Euronoise 2018
Crete

Experimental evaluation of speech enhancement methods in remote microphone systems for hearing aids

Gilles Courtois, Vincent Grimaldi, Hervé Lissek
Swiss Federal Institute of Technology (EPFL), Signal Processing Laboratory (LTS2), Lausanne, Switzerland

Ina Kodrasi
Idiap Research Institute, Martigny, Switzerland

Eleftheria Georganti
Sonova AG, Stäfa, Switzerland

Summary
Enhancing speech intelligibility for hearing-impaired subjects in complex acoustic conditions is still a challenging topic of research. To mitigate the detrimental effects of background noise and reverberation, current hearing instruments incorporate various hardware and software strategies, among which speech enhancement algorithms are of primary importance. In this paper, two algorithms based on the multichannel Wiener filter previously reported in the literature and one proprietary algorithm are experimentally assessed and compared. All of them make use of a remote microphone worn by the speaker of interest. The objective of these algorithms is to improve the speech contribution within the hearing aid microphone signals. The algorithms are assessed in terms of interference reduction performance, speech quality, spatial hearing preservation, and technical requirements. Using a recorded database of audio signals, the effects of the signal-to-noise ratio and of the delay between the remote and hearing aid microphone signals are studied. The results show that the proprietary algorithm provides a good performance and yields the lowest distortion of the binaural localization cues, while being the most efficient in terms of computational cost and wireless usage. The main drawback is the degradation of the output sound quality that is observed when the remote and hearing aid microphone signals are not temporally aligned.

1. Introduction

Understanding speech in complex acoustic conditions is a challenging task for hearing-impaired subjects. Current hearing aids (HAs) incorporate various hardware and software solutions to circumvent the detrimental effects of noise and reverberation, such as using microphone arrays, binaural wireless communication, noise and reverberation reduction, etc. Noise reduction was initially performed independently in both left and right devices using the spectro-temporal properties of the acoustic signals [1]. However, such an approach many times fails to improve speech intelligibility and deteriorates the binaural localization cues, because different gain models are applied on the left and right HAs [2]. Wireless connectivity now allows to share information and stream signals between both devices, yielding new binaural approaches.

Single-channel binaural algorithms usually rely on the computation of the short-term binaural cues so as to discriminate the time-frequency areas dominated by noise and reverberation. Then, real-valued gains are applied to attenuate those undesired components. The computed binaural cues can be the interaural time and level differences (ITD and ILD) [3, 4], as well as the interaural coherence (IC) [5, 6, 7]. These algorithms usually process the left and right signals with identical gain models, which allows to preserve the original binaural localization cues. Improvements in speech intelligibility have been reported in e.g. [7].

Multi-channel binaural algorithms take advantage of the multiple microphones available in each hearing device. One of the most common approaches is the multi-channel Wiener filter (MWF), which provides a minimum mean-square estimate of the desired signal [8]. The output of the MWF is obtained by filtering and summing the input signals coming from every microphone at both HAs. It has been theoretically proven that the MWF preserves the binaural localization cues related to the target speech, but distorts the ones of the noise [9]. In order to partly preserve the binaural cues of the noise, alternative MWF formulations have been proposed in [10, 11, 12].

The implementation of speech enhancement algorithms into HAs demands to take into consideration numerous technical constraints. These include the limited processing power and embedded memory, the real-time framework (i.e. short analysis frames), the robustness against fixed-point resolution and quantization errors, and a rational battery consumption, e.g. by limiting the binaural wireless communication. A few solutions have been proposed towards the observation of these constraints, such as decreasing the bit rate of the wireless streaming [13], exchanging only a limited amount of data between the devices [14], or reducing the occurrence of heavy-computational-load operations like matrix inversions [15]. Thus, the development of speech enhancement techniques is governed by the trade-off that must be found between the performance of the algorithm, the output speech quality, the preservation of the localization cues, and the technical limitations of the device.

When the aforementioned strategies are not sufficient to restore a satisfying speech intelligibility, it is common to resort to remote microphone (RM) systems. A RM system is composed of a small transmitter microphone that picks up the voice of a speaker, and sends the speech signal wirelessly to a radio-frequency receiver plugged or integrated into the HAs of a listener. This gives access to a clean signal, which significantly improves the speech understanding performance of the users [16, 17, 18, 19]. Since the RM signal does not contain any spatial information, it is usually mixed with the noisy and reverberant sound captured by the local HA microphones. However, this reduces the intelligibility gain obtained with RM systems, while being of limited help for sound localization in noisy conditions [20]. Ideally, the RM and local microphone signals should be mixed in an optimal way, ensuring an accurate reproduction of the spatial cues while minimizing the noise and reverberation contribution. Szurley *et al.* [21] used the RM as an input to the MWF and analytically showed that such a system yields a higher noise reduction performance while better preserving the noise-related binaural cues.

The goal of this paper is to compare the results obtained with three different speech enhancement algorithms that use the additional information provided by a RM. The first algorithm is the speech distortion weighted MWF with partial noise estimation (MWF_η) [10] using only the HA microphone signals as input while exploiting the RM signal only to improve the estimation of the signal statistics. The second algorithm is the MWF_η [10] using all signals as input and additionally exploiting the RM signal to better estimate the signal statistics (MWF_η - RM). The third algorithm is a Phonak's proprietary algorithm which implements a single-channel binaural technique. The criteria for analysis include the performance in terms of interference reduction, sound quality, binaural cue preservation, and implementation constraints.

Section 2 reviews the principle of the three algorithms. Section 3 describes the setup used for data acquisition in a reverberant room. Section 4 reports the results obtained from recordings, regarding the effects of the signal-to-noise ratio (SNR), of the transmission delay between the RM and HA microphone signals, and of the length of the analysis frames. Conclusions are drawn in Section 5.

2. Algorithms

This section reviews the three speech enhancement methods that are depicted in Figure 1. In the short-time Fourier transform domain, the signal captured by the m^{th} microphone of the HA at side X (L or R) is expressed as

$$y_{\text{X},m}(k,i) = s(k,i)a_{\text{X},m}(k,i) + v_{\text{X},m}(k,i), \quad (1)$$

where k and i denote the frame and frequency bin indices respectively, $s(k,i)$ is the clean speech, $a_{\text{X},m}(k,i)$ is the acoustic transfer function (ATF) between the microphone m^{th} and the speech source, and $v_{\text{X},m}(k,i)$ denotes the interference components.

Similarly, the signal of the RM is expressed as

$$y_{RM}(k,i) = s(k,i)a_{RM}(k,i) + v_{RM}(k,i). \quad (2)$$

Since the speaker is wearing the microphone, the signal-to-interference ratio at the RM is very high.

In the following, k and i are omitted for conciseness, and a pair of HAs embedding two microphones on each device is considered. Two stacked 2-dimensional vectors containing the signals of the left (X $= L$) and right (X $= R$) HAs are defined as

$$\mathbf{y}_\text{X} = \begin{bmatrix} y_{\text{X},1} \\ y_{\text{X},2} \end{bmatrix}. \quad (3)$$

In addition, a stacked 4-dimensional vector is defined containing all the microphone signals of the left and right devices, i.e.,

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_L \\ \mathbf{y}_R \end{bmatrix}. \quad (4)$$

Algorithm 1. As depicted in Figure 1, the first considered algorithm is the MWF_η [10], which uses as input only the 4-dimensional vector of HA microphone signals $\mathbf{y}$. The MWF_η allows to explicitly control the

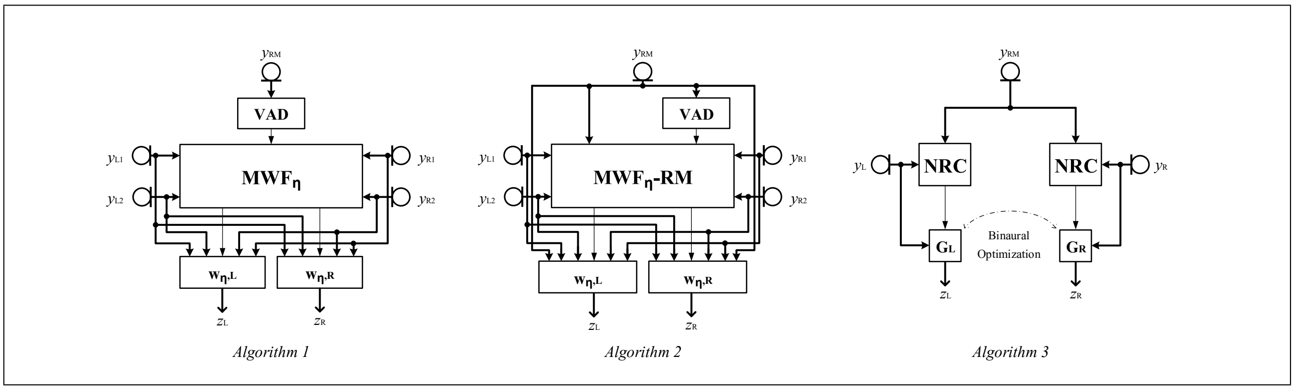

Figure 1. Overview of the three studied algorithms. Thick lines denote signals, thin lines correspond to data.

trade-off between noise reduction and binaural cue preservation, and is given by

$$\mathbf{w}_{\mathrm{X},\eta} = (1-\eta)\frac{\mathbf{R}_{\mathbf{vv}}^{-1}\mathbf{R}_{\mathbf{dd}}\mathbf{e}_{\mathrm{X}}}{\mu + \mathrm{Tr}\{\mathbf{R}_{\mathbf{vv}}^{-1}\mathbf{R}_{\mathbf{dd}}\}} + \eta\mathbf{e}_{\mathrm{X}}, \qquad (5)$$

where $\mathbf{w}_{\mathrm{X},\eta}$ are the complex-valued filter coefficients at the HA at side X, Tr{} is the trace of a matrix, $\mathbf{e}_{\mathrm{X}}$ are the unit vectors equal to 1 at the index of the reference microphone (microphone 1 on both left and right devices) and 0 otherwise, μ is a parameter to control the trade-off between noise reduction and speech distortion ($\mu > 0$), and η is a parameter to control the trade-off between noise reduction and binaural noise cue preservation ($0 \leqslant \eta \leqslant 1$). The speech and noise correlation matrices $\mathbf{R}_{\mathbf{dd}}$ and $\mathbf{R}_{\mathbf{vv}}$ are computed as

$$\mathbf{R}_{\mathbf{dd}} = \mathcal{E}\{\mathbf{dd}^H\}, \quad \mathbf{R}_{\mathbf{vv}} = \mathcal{E}\{\mathbf{vv}^H\}, \qquad (6)$$

where ^{H} is the Hermitian transpose, $\mathcal{E}$ denotes the expected value, and $\mathbf{d} = s\mathbf{a}$ with $\mathbf{a}$ being the vector containing the ATFs of all microphones.

In order to accurately estimate the statistics of the target signal and the interference, the MWF requires the information of a voice activity detector (VAD), the goal of which is to detect the noise-only periods and the speech-and-noise periods. This way, the noise correlation matrix $\mathbf{R}_{\mathbf{vv}}$ can be updated during noise-only periods, whereas the speech correlation matrix $\mathbf{R}_{\mathbf{dd}}$ can be updated during speech-and-noise periods, as

$$\mathbf{R}_{\mathbf{dd}} = \mathbf{R}_{\mathbf{yy}} - \mathbf{R}_{\mathbf{vv}}, \qquad (7)$$

with $\mathbf{R}_{\mathbf{yy}} = \mathcal{E}\{\mathbf{yy}^H\}$. In the following, it is assumed that a perfect VAD is obtained from the RM. The filtered output at each HA z_{X} is then given by

$$z_{\mathrm{X}} = \mathbf{w}_{\mathrm{X},\eta}^H\mathbf{y}. \qquad (8)$$

Algorithm 2. As depicted in Figure 1, the second considered algorithm is the MWF$_\eta$ - RM, where the RM is not only involved in the computation of a precise VAD, but is also used as an additional input to the MWF [21]. Thus, the stacked vector $\mathbf{y}$ becomes a 5-dimensional vector, i.e.,

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_L \\ \mathbf{y}_R \\ y_{\mathrm{RM}} \end{bmatrix}. \qquad (9)$$

The filter coefficients $\mathbf{w}_{\mathrm{X},\eta}$ are computed as in (5), and z_{L} and z_{R} are obtained by a weighted combination of the five input signals.

Algorithm 3. The third algorithm is a Phonak's proprietary algorithm which implements a single-channel speech enhancement method, as depicted in Figure 1. First, the algorithm operates independently in each device, using as input the RM and local HA microphone (bilateral step). The objective is to emphasize the time-frequency components corresponding to the target speech in the local microphone, taking as reference the RM signal. Second, the real-valued gain models computed in the left and right devices are exchanged and adjusted in an optimal way, in order to limit the ILD distortion while preserving the best possible interference reduction performance (binaural step). In the following, this algorithm is referred to as the noise and reverberation canceller (NRC).

3. Measurements

This section reports the experimental setup used to assess the three algorithms. The measurements were conducted in an empty classroom (volume = 262 m^3, background noise = 30 dBA, RT$_{60}$ = 0.59 s at the listener's position). The setup is described in Figure 2. Two manikins were used as the speaker (HATS B&K type 4128) and the listener (G.R.A.S. KEMAR), installed two meters apart. The speaker was wearing a RM (Phonak Roger$^{\mathrm{TM}}$ Touchscreen Mic), the signal of which was captured and demodulated by a wireless receiver (Phonak Roger$^{\mathrm{TM}}$ X) located at the listener's position. The RM system is equipped with three microphones and performs beamforming to focus on the

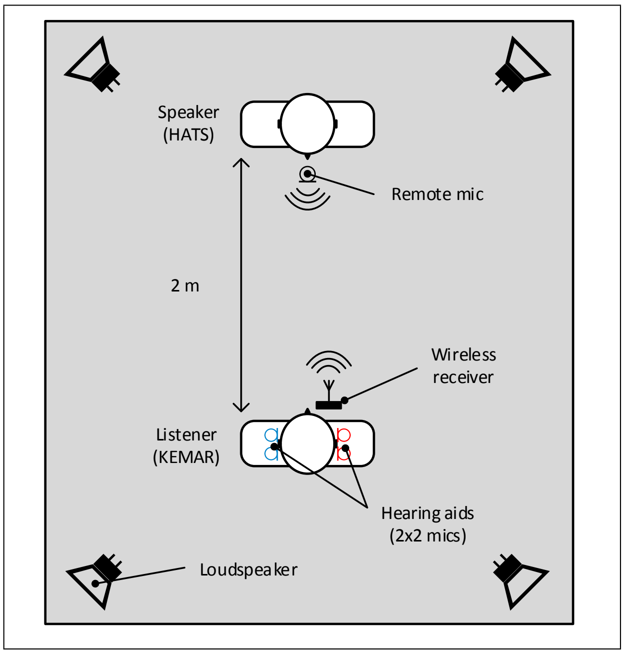

Figure 2. Setup of the measurements as mounted in the classroom.

speaker's voice, while reducing the noise coming from the other directions. It also introduces non-linear dynamic compression and operates at a sample rate of 16 kHz (8-kHz bandwidth). A pair of HAs (Phonak Bolero$^{\mathrm{TM}}$ Q), each equipped with two microphones were worn by the listener. Four loudspeakers (Tannoy Reveal Active) were mounted in the corners of the room and used to reproduce a diffuse babble noise.

The target stimulus was a 15-second excerpt of male speech (EBU SQAM CD, Track 50) that was played through the mouth of the speaker such that the sound level at the RM was 80 dBA (20 cm away from the mouth) to simulate realistic speech conditions. The audio signals from the five microphones (RM + HAs) were recorded at 44.1 kHz using an M-Audio M-Track 8 sound card. They were then downsampled to 22.05 kHz to match the common sampling rate of Phonak HAs. The recordings were conducted twice: once in the speech-only condition (no noise from the loudspeakers) and once in the noise-only condition (no speech from the speaker), such that different SNRs can be generated. In the latter condition, the noise level was adjusted to be identical to the speech level at the listener's location (i.e. SNR = 0 dB).

4. Results and discussion

In this section, the results obtained with the three algorithms are presented, compared and discussed. All algorithms were tested with 128-sample analysis frames (Hanning window) and 75% overlap at a sampling frequency of 22.05 kHz. The algorithms are assessed using the difference in several instrumental measures between the output and input signals of the left HA. The interference reduction performance is assessed by computing the difference between the input and output segmental SNR (segSNR). The spatial hearing preservation is assessed by computing the mean absolute error (MAE) between the short-term ILDs at the input and output of the algorithms. Thus, both the components of speech and interference are taken into account. This is because the role of the HA microphone contribution is to restore the overall spatial impression that is lost with the RM signal, i.e. not only the binaural cues of the speech. The ILD is computed in the 1.5-8 kHz frequency band for the frames showing an interaural coherence equal to or greater than 0.9, as suggested by Schwartz *et al.* [23]. The speech quality is assessed by computing the difference between the input and output PESQ and Hearing-Aid Speech Quality Index (HASQI) [22].

4.1. SNR effect

In order to evaluate the results provided by the three algorithms in various levels of noise, a range of input SNRs from -5 dB to 30 dB (5-dB steps) was tested. Figure 3 displays the results obtained with the three algorithms in terms of ΔsegSNR, MAE ILD, ΔPESQ and ΔHASQI as a function of the input SNR. To ensure a fair comparison, the parameters of the MWF$_\eta$ and NRC algorithms were tuned so that both yielded a similar HASQI (ΔHASQI $\approx$ -0.65) at an input SNR of 10 dB. For the MWF$_\eta$ algorithm, this corresponds to $\mu = 20$ and $\eta = 0.08$. The MWF$_\eta$ - RM algorithm was used with identical values of μ and η. In this section, the signals from the RM and HA microphones were temporally aligned.

The MWF$_\eta$ - RM outperforms the two other approaches in terms of interference suppression performance by an average amount of 1 dB, at the cost of higher ILD distortions. The segSNR increases from +1.9 to +3 dB for SNRs between -5 and 30 dB. This is related to the fact that the output of the MWF$_\eta$ - RM includes a filtered version of the clean RM signals, which is not present in the output of the MWF$_\eta$ and NRC. The SNR measured at the RM output ranges from 17 to 52 dB when the SNR at the HA microphones increases from -5 to 30 dB. The NRC algorithm tends to provide better segSNR improvement than the MWF$_\eta$ at mid and high SNRs, while yielding the least ILD distortion.

The three algorithms present similar trends with respect to ΔHASQI that suggests a modest improvement of the sound quality at lower SNRs and a slight degradation at higher SNRs. The outcomes obtained from PESQ rather indicate a preservation of the sound quality with the MWF$_\eta$ and NRC methods, and an enhancement provided by the MWF$_\eta$ - RM. It is therefore difficult to draw valuable conclusions, and subjective evaluation should be considered.

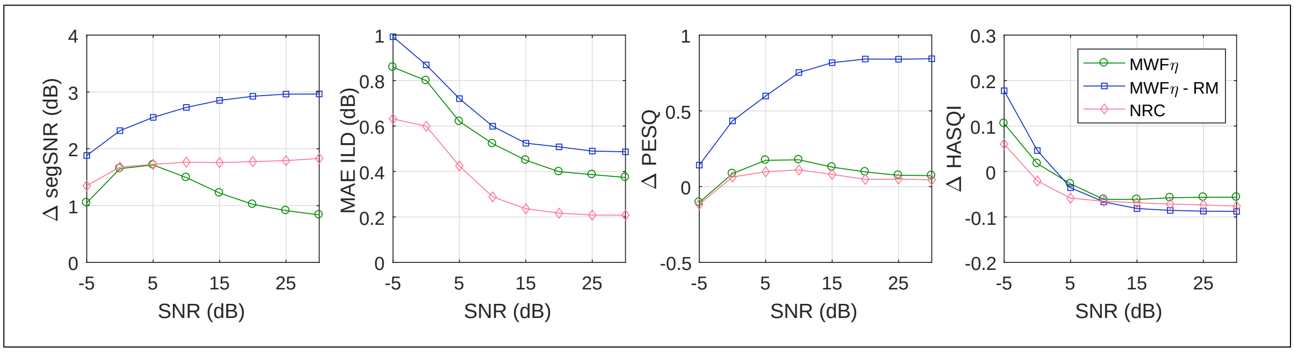

Figure 3. Performance of the three algorithms as a function of the input SNR. From left to right: interference reduction score (segSNR), ILD distortion (MAE), and sound quality (PESQ and HASQI). The results from *Algorithm 1* (MWF$_\eta$) are depicted in green with round markers, those of *Algorithm 2* (MWF$_\eta$ - RM) are in blue with square markers, and those of *Algorithm 3* (NRC) are in pink with diamond markers.

4.2. Delay effect

In the previous section, the RM and HA microphone signals were temporally aligned. In practice, these signals may reach the HA processor at different time instants. While the delay of the HA microphone signals is essentially governed by the time of flight between the speaker and listener, i.e. it is a distance-dependent delay, the delay of the RM signal is fixed and depends on the digital communication protocol used for the wireless transmission [20]. At short distances between the speaker and listener, the HA microphone signals might arrive before the RM-transmitted speech, whereas at larger distances, the RM-transmitted speech might arrive earlier than the HA microphone signals. Since the three tested algorithms rely on the availability of the RM signal in the HAs, it is expected that these delay differences have an effect on the performance of the algorithms.

In order to investigate the impact of these delay differences, the algorithms were run for various delay values between the RM and HA signals. A range of RM-to-HA delay from -15 ms to 15 ms (in steps of 5-ms) was introduced, corresponding to speaker-to-listener distances between 2 and 12 m for the used RM system. Negative delays occur when the RM signal arrives before the HA microphone signals, whereas positive delays occur when the HA microphone signals arrive earlier than the RM signal. The SNR was fixed to 10 dB and the same algorithmic settings as in the previous section were used.

Figure 4 depicts the results in terms of ΔsegSNR, MAE ILD, ΔPESQ and ΔHASQI. Both the MWF$_\eta$ and NRC algorithms appear to provide robust interference suppression performance for the range of the considered delays. The MWF$_\eta$ relies on the RM signal only for the VAD, which exploits the smoothed statistics of the RM signal. A shift by a few frames in the discrimination between speech and non-speech periods seems to have a limited effect on the results. The NRC was designed to be efficient over the range of delays that are likely to occur when using a RM system, which is confirmed here. Due to the delay between the HA and RM signals, it may happen that the instantaneous 128-sample analysis frames of the RM and HA microphone signals do not share any common speech and noise pattern. This appears to significantly impair the interference reduction performance provided by the MWF$_\eta$ - RM algorithms, which falls from +2.5 dB for temporally-aligned signals to +0.5 dB for a delay of -15 ms.

Furthermore, in terms of the ILD distortion and sound quality, the MWF$_\eta$ algorithm is robust for the considered RM-to-HA delay values. Conversely, the results obtained with the NRC and MWF$_\eta$ - RM algorithms are significantly degraded when the RM and HA microphone signals are not temporally aligned. Considering the HASQI scores, both algorithms dramatically decrease the monaural sound quality and distorts the binaural reproduction, preventing the rendering of an accurate spatial hearing. These appear to be strong limitations of the NRC and MWF$_\eta$ - RM approaches.

4.3. Frame-length effect

One way to circumvent the inherent delay issue is to resort to longer analysis frames. For the RM system considered in this study, a speaker-to-listener distance of approximately 3 meters causes the RM signals to reach the HA processor 10 ms after the local microphones, i.e. a shift of 221 samples. Therefore, no common speech and noise pattern are instantaneously present in the RM and HA 128-sample frames. Extending the frame length to 256 samples would allow to reach 14% of common pattern between the RM and HA microphone signals. This can be increased to 57%, 78% and 89% with frame lengths equal to 512, 1024 and 2048 samples respectively. On the other hand, longer analysis frames yield a longer latency (e.g. 92 ms for 2048-sample frames), raise the computational cost of the operations, and may result in algorithms that cannot track the changing statistics of non-stationary noises sufficiently fast.

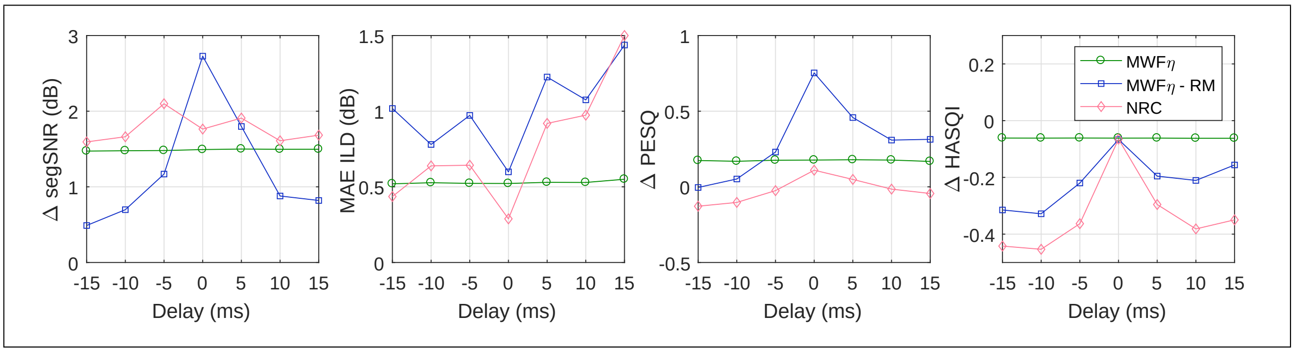

Figure 4. Performance of the three algorithms as a function of the delay between the RM and the HA microphones (SNR = 10 dB). A negative delay means that the RM is the lead. The metrics and color code are similar to Figure 3.

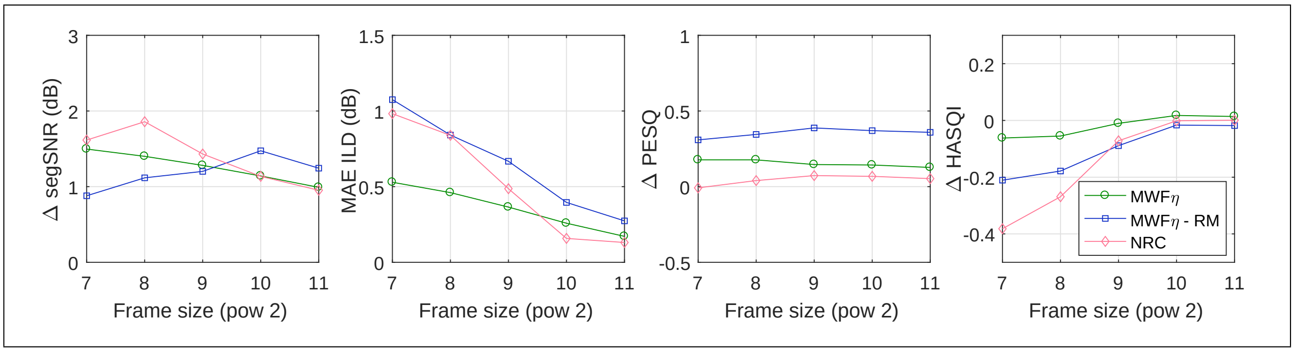

Figure 5. Performance of the three algorithms as a function of the size of the analysis frames (SNR = 10 dB, Delay between RM and HA microphones = 10 ms). The frame length is given as a power of two, corresponding to 128-, 256-, 512-, 1024- and 2048-sample sizes. The metrics and color code are similar to Figure 3.

Figure 5 presents the results obtained for the three algorithms when increasing the frame size from 2^7 (128 samples) to 2^{11} (2048 samples). The SNR was fixed to 10 dB and the RM-to-HA delay was equal to 10 ms. The same algorithmic settings as in the previous section were used for all algorithms. As expected, the MWF$_\eta$ and NRC algorithms, which were not affected by the considered RM-to-HA delay in the previous section, do not benefit from longer analysis frames with respect to interference reduction. On the contrary, the MWF$_\eta$ - RM largely benefits from longer analysis frames and outperforms the other two algorithms when the frame length reaches 512 samples (i.e. 23.2 ms). Interestingly, its performance decreases with the longest frame size (2048 samples). The explanation might be that the MWF$_\eta$ - RM starts performing too slow to follow the variations of the non-stationary babble for such a long frame duration. A similar trend is observed with the MWF$_\eta$ when increasing the frame size from 128 to 2048 samples.

Resorting to longer frames is also beneficial for reducing ILD distortions in both MWF-based approaches, because the successive gain values are less affected by sudden short-term variations. Additionally, the sound quality at the output of the MWF$_\eta$-RM and NRC algorithms gets better when the analysis frames are lengthened. It even improves when frames of 1024 samples (46.4 ms) are used (positive ΔHASQI values). However, real-time processing cannot be ensured with such frame durations.

4.4. Implementation considerations

Despite their relatively good performance, the MWF-based speech enhancement techniques require strong technical capabilities, such as heavy-computational-load operations (e.g. matrix inversions) and audio streaming of all incoming microphone signal frames. In order to make the MWF$_\eta$ and MWF$_\eta$ - RM algorithms more efficient, Szurley *et al.* [21] suggested to assume the presence of a single coherent noise source and additive incoherent noise, which allows to estimate the inverse matrix $\mathbf{R}_{\mathbf{vv}}^{-1}$ in (5) with simple operations.

On the other hand, the NRC algorithm does not require audio streaming between both hearing devices. It also does not involve computationally costly operations and has been designed to run real-time, i.e. with short analysis frames. The NRC is therefore the most efficient algorithm considered in this study.

5. Conclusion

Three RM-based algorithms for speech enhancement for HAs have been compared, with respect to interference reduction performance, speech quality, ILD preservation and technical requirements. The first two algorithms were multi-channel approaches, the so-called MWF_η and MWF_η - RM algorithms, while the latter was a single-channel proprietary algorithm. The algorithms were assessed with measurements recorded in a reverberant classroom. The MWF_η - RM appeared to provide the best interference reduction performance and improvements of speech quality, but it introduced more ILD distortions. Since the primary purpose of the HA microphone contribution in RM systems is to restore the spatial cues that are lost in the RM, the NRC might be the best candidate because it yields the lowest computational cost and provides the least ILD distortion.

The effect of the transmission delay between the RM and HA microphone signals was investigated as well. The interference reduction performance of both MWF_η and NRC algorithms was preserved, contrary to the MWF_η - RM. However, the NRC appeared to significantly deteriorate the sound quality as soon as the RM and HA microphone signals were not temporally aligned. This is considered as the main limitation of the NRC and might be solved by using longer analysis frames, which would limit the real-time performance capabilities. Future research is concerned with subjective evaluations, so as to determine which of the MWF_η - RM or NRC algorithm is preferred by the users of RM systems, and what is the perceived effect of the delay between the RM and HA microphone signals.

Acknowledgments

This study was funded by the Commission for Technology and Innovation (CTI) of the Swiss Federal Department of Economic Affairs, Education and Research (EAER), with grant number 25760.1 PFLS-LS, in collaboration with Sonova AG.

References

[1] R. Bentler and Li-K. Chiou. Digital noise reduction: An overview. *Trends Amplif.*, 10(2):67–82, 2006.

[2] B.C.J. Moore. Hearing aids. In *Cochlear hearing loss: physiological, psychological and technical issues.* John Wiley & Sons, 2007.

[3] S. Harding, J. Barker, and G.J. Brown. Mask estimation for missing data speech recognition based on statistics of binaural interaction. *IEEE Trans. Audio Speech Lang. Process.*, 14(1):58–67, 2006.

[4] S. Srinivasan, N. Roman, and D. Wang. Binary and ratio time-frequency masks for robust speech recognition. *Speech Commun.*, 48(11):1486–1501, 2006.

[5] A. Westermann, J. M. Buchholz, and T. Dau. Binaural dereverberation based on interaural coherence histograms. *J. Acoust. Soc. Am.*, 133(5):2767–2777, 2013.

[6] C.-Y. Yang, W.-S. Chou, K.-C. Chang, C.-W. Liu, T.-S. Chi, and S.-L. Jou. Spatial-cue-based multi-band binaural noise reduction for hearing aids. In *Workshop on Signal Processing Systems (SiPS)*, pages 278–283. IEEE, 2013.

[7] N. Yousefian, P.C. Loizou, and J.H.L. Hansen. A coherence-based noise reduction algorithm for binaural hearing aids. *Speech Commun.*, 58:101–110, 2014.

[8] S. Doclo, S. Gannot, M. Moonen, and A. Spriet. Acoustic beamforming for hearing aid applications. In *Handbook on array processing and sensor networks*, pages 269–302. John Wiley & Sons, 2010.

[9] B. Cornelis, S. Doclo, T. Van dan Bogaert, M. Moonen, and J. Wouters. Theoretical analysis of binaural multimicrophone noise reduction techniques. *IEEE Trans. Acoust. Speech Signal Process.*, 18(2):342–355, 2010.

[10] T. J. Klasen, T. den Bogaert, M. Moonen, and J. Wouters. Binaural noise reduction algorithms for hearing aids that preserve interaural time delay cues. *IEEE Trans. Signal Process.*, 55(4):1579–1585, 2007.

[11] E. Hadad, D. Marquardt, S. Doclo, and S. Gannot. Binaural multichannel Wiener filter with directional interference rejection. In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 644–648. IEEE, 2015.

[12] D. Marquardt, V. Hohmann, and S. Doclo. Interaural coherence preservation in multi-channel Wiener filtering-based noise reduction for binaural hearing aids. *IEEE Trans. Audio Speech Lang. Process.*, 23(12):2162–2176, 2015.

[13] S. Srinivasan, A. Pandharipande, and K. Janse. Beamforming under quantization errors in wireless binaural hearing aids. *J. Audio Speech Music Process.*, 2008(1), 2008.

[14] B. Cornelis, M. Moonen, and J. Wouters. Reduced-bandwidth Multi-channel Wiener Filter based binaural noise reduction and localization cue preservation in binaural hearing aids. *Sig. Process.*, 99:1–16, 2014.

[15] W.-C. Liao, Z.-Q. Luo, I. Merks, and T. Zhang. An effective low complexity binaural beamforming algorithm for hearing aids. In *Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 1–5. IEEE, 2015.

[16] C. J. Crandell and J. J. Smaldino. Improving classroom acoustics: utilizing hearing-assistive technology and communication strategies in the educational setting. *Volta Rev.*, 101(5):47–62, 1999.

[17] M. S. Lewis, C. C. Crandell, M. Valente, and J. E. Horn. Speech perception in noise: Directional microphones versus frequency modulation (FM) systems. *J. Am. Acad. Audiol.*, 15(6):426–439, 2004.

[18] L. Thibodeau. Benefits of adaptive FM systems on speech recognition in noise for listeners who use hearing aids. *Am. J. Audiol.*, 19(1):36–45, 2010.

[19] L. Thibodeau. Comparison of speech recognition with adaptive digital and FM remote microphone hearing assistance technology by listeners who use hearing aids. *Am. J. Audiol.*, 23(2):201–210, 2014.

[20] J. G. Selby, A. Weisser, and E. N. MacDonald. Influence of a remote microphone on localization with hearing aids. In *International Symposium on Auditory and Audiological Research (ISAAR)*, volume 6, pages 405–411, 2017.

[21] J. Szurley, A. Bertrand, B. Van Dijk, and M. Moonen. Binaural noise cue preservation in a binaural noise

reduction system with a remote microphone signal. *IEEE Trans. Audio Speech Lang. Process.*, 24(5):952–966, 2016.

[22] J.M. Kates and K.H. Arehart. The hearing-aid speech quality index (HASQI) version 2. *J. Audio Eng. Soc*, 62(3):99–117, 2014.

[23] B. Schwartz, S. Gannot, and E. A.P. Habets. An online dereverberation algorithm for hearing aids with binaural cues preservation. In *Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 1–5. IEEE, 2015.