

# SUB-BAND ADAPTIVE SPEECH ENHANCEMENT FOR HEARING AIDS

*D R Campbell*

Department of Electrical and Electronic Engineering, University of Paisley,  
High Street, Paisley, Scotland, UK, PA1 2BE  
Tel/Fax: 0141 848 3400/3404 Email: d.r.campbell@paisley.ac.uk

## ABSTRACT

This paper presents: a summary of features of the human auditory system and aspects of SHL that support suspicion of an adaptive sub-band binaural noise-cancellation process; description of a diverse sub-band adaptive processing (DSBAP) approach; experimental results indicating that DSBAP shows promise as a method of speech enhancement for hearing aids.

## 1. INTRODUCTION

A major cause of sensorineural hearing loss (SHL) is the age related cochlear degeneration, presbycusis. European Community demographics indicate that the over 60's will comprise 25% of the population by 2020 AD, and 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%.

A subject with SHL may retain the ability to communicate one-to-one in a quiet low-reverberation environment and benefit from: the considerable redundancy in speech; the presence of many contextual cues; in the case of an elderly person, a lifetimes auditory pattern recognition and language interpretation experience; knowledge of the location of the speaker by visual or other non-acoustic means; the ability to exploit "spectral contrast" and to perform a degree of monaural lateralisation in those with unilateral hearing deficiency; the likelihood that many sufferers may be able to use binaural lateralisation through access to low frequency signals.

In spite of the above, limited unilateral or bilateral cochlear damage can seriously degrade the intelligibility of speech in a complex acoustic environment containing noise and reverberation. In 1978 Plomp [1] reported little success from hearing aids in restoring this ability and it remains today a frequent source of complaint and discomfort to the users of hearing aids [2,3]. Given the current capabilities and pace of development in microelectronics, the major problem is to find successful processing schemes to enhance speech in everyday conditions.

## 2. HUMAN HEARING

The processing functions from the outer ear to the auditory nerve leaving the cochlea are accepted as: transduction, filtering, amplification, non-linear compression, impedance matching and the cochlear operation of spectral analysis. The processing centres of the auditory brainstem receive from the cochlea a two-channel set of time-domain signals in contiguous non-linearly

spaced frequency bands. Current interpretations of experimental evidence [4,5,6,7] identify 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 and levels in the hierarchy.

Plomp [8] found the masked threshold of intelligibility to be about 3dB lower for competing connected speech than for continuous speech-noise. It appears that the windows in either the desired or competing signal can be used to view and estimate the other for interference reduction [9,10]. Experimental evidence suggests that the auditory system is able to model a communications channel [11]. Thus signal enhancement and source location may be aided by estimation of the acoustic transfer function path difference between the interference and the desired signal.

### 2.1 Binaural Signal Processing

Binaural hearing is most obviously connected with the ability to locate (lateralise) signals. Binaural information can also be used to de-emphasise an undesired signal dependent on its binaural correlation properties, as demonstrated by "binaural unmasking" [12] which can lower the hearing threshold by 15dB or more [13, 14]. Moncur and Dirks [15] and Plomp [8] found binaural hearing superior to monaural at maintaining the intelligibility of speech in the presence of reverberation, continuous speech shaped noise, or competing connected speech. This, supported by the results of Carhart et al [13] emphasises the advantage of binaural noise-cancellation over simple coherent addition of the signals at the two ears.

Interaural correlation has been implicated in binaural unmasking [7,16]. This noise cancellation by contralateral inhibition appears located in the auditory brainstem [5,6,17], and there is evidence to suggest it may operate in frequency sub-bands [7,17,18]. However Culling and Summerfield [19] found that masking release was largely independent of the pattern of interaural correlations across frequency, and did not support masking release by source segregation (lateralisation) through grouping frequency components with common interaural time delay (ITD).

Since lateralisation is not necessary for effective binaural masking release [13], how can correlation aid binaural unmasking if not used for grouping by ITD? One possibility is 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, possibly as an adaptive noise cancellation process operating in frequency sub-bands. Adaptive noise cancellation (ANC) [20], is an operation at least superficially analogous to binaural unmasking, and the LMS algorithm offers a relatively simple and robust means of adaption [21]. It is likely that the ability to exploit adaption, intermittency, channel modelling, and lateral inhibition will be necessary features.

### 3. MULTI-MICROPHONE SUB-BAND ADAPTIVE PROCESSING

Speech enhancement combining multi-microphone methods with intermittent adaptive processing and diversity of processing in sub-bands (Fig.1) has been suggested by Toner and Campbell [22].

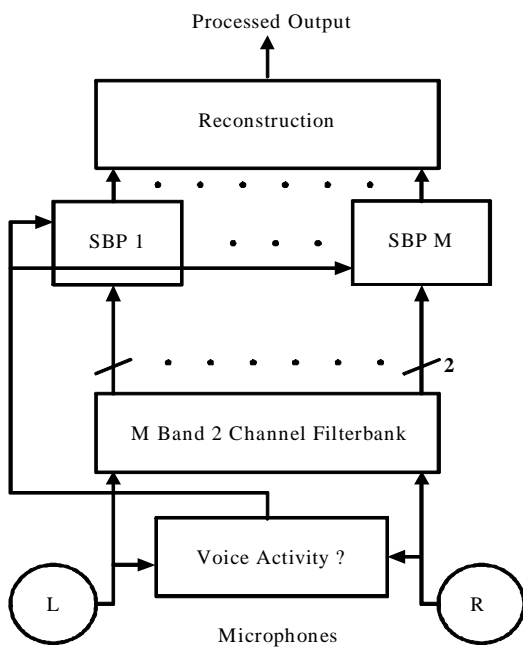


Figure 1: Diverse Sub-band adaptive processing

It allows noise features within sub-bands, such as the noise power, the correlation or coherence between signals from multiple sensors, and the behaviour of the adaptive algorithm (Fig. 2), to influence the subsequent processing during the “noisy speech” period. Close spacing of the sensors reduces the required order (complexity) of the adaptive filter and thus the sub-band computational load both for adaptive and intermittent fixed processing. It also reduces the misadjustment noise when using continuously adapting schemes. Sub-band operation gives faster adaption through the freedom to use different adaptive step-sizes in each band [23,24]. Separate decisions can be made on the appropriate form of processing for each sub-band. The inherent parallelism of the approach allows for future parallel processor implementation.

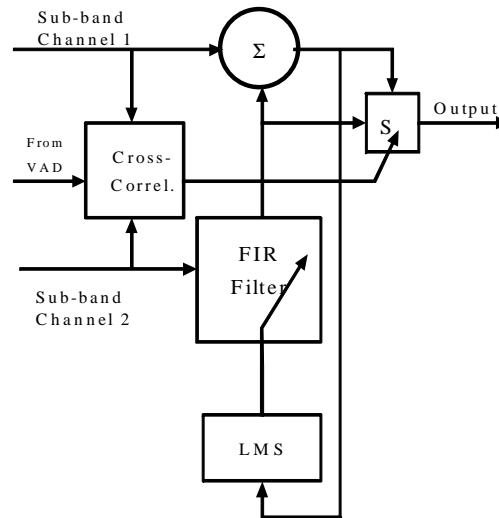


Figure 2: Sub-band processor (SBP)

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

Utilising the experience of previous experiments the form of sub-band processing was selected depending on the cross-correlation of the sub-band noise. Three ANC treatments were compared: Conventional wideband (CLMS); Identical processing in all sub-bands (USBAP); Sub-band processing chosen based on the inter-channel correlation of the noise (DSBAP).

#### 4.1 Simulated Environment

Initial testing of the method was by simulation of an echoic environment. Noise signals were synthesised with a correlation coefficient between L and R channels that varied with frequency (Fig. 3), and added to clean monophonic speech signals.

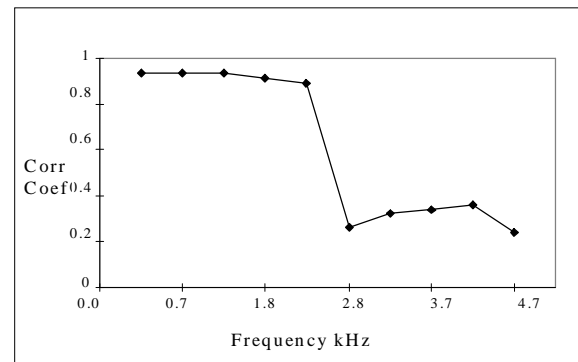


Figure 3: Correlation vs Frequency, simulated room.

The data was then processed by the three treatments. Typical comparative results (Fig. 4) indicate that DSBAP has potential for signal to noise ratio (SNR) improvement over the other methods when noise correlation varies between channels within sub-bands.

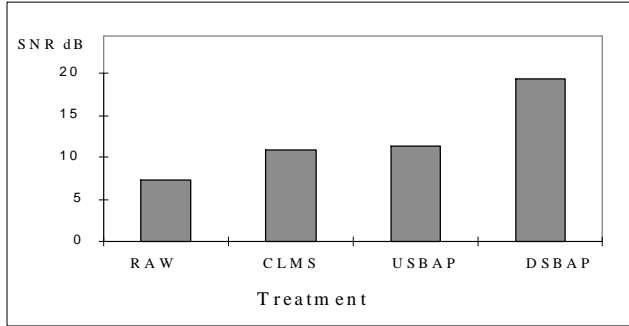


Figure 4: Signal to noise ratio vs Treatment, Simulated room.

## 4.2 Real Environment

Realistic noise data was gathered in a room containing desks, cabinets, computer system etc., using a pair of microphones. Three microphone-to-microphone (MTM) spacings (0.05m, 0.15m and 0.5m) were used. Three noise-to-microphone (NTM) distances and three azimuths (AZ) of the noise source with respect to the microphones were applied (0 deg. i.e. equidistant from the microphones, 45 deg., and 90 deg). The NTM distances were representative of everyday conditions (0.5m, 1.0m and 3.0m). The noise data was combined with clean monophonic speech and processed by the three treatments. Sixteen of the twenty-seven possible configurations for the independent variables NTM, MTM and AZ were used to test the three treatments at raw SNR's ranging from 6.9dB to 11.7dB at the microphone pair.

A Wilcoxon test revealed the raw (unprocessed) SNR's and those of the three treatments as significantly different at the 5% level. Summary statistics of the SNR of the raw, CLMS, USBAP and DSBAP data imply that DSBAP processes more configurations to better effect. Correlation coefficients between the SNR's of the raw data and the three treatments across configurations (Fig 5) indicate that the result of DSBAP tends to maintain the same relationship with configuration as the raw data, and suggest that the results of CLMS and USBAP may tend to be dominated by selected aspects of configuration.

Results for SNR improvement achieved by the DSBAP method compared with the CLMS and USBAP methods for all microphone and noise source configurations tested are shown in

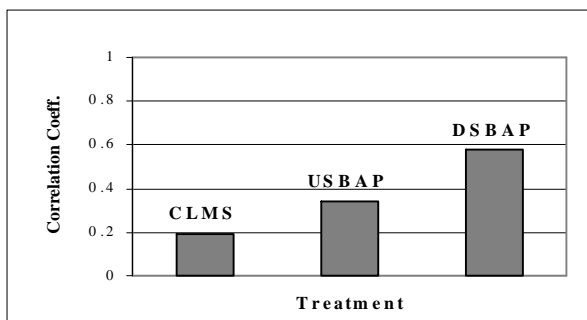


Figure 5: Correlation of Treatment with Raw

Fig. 6. In the case of a high correlation in all sub-bands (MTM small), low reverberation (NTM small) and the noise source relatively closer to one microphone (AZ = 90deg) CLMS performs better than USBAP or DSBAP as expected. Typically, but not uniformly, DSBAP is better than USBAP which is better than CLMS, a conclusion supported by informal listening tests. Comparison of USBAP with DSBAP indicates that the latter is degraded less by reverberation through taking sub-band noise correlation into account. Fig 6 also shows the general trend of SNR improvement decreasing as NTM increases (i.e. as direct-to-reverberant ratio decreases) in agreement with Lu and Clarkson using CLMS [25], with however DSBAP holding up better with increasing NTM.

## 5. CONCLUSION

Binaural unmasking appears to be an important enhancement process for noisy reverberant signals. Sub-band adaptive noise cancellation offers a possibility of performing "binaural unmasking" outwith the body, providing signals of improved SNR to the better ear, a conventional aid or a cochlear implant processor. In comparisons with the conventional wide-band ANC approach, DSBAP has been shown capable of lower susceptibility to the effects of reverberation, and to deliver SNR improvements for the majority of the configurations reported here. Current work is concentrating on defining the importance of adaptive filter order, the number and distribution of sub-bands, metrics for selecting between processing methods, and quantitative listening tests using a standard speech corpus and assessment methods.

## 6. ACKNOWLEDGMENTS

The author acknowledges support from Scottish Office Home and Health Department (K/RED/4/C237), The Leverhume Trust, and The Hearing Research Trust during the writing of this paper; and Prof. S. Gatehouse and Dr. J. Culling of the MRC Institutes of Hearing Research at Glasgow Royal Infirmary and University of Nottingham respectively for their support and advice.

## 7. REFERENCES

1. Plomp, R., "Auditory handicap of hearing impairment and the limited benefit of hearing aids", *J. Acoust. Soc. Am.*, 63(2):533-549, 1978.
2. Engebretson, A M., "Benefits of digital Hearing aids", *IEEE Engineering in Medicine and Biology*, April/May :238-248, 1994.
3. Plomp, R., "Noise, amplification, and compression: Considerations of three main issues in hearing aid design", *Ear Hear.*, 15 :2-12, 1994.
4. Wightman, F L. & Kistler, D J., "The dominant role of low-frequency interaural time differences in sound localization", *J. Acoust. Soc. Am.*, 91(3):1648-1661, 1992.
5. Pickles, J O., "An introduction to the physiology of hearing", Academic Press, San Diego, CA, 1992.
6. Feng, A S., "Information processing in the auditory brainstem", *Current Opinion in Neurobiology*, 2:511-515, 1992.

7. Bernstein L R. & Trahiotis C., "Discrimination of interaural envelope correlation and its relation to binaural unmasking at high frequencies", *J. Acoust. Soc. Am.*, 91(1):306-316, 1992.
8. Plomp, R., "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*, 34:200-211, 1976.
9. Gustafsson, H A. & Arlinger, S D., "Masking of speech by amplitude-modulated noise", *J. Acoust. Soc. Am.*, 95(1):518-529, 1994.
10. van der Heijden, M. & Kohlrausch, A., "The role of envelope fluctuations in spectral masking", *J. Acoust. Soc. Am.*, 97(3):1800-1807, 1995.
11. Darwin, C J., McKeown, J D. & Kirby, D., "Compensation for transmission channel and speaker effects on vowel quality", *Speech Communication*, 8(3):221-234, 1989.
12. Durlach, N I., "Binaural signal detection: Equalization and cancellation theory", in *Foundations of Modern Auditory Theory*, Tobias, J. V. (Ed), Vol II, Academic Press, London, 1972.
13. Carhart, R., Tillman, T W. & Johnson, K R., "Effects of interaural time delays on masking by two competing signals", *J. Acoust. Soc. Am.*, 43(6):1223-1230, 1968.
14. Gilkey R H & Good, M D, "Effect of frequency on free-field masking", *Human Factors*, 37(4):835-843, 1995.
15. J P Moncur & D Dirks, "Binaural and monaural speech intelligibility in reverberation", *J. Speech and Hearing Res.*, 10:186-195, 1967.
16. N I Durlach, K J Gabriel, H S Colburn & C Trahiotis, "Interaural correlation discrimination: II. Relation to binaural unmasking", *J. Acoust. Soc. Am.* 79(5):1548-1557, 1986.
17. Micheyl, C. & Collet, L., "Involvement of the olivocochlear bundle in the detection of tones in noise", *J. Acoust. Soc. Am.*, 99(3):1604-1610, 1996.
18. Jeffres, L A., "Binaural signal detection: Vector theory", in *Foundations of Modern Auditory Theory*, Tobias, J. V. (Ed), Vol II, Academic Press, London, 1972.
19. Culling, J F. & Summerfield, Q., "Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay", *J. Acoust. Soc. Am.*, 98(2):785-797, 1995.
20. Widrow, B. & Stearns, S D., *Adaptive Signal Processing*, Prentice-Hall, Englewood Cliffs, N.J., 1985.
21. Hassibi, B., Sayed, A H. & Kailath, T., "H<sup>∞</sup> Optimality of the LMS algorithm", *IEEE Trans. On Signal Process.*, 44(2):267-280, 1996.
22. Toner, E. & Campbell, D R., "Speech enhancement based conceptually on auditory processing", *ESCA ETRW on Speech Processing in Adverse Conditions*, Cannes-Mandelieu, France, 151-154, 1992.
23. Toner, E. & Campbell, D R., "Speech enhancement using sub-band intermittent adaption", *Speech Communication*, 12:253-259, 1993.
24. Mahalanobis, A., Song, S., Mitra, S K. & Petraglia, M R., "Adaptive FIR filters based on structural subband decomposition for system identification problems", *IEEE Trans. Circuits, Sys. - II: Analog and digital signal processing*, 40(6):375-381, 1993.
25. Lu, M. & Clarkson, P M., "The performance of adaptive noise cancellation systems in reverberant rooms", *J. Acoust. Soc. Am.*, 93(2):1122-1135, 1993.

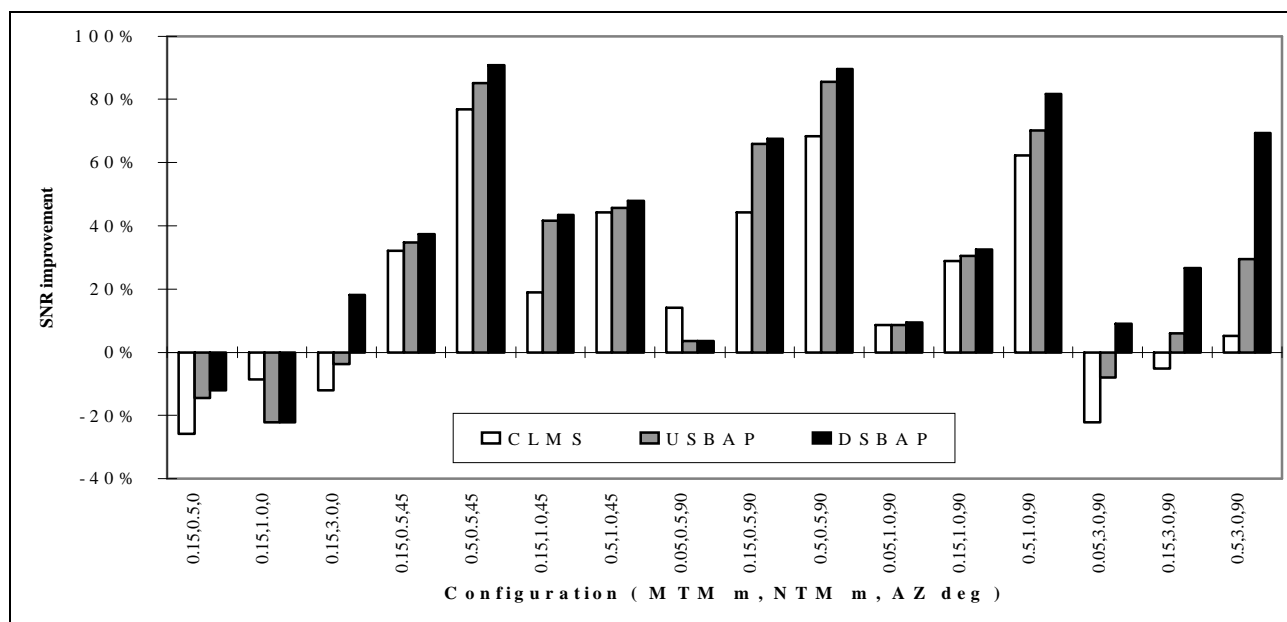


Figure 6: %SNR improvement vs Configuration and Treatment