}_1 = \begin{bmatrix} Y_{1,0} \\ \vdots \\ Y_{1,M-1} \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} \mathbf{Y}_0 \\ \mathbf{Y}_1 \end{bmatrix}.
\] (2)

The signal vector can be written as \( \mathbf{Y} = \mathbf{X} + \mathbf{V} \), with \( \mathbf{X} \) and \( \mathbf{V} \) defined similarly as \( \mathbf{Y} \). In the case of a single speech source, the speech signal vector can be written as \( \mathbf{X} = \mathbf{A}\mathbf{S} \), with the 2M-dimensional steering vector \( \mathbf{A} \) containing the acoustic transfer functions between the speech source and the microphones (including room acoustics, microphone characteristics and head shadow) and \( \mathbf{S} \) the speech signal. The vector \( \mathbf{A} \) is defined similarly as \( \mathbf{Y} \).

\[ \mathbf{Y}_0(\omega) \]
\[ \mathbf{Y}_1(\omega) \]

\[ \mathbf{W}_{00}(\omega) \quad \mathbf{G}_{01}(\omega) \]
\[ \mathbf{Y}_{01}(\omega) \]
\[ \mathbf{W}_{10}(\omega) \quad \mathbf{G}_{10}(\omega) \]
\[ \mathbf{F}_{01}(\omega) \]
\[ \mathbf{Z}_0(\omega) \]
\[ \mathbf{Z}_1(\omega) \]

Figure 1: General binaural processing scheme

*Simon Doclo is a postdoctoral researcher supported by the Fund for Scientific Research - Flanders. This work was carried out in the frame of GOA-AMBIORICS, CoE EF/05/006 and IAP P6/04.
In a binaural processing scheme, collaboration between both hearing aids is achieved by transmitting signals between the hearing aids (e.g., using a wireless link). The signals transmitted from the left (right) hearing aid to the right (left) hearing aid are respectively represented by the \( N \)-dimensional vectors \( Y_{10} \) and \( Y_{01} \), typically with \( N \leq M \). We assume that the transmitted signals are a linear combination of the contralateral microphone signals, i.e.

\[
Y_{10} = F_{10}^H Y_0, \quad Y_{01} = F_{01}^H Y_1, \tag{3}
\]

where \( F_{10} \) and \( F_{01} \) are \( M \times N \)-dimensional complex matrices. The output signals \( Z_0 \) and \( Z_1 \) for the left and the right ear are obtained by filtering and summing the ipsilateral microphone signals and the transmitted signals from the contralateral ear, i.e.,

\[
Z_0 = W_{00}^H Y_0 + G_{01}^H Y_1 = W_{00}^H Y_0 + G_{01}^H F_{01}^H Y_1, \tag{4}
\]

\[
Z_1 = G_{10}^H Y_{10} + W_{11}^H Y_1 = G_{10}^H F_{10}^H Y_0 + W_{11}^H Y_1, \tag{5}
\]

where \( W_{00} \) and \( W_{11} \) are \( M \)-dimensional vectors and \( G_{01} \) and \( G_{10} \) are \( N \)-dimensional vectors. Hence, the output signals can be written as linear combinations of all microphone signals, i.e., \( Z_0 = W_{00}^H Y \) and \( Z_1 = W_{11}^H Y \), where the \( 2M \)-dimensional vectors \( W_0 \) and \( W_1 \) are given as

\[
W_0 = \begin{bmatrix} W_{00} \\ W_{01} \end{bmatrix}, \quad W_1 = \begin{bmatrix} W_{10} \\ W_{11} \end{bmatrix}, \quad F_{10} = \begin{bmatrix} F_{10} G_{10} \\ F_{10} G_{11} \end{bmatrix}.
\]

### 3. BINAURAL MULTI-CHANNEL WIENER FILTER

The binaural MWF (B-MWF) in [3] assumes that all microphone signals are transmitted, i.e. \( F_{10} = F_{01} = I_M \). The binaural MWF produces an MMSE (minimum-mean-square-error) estimate of the speech component in both hearing aids, hence simultaneously performing noise reduction and limiting speech distortion. The MSE cost function for the filter \( W_0 \) estimating the speech component \( X_{0,0} \) in the front microphone of the left hearing aid is equal to

\[
J_{MSE,0}(W_0) = \mathcal{E}\{ |X_{0,0} - W_0^H Y|^2 \}. \tag{6}
\]

In order to provide a trade-off between speech distortion and noise reduction, the speech distortion weighted multi-channel Wiener filter (SDW-MWF) minimises the weighted sum of the residual noise energy and the speech distortion energy [7], i.e.,

\[
J_{0}(W_0) = \mathcal{E}\{ |X_{0,0} - W_0^H X|^2 \} + \mu \mathcal{E}\{ |W_0^H V|^2 \}, \tag{7}
\]

where \( \mu \) is a trade-off parameter. Similarly, the SDW-MWF cost function for the filter \( W_1 \) estimating the speech component \( X_{1,0} \) in the front microphone of the right hearing aid is equal to

\[
J_{1}(W_1) = \mathcal{E}\{ |X_{1,0} - W_1^H X|^2 \} + \mu \mathcal{E}\{ |W_1^H V|^2 \}. \tag{8}
\]

The filters \( W_{00} \) and \( W_{11} \) minimising (7) and (8) are equal to

\[
W_{00}^m = (R_0 + \mu R_v)^{-1} R_v e_0, \tag{9}
\]

\[
W_{11}^m = (R_1 + \mu R_v)^{-1} R_v e_1, \tag{10}
\]

where \( R_s \) and \( R_v \) are the speech and the noise correlation matrix, i.e. \( R_s = \mathcal{E}\{ X X^H \} \) and \( R_v = \mathcal{E}\{ V V^H \} \), and \( e_0 \) and \( e_1 \) are vectors of which only one element is equal to 1 and the other elements are equal to 0, with \( e_0(1) = 1 \) and \( e_1(M+1) = 1 \).

In the case of a single speech source, the speech correlation matrix is a rank-1 matrix, i.e. \( R_s = P_s A A^H \), with \( P_s = \mathcal{E}\{ |S|^2 \} \) the power of the speech signal. Using the matrix inversion lemma, the filters \( W_{00}^m \) and \( W_{11}^m \) are then found to be equal to [4]

\[
W_{00}^m = \frac{R_0^{-1} A}{A^H R_v^{-1} A + \frac{\mu}{\mu}} A_0^0, \tag{11}
\]

\[
W_{11}^m = \frac{R_1^{-1} A}{A^H R_v^{-1} A + \frac{\mu}{\mu}} A_1^0, \tag{12}
\]

with \( A_{0,0} \) and \( A_{1,0} \) elements of \( A \), cf. (2). This implies that

\[
W_{01}^m = a W_{00}^m \quad \text{and} \quad W_{10}^m = a W_{11}^m
\]

where \( a = A_{1,0}^* / A_{0,0}^* \) is the complex conjugate of the interaural transfer function [4] of the speech component.

### 4. REDUCED-BANDWIDTH MWF ALGORITHMS

The binaural MWF in Section 3 exploits all microphone signals, requiring \( 2N = 2M \) signals to be transmitted over the binaural link. However, due to power limitations the bandwidth of the link typically does not allow to transmit all microphone signals. This section presents MWF-based algorithms that use only one signal transmitted from the contralateral ear, i.e. \( N = 1 \), reducing \( F_{01} \) to \( M \)-dimensional vectors and \( G_{01} \) and \( G_{10} \) to scalars. It is still possible to obtain the optimal B-MWF performance, namely if \( F_{01} = W_{00}^m \) and \( F_{10} = W_{10}^m \) (up to a complex scaling), assuming that \( W_{01}^m \) and \( W_{11}^m \) can be computed without all microphone signals being transmitted. First, we present suboptimal solutions, either using the front contralateral microphone signal or the output of a monaural MWF. Although it seems impossible at first sight to obtain the optimal B-MWF performance without transmitting all microphone signals, in Section 4.3 we present an iterative distributed MWF scheme that converges to the optimal B-MWF solution in the case of a single speech source.

#### 4.1. Front contralateral microphone signal (MWF-front)

In this simple scheme, only the front contralateral microphone signals are transmitted, i.e. \( F_{10} = F_{01} = [1 \quad 0 \quad \ldots \quad 0]^T \).

#### 4.2. Contralateral MWF (MWF-contra)

In this scheme, the transmitted signals are the output of a monaural MWF, estimating the contralateral speech component only using the \( M \) contralateral microphone signals. Hence, the filters \( F_{10} \) and \( F_{01} \) are respectively minimising the cost functions

\[
J_0^e(F) = \mathcal{E}\{ |X_{0,0} - F^H X|^2 \} + \mu \mathcal{E}\{ |F^H V|^2 \}, \tag{14}
\]

\[
J_1^e(F) = \mathcal{E}\{ |X_{1,0} - F^H X|^2 \} + \mu \mathcal{E}\{ |F^H V|^2 \}. \tag{15}
\]

The resulting filters can be written, using \( Q_0 = [I_M \quad 0_M] \) and \( Q_1 = [0_M \quad I_M] \), as

\[
F_{10} = [Q_0 (R_s + \mu R_v) Q_1^T]^{-1} Q_0 R_v e_0 \tag{16}
\]

\[
F_{01} = [Q_1 (R_s + \mu R_v) Q_0^T]^{-1} Q_1 R_v e_1. \tag{17}
\]

In general, this solution is suboptimal, since it can be shown that the optimal solution, i.e. \( F_{01} \) being a scaled version of \( W_{01}^m \) and \( F_{10} \) being a scaled version of \( W_{10}^m \), is only obtained in the case of a single speech source and if no correlation exist between the noise components on the left and the right hearing aid. In addition, two MWF solutions need to be computed on each hearing aid, e.g. for the left hearing aid an \( M \)-dimensional MWF for computing \( F_{10} \) and an \( (M+1) \)-dimensional MWF for computing \( W_{00} \) and \( G_{01} \).
4.3. Distributed binaural MWF scheme (dB-MWF)

The distributed binaural MWF scheme is depicted in Figure 2. Basically, in each iteration the filter $F_0$ is equal to $W_{00}$ from the previous iteration, and the filter $F_{11}$ is equal to $W_{11}$ from the previous iteration. If we denote the filters and the signals in the $i$th iteration with superscript $i$, then the iterative procedure runs as:

1. Transmit $Y_{01} = W_{i1}^0Y_1$ to the left hearing aid.
2. Using $Y_{00}$ and $Y_{10}$ as input signals, calculate $W_{00}^i$ and $G_{10}^i$ that minimise the SDW-MWF cost function estimating the speech component in the left front microphone, i.e.
   \[
   J_0(W_{00}^i, G_{10}^i) = E\{[X_{0,0} - (W_{00}^iX_0 + G_{01}^i X_{10})]^2\}
   \]
   \[+\mu E\{|W_{00}^i X_0 + G_{01}^i X_{10}^2\}\].
3. Transmit $Y_{10} = W_{i0}^1Y_0$ to the right hearing aid.
4. Using $Y_{11}$ and $Y_{01}$ as input signals, calculate $W_{11}^{i+1}$ and $G_{10}^{i+1}$ that minimise the SDW-MWF cost function estimating the speech component in the right front microphone, i.e.
   \[
   J_1(W_{11}^{i+1}, G_{10}^{i+1}) = E\{[X_{1,0} - (W_{11}^{i+1}X_1 + G_{10}^{i+1} X_{01})]^2\}
   \]
   \[+\mu E\{|W_{11}^{i+1} X_1 + G_{10}^{i+1} X_{01}^2\}\].

Note that the filters $W_{00}^i$ and $W_{11}^{i+1}$ are hence structured as
\[
W_{00}^i = \begin{bmatrix} W_{00}^i \\ G_{10}^iW_{11}^i \end{bmatrix}, \quad W_{11}^{i+1} = \begin{bmatrix} G_{10}^iW_{00}^i \\ W_{11}^{i+1} \end{bmatrix}, \quad (18)
\]
such that the following holds at convergence, i.e. for $i \to \infty$,
\[
\begin{bmatrix} W_{00}^{\infty} \\ G_{10}^{\infty}W_{11}^{\infty} \end{bmatrix} = \begin{bmatrix} G_{10}^{\infty}W_{00}^{\infty} \\ 1/G_{01}^{\infty}W_{00}^{\infty} \end{bmatrix}. \quad (19)
\]

In the case of a *single speech source*, it can be proven that the SDW-MWF cost functions are decreasing in each iteration, i.e.
\[
J_0(W_{00}^{i+1}) \leq J_0(W_{00}^i), \quad J_1(W_{11}^{i+1}) \leq J_1(W_{11}^i). \quad (20)
\]
Since the optimal filters $W_{00}^m$ and $W_{11}^m$ in (11) and (12) satisfy (19), with $G_{01}^{\infty} = \alpha$ and $G_{10}^{\infty} = 1/\alpha$, the distributed binaural MWF scheme converges to the optimal B-MWF solution in the case of a rank-1 speech correlation matrix. However, in the case of a full-rank speech correlation matrix, the proposed dB-MWF scheme does not converge to the optimal filters $W_{00}^m$ and $W_{11}^m$ in (9) and (10), as these filters do not satisfy (19). Nevertheless, it is shown in Section 5 that this procedure can still be used in practice and approaches the optimal B-MWF performance.

![Figure 2: Distributed binaural MWF scheme (dB-MWF)](image)

5. EXPERIMENTAL RESULTS

5.1. Set-up and performance measures

Two hearing aids with $M = 2$ omni-directional microphones have been mounted on a CORTEX MK2 artificial head in a low reverberant room having a reverberation time $T_{60} \approx 140$ ms. The distance between the microphones on each hearing aid is about 1 cm. Acoustic transfer functions have been measured for positions at a distance of 1 m and at different angles from the head. The sampling frequency is equal to 20.48 kHz. The speech source is positioned in front of the head (0°) and consists of sentences from the HINT database, while multi-talker babble is used as noise source and several noise configurations (single and multiple sources) are considered. For all noise configurations, the input broadband SNR is 0 dB at the front microphone signal of the left hearing aid. The FFT-size used for frequency-domain processing is $L = 96$. Using a perfect voice activity detector, the noise correlation matrices $R_u$ are computed during noise-only periods, the correlation matrices $R_y$ are computed during speech-and-noise periods, and the speech correlation matrices are estimated as $R_x = R_y - R_u$. For all MWF algorithms we have used $\mu = 5$. For the distributed binaural MWF scheme, the number of iterations is $K = 10$, and the filter $W_{11}$ has been initialised as $W_{11}^0 = [1 \ 0 \ldots 0]^T$.

To assess the performance of the different algorithms, the intelligibility weighted SNR improvement [8] between the output and the front microphone signal is used, e.g. for the left hearing aid
\[
\Delta SNR_0 = \sum_i I(\omega_i) \left[ SNR_{\omega_0}(\omega_i) - SNR_{\omega_0,0}(\omega_i) \right], \quad (21)
\]
where $I(\omega_i)$ expresses the importance of the $i$th frequency bin for speech intelligibility. The SNR improvement for the right hearing aid $\Delta SNR_1$ is defined similarly.

5.2. Comparison of SNR improvement and directivity pattern

Figures 3 and 4 plot the SNR improvement at the left and the right hearing aid for several noise configurations for the B-MWF, MWF-front, MWF-contra and dB-MWF algorithms discussed in Sections 3 and 4. In general, for all algorithms the SNR improvement is larger when the speech source and the noise source(s) are spatially more separated, with the largest improvement occurring in the hearing aid where the input SNR is lower.

As expected, the binaural MWF (using 4 microphones) results in the largest SNR improvement for all noise configurations, and MWF-front (using 3 microphones) degrades the performance with 2-4 dB. Although MWF-contra is a suboptimal solution, its performance lies between MWF-front and B-MWF (except for 60° and 300°). The best performance of all reduced-bandwidth algorithms is achieved by the distributed binaural MWF scheme, and compared to MWF-contra a substantial performance benefit is obtained, especially for 60° and 300° and when multiple noise sources are present. However, the performance of dB-MWF does not reach the performance of B-MWF (as theoretically expected for a single speech source), due to the fact that $R_u$ does not have rank 1 because of overlap between adjacent FFT frequency bands and because of estimations errors.

For a noise source at 120°, Figure 5 depicts the SNR improvement of dB-MWF as a function of the number of iterations. Already after two or three iterations the final performance seems to be obtained. For the same noise configuration, Figure 6 plots the fullband spatial directivity pattern of the filter $F_{01}$, i.e. the pattern generated using the right microphone signals and transmitted to the
Performance comparison of MWF-based binaural algorithms

Figure 3: $\Delta$SNR$_0$ for B-MWF, MWF-front, MWF-contra and dB-MWF ($K = 10$) for different noise configurations $\theta_v$

Performance comparison of MWF-based binaural algorithms

Figure 4: $\Delta$SNR$_1$ for B-MWF, MWF-front, MWF-contra and dB-MWF ($K = 10$) for different noise configurations $\theta_v$

left hearing aid. Optimally, i.e., using B-MWF, a null is steered towards the direction of the noise source, implying that a signal with a high SNR should be transmitted. Since this is not the case when transmitting the front microphone signal, the SNR improvement substantially degrades for MWF-front. It can be observed that the directivity patterns obtained with MWF-contra and dB-MWF both also exhibit a null in the direction of the noise source.

6. REFERENCES


Figure 5: SNR improvement at left and right hearing aid with $\theta_v = 120^\circ$ for dB-MWF as a function of the number of iterations

Figure 6: Spatial directivity pattern of $F_{01}$ with $\theta_v = 120^\circ$ for (a) B-MWF, (b) MWF-front, (c) MWF-contra, (d) dB-MWF