MULTI-MICROPHONE SUB-BAND ADAPTIVE SIGNAL PROCESSING FOR IMPROVEMENT OF HEARING AID PERFORMANCE: PRELIMINARY RESULTS USING NORMAL HEARING VOLUNTEERS

Paul W. Shields and Douglas 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 system for the binaural pre-processing of speech signals for input to a standard linear hearing aid has been proposed. The work is based on that of Toner & Campbell [1] which applied the Least Mean Squares (LMS) algorithm in sub-bands to speech signals from various acoustic environments and signal to noise ratios (SNR). The method 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 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 normal hearing 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].

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 [4]. 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].

An alternative approach aims to improve the intelligibility by attempting to increase the SNR e.g., beamforming [3,5], Spectral Subtraction [6], or binaural noise reduction [7]. These methods have proved more successful particularly that of Soede et al [3], who demonstrated an overall improvement of approximately 7dB.

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 experiments reported here aimed to establish whether the measure of SNR improvement translates to a significant intelligibility improvement. This pilot test on normal hearing volunteers should give a useful indication as to the likelihood of success on hearing impaired subjects [8]

2. THE MMSBA PROCESSING SCHEME

The experiment aims to model a realistic scenario in which a person suffering from sensorineural hearing loss would have difficulty understanding speech. 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 1 represents the complete procedure of the simulation and processing mechanism. The section before the summing junction 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 [9,10], which 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

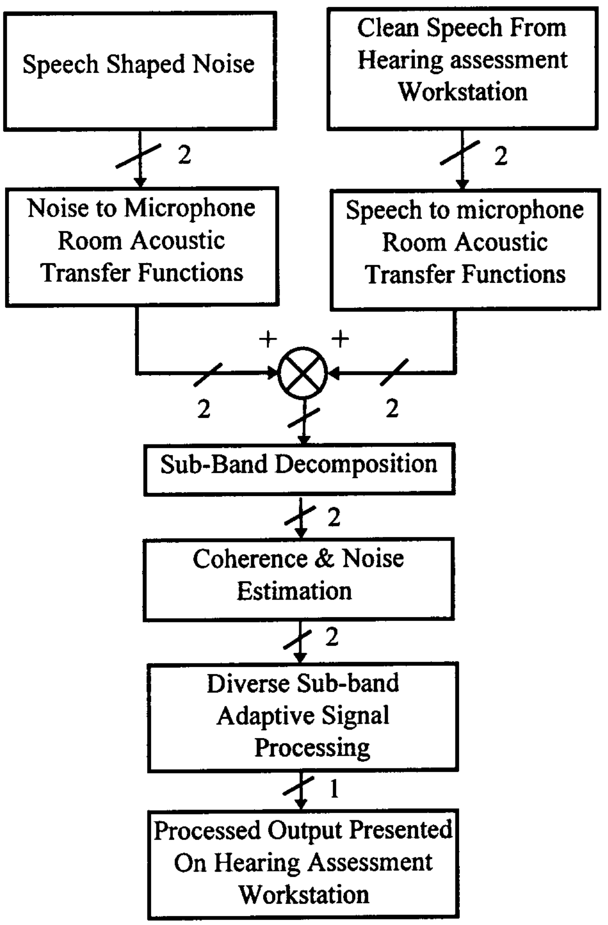

Figure 1: MMSBA Simulation and Processing

aid. The model of the acoustic transfer functions is generated using a program based on the image method [11]. 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. The remainder of the diagram illustrates the adaptive noise cancellation section. 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 noise masker. The adaptive filter algorithm implemented used the LMS algorithm [12]. When speech is present, the weights in the adaptive filter were 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 noise source. This can take advantage of an approach described by Ferrara Widrow [13]. 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 outputs from each sub-band are then summed to provide a full-band noise-reduced output for evaluation by the test subjects.

3. INTELLIGIBILITY TESTING

The intelligibility test involved 10 normal hearing volunteers of between 22 and 55 years age, with hearing levels being established as normal by prior audiometric testing. Subjects were then tested using speech masked by speech shaped noise at four SNRs, three reverberation levels, two sub-band spacings [14,15], and three sub-bands. The subjects were presented with the data in a four choice forced response approach using the FAAF data set [16].

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 the number of clean speech practice sentences required until they were familiar with the procedure.

The reverberant levels chosen were T_{60}=0s, T_{60}=0.35s, and T_{60}=1.8s. These figures are representative of anechoic, typical living room, and concert hall levels of reverberation [17]. The levels of SNR were -11, -9, -7, and -4dBs, chosen by experimentation to elicit a significant number of errors.

4. RESULTS

Figures 2,3 and 4 show unprocessed and processed scores at different levels of reverberation.

Figure 2 incorporates original speech intelligibility scores for the FAAF test [16]. These scores were obtained using 16 normal hearing volunteers; the competing noise being digitally mixed and having the same long term spectrum as the speech. From fig. 2 it can be seen the Foster & Haggard curve and the response

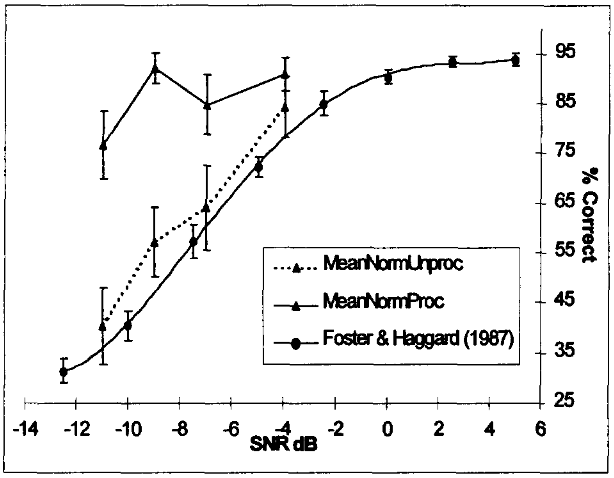

Figure 2: Intelligibility scores: Anechoic condition

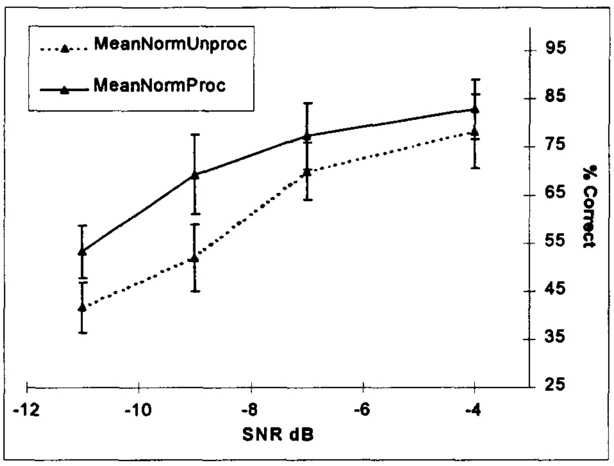

Figure 3: Intelligibility scores: T_{60}=0.35s condition

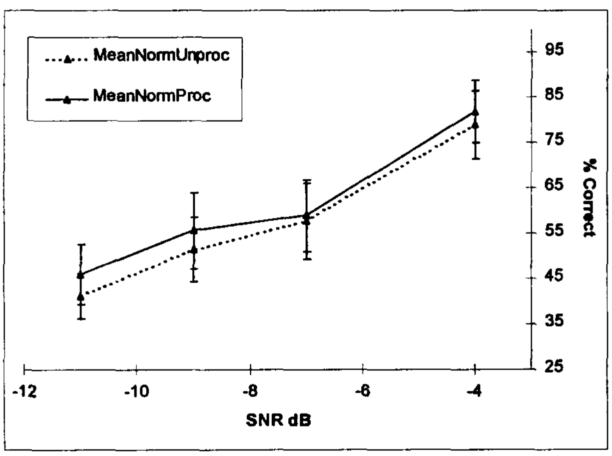

Figure 4: Intelligibility scores: T_{60}=1.8s condition

for the unprocessed anechoic case are consistent within 95% confidence intervals.

From this it can be concluded that the experimental methodology employed is a valid one.

Questions raised when examining the effect on intelligibility of the data are:

- Is there a significant enhancement due to processing, when results are blocked by reverberation?
- Does the processing have a degrading effect on intelligibility, when results are blocked by reverberation?
- At high levels of reverberation, where the unprocessed scores are large, is there any degradation in subject scores?
- Is there any effect of processing using different numbers of sub-bands, and different sub-band spacing?

An analysis of variance (ANOVA) was performed using processing, SNR, reverberation, spacing of sub-bands and number of sub-bands as factors, the only ones which are not significant at the 95% confidence level were the number and spacing of the sub-bands within the processing.

The means of the unprocessed and processed scores across all treatments were 46.38 and 56.68 respectively out of 80.

The improvement in mean score and 95% confidence interval due to processing is: **10.3** $\pm$ 1.95

It is justified therefor to claim an average 10 word increase in score across all reverberation levels.

It can be seen from figure 4, that there is little improvement with processing for the highly reverberant condition T_{60}=1.8s. An ANOVA performed using only the low and moderate levels of reverberation yielded average scores for the unprocessed and processed scores of 47.51 and 61.52 respectively. This results in an improvement in score of: **14.01** $\pm$ 2.22

At relatively high SNRs, where subjects would normally perform relatively well with unprocessed data, we wish to avoid depressing intelligibility. It is therefore interesting to examine the effect of the MMSBA processing scheme at -4dB. A two sample t-test was performed on the -4dB data. The results of which indicate there is no significant degradation in intelligibility at the 95% confidence level. Table 1 presents the mean score improvement and standard deviation for each number of sub-bands

The ANOVA examination of the effect of sub-band spacing and number of sub-bands showed that there was no significant effect due to either factor at the 95% confidence level.

Number of Sub-bands	8	16	32
Mean Score Improvement	13.15	12.81	6.5
Stdev	12.76	12.16	6.8

Table 1: Sub-band analysis.

5. CONCLUSIONS

The MMSBA has been shown to significantly improve the intelligibility of speech corrupted with noise in a low or moderately reverberant environment by on average 17.5%. It has been shown that the processing has no detrimental effect on intelligibility in high SNR, or when the level of reverberation is large.

It appears that a system with either 8 or 16 sub-bands would be sufficient for further tests.

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, and Dr. E. MacArthur from the Department of Mathematics and Statistics at the University of Paisley. 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] 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.

[5] 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

[6] 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.

[7] 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.

[8] 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.

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

[10] 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.

[11] 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.

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

[13] 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.

[14] Darlington, D.,J., Campbell, D.,R., (1996b). '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.

[15] Darlington, D.,J., Campbell, D.,R., (1996a). 'Sub-Band Adaptive Filtering Applied to Speech Enhancement', ESCA Tutorial and Research Workshop on 'The Auditory Basis of Speech Perception', Keele University, United Kingdom, 15-19 July, 241-244.

[16] 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, Vol. 21, Pages 165-174.

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