Quest Journals
Journal of Electronics and Communication Engineering Research
Volume 7 ~ Issue 10 (2021) pp: 01-05
ISSN(Online) : 2321-5941
www.questjournals.org

Research Paper

Affine-Projection Adaptive Filter Speech Enhancement System

Rubén Hernández[1], Julio C. Sosa[2], 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. México.

ABSTRACT
Evaluation of a system for speech enhancement is presented. A two stage dual channel structure of adaptive filters allready proposed [1] is used introducing the implemenation of Affine Projection Algorithm at second stage. 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 against single stage adaptive filter and two stage LMS structure. Using a simple adaptive filter, it is not enough to reduce background noise while maintaining the intelligibility of enhanced speech. Therefore, a noise emphasizing channel is used in order to achieve a better noise canceling.

Received 20 September, 2021; Revised: 03 October, 2021; Accepted 05 October, 2021

I. 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. Many noise reduction schemes have been proposed over the last decades [Kompis 1998, Moisa 2011, Ngo 2011, Davis 2002]. They can be divided into several categories and using many algorithms see [Davis 2002] for a very good review. 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 [Davis 2002]. 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 [Kompis 1998, Moisa 2011, Ngo 2011, Davis 2002]. They can be divided into several categories and using many algorithms see [Davis 2002] 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.

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 [Valente 1996]. 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 [Kompis 1998, Benesty 2009]. 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.

The hearing loss is related to the intelligibility of the speech embedded in noise and the SNR is useful in this case for both phenomena, so it is necessary an SNR to get an acceptable comprehension of the speech, nevertheless 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 [Davis 2002]. 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.

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 [Hernandez 2012].

II. 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 [Widrow 1985, Haykin 2001] (see Figure 1).

2.1. The evaluation model

Figure 1 shows a typical two channel noise canceler [Widrow 1985] 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.

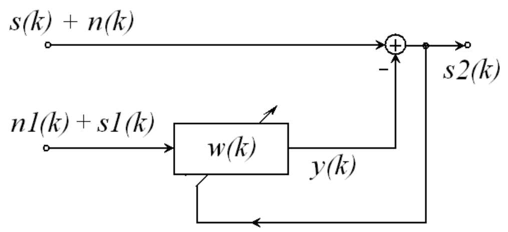


Figure 1. Adaptive filter as a two channel noise canceler.

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 with out 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.

The speech enhancement system evaluated in this work is based on an adaptive filters used as shown in the Figure 2, where the signal at channel 1 consists of the primary signal *s(k)* contaminated with the noise from other sources *n(k).* While the channel 2 provides the noise from all the sources and is considered as the reference noise *n1(k).*

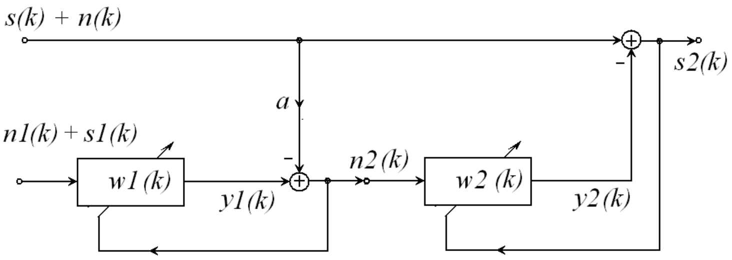


Figure 2. Two stage adaptive filter as a two channel noise canceler

The error signal *n2(k)* is evaluated by the first adaptive algorithm (LMS) for modifying the filter weights *w1(k)* so that *y(k)* be almost the replica of *as(k)* then *n2(k)* be a signal very likely to *n(k)*. This way the first stage emphasize the babbling noise so that the second adaptive algorithm (AP) modify the filter weights *w2(k)* for having a result *s2(k)* much more similar to *s(k)* than what can be obtained using a single stage noise canceler as the one shown in Figure 1.

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.

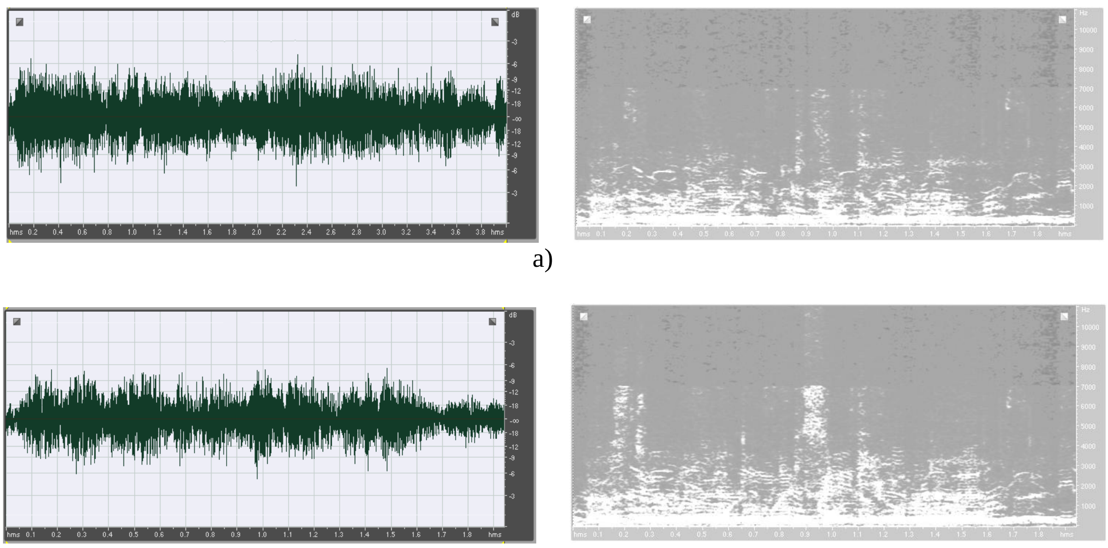
III. EVALUATION

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:

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

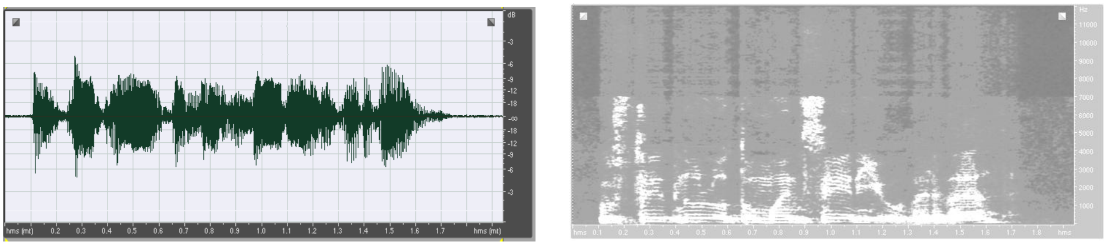
where *d(k) = s2(k) – s(k)*

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.

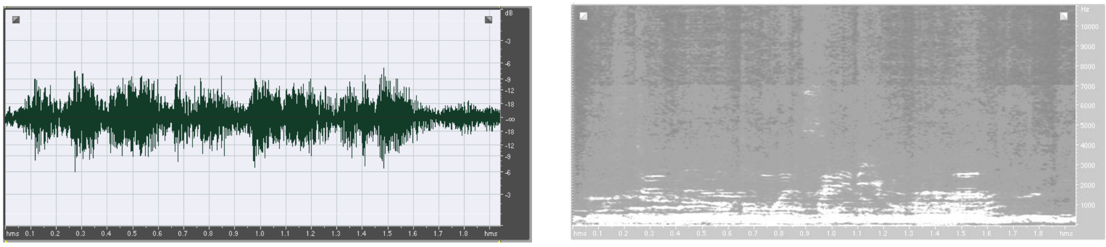


a) b)

Figure 3. 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))*



a)



b)

Figure 4. Signals in time and their respective spectrograms: a) speech signal, b) enhanced speech signal

Figure 3 and 4 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 very related to the listening tests, due to the low SNR the intelligibility is almost loss but after the enhancement it seems recovered, this for the two stage system. 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 1: 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	5852	5731
6 dB	4199	3610	3452
9 dB	2531	2911	1808
12 dB	1437	1123	1182
15 dB	1234	754	787

IV. CONCLUSION

In this work the evaluation of a two channel adaptive filter structure has been carried out. The use of two adaptive filtering stages allows controlling both the background noise as the babbling. 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 a 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.

REFERENCES

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

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

[3]. Hernandez, R., 2012, Evaluation of Dual Channel Adaptive Noise Canceller for Speech Enhancement, International Journal of Tomography & Statistics, Vol. 19, No. 1.

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

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

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

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

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

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