Global Journal of Engineering and Technology Advances

eISSN: 2582-5003

Cross Ref DOI: 10.30574/gjeta

Journal homepage: https://gjeta.com/

(Research Article)

Check for updates

Comparison of affine-projection adaptive filter vs LMS speech enhancement systems

Rubén Hernández [1,*], Julio C Sosa [2] and Víctor H García [2]

[1] *Unidad Profesional Interdisciplinaria en Ingeniería y Tecnologías Avanzadas Del Instituto Politécnico Nacional Av. Instituto Politécnico Nacional No. 2580, Col. Barrio la Laguna Ticomán. CP 07340 Gustavo A. Madero, D. F., México.*

[2] *Escuela Superior de Cómputo Del Instituto Politécnico Nacional Av. Juan de Dios Batiz s/n, Col. Lindavista, CP 07738. Gustavo. A. Madero, D.F. Méxic.*

Global Journal of Engineering and Technology Advances, 2021, 09(01), 065–069

Publication history: Received on 22 August 2021; revised on 01 October 2021; accepted on 03 October 2021

Article DOI: https://doi.org/10.30574/gjeta.2021.9.1.0134

Abstract

Evaluation of a two systems for speech enhancement is presented. A clasical Widrow dual channel noise canceller adaptive filter is used introducing the implemenation of Affine Projection Algorithm and comparing results versus traditional LMS implementation. Clasical Widrow noise canceller concept is used under reallistic input considerations, where signal of interest and noise are present at both inputs. Simple Root Mean Square Error (RMSE), Itakura-Saito Distance (ISD) and subjective evaluation tests are made in order to ease comparisons.

Keywords: Speech enhancement; Noise canceller; LMS adaptive filter; Affine projection

Introduction

Quality in a signal coming from a microphone can be poor if enviroment is noisy, this is common for most of personal devices such as hearing aids. Out of a noise cotrolled place, there is ambient noise, where close to people babbling is the main noise, and this is the case of more interest when a hearing aid is used. Esporadic ambient noise such that from the street can represent a similar problem to normal people as hearing aid users, nevertheless babbling noise in crowded places affect more to hearing aid users.

The hearing loss is related to the intelligibility of the speech embedded in noise and the SNR is useful in this case for cuantifying of both phenomena, so it is necessary a high value of SNR to get an acceptable comprehension of the speech, a person with hearing loss has a double trouble due to a loss in the intelligibility. Degradation in hearing loss of 10 dB is accompanied with degradation in intelligibility within 1 to 1.5 dB [2]. This means that a person with hearing loss has more difficulties for getting a good comprehension than a normal person when the conversation is done in a noisy environment.

This has made necessary to design speech enhancement systems for improving the intelligibility and quality of speech. Many noise reduction schemes have been proposed over the last decades [5, 6, 7, 2]. They can be divided into several categories and using many algorithms see [2] for a very good review.

We are interested on hearing aids because there is a need for small devices with few resources and a very good noise cancelling with great intelligibility.

* Corresponding author: Rubén Hernández

Unidad Profesional Interdisciplinaria en Ingeniería y Tecnologías Avanzadas Del Instituto Politécnico Nacional Av. Instituto Politécnico Nacional No. 2580, Col. Barrio la Laguna Ticomán. CP 07340 Gustavo A. Madero, D. F., México.

Some studies have shown that single-channel speech enhancement systems are still unable to improve intelligibility, even if they can now at least enhance signal quality without reducing intelligibility [9]. In contrast, multiple-microphone noise reduction schemes have been shown repeatedly to increase speech intelligibility even if there remain some theoretical and practical issues to be solved [5, 1]. The performances of multiple channel speech enhancement algorithms improve with an increasing number of microphones. However, a larger number of microphones imply higher costs and increasing demands in computational load, so they are not convenient for implementation on small devices.

There are many implementations for speech enhancement systems embedded in hearing aids. Multimicrophone arrays offer very good results considering intelligibility, nevertheless are not aesthetically convenient. The user of hearing aids needs small and discrete systems. This work presents a proposal including several benefits, such as the using of two microphones and a dual channel enhancement system based on affine projection adaptive filter so it can be programmed on a common hearing aid dsp. This work is one of a series we are developing in pursuit of the best performance system [3].

The two channel adaptive noise canceler

The digital adaptive filtering process considers the availability of a reference signal as a sample of the signal noise which is going to be diminished or cancelled. It is possible to modify the digital filter coefficients *w (k)* dynamically according to the behaviour of the involved signals. The aim is to obtain an approximated solution *y (k)* in order to minimize an output signal called error *s2 (k)* by using an iterative process [9, 4] (see Figure 1).

1.1. The evaluation model

Figure 1 shows a typical two channel noise canceler [9] and as the microphones are different for each channel we say is non balanced input, so that one channel receives the desired signal s(k) with additive noise while the other receives a noise signal n1(k) including a low power desired signal, that is the composition of signals at both inputs includes the desired signal and the noise signal but measuring different levels of SNR. And this SNR has to be always higher for the first channel. This is achieved by using different microphones, one highly directive for the first channel and one omni directional for the second channel. This way the omni directional microphone receives the voices nearby with the same gain than the desired signal and the directional microphone receives the desired signal with a higher gain than signals from other directions. So if the sources are equally distant the SNR is ≤ 0 for the omni directional and ≥ 0 for the highly directive.

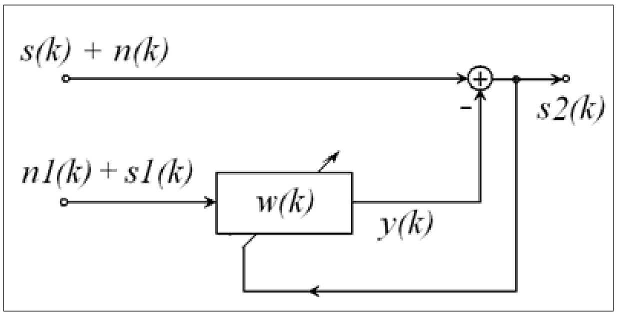

Figure 1 Adaptive filter as a two channel noise canceler

In a real environment there is the source of speech mixed with noise. The goal of a noise canceler is to estimate the signal of interest that is the primary source without noise. In an evaluation model the primary source has to be mixed with a noise signal for simulate a noisy environment, so it is possible to have control of the different levels of SNR and to make a good comparison between the estimated signal result of the noise canceler and the primary source without noise.

The recorded signals are a couple of phrases recorded with a directional microphone without noise and babbling recorded with omni directional microphone in a hospital waiting room.

In a future, we will test several systems in real environments using the above described configuration; in the assumption that the omni directional microphone has lower gain than the directional one and the SNR resulting is also lower for this channel so it can be used as a noise reference.

1.2. The LMS Algorithm

The Least Mean Squares Algorithm is very well known and simple the adaptive equation allows to calculate the optimum weigth coefficients as:

$$W(k + 1) = w(k) + \mu\, e(k)\, x(k)$$

Where w (k + 1) is the next weigth vector, w (k) is the current weigth vector, μ is the step size or local minimun error, e (k) is the error signal and x (k) is the input signal.

1.3. The Affine-Projection Algorithm

Reusing past data is a procedure that can improve the convergence of adaptive algorithms. The affine-projection algorithm is an LMS-based reusing algorithm, in that it includes L older input vectors in the updating rule and results in a generalization of the normalized LMS algorithm. Its behaviour depends mainly of number L. So as higher is the number of input vectors the algorithm is faster. The ecuation for weight vector calculation is:

$$W(k + 1) = w(k) + \mu X(k) (X^T(k) X(n) + \delta I_M)^{-1} e(k)$$

where where w (k + 1) is the next weigth vector, w (k) is the current weigth vector, μ is the step size or local minimun error, e (k) is the error signal, X (k) is the matrix of L input vectors, δ is a regularization parameter an d I_M is an identity matrix.

Evaluation

In order to know the performance of the proposed model to enhance noisy speech signals and compare against the performance of a single stage noise canceler we have done several simulations.

An RMSE (expression 1) measure is done to the difference between the enhanced signal *s2 (k)* and the original *s (k)*, this is a tough approximation to the noise level present at the output of the system and is determined by:

$$RMSE = \sqrt{\frac{1}{N}\sum_{k=0}^{M} d(k)^2} \quad (1)$$

Where d (k) = s2 (k) – s (k)

This measure allows a fast and easy comparison for the several simulations and it is well related to listening tests. The Itakura-Saito Distance was measured only for some simulation results.

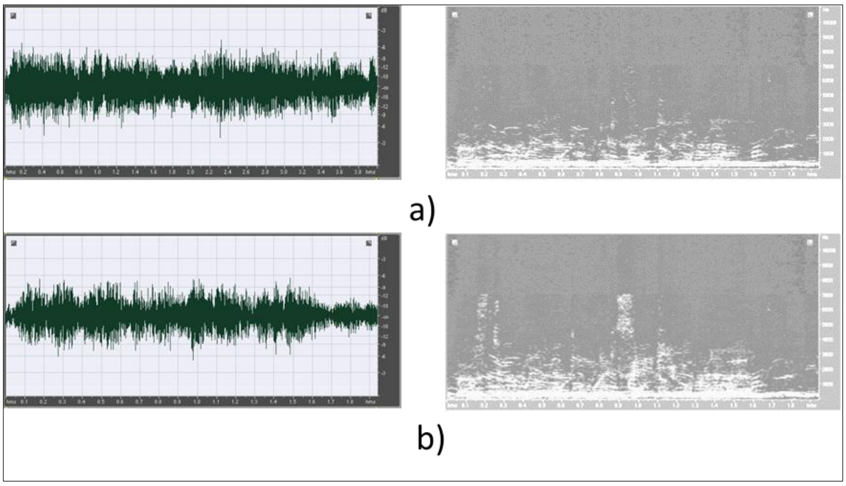

Figure 2 Signals in time and their respective spectrograms: a) Babbling mixed signals 0 dB SNR (n (k) +s1 (k)), b) Mixed signals 9 dB SNR (s (k) + n (k))

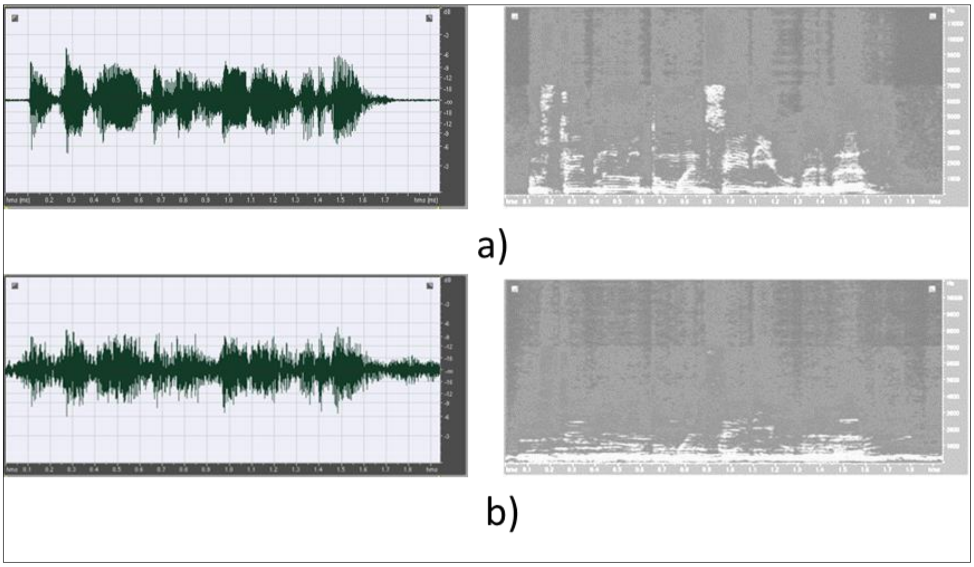

Figure 3 Signals in time and their respective spectrograms: a) speech signal, b) enhanced speech signal

Figure 2 and 3 shows snapshots of signals in time and their spectrograms, this were taken for one simulation corresponding to 9 dB of SNR for channel 1 and 0 dB for channel 2.

Table 1 shows the measures of the Itakura-Saito Distance for some simulations under the conditions established above, similarity of results for the higher SNR can be noticed, this may be because our signal noise is babbling and has almost the same statistics as the desired signal. The ISD works better when noise is gaussian.

Table 1 Measures of the Itakura-Saito Distance

SNR main input channel *s(k)+ n(k)*	**ISD main input channel *s(k)+ n(k)***	**ISD output (LMS) *s2(k)***	**ISD output (A-P) *s2(k)***
3 dB	2.52	2.22	2.14
6 dB	2.22	2.05	2.01
9 dB	1.53	1.56	1.54
12 dB	1.41	1.44	1.42
15 dB	0.93	0.77	1.01

Some results from measuring the RMSE level of noise are shown in Table 2. They were carried out for the same simulations of Table 1. The difference between the LMS algorithm and Affine-Projection Algorithm results under low SNR are very interesting because they are much related to the listening tests. This is very important so the comparisons among different simulations can be made easier and faster trusting just on the RMSE measure and for low SNR.

Table 2 Measures of the RMSE

SNR main input channel *s(k)+ n(k)*	**RMSE main input channel *s(k)+ n(k)***	**RMSE output (LMS) *s2(k)***	**RMSE output (A-P) *s2(k)***
3 dB	6318	5774	5620
6 dB	4199	3730	3445
9 dB	2531	2332	1976
12 dB	1437	1113	1084
15 dB	1234	802	776

Conclusion

In this work the evaluation of a two channel adaptive filter structure has been carried out. It is convenient to focus on the results for the evaluation procedures, so that using two channels the enhancement of the speech is good in quality as intelligibility and the simplest procedure RMSE is related in proportion to the listening test, despite RMSE only evaluates the quality, it will serve as a good tool of comparison, considering its simplicity, among the several structures we are going to evaluate in our pursuit for a small adaptive filter structure. In this case the use of affine Projection algorithm results in no reduction of generalized RMSE compared with the simple LMS and clearly it is a very good option for implementing this speech enhancement system. Of course when we obtain the best results for certain structures will be necessary to perform the normalized intelligibility tests. The unbalanced input has shown to contribute to the good results even for the single stage as when simulations where made with similar level of SNR on both channels (considering two identic microphones) the performance were not the same as reported here for none of the three tests used ISD, RMSE and listening.

Compliance with ethical standards

Acknowledgments

We acknowledge Instituto Politecnico Nacional for supporting this work.

Disclosure of conflict of interest

There is no conflict of interest to declare in this study.

References

[1] Benesty J, et al. Noise Reduction in Speech Processing, Springer, Berlin. 2009.

[2] Davis G. Noise reduction in speech applications, CRC Press, FL. 2002.

[3] Hernandez R. Evaluation of Dual Channel Adaptive Noise Canceller for Speech Enhancement, International Journal of Tomography & Statistics. 2012; 19(1).

[4] Haykin S. Adaptive Filter Theory, Prentice Hall, NJ. 2001.

[5] Kompis M. "Improving speech intelligibility with multi-microphone noise reduction systems for hearing aids", Acoustical Research. 1998; 2.

[6] Moisa C, et al. "Voice Signal's Noise reduction Using Adaptive/reconfigurable Filtersfor the Command of an Industrial Robot", Proceedings of the 11th RASPCGST. 2011.

[7] Ngo K. et al. "A flexible speech distortion weighted multi-channel Wiener filter for noise reduction in hearing aids", IEEE ICASSP. 2011; 2528-2531.

[8] Valente M. Hearing aids: standards, options and limitations, Thieme Medical Publisher, New York. 1996.

[9] Widrow B, Stearns S. Adaptive signal processing, Prentice Hall, NJ. 1985.