TWO MICROPHONES SPEECH ENHANCEMENT SYSTEM BASED ON A DOUBLE FAST RECURSIVE LEAST SQUARES (DFRLS) ALGORITHM

M. Gabrea*, E. Mandridake*, M. Menez+, M. Najim* and A. VallauriT
* Equipe Signal et Image, ENSERB and GDR-134, CNRS
BP 99, 33 402 Talence, France
+ LASSY-I3S Nice, France
T Texas-Instruments, Villeneuve-Loubet, France
email: limby@goelette.tsi.u-bordeaux.fr

ABSTRACT

In this paper a symmetric feedback implementation scheme of a two microphones speech enhancement is presented. We consider the coupling systems modelled as a linear time-invariant Finite Impulse Response (FIR) filters and propose a new recursive-based adaptive filter solution to enhance the noisy speech . The optimum filter weight adaptation is based on a Double Fast Recursive Least Squares (DFRLS) algorithm. This approach can be extended for a subclass of signal separations where the direct link is stronger than the interference link in the both channels. A comparative study with other adaptive algorithms shows the superiority of the DFRLS in SNR performance improvement.

1 INTRODUCTION

Let us consider the system modeled by the diagram represented in the figure 1. The purpose is to recover the free noise speech signal s(n) from the two available observations $p_1(n)$ and $p_2(n)$ in the presence of the noise signal b(n).

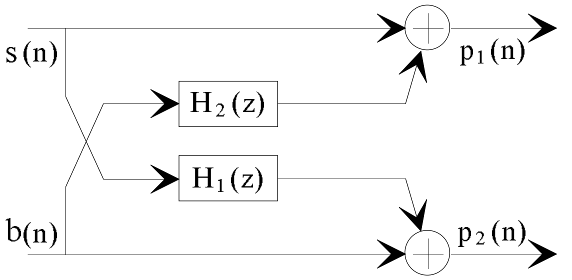

Fig. 1: Signal model for noise cancellation

The basic scheme of adaptive noise canceller given in [1] uses an adaptive filter based on the LMS algorithm for estimating the additive noise. In the simplified case where all filters are assumed to be single tap another system called Symmetric Adaptive Decorrelation (SAD) using two adaptive filters, as an extension of the classical LMS acoustic noise canceller, has been presented in [2]. In [3] the case in which the coupling systems are unknown FIR filters was considered. In this current case the signal is assumed to be non-Gaussian having an asymmetric probability distribution and the noise may be a Gaussian.

In this paper we present a new feedback implementation of a noise canceller based on the DFRLS algorithm. We only suppose that the speech signal and the noise are statistically independents and we consider the coupling systems being FIR filters. This algorithm can also be used for a subclass of signal separations where the direct link must be stronger than the interference link in the both channels. A comparative performance study is presented in the framework of noise cancellation.

The remainder of the paper is organized as follows. In the next section we present the Double Fast Recursive Least Squares algorithm. A comparative experimental study of different schemes and algorithms including the Normalized LMS and the Symmetric Adaptive Decorrelation is presented in section 3. We conclude by evaluating the performance of the proposed system.

2 THE DFRLS ALGORITHM

Figure 2 shows the feedback implementation of the noise canceller. $W_1(z)$ and $W_2(z)$ are two adaptive filters. Each one has as input the output error signal of the other filter.

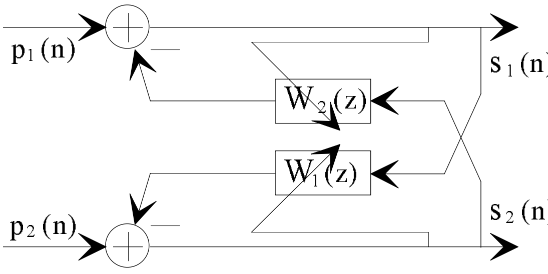

Fig. 2: Feedback implementation of the noise canceller

The optimum value, in the Wiener sense, of the tap-weight $W_i(z)$ (i=1,2) is obtained by minimising the cost function:

$$\varepsilon_i(n) = \sum_{j=1}^{n} \lambda^{n-j} \left| \mathbf{s}_i(\mathbf{j}) \right|^2 \tag{1}$$

where λ is a forgetting factor.

The Wiener-Hopf equations for the two filters are [1][6]:

$$\mathbf{W}_i(\mathbf{z}) = \frac{\mathbf{S}_{p_j s_i}(\mathbf{z})}{\mathbf{S}_{s_i s_i}(\mathbf{z})} \qquad i=1,2 \;\; j=1,2 \;\; i\neq j \tag{2}$$

where $S_{ab}(z)$ represents the crossing power density spectrum of the signals a(n) and b(n).
If the speech signal and the noise are statistically independents then optimum solutions are provided by the equations:

$$(\mathbf{W}_i(\mathbf{z}) - \mathbf{H}_i(\mathbf{z}))(1 - \frac{\mathbf{N(z)}}{\mathbf{D(z)}}) = 0 \quad i=1,2 \tag{3}$$

where:

$$\frac{\mathbf{N}(\mathbf{z})}{\mathbf{D}(\mathbf{z})} = \frac{\mathbf{S_{pp}}(\mathbf{z})\mathbf{S_{bb}}(\mathbf{z})\mathbf{N}_W\mathbf{N}_H}{(\mathbf{D}_{b_1} + \mathbf{D}_{p_1})(\mathbf{D}_{b_2} + \mathbf{D}_{p_2})} \tag{4}$$

where

$$\mathbf{N}_H = \left|1 - \mathbf{H}_1(\mathbf{z})\mathbf{H}_2(\mathbf{z})\right|^2$$

$$\mathbf{N}_W = \left|1 - \mathbf{W}_1(\mathbf{z})\mathbf{W}_2(\mathbf{z})\right|^2$$

$$\mathbf{D}_{b_1} = \mathbf{S_{bb}}(\mathbf{z})\left|1 - \mathbf{W}_1(\mathbf{z})\mathbf{H}_2(\mathbf{z})\right|^2$$

$$\mathbf{D}_{p_1} = \mathbf{S_{pp}}(\mathbf{z})\left|\mathbf{W}_1(\mathbf{z}) - \mathbf{H}_1(\mathbf{z})\right|^2$$

$$\mathbf{D}_{b_2} = \mathbf{S_{pp}}(\mathbf{z})\left|1 - \mathbf{W}_2(\mathbf{z})\mathbf{H}_1(\mathbf{z})\right|^2$$

$$\mathbf{D}_{p_2} = \mathbf{S_{bb}}(\mathbf{z})\left|\mathbf{W}_2(\mathbf{z}) - \mathbf{H}_2(\mathbf{z})\right|^2$$

$S_{pp}(z)$ and $S_{bb}(z)$ represent respectively the power density spectrum of p(n) and b(n).
We can see that the equations (4) provide multiple solutions. Among all these solutions we can find the desired solution $W_i(z) = H_i(z)$ (i=1,2). In this case it is easy to verify that $s_1(n)=s(n)$ and $s_2(n)=b(n)$ and it is possible to recover the signals that would have been measured at each microphone in the absence of the other source signal. If for each generating filter:

$$\sum \mathbf{h}_i^2(n) < 1 \qquad i=1,2 \tag{5}$$

then the filters $W_i(z)$ (i=1,2) converge to the desired solutions.
These desired solutions can be reached using a weight adaptive filters updating based on the LMS or RLS algorithm. We propose to use the Fast RLS (FRLS) algorithm for the following reasons. RLS algorithm has a rate of convergence typically an order of magnitude faster than the LMS algorithm. Among the existing versions of the FRLS algorithm which have appreciably lower computational complexity than RLS we have opted for the numerically stabilized [4] version well adapted to non-stationary input signal like speech. On the other hand the behaviour of these algorithms for real time applications and DSP implementations has been mastered [5].

3 SIMULATION RESULTS

The speech signal and the noise have been separately recorded in a car moving at a speed of 130 Km/h with closed windows. So, the noises are considered as ″road traffic″ noises. The two microphones are placed in front of the driver at about 40 cm from each other. The noises have been artificially added to the noise-free speech so that one would master the SNR input.
The coupling systems are 10 taps two FIR filters with. One considers the eventual delay between the two observation signals taken into account by one of the FIR filters.
The signal captured by the first microphone $p_1(n)$ and the second microphone $p_2(n)$ are respectively shown in figure 3a and 3b. The SNR of $p_1(n)$ and $p_2(n)$ are respectively 2.12 dB and -1.83 dB. The output and the cepstrogram signal of the noise canceller system using the DFRLS algorithm are shown in the figure 4. The first original desired speech signal $s_1(n)$ is shown in the figure 5. A comparative SNR output gain the between the Normalized LMS, the Double SAD and the DFRLS algorithms is provided in table 1. This table shows the superiority of the noise canceller DFRLS based algorithm.

Table 1: The SNR of $s_1(n)$ for different algorithms

Algorithm	NLMS	SAD	DFRLS
SNR $s_1(n)$ dB	3.5	8.68	13.18
Gain dB	1.38	6.56	11.06

This global performance behaviour is confirmed by the figure 6 which shows, frame by frame, the SNR output of the 3 different noise cancellers. The SAD algorithm takes more time before handling the noise field after which its segmental SNR behaviour is close to the segmental behaviour of the DFRLS algorithm. This fact is illustrated by the figure 7. As for the Normalized LMS, its segmental SNR behaviour is always lower since it is penalized by its slow convergence and therefore can't track the statistical change of the noise between two successive frames. However, during high energy regions, its behaviour is close to the two other algorithms. The reason is that noise is masked by the high energy speech regions, and hence does not require complexe treatement.

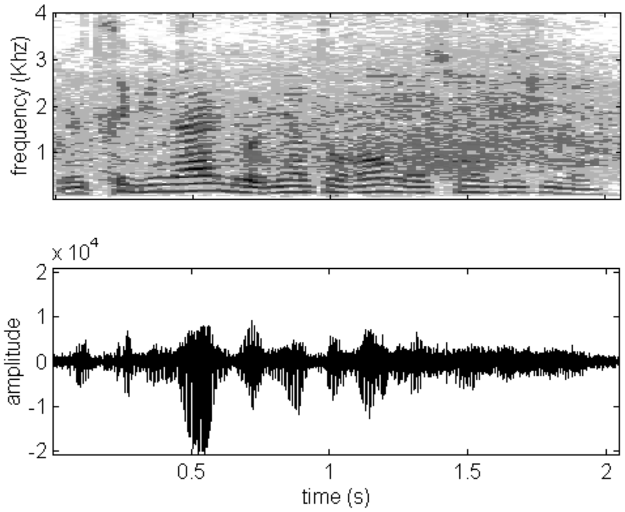

Fig. 3a: the signal $p_1(n)$ captured by the first microphone

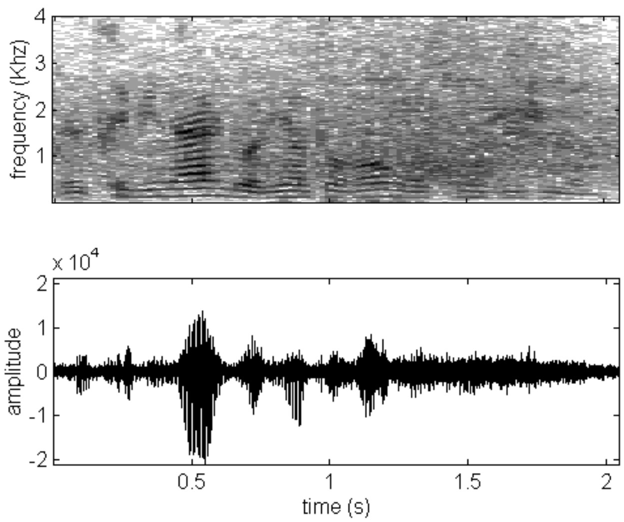

Fig. 3b: the signal $p_2(n)$ captured by the second microphone

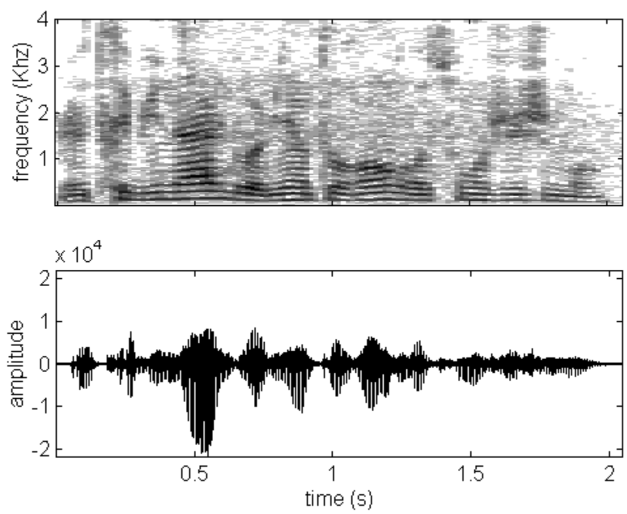

Fig. 4: Enhanced speech obtained with the noise canceller system based DFRLS algorithm

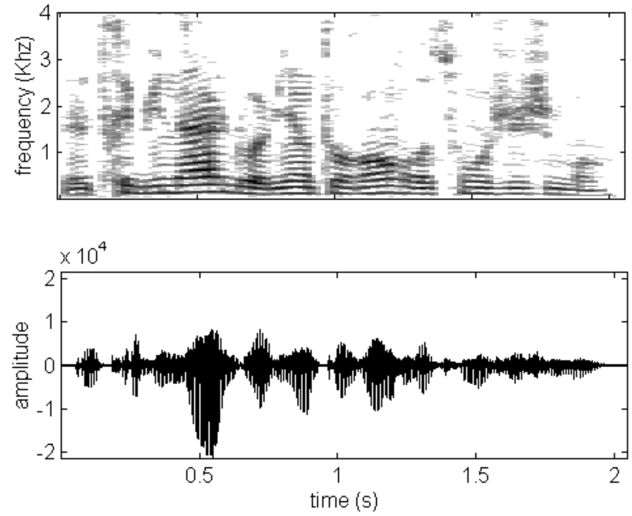

Fig. 5: The original speech signal s(n)

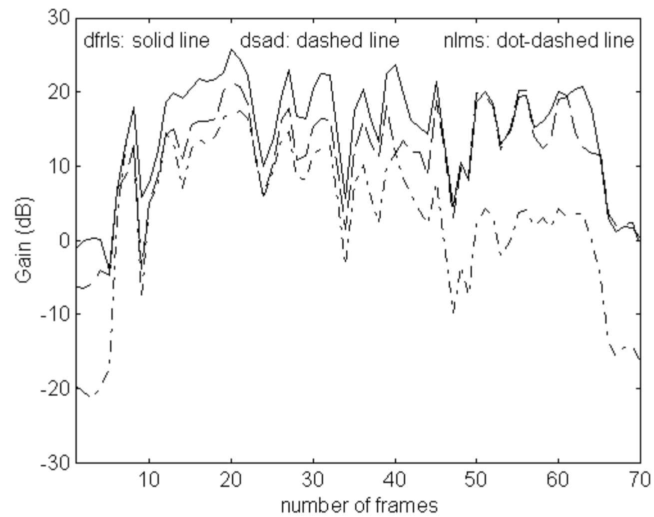

Fig. 6: The SNR of different schemes and algorithms, for successive frames (1 frame = 256 samples).

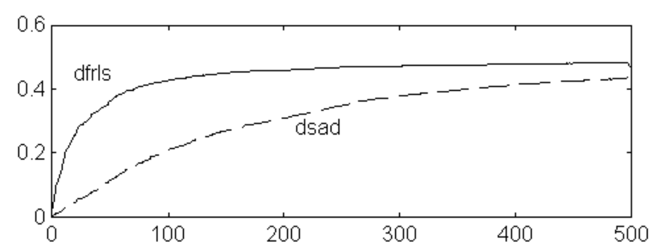

Fig. 7: Comparative convergence behaviour of the DFRLS and DSAD algorithm

4 CONCLUSION

In conclusion, in this paper we have presented a noise canceller system based on the Double Fast Recursive Least Square algorithm. Different aspects, such as the convergence, global and segmental SNR and subjective quality, have been considerd. We have shown the superiority of the presented algorithm compared to the Double SAD and the Normalized LMS algorithm. Furthermore, the structure based on the coupling FIR filters permits the DFRLS algorithm to be also used as a signal separators or a signal deconvolvers rather than only a simple noise canceller.

Informal quality and intelligibility tests indicate also significant superiority of such algorithm to enhance speech signal.

We should remark that the discussions about a complete mathematical analysis is on way [7]. In this short paper we have preferred to focalize the presentation on the case where the physical solutions of the equations (4) are possible. The case where the desired solutions are $W_i(z) = H_i(z)$ (i=1,2) and where the outputs verify $s_1(n)=s(n)$ and $s_2(n)=b(n)$.

ACKNOWLEDGEMENTS

We would like to thank the ″Ministère des Affaires Etrangères (France) ″ for supporting one of the authors (M. Gabrea). We would also like to acknowledge MATRA Communication Company (Paris) who has gently provided the recorded speech signals.

REFERENCES

[1] B. Widrow and al, "Adaptive Noise Cancelling: Principles and Applications", Proc. IEEE, vol. 63, no. 3, pp. 1692-1716, Dec. 1975.

[2] S. V. Gerven and D. V. Compernolle, "Signal Separation by Symmetric Adaptive Decorrelation: Stability, Convergence, and Uniqueness", IEEE Trans. on Signal Processing, vol. 43, no. 7, pp. 1602-1612, July 1995.

[3] D. Yellin and E. Weinstein, "Criteria for Multi-Channel Signal Separation", IEEE Trans. on Signal Processing, vol. 42, no. 8, pp. 2158-2168, August 1994.

[4] A. Benallal and A. Gilloire, "Instabilité et stabilité numérique des algorithmes des moindres carrés transversaux rapides excités par la parole", Proc. GRETSI 89, pp. 509-512, Juin 1989 (in French).

[5] R. Atay, P. Baylou and M.Najim, "Implementation of the Stabilized Fast Transversal Filter algorithm on Fixed Point DSP", Proc. of ICASSP 92, vol. IV, pp. 249-252.

[6] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*, Prentice-Hall, New Jersey 1992.

[7] M. Gabrea, *Ph. Dissertation*, to be defended (1997) Bordeaux-France