ECE–3151 Linear Systems Lab
Echo Cancellation

Chad Chapnick

February 16, 2017

I. Introduction

In this lab, we explore the concept of removing an echo from a signal. Mathematically, an echo can be thought of as an undesired convolution of a given audio signal. He were model the system using:

$$y(t) = x(t) + \alpha x(t - \tau)$$

Applying the superposition priniciple, the equation above suggests that an echo is created from the sum of the original audio signal and a version of that signal delayed by τ, and attenuated by the factor α (ie. $\alpha < 1$). Upon closer inspection, the echo system is a *feedforward comb filter* [1]. In `background.pdf`, we were provided with the system equation for the echo removal system, which takes the form of a *feedback comb filter* [1]:

$$z(t) = y(t) - \alpha z(t - \tau)$$

The impulse response of this system is an infinite series of unit impulse functions with a weight of $(-\alpha)^k$, where k corresponds to the term number (see examlpe calculations in Appendix) [2]. With this in mind, we developed an echo removal system that implements deconvolution to remove the echo from an audio signal. The impulse response of our echo removal system is:

$$h_i(t) = \delta(t) + \sum_{k=1}^{7} (-\alpha)^k \delta(t - k\tau)$$

The main question which arises is how to estimate the parameters τ and α. We achieved this using the autocorrelation function for an energy signal $x(t)$ which is given by:

$$R_{xx}(\tau) = \int_{-\infty}^{\infty} x(\lambda) x(\tau + \lambda)$$

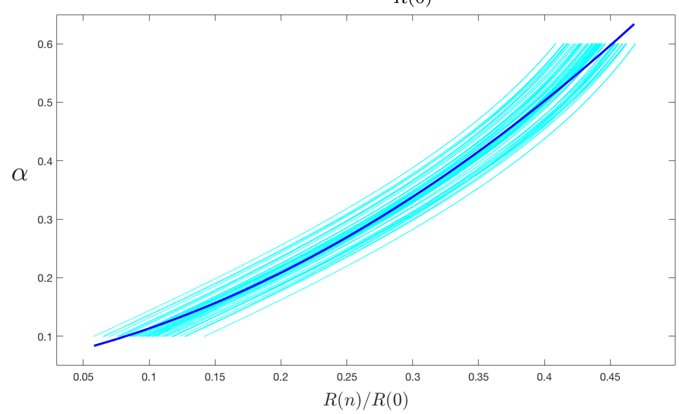
Taking advatange of resulting signal's even symmetry, we were able to extract information about τ and α by looking at the dominant peaks for positive values of τ. The first peak at $R_{xx}(\tau = 0) = R(0)$ gave the total energy of the audio signal. The second peak, which we denote as $R(n)$, gave insight into the parameters of interest. In particular, the value of τ could be estimated by finding the maximum value, $R(n)$, within the known range of τ, then using the index of this peak to find the corresponding value of τ. Estimating the value of α was a more intimate process. Through simulation, we found that α has a nonlinear relationship with $R(n)/R(0)$ which can be modeled by quadratic growth. Using MATLAB, we found the coefficients for this model which allowed us to reasonably estimate the echo gain for a given audio signal. Once there parameters were obtained, they were substituted into $h_i(t)$ and deconvolution was performed to practically eliminate the echo from the audio signal.

II. Results

A. Mapping $R(n)/R(0)$ to α

For our approach, we ran a test that loops through values of α ranging from 0.1 to 0.6, in steps of 0.001. Within this loop, we also iterated through values of τ ranging from 0.05 to .5, in steps of 0.01. For each pair of τ and α values, we obtained a clean audio signal from `probe5.m`, then added an echo to the signal and found the ratio between the second highest peak and the total energy in the signal. This simulation gave over 23,000 values for $R(n)/R(0)$. This is the reason that the cyan dots seem nearly continuous in figure 1. With this fact, we are reasonably confident in our estimation of the coefficients for the quadratic equation which allows us to map $\frac{R(n)}{R(0)}$ to α.

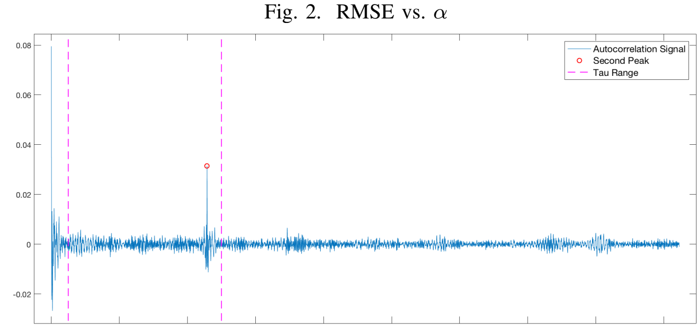
Fig. 1. α vs. $\frac{R(n)}{R(0)}$



B. Time Shift Estimations

In general our algorithm was highly effecitvely in it's ability to estimate the τ parameter. In fact, most of the time the τ estimations gave a percent error of 0.00%. With that in mind, we found that the error associated with our estimations of α were relatively high in comparison. Quantitatively speaking, the average error for out α estimations was 1.92 %. This error may be attributed to the relatively high variance in the values of $R(n)/R(0)$, which can be seen in figure 1.

Fig. 2. RMSE vs. α



C. Audio Quality

We tested the quality of our echo removal system using the signal processing ability of our auditory cortex. To do this, we obtained a sample audio signal defined by a time vector `t`, and samples vector `y`, which contained an echo. After passing these parameters to our function `Remove_echo.m`, we found that the root–mean squared error between the original audio signal and the samples vector returned from our function was approximately 0.0055. Next, we made use of the function `myplay.m` to qualitatively compare the original audio signal, the echo–containing signal, and the signal returned from our echo removal system.

On the whole, we found that our algorithm greatly reduced the echo, however, it was not as crisp as the original audio signal, `x`. The first plot in figure 3 shows an example of the echo and echoless audio signals generated from `probe5.m`. The second plot in figure 3 shows how well the signal returned from our echo removal system resembles the original audio signal. From figure 3, you can see that the echo–containing signal (shown in red) generated from the probe is similar to the original audio signal, but contains many portions of the signal where the amplitdue exceedes that of the original signal, possibly due to a form of constructive interference. On the other hand, after passing the noisy signal through our echo removal system, is clear that the resulting signal is very simliar to the original, however, there are still discrepancies in their amplitudes. On playback, this sounds like a form of static, however the echo does seem to have been removed. On a scale of 1 to 10, an average score of 9 out of 10 seems appropriate.

III. Conclusion

We found our echo removal system to be highly effective in eliminating echos from audio signals generated by the probe function for this lab. With clear parameterization of the echo generation system and knowledge of the finite bounds for the time delay and echo gain values, the complexity of our algorithm was greatly reduced. Applying the autocorrelation function and simple peak detection allowed us to create a model which maps $R(n)/R(0)$ to α and obtain highly accurate estimates for the time delay (τ) of an echo. One area in which our system failed was the ability to eliminate noise from the input signal entirely, resulting in a somewhat grainy audio quality relative to the source. Further work could explore additional filters to improve this part of our system. In conclusion, our echo removal system can reliably reconstruct audio signals using straightforward computational techniques.

References

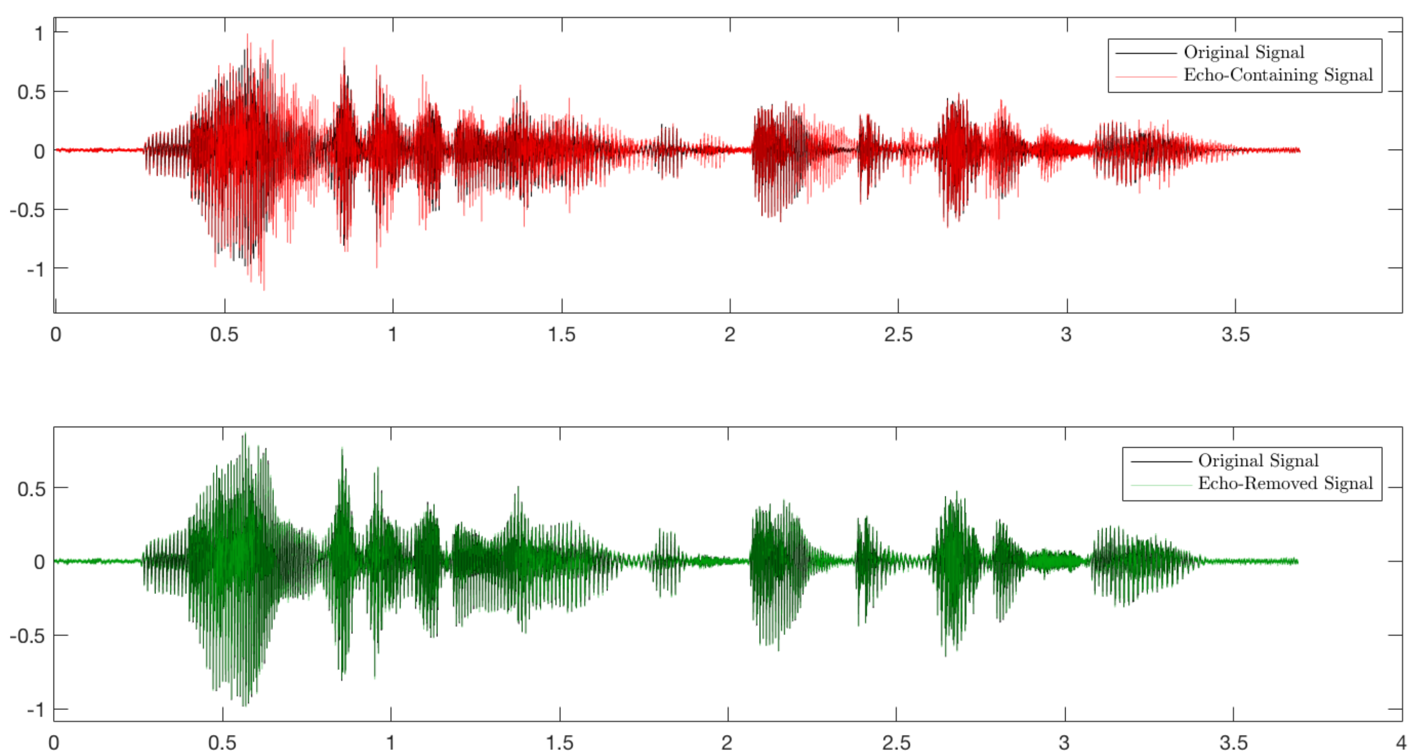
[1] D. Havelock, *Handbook of Signal Processing in Acoustics*. Springer, 2008, no. v. 1-2. [Online]. Available: https://books.google.com/books?id=TDPIO01DLSUC

[2] T. Smyth. (2012) Comb filter impulse response. [Online]. Available: \url{https://www.cs.sfu.ca/~tamaras/delayEffects/Comb_Filter_Impulse.html}

[3] *Signals & Systems: Continuous And Discrete, 4/E*. Pearson Education, 1998. [Online]. Available: https://books.google.com/books?id=F-RSrV8oX_4C

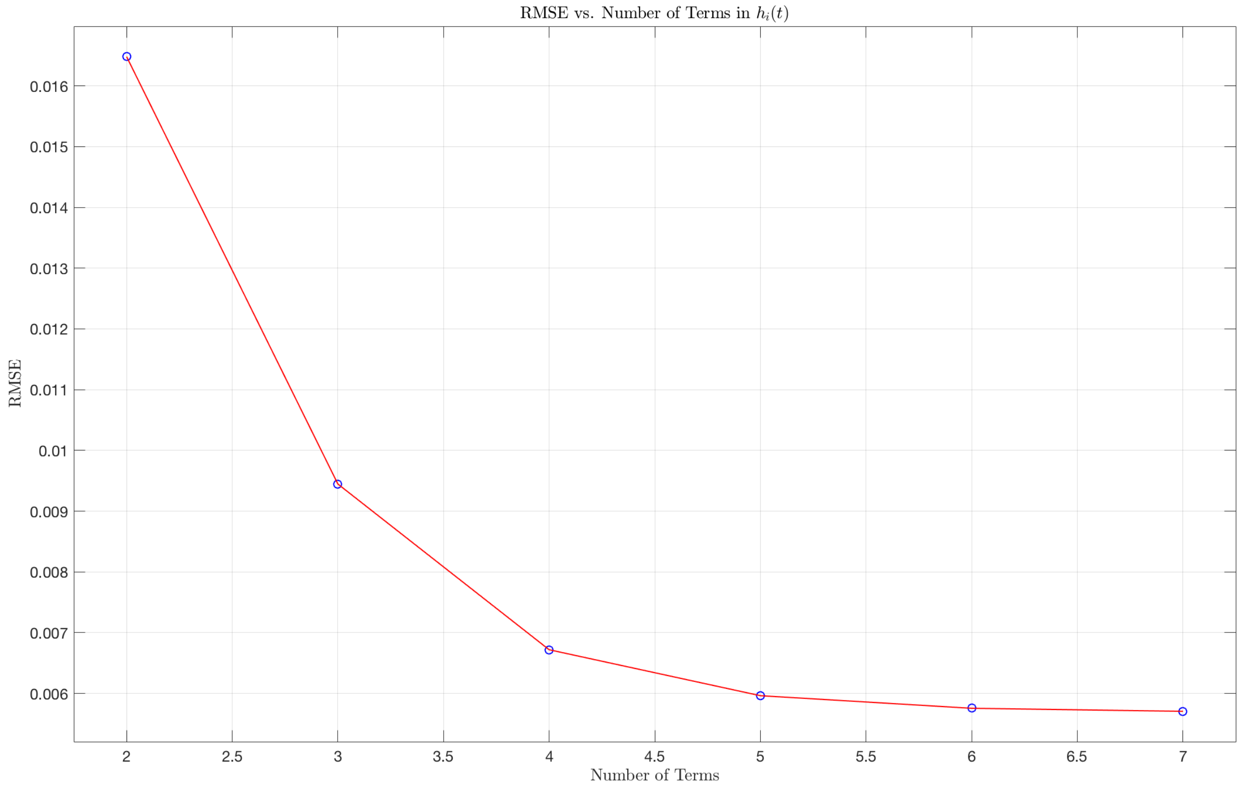
[4] P. William J. Ebel. (2016) background.pdf. [Online]. Available: https://csss.slu.edu/ebel/

Fig. 3. `probe5.m` Example



IV. Appendix

Fig. 4. RMSE and Terms in Impulse Response of Echo Removal System



Infinite Impulse Response Expansion

$$
\begin{aligned}
h_i(t) &= \delta(t) - \alpha h_i(t-\tau) \\
&= \delta(t) - \alpha \left[\delta(t-\tau) - \alpha h_i(t-2\tau)\right] \\
&= \delta(t) - \alpha\delta(t-\tau) + \alpha^2 h_i(t-2\tau) \\
&= \delta(t) - \alpha\delta(t-\tau) + \alpha^2 \left[\delta(t-2\tau) - \alpha h_i(t-3\tau)\right] \\
&= \delta(t) - \alpha\delta(t-\tau) + \alpha^2\delta(t-2\tau) - \alpha^3 h_i(t-3\tau) \\
&\;\;\vdots
\end{aligned}
$$