EUROSPEECH '97
5th European Conference on Speech
Communication and Technology
Rhodes, Greece, September 22-25, 1997

MULTI-MICROPHONE SUB-BAND ADAPTIVE SIGNAL PROCESSING FOR IMPROVEMENT OF HEARING AID PERFORMANCE

P. W. Shields and D. R. Campbell
Department of Electronic Engineering and Physics, University of Paisley,
High Street Paisley, Renfrewshire, PA1 2BE, SCOTLAND, U.K.
paul@diana22.paisley.ac.uk

10.21437/Eurospeech.1997-308

ABSTRACT

A scheme for binaural pre-processing of speech signals for input to a standard linear hearing aid has been proposed. The system is based on that of Toner & Campbell [1] who applied the Least Mean Squares (LMS) algorithm in sub-bands to speech signals from various acoustic environments and signal to noise ratios (SNR). The processing scheme attempts to take advantage of the multiple inputs to perform noise cancellation. The use of sub-bands enables a diverse processing mechanism to be employed, where the wide-band signal is split into smaller frequency limited sub-bands, which can subsequently be processed according to their signal characteristics. The results of a series of intelligibility tests are presented from experiments in which acoustic speech and noise data, generated in a simulated room was tested on hearing impaired volunteers.

1. INTRODUCTION

Many of the sensorineural hearing impaired suffer considerable difficulty understanding speech in the presence of medium to high reverberation or background noise, particularly from competing speakers. The difficulties occur at SNR around and below 6dB which would cause few problems for normal hearing listeners. Subjects with sensorineural hearing loss may require 5dB to 15dB greater SNR [2], and aided subjects may exhibit an SRT (Speech Reception Threshold; 50% correct recognition level) around 8dB worse, than normal hearing subjects [3].

It is a criticism of much research into the enhancement of speech signals corrupted with noise and/or reverberation, that too much emphasis is placed on the measure of SNR improvement or Speech Transmission Index, rather than a quantitative analysis of the improvement in terms of intelligibility [4].

Current signal processing research into improving the intelligibility of noisy speech has taken various approaches. One approach has been to attempt to emphasise certain signal characteristics e.g., increase the spectral contrast of the speech signal [5]. This type of approach has not yet yielded the level of improvement in SNR or intelligibility that was deemed necessary by Plomp [2] or Soede *et al* [3].

Alternative approaches aim to improve the intelligibility by attempting to increase the SNR e.g., beamforming [3,6], Spectral Subtraction [7], or binaural noise reduction [8]. These methods have proved more successful particularly that of Soede et al [3] and Debrunner & McKinney [9], who demonstrated an overall improvement of approximately 7dB and 12dB respectively.

The Multi-Microphone Sub-Band Adaptive (MMSBA) signal processing scheme falls into the latter category.

The process has been shown in simulation to improve, by up to 16 dB, the SNR of a speech signal corrupted with speech shaped noise. The MMSBA processing scheme has also been shown to significantly improve intelligibility for normal hearing listeners [10].

It is extremely important when assessing a speech intelligibility enhancement scheme to use realistic test signals. The use of simple additive noise is often not indicative of the systems performance in a real acoustic environment. Therefore, the intelligibility experiment presented here employs simulated reverberant convolutional noise. The use of simulated acoustics for evaluation of the MMSBA processing scheme was justified by Shields & Campbell [11].

From the results published here it will be possible to establish whether the measure of SNR improvement translates to a significant intelligibility improvement for hearing impaired listeners.

2. THE MMSBA PROCESSING SCHEME

2.1. Acoustic Model

The experiment aims to model a realistic scenario in which a person suffering from sensorineural hearing loss would experience difficulty with speech intelligibility. This is achieved by computer simulation of a rectangular room containing a speech source at a distance of 0.5m directly in front (0 degrees azimuth) of the input microphones (omnidirectional and placed at opposite points of a spherical simulated head of diameter 18cm), and a masking source of speech shaped noise at 135 degrees azimuth, and a distance of 4m. Figure 1a represents the acoustic model depicted above. Both speech, S, and noise, N, pass through their respective left and right acoustic FIR transfer functions, H_{11}, H_{12}, H_{22}, H_{21}, before forming the Primary, P, and reference, R, inputs to the MMSBA processing scheme. This illustrates the binaural speech and noise paths from their respective point sources to the input microphones of the system through the room acoustic transfer functions.

This approach should enable the system to simulate the binaural unmasking effect [12,13], that allows subjects listening binaurally to perform better, in speech intelligibility testing in noise, than subjects auditioning monaurally. The multi-microphone approach to noise reduction should enable a similar advantage over systems which only have one input, such as a standard linear hearing aid.

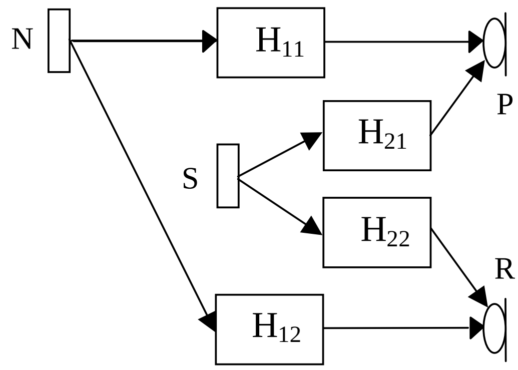

Figure 1a: Acoustic Model.

The model of the acoustic transfer functions is generated using a program based on the image method [14]. This computes an FIR filter which models the impulse response between the signal source and the microphone position, within an empty rectangular room, including the diffraction effect of the head. For the purpose of this study a filter length of 2048 points was established experimentally as being adequate for the acoustic transfer functions.
The speech and noise signals were sampled at 20 kHz., and convolved with their respective FIR acoustic transfer functions. The convolved speech and noise data were then summed at each microphone position to generate the desired SNR

2.2. Sub-band Decomposition

Figure 1b illustrates the sub-band decomposition process. This involves taking the 256 point FFT of each frame from both input channels, and reconstructing the signal in the time domain with 8, 16 or 32 sub-bands, with either linear or cochlear spacing. Using this approach the decon/reconstruction error is in the order of 10^{-6}.

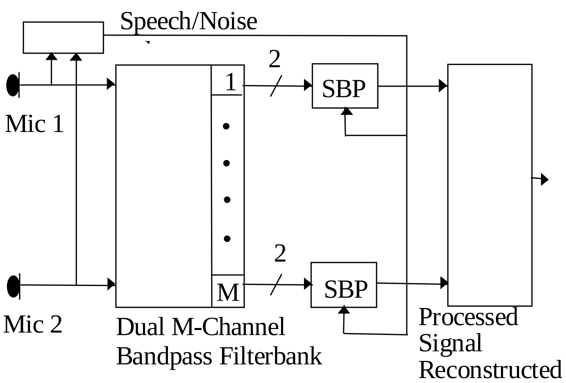

Figure 1b: Sub-band Decomposition

2.3. Sub-band Processing Scheme

Figure 1c represents the adaptive noise cancellation scheme. The processing method employed depends on the cross-correlation/coherence between the channels. This allows the lower frequency bands which generally have high coherence (> 0.7), to use an adapt and freeze strategy during a predetermined noise alone period (~ 0.4 second), to adapt to the differential acoustic transfer function of the noise masker. The adaptive filter algorithm implemented was the LMS algorithm [15]. When speech is present, the weights in the adaptive filter are frozen, to allow the filtering out of the noise signal, leaving ideally only desired speech at the output E. In some of the higher frequency bands the speech information generally has a higher coherence than the noise source. This can take advantage of an approach described by Ferrara Widrow [16]. In these bands the system is continually adapted, filter output Y, to enhance the correlated component of the signal in each sub-band, which should emphasise the desired speech signal. The output from each sub-band is then summed to provide a full-band noise-reduced output for evaluation by the test subjects.

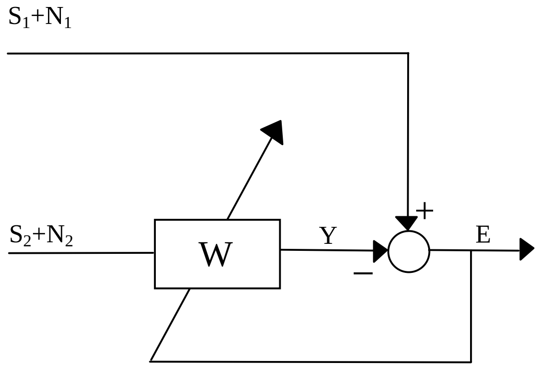

Figure 1c: LMS Sub-band Processor (SBP)

3. INTELLIGIBILITY TESTING

This paper is presenting interim results based on 11 subjects from an experiment involving 15 hearng-impaired volunteers of between 40 and 77 years of age. All subjects had moderate sensorineural hearing loss which had been established through prior audiometric testing. Each subject had his or her hearing aid NAL curve [17] matched using an eight-band graphic equaliser. Subjects were tested using speech masked by speech shaped noise at six SNRs, two sub-band spacings [18], and three sub-band distributions. The subjects were presented with the data in a four choice forced response approach using the FAAF data set [19].

The subjects were asked to identify each of 80 keywords " **** " from a sentence;

*'Can you hear **** clearly ?'.*

The options visually presented to the subjects differed by only one phoneme e.g. TIN, BIN, PIN, and DIN. The acoustic and visual presentations and monitoring of subject responses were under the control of a PC based Hearing Assessment Workstation. Each subject was given a number of clean speech practice sentences required until they were familiar with the procedure.

The reverberant level chosen was T_{60}=0.35s. This is representative of a typical living room level of reverberation [19]. The SNR's were -6, -3, 0, +3, +6, and +9dBs, chosen by experimentation to elicit a significant number of errors. Comparing results from Shields & Campbell [10] with those of the original authors of the FAAF test, Foster & Haggard [20], verified the experimental methodology employed.

4. RESULTS

Analysis of the experiment aims to answer the following questions regarding the intelligibility scores:

- Is there a significant enhancement due to processing when results are blocked by SNR?
- Does the processing have a degrading effect on intelligibility when the SNR is high?
- Is there any significant effect of processing using different sub-band spacing?
- Is there any significant effect of processing using different numbers of sub-bands.

Figure 2 shows unprocessed and best processed scores averaged across 11 subjects at different levels of SNR including the 95% confidence intervals.

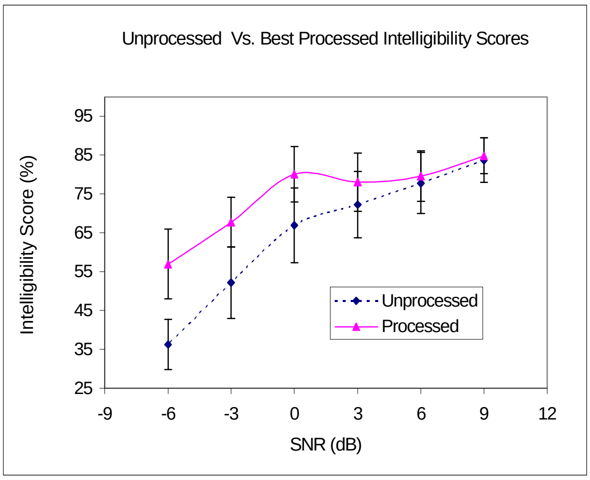

Figure 2: Unprocessed vs. Best Processed Intelligibility scores.

It can be seen from figure 2 that for all treatments the intelligibility score after the MMSBA processing scheme was higher than the unprocessed case. The scores are statistically significant at the 95% confidence level for –3dB and –6dB, with improvements of **15.6%** and **21.0%** respectively. The scores from the treatments with SNR between 0dB and 9dB SNR are not significant at the 95% level. This can be attributed to the high initial subject scores for the unprocessed higher SNR treatments.

It is thought that the reason for the slight dip in the processed curve at 3dB could be due to the incomplete set of factors used to analyse the experiment. Once all the data becomes available for a full-factor analysis the curve may flatten out, with the unprocessed and processed curves expected to merge as the scores reach saturation.

It is important to show that while at low SNR there is a significant improvement due to processing, there is also no degradation of intelligibility at high SNR due to processing. From figure 1, the 95% confidence intervals between 0dB and 9dB are all overlapping. This indicates that there is no significant reduction in processed speech quality for instances when intelligibility scores for the hearing impaired subjects are already high.

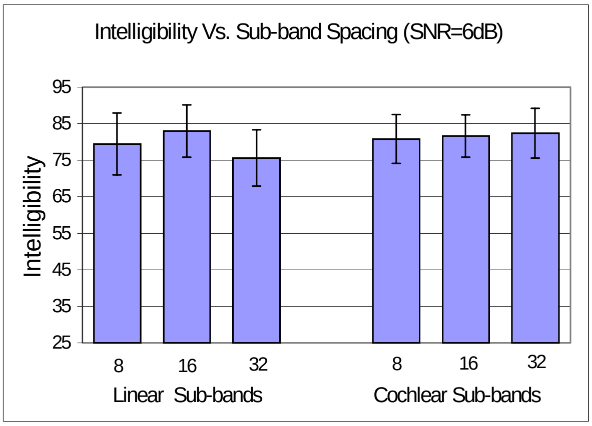

Figure 3: Intelligibility Vs. Sub-band Spacing (SNR=6dB)

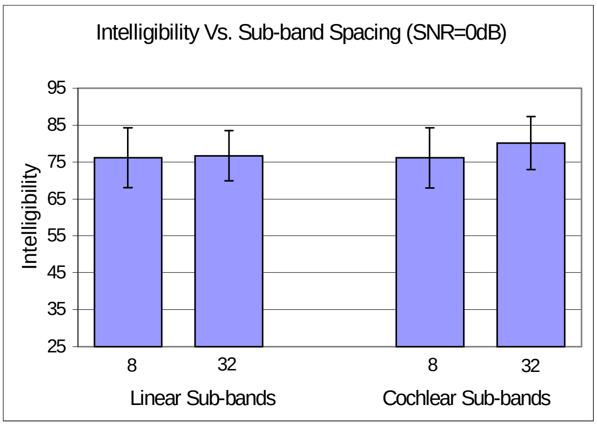

Figure 4: Intelligibility Vs. Sub-band Spacing (SNR=0dB)

Figures 3, 4 and 5 show a comparison of Linear verses Cochlear sub-band spacing at 6dB, 0dB and –3dB SNR with different sub-band distributions

There is little variation between intelligibility scores with all 95% confidence intervals overlapping. From this information it can be concluded that neither sub-band distribution or spacing are significant factors within the experiment.

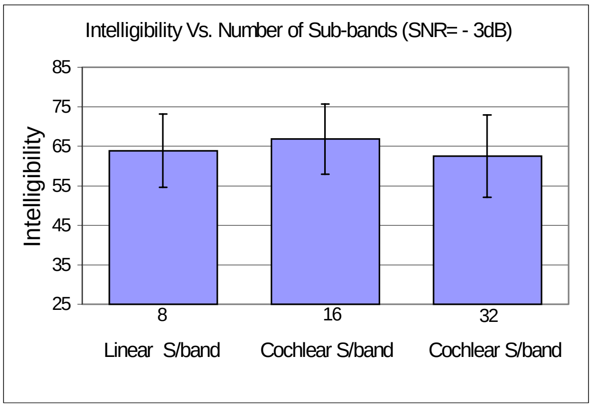

Figure 5: Intelligibility Vs. Sub-band Spacing (SNR= -3dB)

5. CONCLUSIONS

The MMSBA processing scheme has been shown to significantly improve the intelligibility of speech corrupted with noise in a moderately reverberant environment by up to **21.0%**. It has been shown that the processing has no detrimental effect on intelligibility at high SNR where unaided intelligibility scores are large. It appears that neither the sub-band spacing or sub-band distribution are important factors.

ACKNOWLEDGEMENTS

The authors thank for their advice and support: Dr. J. Culling then with the Institute of Hearing Research (IHR) at the University of Nottingham for the modified room simulation software, and Professor S. Gatehouse of the IHR at Glasgow Royal Infirmary who supplied the test data. This investigation is supported by the Scottish Office Department of Health grant number K/RED/4/237. D.R Campbell was supported by a Leverhume Trust fellowship during the writing of this paper.

REFERENCES

[1] Toner,E., Campbell,D.R., (1993**),** 'Speech Enhancement using Sub-Band intermittent adaptation', Speech Communication, **12**, 253-259.
[2] Plomp, R., (1986). 'A Signal to Noise Ratio Model for Speech Reception Threshold of the Hearing Impaired', Journal Speech, Hearing Research , **29**, 146-154.
[3] Soede, W., Berkhout, A.J, Bilsen, F., (1993). "Development of a directional hearing instrument based on array technology", J. Acoustical Society Of America, **94**, 785-798.
[4] Ludvigsen, C., Elberling, C., Keidser, G., (1993). 'Evaluation of a Noise Reduction Method - Comparison between Observed Scores and Scores Predicted from STI', Scandanavian Audiology, Suppl. **38**, 50-55.
[5] Baer, T., Moore ,C.J., Gatehouse, S., (1993). 'Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: effects on intelligibility, quality and response times.', J. of Rehabilitation Research and Development, **30**(1), 49-72.
[6] Greenberg, J.E., Zurek, P.M., (1992). 'Evaluation of an Adaptive Beamforming Method for Hearing Aids', Journal of the Acoustical Society of America, **91**(3), 1762-1676
[7] Elberling, C., Ludvigsen,C., Keidser, (1993). 'The Design and Testing of a Noise Reduction Algorithm Based on Spectral Subtraction.', Scandanavian Audiology, Suppl. 38, August 1993, 39-49.
[8] Kollmeier, J.P., Peissig, J., Hohmann V., (1993). 'Binaural Noise-Reduction Hearing Aid Scheme with Real-Time Processing in the Frequency Domain', Scandanavian Audiology, Suppl. **38**, 28-38.
[9] DeBrunner, V.E., McKinney, E.D., , (1995) 'A Directional Adaptive Least-Mean-Square Acoustic Array for Hearing Aid Enhancement', J. Acoustical Society Of America, **98**, 437-444.
[10] Shields, P.W., Campbell D.R.,,(1997)' Multi-Microphone Sub-Band Adaptive Signal Processing For Improvement of Hearing Aid Performance: Preliminary Results Using Normal Hearing Volunteers', Proc. ICASSP-97, I.E.E.E Conference on Acoustics, Speech and Signal Processing, **1**, pp 415-418.
[11] Shields, P.W., Campbell D.R., (1997) 'Intelligibility Improvements obtained by Multi-Microphone Sub-Band Adaptive Signal Processing Applied to Speech Enhancement' ESCA-NATO Workshop on Robust communication Channels, April 17-18, 1997, PONT-a-MOUSSON, FRANCE
[12] Moncur, J.P., Dirks, D., (1967**).** 'Binaural and Monaural Speech Intelligibility in Reverberation', J. of Speech and Hearing Research, **10**, 186-195.
[13] Plomp, R., (1976). 'Binaural and Monaural Speech Intelligibility of Connected Discourse in Reverberation as a Function of Azimuth of a Single Sound Source (Speech or Noise), Acoustica, **31**, 200-211
[14] Allen, J.B., Berkley, D.A., (1979). 'Image Method for Simulating Small-Room Acoustics', Journal of the Acoustical Society of America, **65**(4), 943-950.
[15] Stearns, S. D., David, R A., *Signal Processing Algorithms*, New Jersey, Prentice-Hall, 1988.
[16] Ferrara, E.R., Widrow, B., (1981). 'Multichannel Adaptive Filtering for Signal Enhancement', IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-29, **3**, 756-770.
[17] Byrne, D., Harvey, D., (1986**),** 'The National Laboratories' (NAL) New Procedure for Selecting the Gain and Frequency Response of a Hearing Aid', Ear and Hearing, Vol. 7, **4**, Pages 257-264.
[18] Darlington, D.,J., Campbell, D.,R., (1996a). 'The Effect of Modified Filter Distribution on an Adaptive, Sub-Band Speech Enhancement Method', IEEE Digital Signal Processing Workshop, Leon, Norway, 1-4 September 1996.
[19] Mackenzie, G.W,. *Acoustics*, London & New York, Focal Press, 1964
[20] Foster, J.,R. ,Haggard, M.P, (1987), 'The Four Alternative Auditory Feature Test (FAAF) - Linguistic and Psychometric Properties of the Material with Normative Data in Noise', Journal of British Audiology, **21**, Pages 165-174.