SPEECH ENHANCEMENT FOR HEARING AIDS

Douglas R. Campbell

Department of Electrical and Electronic Engineering, University of Paisley,
High Street, Paisley, Scotland, UK, PA1 2BE

Tel/Fax: +44 (0)141 848 3400/3404, email: d.r.campbell@paisley.ac.uk

ABSTRACT

The performance of hearing aids in noisy reverberant surroundings remains a major source of complaint and discomfort to wearers. Given the current capabilities and pace of development in microelectronics, the major problem is to find successful speech enhancement schemes. "Binaural unmasking" experiments demonstrate an enhancement advantage, due to binaural correlation properties, which can lower the hearing threshold in noise and there is evidence that this may operate in frequency sub-bands. The performance is presented of an adaptive sub-band noise cancellation scheme which supports the possibility of performing "binaural unmasking" outwith the body, and is shown to be capable of out-performing a standard noise-cancellation scheme in the presence of reverberation.

1. INTRODUCTION

The separation of desired from undesired sounds, especially in reverberant environments, is degraded by sensorineural hearing loss (SHL). The major cause of SHL is ageing of the cochlear structures and European Community demographics indicate that the over 60's will increase to 25% of the population by 2020 AD. Government estimates predict that in the UK by 2025 the number of over-65s will have grown from the present 9m to nearly 12m, and the number of over-75s by nearly 40%. Particularly important for education, safety and social interaction, is the separation of desired speech from background noise or competing speech in a reverberant environment. Available hearing aids and cochlear implants have little success in restoring this ability [1] and it remains a frequent source of complaint and discomfort to their users. The problem is one of separating two (or more) signals whose spectra overlap. Given the current pace of development in microelectronics, the major problem is to find successful processing schemes to enhance speech.

2. THE ACOUSTIC ENVIRONMENT

The correlation or coherence of the acoustic signals at two omnidirectional microphones, separated by a fixed distance in a room, varies with frequency. In the case of two microphones separated by free space in a steady sound field created by a sinusoidal source, the variation with frequency of the spatially averaged correlation exhibits a sinc form, with its first zero inversely proportional to the microphone spacing. The presence of a "head" between the two microphones (ears) modifies this correlation and shifts the first zero of the sinc function to a lower frequency [2]. Although of analytic interest, this is a highly artificial circumstance for a human listener, steady sinusoidal sound fields are uncommon and spatial localisation not averaging is the norm. The geometry, structures, and materials affect the acoustic characteristics of a room including its reverberation time T, and further modify the frequency dependent relationship between the microphone signals. Fig 1 shows the effect of the presence or absence of a simulated head (0.18m sphere) on the magnitude squared coherence between two microphones at a fixed location in a simulated reverberant room (W=5m, L=6m, H=4m, T=0.35s) excited by a noise source 1m distant, at 135deg azimuth, shaped to simulate a speech spectrum obtained from [3]. The presence of a "head" between the two "ears" can be seen to reduce the coherence of signals above about 1.5kHz. This implies that with a speaker from the front, on the sagital axis (0 deg), the high frequency low amplitude sibilant components of the received speech will have a different coherence from interfering speech or noise from off-sagital axis signals.

The effect on coherence of an apparently simple change in geometry caused by changing the relative distance of the noise source to the "head" is shown in Fig 2. It is obvious that over the distances humans normally converse directly, that these are likely to be significant effects.

3. HUMAN HEARING

Current evidence suggests the auditory brainstem receives from the cochlea a two-channel set of time-domain signals in contiguous non-linearly spaced frequency bands, and supports the separation of: left from right ear signals, low from high frequency signals, timing from intensity information; and their re-integration at various processing centres in the hierarchy. Competing connected speech has a lower masked threshold of intelligibility than continuous speech-noise [4] suggesting that the gaps in signals aid interference reduction. Experimental evidence also suggests that the auditory system is able to model a communications channel [5]. Thus enhancement and source location may be aided by estimation of the acoustic transfer function path difference between the interference and the desired signal. Binaural hearing has been found superior to monaural at maintaining the intelligibility of speech in the presence of reverberation, continuous speech shaped noise, or competing connected speech [4]. "Binaural unmasking" experiments demonstrate that binaural correlation properties can lower the hearing threshold in noise [6] and there is evidence that this may operate in frequency sub-bands [7]. However, masking release appears largely independent of the pattern of interaural correlations across frequency [8], and does not support masking release by source segregation through grouping frequency components with common interaural time delay (ITD). Since lateralisation is not necessary for effective binaural masking release [6] it is possible that the sub-band signals are being grouped for selective processing dependent on their degree of interaural correlation rather than their ITD value. This suggests that the main enhancement advantage of binaural hearing may be in the ability to perform “binaural unmasking”. An engineering implementation of this, could offer the possibility of performing “binaural unmasking” outwith the body.

4. EXPERIMENTS WITH DIVERSE SUB-BAND ADAPTIVE PROCESSING (DSBAP)

Speech enhancement combining multi-microphone methods with intermittent adaptive processing and diversity of processing within sub-bands (Figs 3 & 4) allows noise features within sub-bands to influence the subsequent processing. Sub-band operation gives faster adaption through the freedom to use different adaptive step-sizes in each band [9,10]. Separate decisions can be made on the appropriate form of processing for each sub-band. The inherent parallelism of the approach allows for implementation by parallel processors. When noise correlation between channels varies over sub-bands, DSBAP has shown signal to noise ratio improvement (SNRI) over conventional wide-band LMS (CLMS) and uniform sub-band adaptive processing (USBAP) [11]. Extending this work, realistic noise data was collected in a room containing desks, cabinets, computer system etc., using a pair of microphones. Sixteen different configurations representative of everyday conditions for the microphone-to-microphone (MTM) spacing, noise-to-microphone (NTM) distance and azimuths (AZ) of the noise source with respect to the microphones were applied, at raw SNR’s ranging from 6.9dB to 11.7dB at the microphone pair. The NTM distances were chosen to provide, a near-field case where the direct radiation would dominate, an intermediate case, and a case with significant reverberation. This acoustic noise data was then combined with clean monophonic speech and processed using CLMS, USBAP and DSBAP.

Summary statistics and a Kolmogorov-Smirnov test indicated that the treatment samples were not statistically normal. The raw unprocessed data and the three treatments were revealed as significantly different at the 5% level (Table 1) by a Wilcoxon matched pairs signed-ranks test.

Summary statistics of the SNR of the raw, CLMS, USBAP and DSBAP data show increasing mean value and spread, with the trend being to stretch the upper end of the distributions. This implies that with DSBAP, more configurations are being processed to better effect. Table 1 also presents correlation coefficients calculated across configurations between the SNR’s of the raw data and the CLMS, USBAP and DSBAP treatments. The strong correlations of CLMS with USBAP, and USBAP with DSBAP are expected since USBAP essentially applies CLMS in each sub-band, while DSBAP employs CLMS in some sub-bands. DSBAP obtained the highest correlation with the raw SNR (0.58) indicating that the result of DSBAP tends to maintain the same relationship with configuration as the raw data. Graphical comparison (Fig 5) of SNRI achieved by the DSBAP method against the CLMS and USBAP treatments indicate that useful SNRI are delivered by the DSBAP method for many of configurations tested, especially those with higher reverberation, a conclusion supported by informal listening tests.

5. CONCLUSION

In comparisons with the conventional wide-band ANC approach, DSBAP has been shown to be less susceptible to the effects of reverberation, and to deliver larger SNRI over a majority of the configurations reported here. This offers a possibility of performing the human hearing process of “binaural unmasking” outwith the body, providing signals of improved SNR to the better ear, a conventional aid or a cochlear implant processor.

6. ACKNOWLEDGMENTS

The author acknowledges support from Scottish Office Home and Health Department project grant K/RED/4/C237 and The Leverhume Trust during the writing of this paper; Prof. S. Gatehouse and Dr. J. Culling of the MRC Institute of Hearing Research at Glasgow Royal Infirmary and University of Nottingham respectively for their support and advice; and Mr P Shields and Mr D J Darlington for the room simulation data.

REFERENCES

[1] **Plomp, R. (1994):** Noise, amplification, and compression: Considerations of three main issues in hearing aid design, *Ear Hear.*, 15, pp2-12.

[2] **Lindevald, I.M. and Benade, A.H. (1986):** Two ear correlation in the statistical sound fields of rooms, *J. Acoust. Soc. Am.*, 80(2), pp661-664.

[3] **Byrne, D et al (1994):** An international comparison of long-term average speech spectra, *J. Acoust. Soc. Am.*, 96(4), pp2018-2120.

[4] **Plomp, R. (1976):** Binaural and monaural speech intelligibility of connected discourse in reverberation as a function of azimuth of a single competing sound source (speech or noise), *Acustica*, 34, pp200-211.

[5] **Darwin, C. J., McKeown, J. D. and Kirby, D. (1989):** Compensation for transmission channel and speaker effects on vowel quality, *Speech Communication*, 8(3), pp221-234.

[6] **Carhart, R., Tillman, T. W. and Johnson, K. R. (1968):** Effects of interaural time delays on masking by two competing signals, *J. Acoust. Soc. Am.*, 43(6), pp1223-1230.

[7] **Bernstein, L. R. and Trahiotis, C. (1992):** Discrimination of interaural envelope correlation and its relation to binaural unmasking at high frequencies, *J. Acoust. Soc. Am.*, 91(1), pp306-316.

[8] **Culling, J. F. and Summerfield, Q. (1995):** Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay, *J. Acoust. Soc. Am.*, 98(2), pp785-797.

[9] **Toner, E. and Campbell, D. R. (1993):** Speech enhancement using sub-band intermittent adaption, *Speech Communication*, 12, pp253-259.

[10] **Mahalanobis, A., Song, S., Mitra, S. K. & Petraglia, M. R. (1993):** Adaptive FIR filters based on structural subband decomposition for system identification problems, *IEEE Trans. Circs. Syts. - II: Analog and digital signal processing*, 40(6), pp375-381.

[11] **Campbell, D.R. (1994):** Adaptive speech enhancement with diverse sub-band processing, *Signal Processing VII, Theories and Applications*, **II**, pp1210-1213.

Treatment	Two-tailed Probability Values			Correlation Coefficients		
	Raw	USBAP	DSBAP	Raw	CLMS	USBAP
CLMS	0.044	0.008	0.004	0.19	-	-
DSBAP	0.002	0.001	-	0.58	0.83	0.94
USBAP	0.013	-	-	0.34	0.94	-

Table 1. Two-tailed Probability Values and Correlation Coefficients Across Configurations

Fig 1 Head effect on MSC, Room 5*6*4m, T=0.35s, NTM=1m

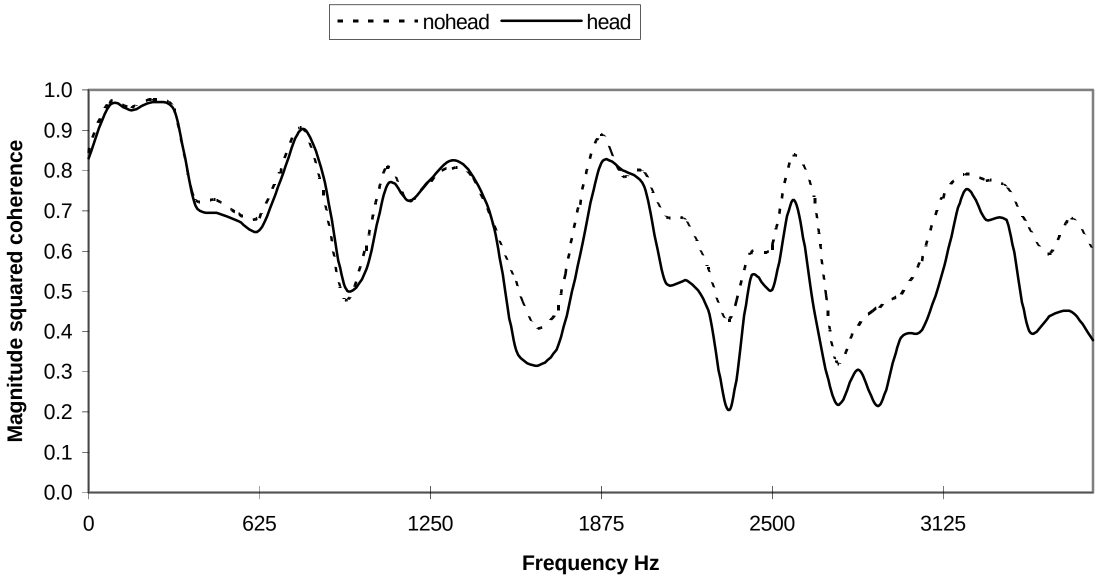

Fig 2 Effect of NTM on MSC: Room 5*6*4 m, T=0.35s

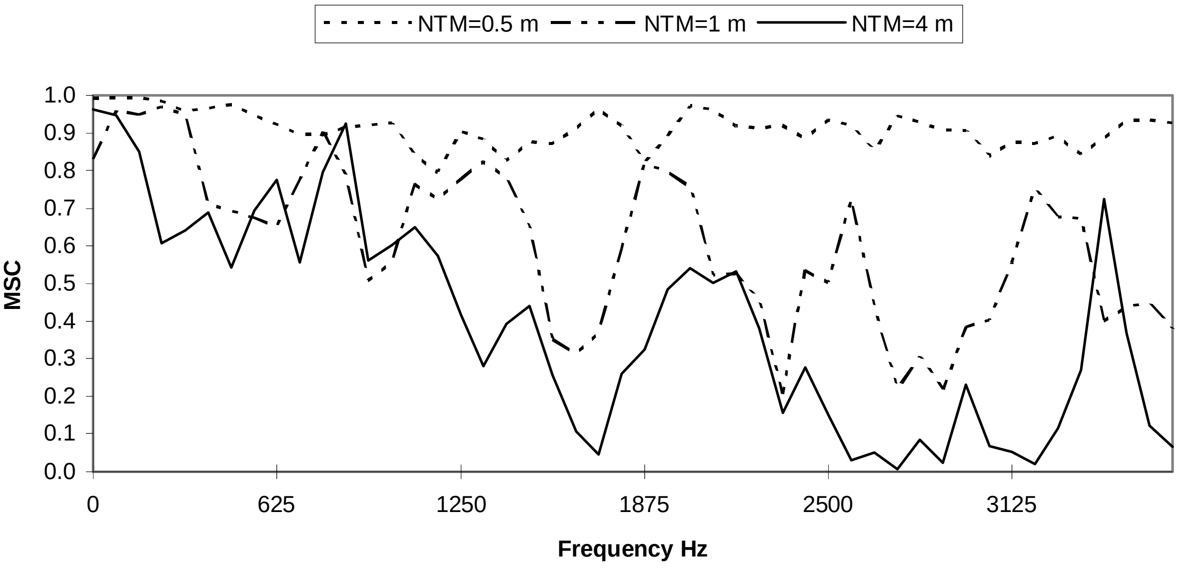

Fig 3 Diverse Sub-band adaptive processing

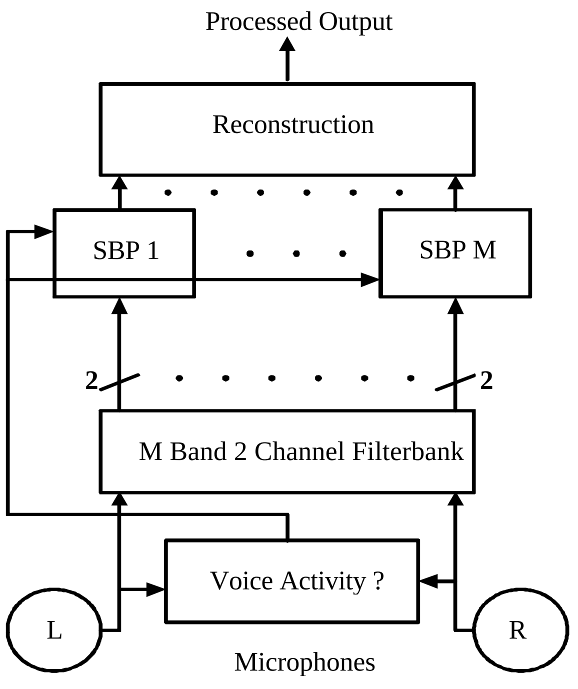

Fig 4 Sub-band processor (SBP)

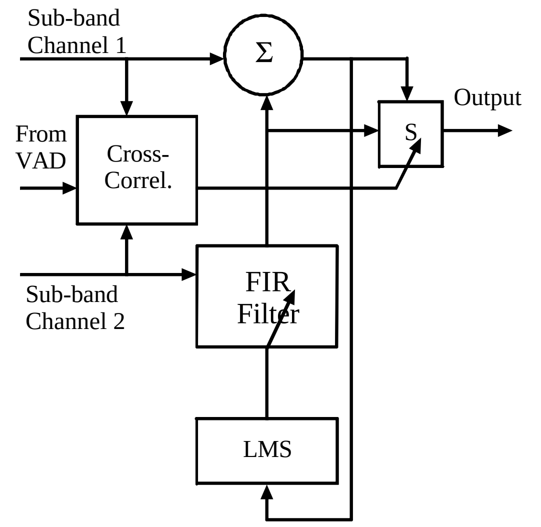

Fig 5. % SNRI vs Configuration and Treatment