# CS578 Speech Processing
# Laboratory 4
# Speech Enhancement


George Manos
csd4333@csd.uoc.gr

Alexandros Angelakis
csd4334@csd.uoc.gr


20 January 2023

## 1 Implementation Details

First of all, we already know that the signal we want to enhance had 1000 zeros before adding the white noise. Thus, we can extract those 1000 samples of white noise and compute its power spectra. We achieve that by doing a frame-by-frame analysis on these samples, we take the square of the magnitude vector of each frame's Fourier Transform $|S_b|^2$ and finally, in order to approach the power spectra of the noise, we calculate the average of these vectors. After calculating the noise power spectra, we use the two enhancement algorithms, Spectral Subtraction and Wiener filtering to clean the signal from the noise as much as possible.

We compare both the initial signal without any noise and the one with noise with our enhanced signal, by plotting them and listening to them, printing the computed Mean Squared Error as well*. On our experiments, we will be comparing the enhanced signal resulting from either method to both the input, noisy one, and the original signal, which we will consider as the ground truth.

Comparing the signal with the white noise and the ground truth, we can see their difference and a high MSE score presented in the following figure. All enhancement methods that we will discuss should have a smaller MSE score to the original signal, if they are to work correctly.

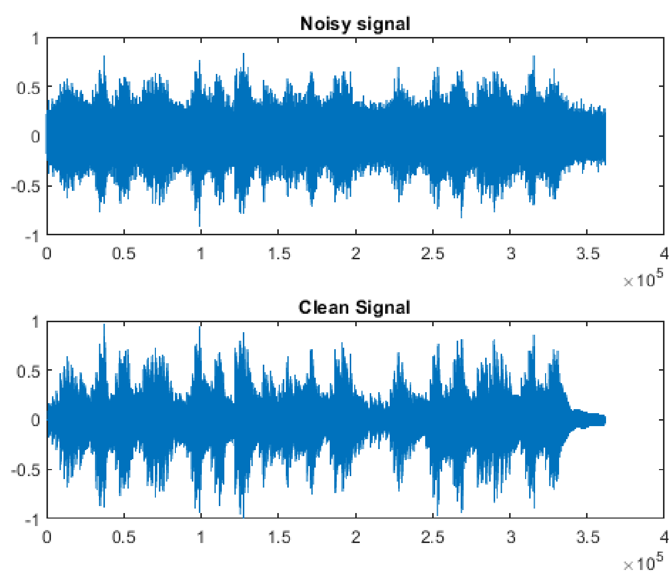


Figure 1: Enhanced signal using Spectral Subtraction and Noisy, with MSE: 0.009705

*Maybe it would have been better if the output signal had been normalized w.r.t. the input signal energy. However, for our case this was unnecessary as we mostly want to compare the 2 methods with themselves

## 2 Spectral Subtraction

### 2.1 Details

For the Spectral Subtraction, we do a frame-by-frame analysis-synthesis, we calculate the magnitude and the phase of the frame's Fourier Transform and we simply subtract the squared magnitude of the frame with the power spectra of the noise that we've calculated. Since we don't want any negative values for the Spectral Subtraction, because we will have to take the square root of it in order to find the inverse Fourier Transform, we set every negative value found to 0. Then we calculate the inverse Fourier transform of the frame, by taking the square root of the Spectral Subtraction we calculated as the magnitude and the phase of the frame and writing it in polar form. Finally, we perform the Overlap And Add method, in order to construct the enhanced signal.

### 2.2 Results

The Spectral Subtraction method is not performing well, it reduces a perceivable amount of noise from the signal but it is still very loud. We can also see the differences between the enhanced signal and the clean signal, noisy signals.

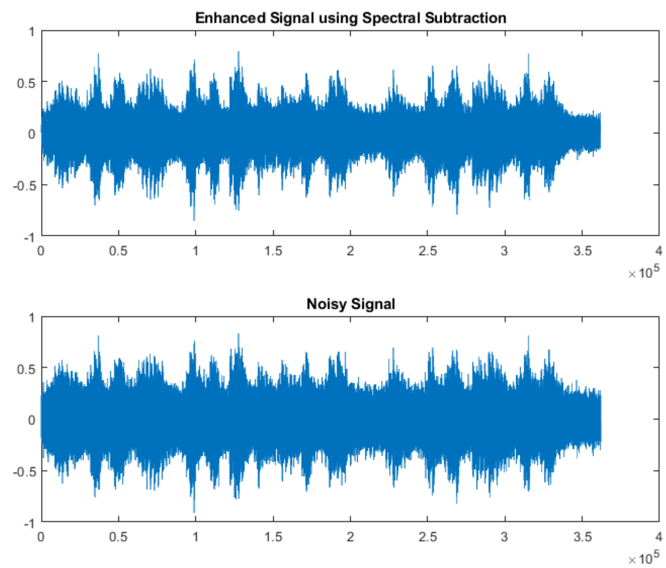


Figure 2: Enhanced signal using Spectral Subtraction and Noisy, with MSE: 0.001973

We can notice that the noise has been reduced, by looking at the middle and at the end of the signal. The Mean Squared Error between them could be represented as their difference - as a metric it does not represent only the noise removed since useful information is also subtracted from the original signal.

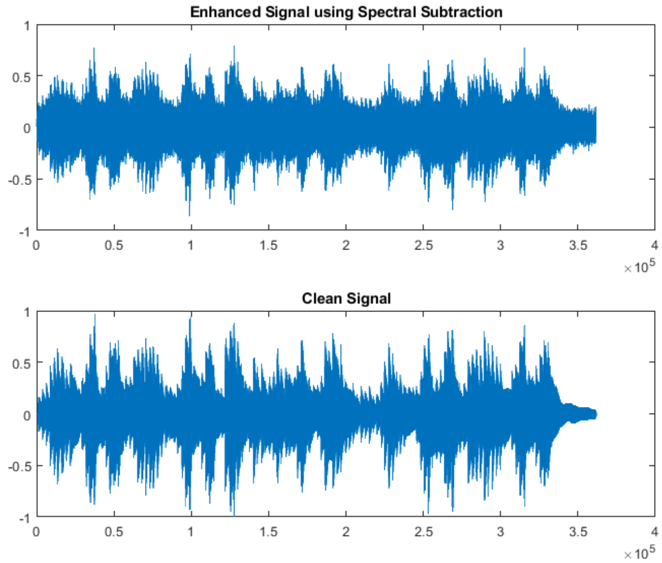


Figure 3: Enhanced signal using Spectral Subtraction and Clean, with MSE: 0.006163

By looking at plots, we can see that the enhanced signal is still quite far from the actual clean signal. However, the MSE score has dropped significantly from the original noisy signal.

# 3 Wiener filtering

## 3.1 Details

For the Wiener filtering, we also do a frame-by-frame analysis-synthesis. We first initialize the Wiener filter with spectral subtraction and we approximate the Fourier Transform of the clean signal $X$, by multiplying the Fourier Transform of the noisy signal $Y$ with the previous frame's Wiener filter $H(p-1)$, $X(p) = Y(p) \cdot H(p-1)$. Then we calculate the new Wiener filter by using the new smoothed power spectra of the frame. Finally, we take the magnitude of $X$ and the phase of the frame, we take the inverse Fourier Transform using this information in polar form and we perform the Overlap And Add method to construct the enhanced signal.

## 3.2 Results

The Wiener filtering is performing pretty well, a lot better than the Spectral Subtraction. It reduces a good amount of noise, making the enhanced signal bearable to the ear. We can visually see the differences between the enhanced signal and the clean, noisy signals.

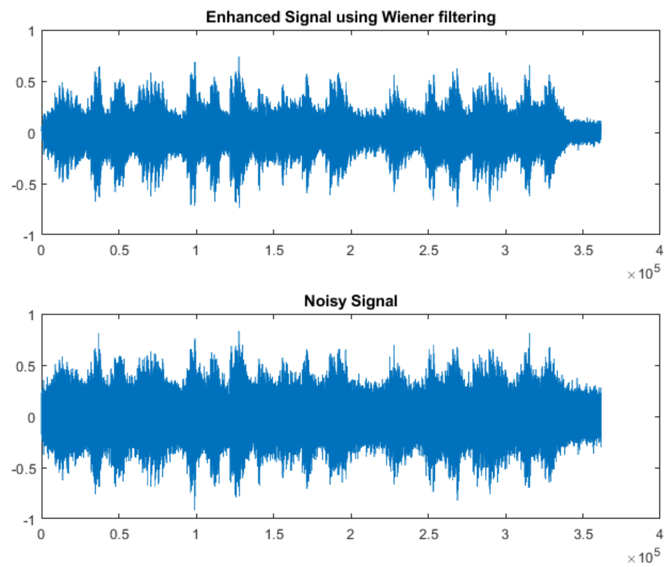


Figure 4: Enhanced signal using Wiener filtering and Noisy, with MSE: 0.003640

We can notice that the noise has been significantly reduced. The Mean Squared Error between them is higher than the spectral subtraction one, meaning that the resulting signal is much different.

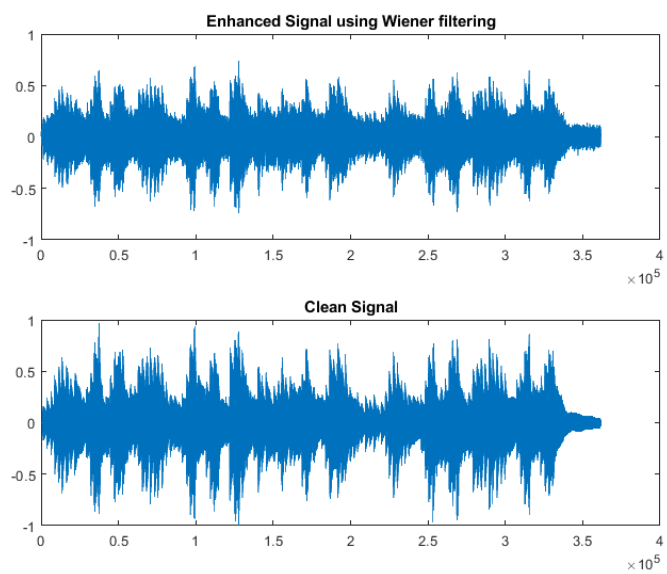


Figure 5: Enhanced signal using Wiener filtering and Clean, with MSE: 0.004769

Looking at plots, we can see that the enhanced signal is quite similar with the clean signal. The method's Mean Squared Error is even lower. Here, we can finally confirm that the Wiener filter indeed performs better than the Spectral Subtraction method.

We also experimented with the smoothing parameter $\tau$, initially we set it to 0.5. By increasing $\tau$ we noticed that the the error between the enhanced and the clean signal is a little bit higher than reducing $\tau$. This happens because the higher value $\tau$ has, the power spectra of the frame has more impact on the Wiener filter than the smoothed power spectra. For example by using $\tau$ = 0.2, the MSE is 0.004738 and by using $\tau$ = 0.8 the MSE is 0.004842. Visually and aurally, there is not a big difference between them.

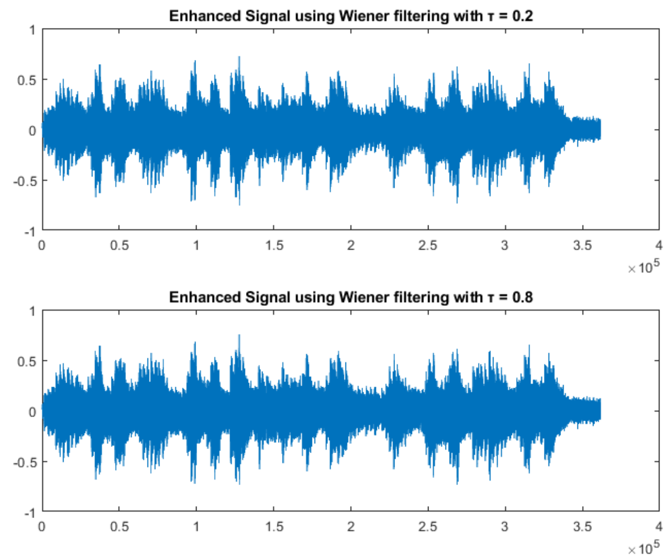


Figure 6: Enhanced signals with different smoothing parameters

# 4 Our voices

We added to our voices 1000 zeros in the beginning and a white noise, in order to enhance and test the two enhancement methods. We can see and hear that the Wiener filtering is again the best option here, it reduces a good amount of noise from the signal and the error between the original and the enhanced is also low.

For George's voice, the MSE between the enhanced signal using Spectral Subtraction and the clean signal is 0.010990 and using Wiener filtering is 0.009033.

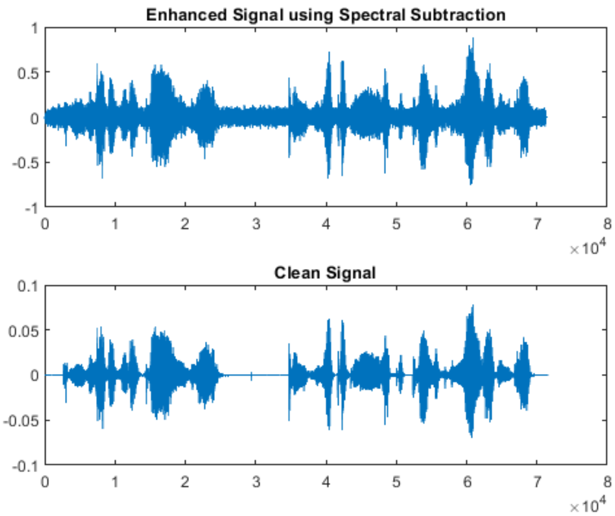


Figure 7: Speech Enhancement using Spectral Subtraction on George's voice

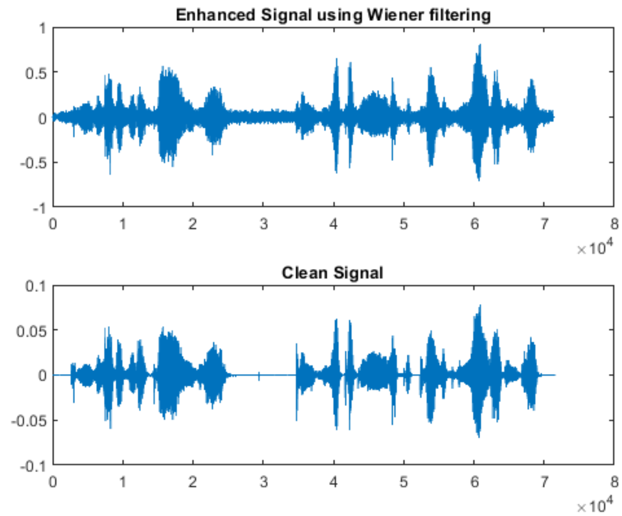


Figure 8: Speech Enhancement using Wiener filtering on George's voice

For Alex's voice, the MSE between the enhanced signal using Spectral Subtraction and the clean signal is 0.004616 and using Wiener filtering is 0.003553.

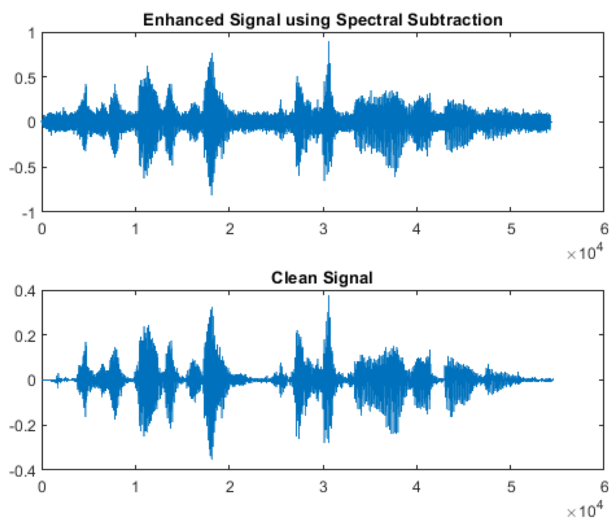


Figure 9: Speech Enhancement using Spectral Subtraction on Alex's voice

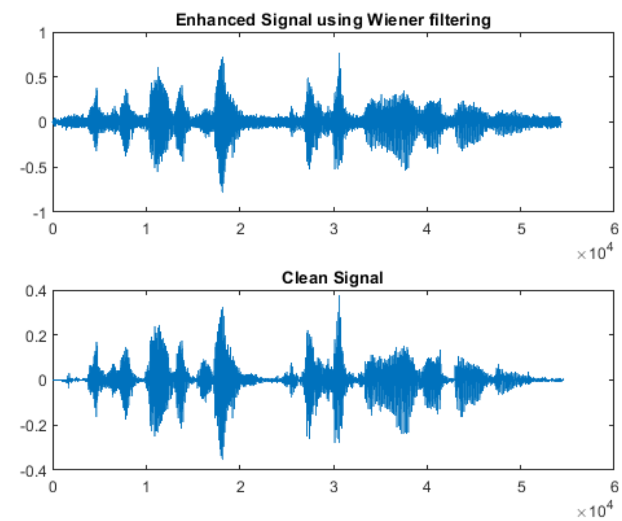


Figure 10: Speech Enhancement using Wiener filtering on Alex's voice

# 5 Appendix

I will presenting the best application the literature has come across of the Wiener filter.

Cranberry-Lemon — Journal of Astrological Big Data Ecology

## Optical Turbulence Characterization and Wiener Filtering of Hot New Intern Induced Temperature Gradients

**Lisa Smithers**[1]

[1] Department of Optical Engineering Cranberry-Lemon University, Pittsburgh, PA, USA

**Abstract**

Ever since Todd began his internship in our LIDAR lab, our system accuracy has taken a nosedive. As LIDAR advances in accuracy are becoming increasingly important in applications such as Autonomous Vehicles, Precision Agriculture, Aerial inspection/surveying/mapping, robotics and much more, our sudden degradation halted research. At first, I thought I was distracted trying to figure out if Todd was flirting with me, or I broke some of our equipment, or I had a bug in our translation software, but after three new LIDARs we determined that Todd is just too FREAKIN HOT! His temperature gradients create too much hot steamy turbulence that it's breaking all our optical systems! Especially when he flicks his hair! This paper introduces a novel method to characterize his unique Hawtness induced optical turbulence and a Wiener filter mitigation method. It helped for a while, then he brought a guitar into the lab.

Keywords: Optical Turbulence, Todd, LIDAR, Hair, Optics, Atmospheric Characterization, Todd, Noise Modeling

**1. Introduction**

In any optical system, turbulence is often a major source of error, range, and intensity. As Todd is the hottest thing at Cranberry-Lemon University [1], some say maybe even the greater Pittsburgh area [2]. He is a walking temperature gradient and the second he steps in our lab, the $C_N^2$ shoots through the roof! Like good Lord! Our LIDARs were designed for most engineers who I've never seen create turbulence let alone turn any heads. I mean not even on a good week when their skin is clear. Shoot, I thought I was hot even John, but not nearly on his level. Woof…

Because I really need some presentable results for my thesis, we really need to mitigate his turbulence. I would ask him to leave the lab, but it just seemed too rude, I mean, to my advisor who got me the funding for an intern, not Todd, I'm sure he's got plenty of labs he could work in. Additionally, ever since he started working with us, I've never collected so much data in my life! As shown in [3], there is a method to mitigate such turbulence using a Wiener filter. It just came up while I was googling solutions.

**2. Background**

Optical turbulence occurs when pockets or surfaces of airs have differentials in temperature. These can be especially caused by hot surfaces such as sunbaked dirt in the desert or asphalt in the summer. Hot surfaces like Todd's pose optical design challenges due to the intensity and unpredictability of turbulence never yet observed by optics engineers. Todd is a senior Electrical Engineering undergrad at CLU that began working in our LIDAR lab a few weeks ago and he is one hot surface. I mean, just look at figure 1…

Figure 1: Todd

A previous CLU study showed that some turbulence may be corrected using a healing crystal based adaptive optics technique [4]. Unfortunately, he's a Leo…There are three

1 — Jabde.com

Figure 11: Source: Journal of Astrological Big Data Ecology