ALGORITHMS TO IMPROVE LISTENING IN NOISE FOR COCHLEAR IMPLANT USERS

Adam A. Hersbach[1,2], *Stefan J. Mauger*[2], *David B. Grayden*[1,3,4], *James B. Fallon*[3,5] *& Hugh J. McDermott*[3,5]

[1]NeuroEngineering Laboratory, Department of Electrical and Electronic Engineering, University of Melbourne, Australia
[2]Cochlear Limited, East Melbourne, Australia
[3]Bionics Institute, East Melbourne, Australia
[4]Centre for Neural Engineering, University of Melbourne, Australia
[5]Departments of Otolaryngology & Medical Bionics, University of Melbourne, Australia

ABSTRACT

In this study, a two-microphone noise reduction algorithm that aims to improve directional beamformer performance for cochlear implant (CI) users is presented. The algorithm is computationally inexpensive and estimates a spatially-based signal-to-noise ratio (SNR), attenuating time-frequency elements that have poor SNR. The attenuation function was specifically tuned for application to CI. Using a real-time implementation of the algorithm, evaluation took place in a variety of noisy situations in which the competing speech sources were spatially separated and changed location during the test. Speech intelligibility tests with CI users revealed the new algorithm improved speech performance by 4.6 dB in speech reception threshold (SRT) compared to a commercially available adaptive beamformer.

Index Terms— noise reduction, speech enhancement, beamformer, cochlear implant, signal-to-noise ratio, post-filter

1. INTRODUCTION

Cochlear implant (CI) systems consist of an externally worn sound processor and an internally implanted stimulator used to electrically excite the neural population of an impaired ear. The sound processor contains the power source, microphone(s), digital signal processing (DSP) unit, and radio-frequency transmission system for transmitting data and power across the skin to the implant. The implant decodes the transmission and generates electrical impulses that are delivered via the multi-electrode array implanted in the cochlea. Noise reduction is used in CI sound processing to enhance the signal before it is encoded for electrical stimulation.

Single-microphone noise reduction techniques have been applied to CI sound processing and have demonstrated speech intelligibility improvements [1-4]. These algorithms are often modulation-based systems that typically fail to provide intelligibility improvement in acoustic applications such as hearing aids [5]. One key difference between CI and acoustic applications is the relative trade-off between speech distortion and noise reduction that is suitable for CI users. CI users have been found to perform better with more noise reduction (and hence more speech distortion) than is tolerated by acoustic listeners [6]. Therefore, algorithm tuning is usually different for CI applications [7], and acoustic-based predictors do not serve well to predict CI performance [8]. Speech intelligibility testing with CI users is the preferred evaluation method.

The maximum benefits of single-channel noise reduction algorithms have been demonstrated in stationary noise with improvements of around 2.5 dB in speech reception threshold (SRT) or, equivalently, 25 percentage point improvement in word recognition [1, 2, 4]. The benefit of single-channel noise reduction is reduced when the interfering noise is more dynamic, such as when the interfering noise is due to competing talkers.

Multi-microphone noise reduction can offer further performance benefits when the target and noise are spatially separated. The physical separation of the microphones provides an acoustic path difference for impinging sound, dependent on the direction of arrival relative to the microphone array. This feature is exploited to spatially filter the signal and reduce the noise.

Using small dual-microphone arrays on a single sound processor, large improvements have been demonstrated for CI users with adaptive beamforming, especially in listening situations involving a single interfering noise [9-12]. However, it is well known that performance degrades with an increase in the number of spatially separated noise sources [10, 11, 13], and with increased reverberation [9, 12]. The inability to cancel multiple noise sources simultaneously is a limitation of the 2-microphone adaptive beamformer, which steers a null in the direction of the noise to attenuate it, but is unable to provide strong attenuation in more than one direction at the same time.

In order to enhance beamformer performance, approaches to the design of a so-called post-filter have been explored from a theoretical perspective [14-17]. Different theoretical assumptions have been made about the sound field to derive an optimal filter. For example, Zelinski [14] assumed a perfectly diffuse noise field, while McCowan and Bourlard [16] assumed knowledge of the noise coherence across the microphone array. However, these assumptions generally do not hold for many real situations and therefore provide limited practical solutions to improve beamformer performance in CI devices. Yousefian and Loizou [18] proposed a coherence-based noise reduction algorithm for CI and demonstrated 5-10 dB SRT improvement over a fixed directional microphone in an anechoic room with one or two competing talkers. The algorithm assumed that speech and noise were coherent across the microphone array, and was also evaluated by normal hearing listeners in various degrees of reverberation. The benefit was substantially reduced to 0-2 dB SRT when evaluated in a moderately reverberant room (T60 = 465 ms) suggesting the coherence assumption was not as strong under these conditions.

As an alternative to relying on assumptions of the speech and noise coherence, post-filters based purely on spatial filtering have been proposed [19, 20]. Although not evaluated with human listeners, these studies provide evidence that the use of a spatial post-filter can improve beamformer output [19], and improve it in a way that is superior to the theoretically optimal Wiener post-filter [20, 21].

In this study, we aim to demonstrate that a beamformer post-filter based on spatial filtering can be adapted for application to CI sound processing, where the microphone array, DSP processing power, and algorithm tuning are taken into account specifically for this application. This research has two main aims: 1) to develop a beamformer post-filter algorithm designed to improve beamformer noise reduction performance for CI users; and 2) to evaluate the post-filter algorithm by measuring speech intelligibility of CI users.

2. ALGORITHM DESCRIPTION

The algorithm ("SpatialNR") is a multi-microphone noise reduction algorithm formulated as a beamformer post-filter. It is designed for real-time, low complexity implementation for a small-separation (< 20 mm), dual-microphone array typically used in hearing aids and CIs. The underlying speech and noise power distributions are not assumed, nor is the coherence of the sound field. A target spatial location is assumed to be in front of the listener.

The algorithm estimates the SNR by analyzing the signals from front-facing and rear-facing fixed directional microphones. Processing is performed independently in each sub-band in the frequency domain. A fast Fourier transform (FFT) is used and FFT bins are combined to form filters with approximate equivalent rectangular bandwidth (ERB), an approximation to the bandwidths of the filters in human hearing. The SNR calculated in each sub-band is used to attenuate frequency channels that are noisy. The resulting spatial post-filter attenuates sound from behind and beside the listener while passing sounds from the front.

Two omni-directional microphone signals are used to form the two fixed directional microphone responses. The front-facing response is configured to have maximum directivity index when worn on the head. The rear-facing response is configured to have a fixed null in the target direction. The frequency-domain representations, $S_k[n]$ (front-facing) and $N_k[n]$ (rear-facing), where k is the frequency index and n is the time index of overlapping FFT windows, are transformed to the log domain and filtered using first-order IIR filters

$$\overline{S^{dB}}_k[n] = \beta S^{dB}{}_k[n] + (1-\beta)\overline{S^{dB}}_k[n-1], \tag{1a}$$

$$\overline{N^{dB}}_k[n] = \beta N^{dB}{}_k[n] + (1-\beta)\overline{N^{dB}}_k[n-1]. \tag{1b}$$

With $\beta = 0.185$, the time constant is 10 ms using a 489 Hz rate of successive FFT frames. Smoothing in the log domain is used because it relates more closely to perceptual loudness. The smoothed signals are used to estimate the instantaneous SNR, $\xi_k[n]$,

$$\xi^{dB}{}_k[n] = \overline{S^{dB}}_k[n] - \overline{N^{dB}}_k[n]. \tag{2}$$

The SNR estimate $\xi_k[n]$ is used as input to a parametric Wiener-like gain function, $H_k[n]$, with adjustable bias, $\alpha > 0$,

$$H_k[n] = \frac{\xi_k[n]}{\alpha + \xi_k[n]}. \tag{3}$$

The noise-reduced output signal, $\hat{X}_k[n]$, is created by applying the filter gain, $H_k[n]$, to the output of the beamformer stage,

$$\hat{X}_k[n] = H_k[n] S_k[n]. \tag{4}$$

3. EVALUATION

Evaluation was performed via speech intelligibility testing in noise with two groups of adult CI users. All subjects were current users of the Cochlear Nucleus® implant system. A first group of 12 subjects were evaluated in a sound-treated room with essentially no reverberation ($T_{30} \approx 70$ ms) in one noise configuration, (a), whereas a second group of 8 subjects were evaluated in a reverberant room ($T_{30} \approx 520$ ms) in two noise configurations, (b) and (c), shown in figure 1.

3.1. Test protocol

The study used a repeated-measure, single-subject design, in which each listener served as his/her own control. The Australian sentence test in noise (AuSTIN) [22] was used to evaluate speech intelligibility. It is an adaptive SRT task that involved listening to sentences embedded in background noise. The noise level was adapted during the test to find the SRT defined as the SNR where 50% of the morphemes in the sentence were correctly understood. A lower SRT is interpreted as better performance since intelligibility of 50% is maintained with more noise. For CI users performing the AuSTIN, a change of 1 dB SRT can be interpreted as a change in intelligibility of approximately 11 percentage points [22]. The test order of processing conditions was randomized for each subject, and counterbalanced by reversing the order in a second test session. According to the AuSTIN adaptive procedure, the noise level was increased if 50% or more morphemes were correct, otherwise it was reduced. The noise level was adjusted by 4 dB for the first four sentences and by 2 dB thereafter for a total of 20 sentences. The SRT was calculated as the mean SNR for the final 16 sentences. Sentences were randomly selected from a set of 1264 and no sentence was repeated for any subject.

3.2. Processing Conditions

The SpatialNR algorithm was adjusted using the bias value, α, and the signal smoothing parameter, β. In the no-reverb room, β with time constant of 10 ms, and α corresponding to biasing ξ^{dB} by -3, 0, +3, and +6 dB was evaluated. In the room with reverberation, β with time constant of 10 ms was evaluated with a bias value of +3 dB, such that there was a condition to allow a comparison with the no-reverb room. In order to improve sound quality, a variable β was also evaluated in reverberation. A time constant of 5 ms was used for signals increasing in amplitude (compared to the previous sample) and 50 ms for signals decreasing in amplitude. The variable β was evaluated with bias values of +3, +6, and +9 dB.

"Beam", an adaptive generalized sidelobe canceller (GSC) with energy-based voice activity detection (VAD), is commercially available in the Nucleus® CP810 speech processor [10]. It was used as a baseline for comparison because it has superior noise reduction performance to other directional microphone options available in the device [4]. In addition, under conditions of well-matched microphones as used in this study, it has been shown to perform equivalently to the spatially pre-processed speech distortion weighted multi-channel Wiener filter (SP-SDW-MWF) [23].

3.3. Test environment

The test environments were designed to represent real-life situations where talkers are spatially separated and involved in conversation [4]. The interfering maskers had different locations that were assigned randomly for each sentence presentation. The algorithm under evaluation was thus forced to adapt to the varying environment.

The target speech was always presented from the front direction. Noise environments were created using multi-talker babble as the interfering noise in two different rooms with different levels of reverberation. Three environments used for evaluation were (a) 4-talker rear-half no-reverb, (b) 4-talker rear-half with reverb, and (c) 20-talker full-circle with reverb.

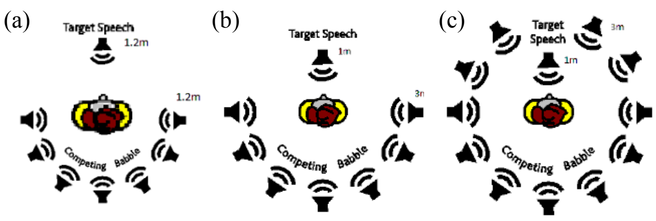

Figure 1: The three test environments used for evaluation in (a) 4-talker rear-half no-reverb, (b) 4-talker rear-half with reverb and (c) 20-talker full-circle with reverb. The independent competing talker locations were randomized among the loudspeaker positions shown.

In the no-reverb room, 4-talker babble was presented from a distance of 1.2 m, the same distance as the target speech. In the reverberant room, both 4-talker and 20-talker babble were used, presented from a distance of 3 m while the target speech was presented from 1 m. In the reverberant room, a close distance was used for the target speech to emulate a real-world situation in which the subject might be involved in a conversation with one other person. The competing babble was located further away to emulate competing conversations within the same room, but not in the immediate vicinity of the subject. In all test environments, the independent competing talkers were randomly assigned to loudspeaker locations selected from those shown in figure 1. The selections were made at the start of each sentence during the test, and multiple competing talkers per loudspeaker were allowed.

4. RESULTS

The SRT benefit relative to the Beam baseline is shown in figure 2 for each of the SpatialNR processing conditions summarized for the two groups of subjects. The benefit relative to Beam was computed by subtracting the SpatialNR SRT from the Beam SRT for each individual and calculating the mean benefit for the group. Analysis of the results was performed separately for each of the three test environments, (a)-(c), using a one-way repeated-measures ANOVA with processing condition as the factor. In all test environments, a significant effect of processing condition was found ((a) $F[4,11] = 48.61$, $P < 0.001$, (b) $F[4,7] = 23.98$, $P < 0.001$, and (c) $F[4,7] = 4.76$, $P = 0.005$).

Post-hoc pairwise analysis (Student-Neuman-Keuls) showed that in 4-talker rear-half (a,b), all SpatialNR conditions were significantly better than Beam (all $P < 0.001$) regardless of the bias parameter value, β smoothing value, or level of reverberation. In 20-talker full-circle babble with reverb (c), the SpatialNR condition with bias value of +6 dB and variable β was significantly better than Beam ($P = 0.002$).

The largest improvement over the Beam condition in (a) was 4.6 dB (range 2.6 to 6.6 dB) obtained with a bias value of +3 dB. In (b), the maximum benefit of SpatialNR was 4.2 dB (range 1.7 to 6.6 dB) obtained with bias parameter value of +3 dB and variable β. In (c), the maximum benefit was 1.7 dB (range 0.8 to 4.0 dB) obtained with a bias parameter of +6 dB and variable β.

The effect of β was assessed via post-hoc comparisons in reverberation between conditions with bias value of +3 dB. No significant effect of the β smoothing value was found ((b) $P = 0.072$, (c) $P = 0.941$).

The effect of reverberation in 4-talker rear-half was tested by comparing the two groups of subjects tested in different rooms. An independent t-test comparing the benefit of SpatialNR with bias value of +3 dB and fixed β revealed a significant decline in performance of 1.2 dB due to reverberation (two-tailed $P = 0.022$).

SRT results in (a) at the individual level revealed a range of CI performers in the group, with baseline Beam SRTs ranging 15 dB, from -4.4 to +10.5 dB. A Pearson product moment correlation between the baseline Beam SRT and the improvement due to SpatialNR with each bias parameter value α revealed no statistically significant relationship ($P = 0.137$, 0.067, 0.452 and 0.895 for α = -3, 0, +3 and +6 dB respectively). Similarly, no significant correlation was found between the baseline Beam SRT and the SpatialNR bias value that produced the largest improvement for each subject ($P = 0.189$).

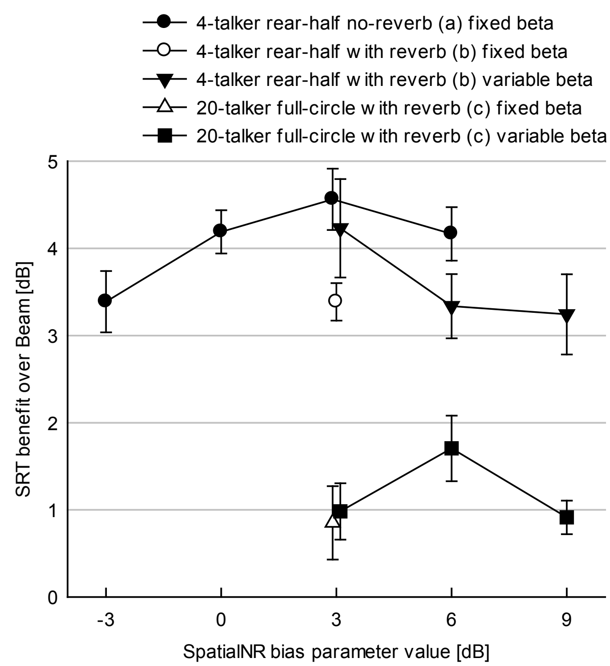

Figure 2: SRT benefit in dB relative to the Beam baseline condition for SpatialNR processing with different bias parameter settings. Evaluation was performed in three different noise environments, (a)-(c), and used two different groups of subjects. Error bars indicate the standard error of the mean.

5. DISCUSSION

The SpatialNR post-filter was implemented in real-time and demonstrated speech intelligibility improvement over Beam of 4.6 dB on average for a group of 12 CI users when tested in the 4-talker, rear-half, no-reverb environment. In the same environment and with a different group of subjects, Beam was previously shown to provide 5.0 dB improvement over the standard directionality setting [4]. Therefore, the SpatialNR algorithm is expected to provide approximately 10 dB benefit in total, enough to increase speech intelligibility from near 0% to near 100% [22] in the noise environment used for evaluation. Direct comparison against other similar algorithms is difficult due to a lack of speech intelligibility data with human listeners, different number of microphones, and different noise environments used for evaluation. However, in a system with four microphones, Wolff and Buck [20] measured computer word recognition rate in real recordings of café and train station noise. They found benefit of 20 to 40 percentage points depending on input SNR and noise type. The benefit due to SpatialNR of 4.6 dB SRT is an improvement of approximately 49 percentage points [22], suggesting SpatialNR with two microphones is delivering similar benefit.

The main reason for superior performance of SpatialNR over Beam was most likely due to the ability to remove multiple interfering noise sources at different locations simultaneously. Beam must adapt a null-steering filter and, while it is possible to obtain very large attenuation of a single masker, the effectiveness is reduced when multiple noise sources are presented simultaneously from different locations. In contrast, the SpatialNR algorithm can attenuate many sources simultaneously from many locations because all sources contribute to a common SNR estimate, which is used to reduce the noise.

The performance benefit of SpatialNR in reverberation was slightly reduced (with fixed β). Nevertheless, processing with variable β seemed to mitigate this decline (at bias value of +3 dB) and some robustness of the SpatialNR algorithm to the reverberation conditions evaluated in the study was demonstrated. The degree of smoothing has been shown to be an important factor for similar noise reduction processing [24], and for CI speech quality and intelligibility [7], and further testing across a wide variety of situations is warranted.

The benefit of SpatialNR over Beam was significant in 20-talker babble for bias setting of +6 dB, but was reduced compared to 4-talker babble. Although a portion of the difference could be attributed to the sheer number of competing talkers, a substantial portion is likely due to the difference in spatial locations of the noise, in particular that maskers were located closer to the target in the 20-talker babble case.

The SpatialNR bias value has the effect of changing the aggressiveness of noise reduction. As the bias value is increased, the amount of noise reduction is increased, but as a consequence more speech is removed, resulting in speech distortion. Increasing the bias value too far is likely to lead to decreased performance due to excessive speech distortion, indicated by the reduced benefit at high bias values. A bias value of +6 dB provided most benefit in 20-talker full-circle noise (compared to +3 dB in 4-talker rear-half noise), where there were comparatively more interfering maskers that were closer to the target signal. This indicates that the bias value may have some relationship to the number of noise sources and the spatial configuration of the noise sources. In general, different bias values may produce optimal performance in different acoustic environments.

During adaptive SRT testing, the SpatialNR algorithm was evaluated across a wide range of input SNRs spanning more than 15 dB. The lack of significant correlation between baseline Beam SRT and SpatialNR benefit indicates that SpatialNR benefit was not affected by the test SNR. In addition, the value of bias parameter that produced best results for each subject was not related to the test SNR. This supports the suitability of the algorithm over a range of listening environments since performance was robust to changes in input SNR.

The SpatialNR algorithm has been developed specifically for use in CI sound processing, but may be transferrable to acoustic applications such as hearing aids. The bias parameter used to tune the aggressiveness of noise reduction, which changes the relative amounts of speech distortion and noise reduction, would probably need to be tuned differently for acoustic applications. For example, [25] showed that normal-hearing listeners perform best with a negative gain threshold, whereas CI users perform best with a positive gain threshold [6]. This indicates that acoustic applications may need a lower bias value, maintaining speech quality at the expense of noise reduction performance.

6. CONCLUSION

A two-microphone noise reduction algorithm (SpatialNR) based on spatial filtering was formulated as a beamformer post-filter. Evaluation was performed with a real-time implementation in complex noisy environments including reverberation. Speech intelligibility tests with CI users revealed a significant benefit of SpatialNR processing compared to a commercially available adaptive beamformer. The competing noise sources changed location during the test to simulate real-world situations. The benefit was likely due to the design of SpatialNR, which allowed simultaneous attenuation of multiple noise sources that is not possible with the adaptive null-steering beamformer. The SpatialNR algorithm was tuned with a bias parameter that provided a mechanism for trading-off speech distortion and noise reduction, and was used to control the aggressiveness of noise reduction. The benefit obtained with the SpatialNR algorithm was up to 4.6 dB SRT compared to a commercially available adaptive beamformer.

7. ACKNOWLEDGMENT

This research was financially supported in part by the HEARing CRC established and supported under the Australian Government's Cooperative Research Centres program. The Bionics Institute acknowledges the support it receives from the Victorian Government through its Operational Infrastructure Support Program.

8. RELATION TO PRIOR WORK

The work presented here has focused on the application of a beamformer post-filter for noise reduction in CI sound processing utilizing dual microphones on a single sound processor. Although beamformers have been well researched in the CI field, post-filtering techniques have not been widely studied for this application. Yousefian and Loizou [26] proposed an algorithm suitable for small dual-microphone arrays used in CIs, and assumed the signal and noise were coherent. In contrast, we make no assumption on the noise field coherence. Others have proposed algorithms based on spatial filtering similar to ours but used seven microphones [19], and more complex algorithms [20, 24]. None performed evaluation with CI users. In this study, we take advantage of the specific requirements for the CI application which requires a small two-microphone array with low computational complexity. We pay specific attention to investigating the effect of parameter tuning suitable for CI users within representative real-world acoustic environments.

REFERENCES

[1] P. W. Dawson, S. J. Mauger, and A. A. Hersbach, "Clinical evaluation of signal-to-noise ratio-based noise reduction in Nucleus(R) cochlear implant recipients," *Ear Hear,* vol. 32, pp. 382-90, May-Jun 2011.

[2] A. Buechner, M. Brendel, H. Saalfeld, L. Litvak, C. Frohne-Buechner, and T. Lenarz, "Results of a pilot study with a signal enhancement algorithm for HiRes 120 cochlear implant users," *Otol Neurotol,* vol. 31, pp. 1386-90, Dec 2010.

[3] Y. Hu and P. C. Loizou, "Environment-specific noise suppression for improved speech intelligibility by cochlear implant users," *J Acoust Soc Am,* vol. 127, pp. 3689-95, Jun 2010.

[4] A. A. Hersbach, K. Arora, S. J. Mauger, and P. W. Dawson, "Combining directional microphone and single-channel noise reduction algorithms: a clinical evaluation in difficult listening conditions with cochlear implant users," *Ear Hear,* vol. 33, pp. e13-23, Jul-Aug 2012.

[5] P. C. Loizou, *Speech Enhancement, Theory and Practice*, illustrated ed. vol. 30. Boca Raton: CRC Press, 2007.

[6] S. J. Mauger, P. W. Dawson, and A. A. Hersbach, "Perceptually optimized gain function for cochlear implant signal-to-noise ratio based noise reduction," *J Acoust Soc Am,* vol. 131, pp. 327-36, Jan 2012.

[7] S. Mauger, K. Arora, and P. Dawson, "Cochlear implant specific noise reduction," *J Nueral Eng,* vol. 9, 2012.

[8] R. L. Goldsworthy and J. E. Greenberg, "Analysis of speech-based Speech Transmission Index methods with implications for nonlinear operations," *J Acoust Soc Am,* vol. 116, pp. 3679-89, Dec 2004.

[9] J. Wouters and J. Vanden Berghe, "Speech recognition in noise for cochlear implantees with a two-microphone monaural adaptive noise reduction system," *Ear Hear,* vol. 22, pp. 420-30, Oct 2001.

[10] A. Spriet, L. Van Deun, K. Eftaxiadis, J. Laneau, M. Moonen, B. van Dijk, A. van Wieringen, and J. Wouters, "Speech understanding in background noise with the two-microphone adaptive beamformer BEAM in the Nucleus Freedom Cochlear Implant System," *Ear Hear,* vol. 28, pp. 62-72, Feb 2007.

[11] K. Chung and F. G. Zeng, "Using hearing aid adaptive directional microphones to enhance cochlear implant performance," *Hear Res,* vol. 250, pp. 27-37, Apr 2009.

[12] M. Kompis, M. Bertram, P. Senn, J. Muller, M. Pelizzone, and R. Hausler, "A two-microphone noise reduction system for cochlear implant users with nearby microphones - Part II: Performance evaluation," *EURASIP Journal on Advances in Signal Processing,* 2008.

[13] K. Chung, F. G. Zeng, and K. N. Acker, "Effects of directional microphone and adaptive multichannel noise reduction algorithm on cochlear implant performance," *J Acoust Soc Am,* vol. 120, pp. 2216-27, Oct 2006.

[14] R. Zelinski, "A microphone array with adaptive post-filtering for noise reduction in reverberant rooms," presented at the IEEE International Conference on Acoustics, Speech and Signal Processing Las Vegas, Nevada, 1988.

[15] J. Meyer and K. U. Simmer, "Multi-channel speech enhancement in a car environment using Wiener filtering and spectral subtraction," presented at the IEEE International Conference on Acoustics, Speech and Signal Processing, Munich, Germany, 1997.

[16] I. A. McCowan and H. Bourlard, "Microphone array post-filter based on noise field coherence," *IEEE Trans Speech Audio Process,* vol. 11, pp. 709-716, 2003.

[17] S. Leukimmiatis, D. Dimitriadis, and P. Maragos, "An Optimum Microphone Array Post-Filter for Speech Applications," presented at the International Conference on Spoken Language Processing, Pittsburgh, Pennsylvania, 2006.

[18] N. Yousefian and P. C. Loizou, "A Dual-Microphone Algorithm That Can Cope With Competing-Talker Scenarios," *IEEE Trans Audio Speech Lang Process,* vol. 21, pp. 145-155, Jan 2013.

[19] Y. C. Cao, S. Sridharan, and M. Moody, "Speech enhancement using microphone array with multi-stage processing," *IEICE Trans Fundamentals,* vol. E79A, p. 386, 1996.

[20] T. Wolff and M. Buck, "Spatial Maximum a Posteriori Post-Filtering for Arbitrary Beamforming," presented at the Joint Workshop on Hands-free Speech Communication and Microphone Arrays, Trento, Italy, 2008.

[21] M. Buck, T. Wolff, T. Haulick, and G. Schmidt, "A compact microphone array system with spatial post-filtering for automotive applications," presented at the IEEE International Conference on Acoustics, Speech and Signal Processing, New York, 2009.

[22] P. Dawson, A. Hersbach, and B. Swanson, "Development of the Australian Speech Test in Noise (AuSTIN)," *Ear Hear,* in press 2013.

[23] S. Doclo, A. Spriet, J. Wouters, and M. Moonen, "Frequency-domain criterion for the speech distortion weighted multichannel Wiener filter for robust noise reduction," *Speech Commun,* vol. 49, pp. 636-656, 2007.

[24] M.-S. Choi and H.-G. Kang, "A Two-Channel Noise Estimator for Speech Enhancement in a Highly Nonstationary Environment," *IEEE Trans Audio Speech Lang Process,* vol. 19, pp. 905-915, 2011.

[25] D. S. Brungart, P. S. Chang, B. D. Simpson, and D. Wang, "Isolating the energetic component of speech-on-speech masking with ideal time-frequency segregation," *J Acoust Soc Am,* vol. 120, pp. 4007-18, Dec 2006.

[26] N. Yousefian and P. C. Loizou, "A Dual-Microphone Speech Enhancement Algorithm Based on the Coherence Function," *IEEE Trans Audio Speech Lang Process,* vol. 20, pp. 599-609, Jul 18 2011.