Speech Enhancement Using Adaptive Algorithms

Filip Gummesson
dat11fgu@student.lu.se

Introduction

Speech signals are often corrupted by various types of noise. Sources of noise can be that a microphone picks up sound from the background or distortions in a transmitted signal; e. g. phone calls, VoIP, etc.
To remove noise from a degraded signal and increase the quality of the speech using signal processing is known as speech enhancement.[1] The objective is to improve the signal-to-noise ratio (SNR) without losing Intelligibility, which is the measure of how comprehensible speech is.

The objective of this project is to use adaptive filtering to achieve an increase in SNR of a speech signal. The signal that is used through out the project is the transmitted signal from the Moon as Neil Armstrong uttered his well known line

> *"That's one small step for a man, one giant leap for mankind".*

The adaptive algorithms applied to the signal are Least Mean Squares (LMS), Normalized Least Mean Squares (NLMS) and Leaky LMS.

Method

Since the specifications of the noise carried with the signal is unknown, a time invariant filter might not satisfy the requirements of of the filtering. Instead an adaptive filtering method is used, which gives the ability of updating filter weights as an effect of previous samples.

---

[1] J. Benesty, S. Makino, J. Chen (ed). Speech Enhancement. pp.1-8. Springer, 2005.

The LMS and NLMS are well known adaptive algorithms and often used in noise canceling applications. In these algorithms, one or more filter coefficients are updated every adaptation cycle.[2] These two algorithms and the Leaky LMS will be applied and analyzed in this project.
The project uses an system prediction structure to predict the present value of an unknown signal. The block scheme of this structure is presented in figure 1, [3]

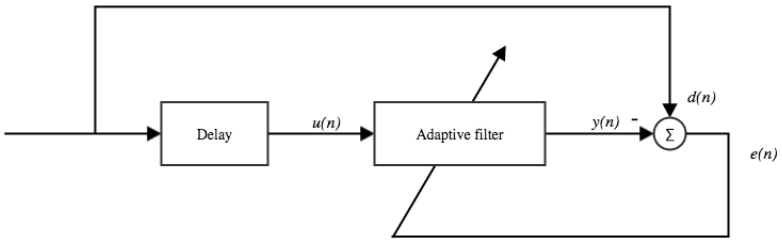

Figure 1: System prediction structure is used to predict present values of a signal. Past values will be the input of the adaptive algorithm.

where n is the iteration number. $d(n)$ denotes the desired signal, where $d(n)$ contains both the actual speech and noise, $v(n)$, as

$$d(n) = s(n) + v(n).$$

The signal to the adaptive filter $u(n)$, is a delayed version of $d(n)$ and can be written as

$$u(n) = d(n - \Delta)$$

when $u(n)$ is delayed Δ samples. Since this setup uses the same signal both as the desired and input signal to the adaptive filter, it is needed to remove the correlation between the noise $v(n)$ and the noise at time $n - \Delta$.
The adaptive filter with finite-duration impulse response (FIR) of length M are defined as $\overline{w} = [w_0, w_1, \dots, w_{M-1}]$ and
$y(n) = u(n) * w$.
$e(n) = d(n) - y(n)$ is the estimation error of the adaptive algorithm.

Lets get back to the algorithms. As mentioned earlier LMS, NLMS and Leaky LMS will be used as algorithms which updates the

---

[2] S. Haykin, "Adaptive Filter Theory" 5th edition, pp 267- 289, Pearson, Edinburgh Gate 2014.
[3] S. Haykin, "Adaptive Filter Theory" 5th edition, pp 35-37, Pearson, Edinburgh Gate 2014.

filter tap weights, and gives the filter an estimate whose estimated value may come close to the Wiener solution w_0.

The LMS algorithm is one of the most used, and also the simplest of the adaptive algorithms.[4] It is defined as

$$\hat{w}(n+1) \ = \ \hat{w}(n) \ + \ \mu \, \overline{u}(n) \, e(n)$$

where $\hat{w}(n)$ is an estimate of the unknown filter coefficients vector, $\overline{w}(n)$. The algorithm is deterministic if the step-size parameter $0 < \mu < \frac{2}{\ _{max}}$, where $_{max}$ is the greatest eigenvalue of the correlation matrix $R \ = \ E[\,\overline{u}(n) \ \overline{u}^{\,T}(n)\,]$. This algorithm computational complexity scales linearly with the number of filter taps in the adaptive filter.

The main disadvantage of the LMS algorithm is that it might suffer from a gradient noise amplification.[5] This will be noticeable when $\overline{u}(n)$ is large. Instead, the use of the NLMS algorithm will come in handy. The only difference between NLMS and standard LMS is that the product vector $\overline{u}(n)e(n)$ are weighted with the squared Euclidian norm of $\overline{u}(n)$. In mathematical terms, the algorithm look like this:

$$\hat{w}(n+1) \ = \ \hat{w}(n) \ + \ \frac{\mu}{a + \|\overline{u}(n)\|^2}\overline{u}(n)e(n)$$

where $a > 0$ is a safety feature when the norm gets to small and introduces numerical faults. This algorithm shall according to theory, have a faster adoption rate than standard LMS.

The Leaky LMS is another modified version of standard LMS.[6] As the name states, this algorithm leaks energy from filter taps. This makes the algorithm forget about past results, which would be in favor since speech may vary a lot from one sample to another. Leaky LMS updates the filter taps as

$$\hat{w}(n+1) \ = \ (1-\alpha)\,\hat{w}(n) \ + \ \mu \, \overline{u}(n)e(n)$$

[4] W. Harrison, J. Lim, E. Singer, "A new application of adaptive noise cancellation", IEEE Trans. Acoustic Speech Signal Processing, Vol. 34, pp 21-27, 1986.

[5] S. Haykin, "Adaptive Filter Theory" 5th edition, pp 333-337, Pearson, Edinburgh Gate 2014.

[6] S. Haykin, "Adaptive Filter Theory" 5th edition, pp 505, Pearson, Edinburgh Gate 2014.

where $(1-\alpha\mu)$ is called the leakage factor. $0 \le \alpha < 1$.

In figure 2, the plot of the desired signal $d(n)$ is shown. It contains both speech and a lot of noise. As stated, the objective is to reduce this noise and make the speech more comprehensive by run the signal through an adaptive filter .

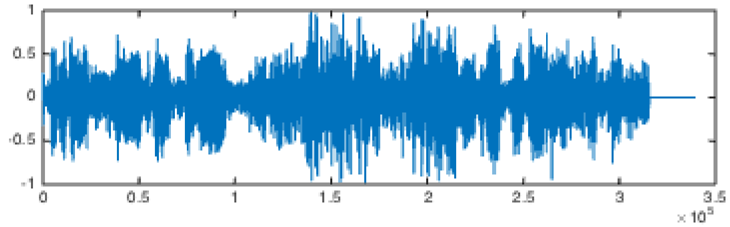

Figure 2: plot of the desired signal $d(n)$. running this signal through an adaptive filter should reduce the noise.

All files that been used during this project can be found at github.com/FGummesson/enhance-speech

Result

After passing the desired signal through the adaptive filter, the $y(n)$ signal is received. This signal is plotted in figures 3 - 5 for every adaptive algorithm. The filter length M was set to 8 and $\mu = 0.002$ for every algorithm. In figure 4, a = 1 to assure no numerical faults.

Figure 5 shows $y(n)$ calculated by the leaky algorithm when the leakage factor is set to 0.998.

Figure 6 shows the learning curves (mean square error) of the three different algorithms.

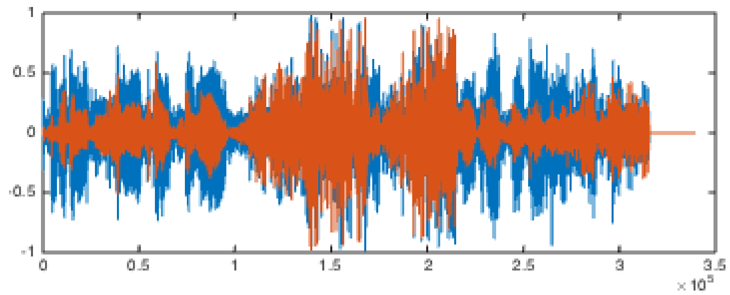

Figure 3: $y(n)$ plotted on top of $d(n)$, when the adaptive algorithm used was standard LMS.

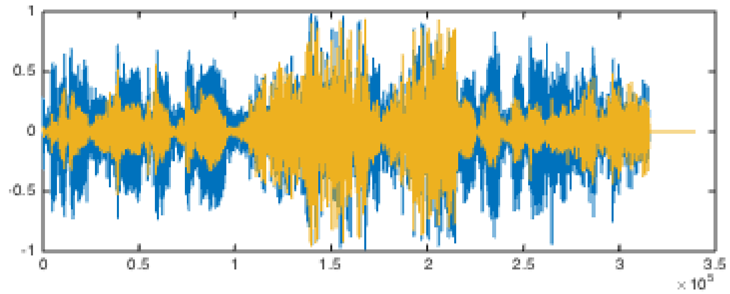

Figure 4: $y(n)$ plotted on top of $d(n)$, when the adaptive algorithm used NLMS.

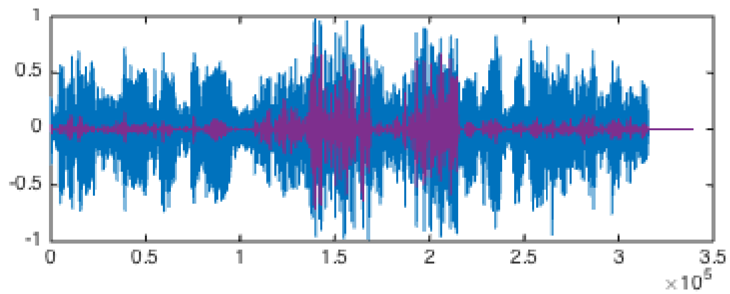

Figure 5: $y(n)$ plotted on top of $d(n)$, when the adaptive algorithm used was the leaky LMS.

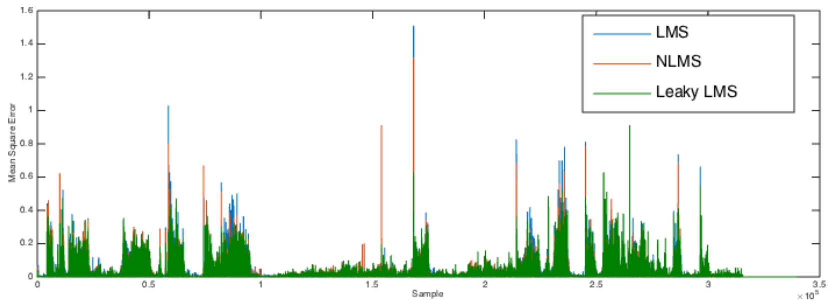

Figure 6: Mean Square Error of all algorithms.

Discussion

The chosen algorithms performs as expected. Standard LMS computationally fast and converges if the step-size parameter is chosen wisely. The performance in imitate the speech signal is not as good as the other. This can be seen by comparing figure 3 with 4 and 5, and it is even noticeable for the human ear.

In the middle of the audio file, there was a pause in speech and just noise was heard. This part is still noisy due to the filter adapts to the noise.

In figure 6 the learning curves of the different algorithms are shown. They show the adaption rate, and that there is almost no difference in adaptation cycles for the algorithms to adapt to the Wiener solution $\overline{w_0}$. It might be that this changes when the design variables are changed.

Reference

1. J. Benesty, S. Makino, J. Chen (ed). Speech Enhancement. pp.1-8. Springer, 2005.

2. S. Haykin, “Adaptive Filter Theory” 5th edition, pp 267-289, Pearson, Edinburgh Gate 2014.

3. S. Haykin, “Adaptive Filter Theory” 5th edition, pp 35-37, Pearson, Edinburgh Gate 2014.

4. W. Harrison, J. Lim, E. Singer, “A new application of adaptive noise cancellation”, IEEE Trans. Acoustic Speech Signal Processing, Vol. 34, pp 21-27, 1986.

5. S. Haykin, “Adaptive Filter Theory” 5th edition, pp 333-337, Pearson, Edinburgh Gate 2014.

6. S. Haykin, “Adaptive Filter Theory” 5th edition, pp 505, Pearson, Edinburgh Gate 2014.