MULTI-MICROPHONE NOISE CANCELLATION 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

ABSTRACT

A scheme for binaural pre-processing of speech signals for input to a standard linear hearing aid has been investigated. 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 large scale series of intelligibility tests are presented from experiments in which acoustic speech and noise data, generated using simulated and real-room acoustics 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 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].

The Multi-Microphone Sub-Band Adaptive (MMSBA) signal processing scheme 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 [5].

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 and real-room acoustics.

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.

A room of similar dimensions to that of the simulated room above was also used to make real room recordings. Unlike the room used for simulation furnishings were present, in the form of tables, chairs etc. These features were included to create as 'typical' a living room situation as possible. The recordings were made using a KEMAR manikin with the microphones connected within the ear canal. Hence, allowing for the inclusion of the head shadow effect present within the simulated room. The orientation of the speech and noise sources were as implemented for the simulation.

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 [6,7], 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 that only have one input, such as a standard linear hearing aid.

The simulated acoustic transfer function is generated using a program based on the image method [8]. 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.

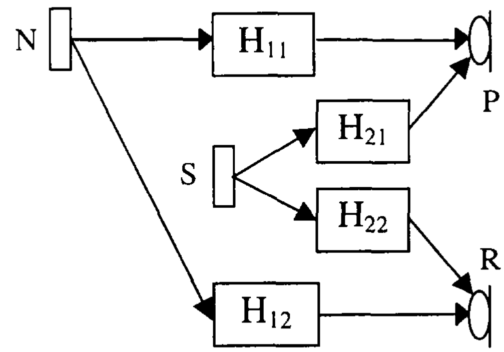

Figure 1a: Acoustic Model.

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 for both acoustic conditions were 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}.

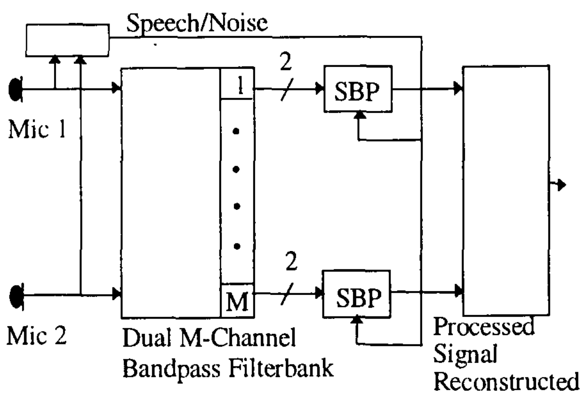

Figure 1b: Sub-band Decomposition

2.3. Sub-band Processing Scheme

Each Sub-band Processing unit (SBP) has an LMS adaptive filter to perform 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 relating the location of the noise masker. The adaptive filter algorithm implemented was the LMS algorithm [9]. 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. In some of the higher frequency bands the speech information generally has a higher coherence than the more distinct noise source. This can take advantage of an approach described by Ferrara Widrow [10]. In these bands the system is continually adapted 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.

3. INTELLIGIBILITY TESTING

The results being presented are from an experiment involving 15 hearing-impaired volunteers of between 40 and 77 years of age. All subjects had moderate sensorineural hearing loss that had been established through prior audiometric testing. Each subject had his or her hearing aid NAL curve [11] matched using an eight-band graphic equaliser. Subjects were tested using speech masked by speech shaped noise at four SNRs, two sub-band spacings [12], and three sub-band distributions. The subjects were presented with the data in a four choice forced response approach using the FAAF data set [13] that yields a "chance level" of 20.

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 reverberation levels involved were T_{60}=0.35s for the simulated room, and approximately T_{60}=0.3s from measurements of the real-room. These are representative of a typical living room level of reverberation [14]. The SNR's were -6, -3, 0, and +3dBs, chosen by experimentation to elicit a significant number of errors. Comparing results from Shields & Campbell [5] with those of the original authors of the FAAF test, Foster & Haggard [12], verified the experimental methodology employed.

4. RESULTS

The analysis performed on the experiment aims to answer the following questions regarding the intelligibility scores:

- Is there any significant difference in intelligibility scores when comparing the simulated room acoustics to the real-room acoustics?
- Is there a significant intelligibility improvement due to processing?
- Does the processing have a degrading effect on intelligibility when the SNR is high?
- Is there any significant effect of processing using different numbers of sub-bands?
- Is there any significant effect of processing using different sub-band spacing?

Figure 2 shows unprocessed scores for simulated and real-room acoustics averaged across 15 subjects at four different levels of SNR including the 95% confidence intervals. For all values of SNR the 95% confidence intervals are overlapping.

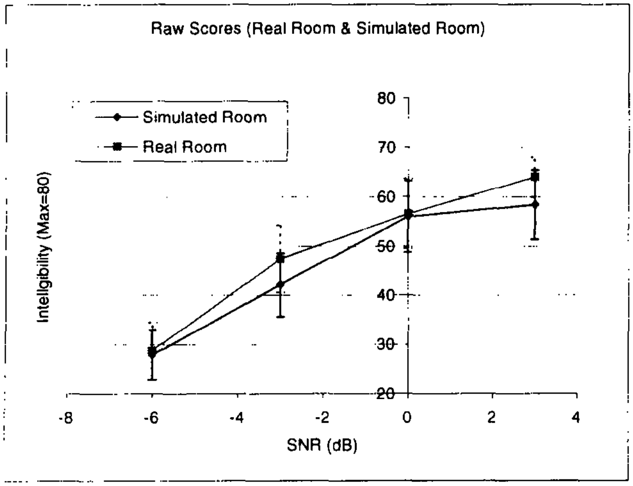

Figure 2 Raw intelligibility scores for simulated and real-room acoustics.

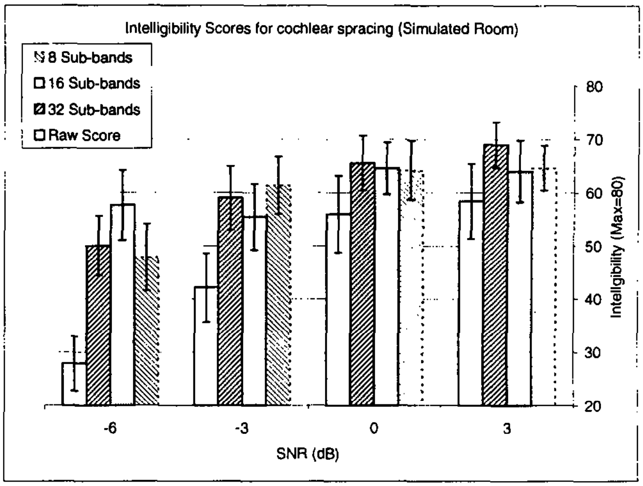

Figure 4 Intelligibility scores for simulated room acoustics and cochlear spacing.

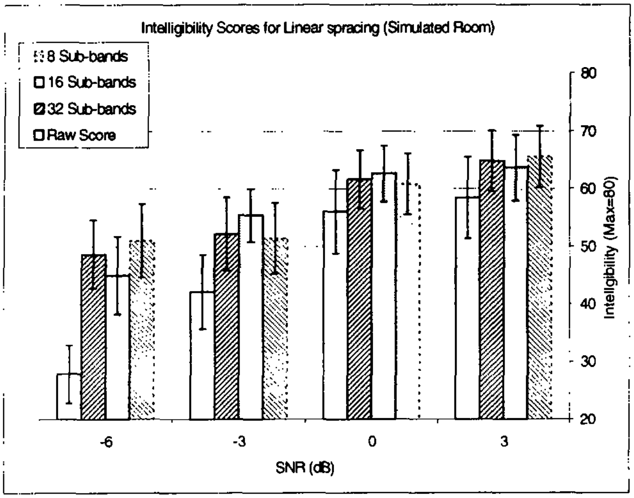

Figure 3 Intelligibility scores for simulated room acoustics and linear spacing.

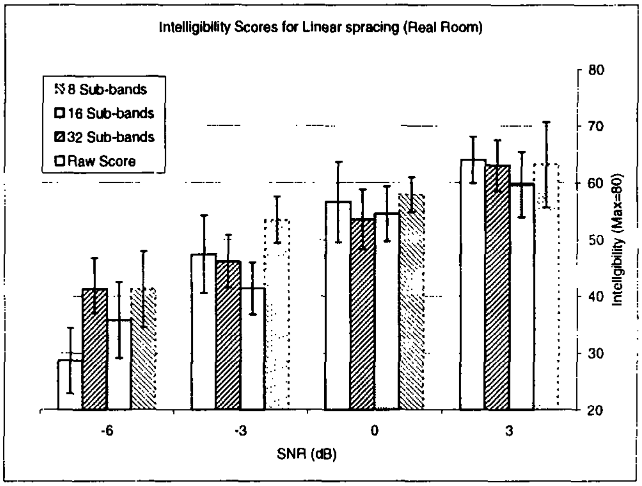

Figure 5 Intelligibility scores for real-room acoustics and linear spacing.

Therefore, the simulated and real-room acoustical raw conditions are statistically inseparable.

Figures 2,3,4,5,6,7 and 8 it can be seen that in most cases there is a significant improvement in speech intelligibility with processing at −6 and −3dB SNR where the unprocessed scores are initially low. This demonstrates a consistent improvement when processing at the 95% confidence level.

It can be seen from examination of figures 3,4,5 and 6 that when the SNR is high, e.g. plus 3 dB, there is no significant degradation to intelligibility due to processing, where the 95% confidence intervals from the raw and processed conditions are overlapping for all conditions. 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.

From figures 3,4,5 and 6 it can also be concluded that the number of sub-bands has not been demonstrated to be a dominant factor in the processing scheme, in all instances there is no significant separation at the 95% confidence level.

Figures 7 and 8 show the effect of sub-band spacing usingboth simulated and real-room acoustics. In both instances the cochlear spaced sub-band processing is consistently better at all SNR's. This effect is not statistically significant at the 95% confidence level. Further analysis is currently being performed with a more refined series of tests to examine the effect of all factors on the processing scheme.

5. CONCLUSIONS

The MMSBA processing scheme has been shown to significantly improve the intelligibility of speech corrupted with noise in both simulated and real-room moderately reverberant environment by up to **37.25%**. 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 cochlear spacing performs better than linear spacing. The number of sub-bands has not been demonstrated to be a significant factor.

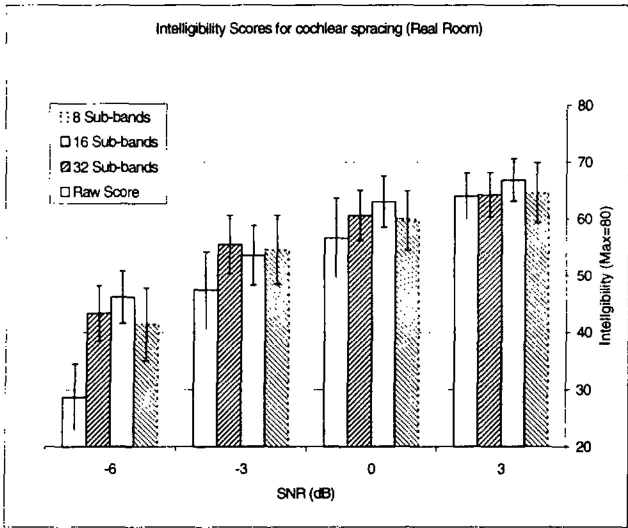

Figure 6 Intelligibility scores for real-room acoustics and cochlear spacing.

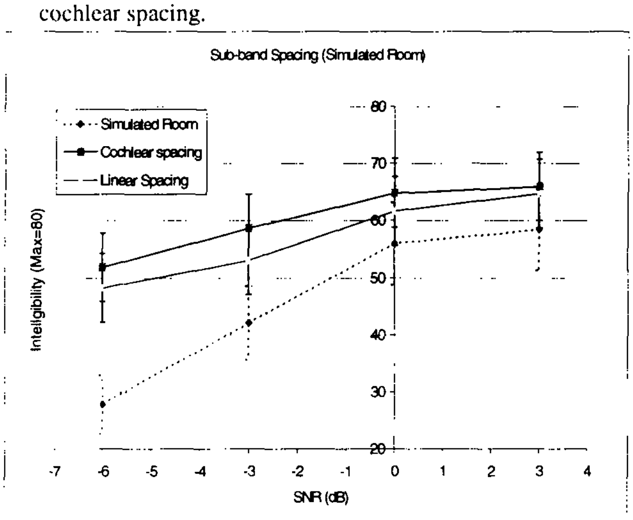

Figure 7. Intelligibility scores examining sub-band spacing for simulated room acoustics.

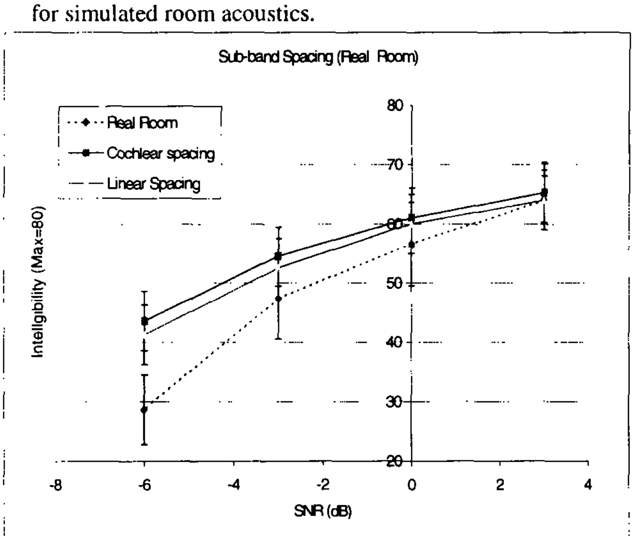

Figure 8. Intelligibility scores examining sub-band spacing for real-room acoustics.

6. 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 Leverhulme Trust fellowship during the writing of this paper.

7. 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] 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.

[6] Moncur, J.P., Dirks, D., (1967). 'Binaural and Monaural Speech Intelligibility in Reverberation'. J. of Speech and Hearing Research, **10**, 186-195.

[7] 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

[8] 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.

[9] Stearns, S. D., David, R A., *Signal Processing Algorithms*, New Jersey, Prentice-Hall, 1988.

[10] Ferrara, E.R., Widrow, B., (1981). 'Multichannel Adaptive Filtering for Signal Enhancement', IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. ASSP-29, **3**, 756-770.

[11] 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**,257-264.

[12] 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.

[13] 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.

[14] Mackenzie, G.W,. *Acoustics*, London & New York. Focal Press, 1964