Efficient blind deconvolution of audio-frequency signal

James N. Caron*†

*Research Support Instruments, 4325-B Forbes Boulevard,
Lanham, MD 20706
†Quarktet, 205 Indian Spring Drive, Silver Spring, MD 20901

Abstract

A signal processing algorithm has been developed in which a filter function is extracted from degraded data through mathematical operations. The filter function can be used to restore much of the degraded content of the data through use of a deconvolution process. The operation can be performed without prior knowledge of the detection system, a technique known as blind deconvolution. The extraction process, designated Self-deconvolving Data Reconstruction Algorithm (SeDDaRA), is applied here to audio-frequency signals showing significant qualitative improvement. Degradation arising from the process of electronic recording and reproduction is significantly reduced.

PACS numbers: 43.60.Ac, 43.60.Dh

1 INTRODUCTION

The self-deconvolving data restoration algorithm, or SeDDaRA, [1, 2, 3] enables efficient restoration and enhancement of degraded data by identification of an impulse function. The effect of the impulse function can be removed from the data using a deconvolution process. In this study, the process is applied to empirical audio-frequency signals that were degraded by electronic reproduction. This action imposes an undetermined frequency response upon the signals. The algorithm compares this frequency response to that of a non-degraded, or 'reference' signal and derives a filter function. The filter function can then be used to restore the signal, and others produced by the same system, using a deconvolution process.

Signal processing [4, 5] using blind deconvolution is an active field of study for a broad spectrum of applications. [6, 7, 8, 9, 11, 10] Blind deconvolution is required in situations where the impulse function cannot be accurately measured or modeled. Blind deconvolution techniques can be classified as iterative or non-iterative. For real-time applications, non-iterative techniques are more desirable, but often have limited applicability. Iterative techniques can be more effective, but are more computationally intensive. [1]

The SeDDaRA process is a non-iterative technique, and has proven to be quite robust in a broad spectrum of applications. Application requires little user input and is computer efficient, in contrast to iterative techniques. The process compares the magnitude of the data in Fourier space to the same quality of a specified reference data set. A filter function is derived from the comparison and used as a transfer function for restoring the original data.

In this study, the technique is applied to acoustic signals as verification that the technique can be applied to audio-frequency signals. An acoustic signal is produced by a conventional computer speaker, and simultaneously recorded by the computer microphone. This action imposes an unknown frequency response upon the signal. With SeDDaRA, much of the degraded frequency response can be restored. This is demonstrated by the significant qualitative improvement in both the signals and their frequency distributions.

2 THEORY

This section summarizes the theoretical development of the blind deconvolution process. A more complete description can be found in the references.[1]

2.1 Deconvolution

The goal of data restoration is to remove degradation from a signal that, with an ideal detection system, would not be present. Once the form of the degradation is known, a class of deconvolution processes, [12, 13, 14, 15] such as non-negative least squares and the Wiener filter, can be used to remove the defect as best as possible. A mathematical representation of the degraded data

$g(t)$ is

$$g(t) = f(t) * d(t) + w(t) \tag{1}$$

where $d(t)$ is the impulse function, $f(t)$ is the truth or non-degraded data, $w(t)$ is a noise term, and '$*$' denotes the convolution. The objective is to find the best estimate of $f(t)$ from $g(t)$ when $d(t)$ and $w(t)$ are unknown. This relationship is simplified by transferring Eq. (1) into frequency space via application of a Fourier transform, yielding

$$G(\nu) = F(\nu)\ D(\nu) + W(\nu) \tag{2}$$

where ν is the coordinate in frequency space, and transformed functions are represented by capital letters.

For known $d(t)$, a deconvolution process can be applied to $g(t)$ to approximate $f(t)$. A pseudo-inverse filter, an approximation of the Wiener filter, is employed here. The deconvolution is produced from

$$F(\nu) \approx \frac{G(\nu)\ D^*(\nu)}{|D(\nu)|^2 + C_2} \tag{3}$$

where the parameter C_2 prevents amplification of noise and can be chosen by trial and error.

2.2 Blind Deconvolution

For unknown $d(t)$, the transfer function $D(\nu)$ is derived from $G(\nu)$ for use in a deconvolution algorithm. To this end, the degradation is assumed invariant, $D(\nu)$ is assumed real and has the form

$$D(u) = [K_G \mathcal{S}\{|G(\nu) - W(\nu)|\}]^{\alpha(\nu)} \tag{4}$$

where $\alpha(\nu)$ is a tuning parameter and K_G is a real, positive scalar chosen to ensure

$$|D(\nu)| \leq 1. \tag{5}$$

Application of the smoothing filter, usually a median filter, $\mathcal{S}\{...\}$ assumes that $D(\nu)$ is a slowly varying function.

Equation (5) demonstrates the difference between the application of SeDDaRA on sounds and images. For images, it is usually sufficient to assume that the degradation only occurs as a result of reduction of specific frequencies. Thus, the criteria for images states that $0 \leq |D(\nu)| \leq 1$. In acoustic settings, the possibility of amplification cannot be ruled out.

Equation (4) is subject to the conditions that the smoothing filter $\mathcal{S}\{...\}$ is separable, and $F(\nu)$ and $W(\nu)$ are uncorrelated. Equation (4) states that application of a smoothing filter and power law to the power spectrum of the reference data, when chosen correctly, will produce the impulse function. Although stated as an equality, in practice this is an approximation stemming from separability condition. For simplicity, the noise term $W(\nu)$ will be assumed negligible.

Since $D(\nu)$ is assumed real, Eq. (2) can be restated as

$$D(\nu) = \frac{\mathcal{S}\{|G(\nu)|\}}{\mathcal{S}\{|F(\nu)|\}} \tag{6}$$

where the smoothing operator has been applied. Since $D(\nu)$ is a slowly varying function, it can be removed from the influence of the smoothing operator.

Equation (6) is substituted into Eq. (4),

$$\frac{\mathcal{S}\{|G(\nu)|\}}{\mathcal{S}\{|F(\nu)|\}} \approx \frac{K_G \mathcal{S}\{|G(\nu)|\}}{K_{F'} \mathcal{S}\{|F'(\nu)|\}} \approx [K_G \mathcal{S}\{|G(\nu)|\}]^{\alpha(\nu)} . \tag{7}$$

Since the truth data $F(\nu)$ is unknown, we have replaced it with a data set $F'(\nu)$ that contains the desired characteristic frequency spectrum, where $K_{F'}$ is another scaling parameter. With the smoothing filter, the replacement data set needs to satisfy

$$K_{F'} \mathcal{S}\{|F'(\nu)|\} \approx K_F \mathcal{S}\{|F(\nu)|\}. \tag{8}$$

Preferably, this function would be a theoretical model of the anticipated result. However, since modeling a detection system is often complicated, using a fair representation of the truth data can be more efficient.

The function $|F'(\nu)|$ is the key to the success of the process. It is a representation of the frequencies that one expects to achieve. With the presence of the smoothing filter, this function need only be similar to the spectrum of the actual truth $|F(\nu)|$. For example, a good quality recording a man's voice may be used to restore a degraded recording of another man's voice. In practice, finding a suitable $|F'(\nu)|$ is not a difficult task.

Solving for $\alpha(\nu)$ produces

$$\alpha(\nu) \approx \frac{\text{Ln}[K_G \mathcal{S}|G(\nu)|] - \text{Ln}[K_{F'} \mathcal{S}\{|F'(\nu)|\}]}{\text{Ln}[K_G \mathcal{S}|G(\nu)|]}. \tag{9}$$

In this relation, K_G and $K_{F'}$ must be determined such that $|D(\nu)| \leq 1$. This condition is satisfied if we set $K_G = 1/\text{Max}[\mathcal{S}\{|G(\nu)|\}]$ and $K_{F'} = 1/\text{Max}[\mathcal{S}\{|F'(\nu)|\}]$.

It follows that

$$D(\nu) = \{K_G \mathcal{S}\{|G(\nu)|\}\}^{\alpha(\nu)} \tag{10}$$

where $\alpha(\nu)$ is given by Eq. (9).

Substitution of Eq. (9) into Eq. (10) produces an approximation of Eq. (6), providing a more concise result.

SeDDaRA is applicable when the degradation is invariant, as it estimates the impulse function from the entirety of the data. In cases where the degradation is not uniform across the data set, some restoration is still possible. However, this may produce non-physical artifacts. The algorithm has also been applied successfully to signals that contain significant noise. A formal study on the influence of noise has not yet been conducted.

3 EMPIRICAL APPLICATION

3.1 Deconvolution with known non-degraded signal

The application of SeDDaRA to sound was first verified by restoring a degraded sound wave using the non-degraded sound wave for the reference data. The synthetic sound wave, sweeping the frequency range from 100 to 2000 Hz for a period of three seconds, is displayed in Fig. 1 (top). The wave is sinusoidal in shape, but due to space constrictions, the waveform appears as a block in the graph. This sound wave was played out to the computer's audio speakers and simultaneously recorded on the computer microphone, Fig. 1 (center), located several inches away. A degraded and unknown frequency response is forced onto the sound wave by both the audio speakers and microphone. The recorded sound wave was processed with SeDDaRA using the synthetic waveform as the reference wave.

The synthetic, recorded, and processed waveforms are shown in Fig. 1. The recorded waveform has a considerable low-frequency component that could be attributed to air currents near the microphone. The processed waveform shows obvious amplification of noise, but removes the low-frequency components and restores the amplitude levels to approximately the original values.

The frequency distributions are displayed in Fig. 2. The low-frequency component dominates the sound wave as evident in the frequency spectrum of the recorded sound wave, Fig. 2 (center). The restoration, however, closely resembles the original spectrum. The highest frequencies (greater than 19 kHz) appear diminished. Some loss may have resulted from the signal being clipped at high amplitudes during the sampling process.

3.2 Deconvolution with derived impulse function

The utility of the previous experiment is that it provides a function that allows for the restoration of any sound wave that traverses the same path as the sound wave tested above. In a broader sense, a correction function can be measured for any system where a signal with a known frequency function can be generated.

Another sound wave, a segment of a digital recording of a bassoon quartet, [16] was played through the system. The recorded wave was restored using the $\alpha(\nu)$ from section 3.1. The function $D(\nu)$ was calculated by following the procedure of section 3.1.

The results of the experiment are shown in Fig. 3. The sound wave is clearly degraded by the playback and recording operation. As shown, application of the pseudo-inverse filter with a derived $D(\nu)$ restores the sound wave. The frequency distributions are shown in Fig. 4.

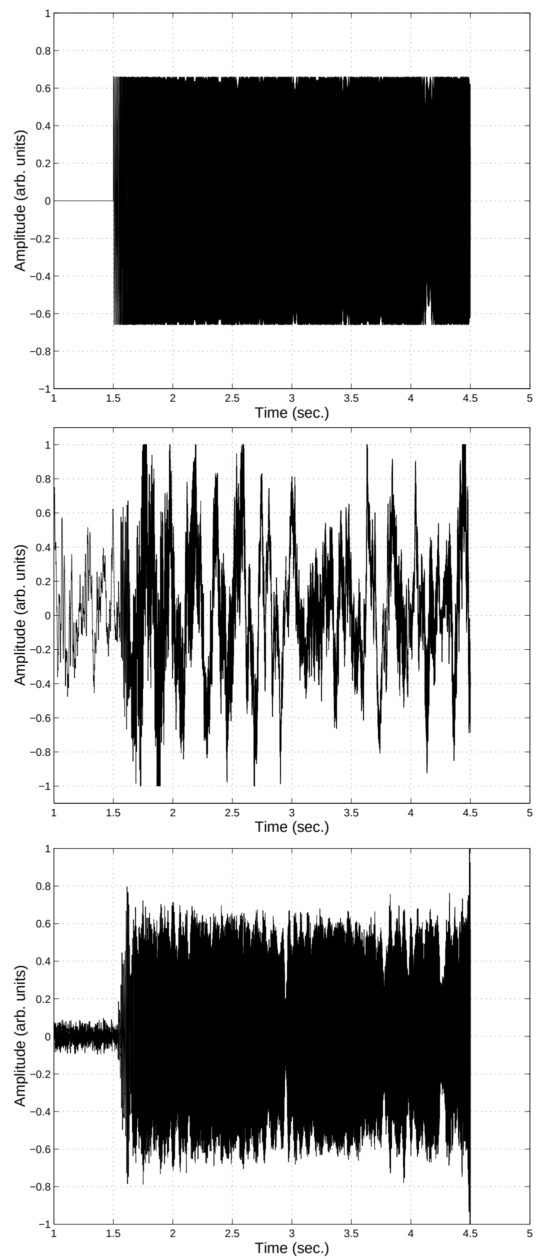

Figure 1: Top: A synthetic waveform was generated that sweeps the frequency range from 100 to 2000 Hz in three seconds with an equal amplitude of $A = 0.66$. The shape of the waves are sinusoidal, but owing to image resolution, cannot be viewed here. Center: The sound wave after it has been played through the computer audio speaker and recorded on the computer microphone. A significant amount of distortion can be seen. Bottom: The waveform after processing with SeDDaRA. Although noise has been amplified, the low-frequency component has been removed and the amplitude for most frequencies has been regained.

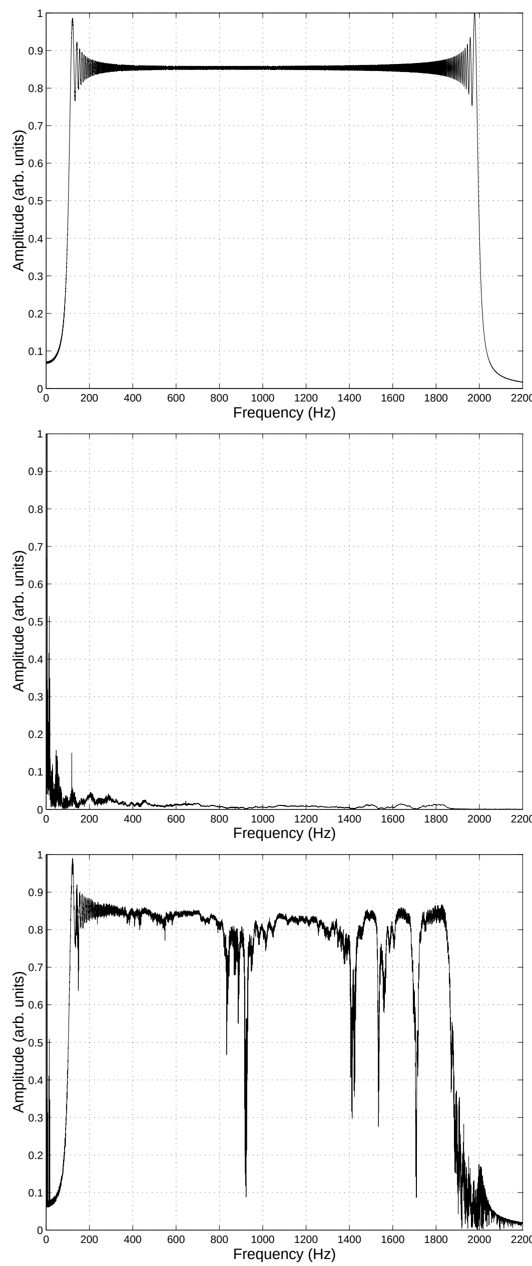

Figure 2: Top: Frequency of the synthetic waveform sweeping the frequency range from 100 to 2000 Hz. Center: The frequency spectrum of the sound wave after it has been played through the computer audio speaker and recorded on the computer microphone. Bottom: The frequency spectrum of the waveform after processing with the SeDDaRA technique. Although some signal loss is apparent, the spectrum closely matches that of the original signal.

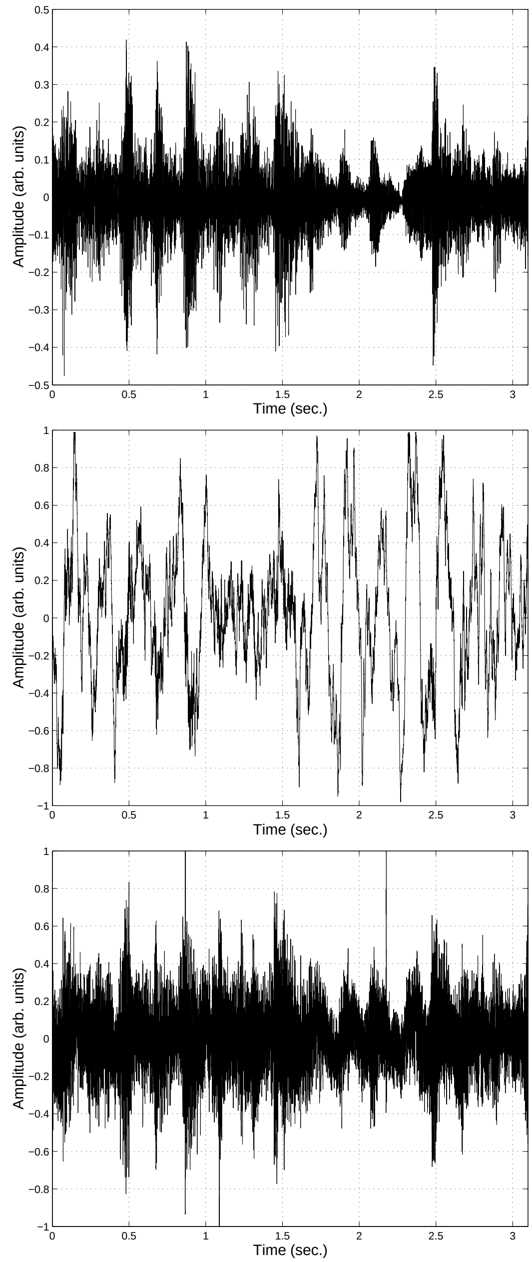

Figure 3: Top: The original segment of the sound clip. Center: The sound wave after it has been played through the computer audio speaker and recorded on the computer microphone. A significant amount of distortion can be seen. Bottom: The waveform after processing with the SeDDaRA technique.

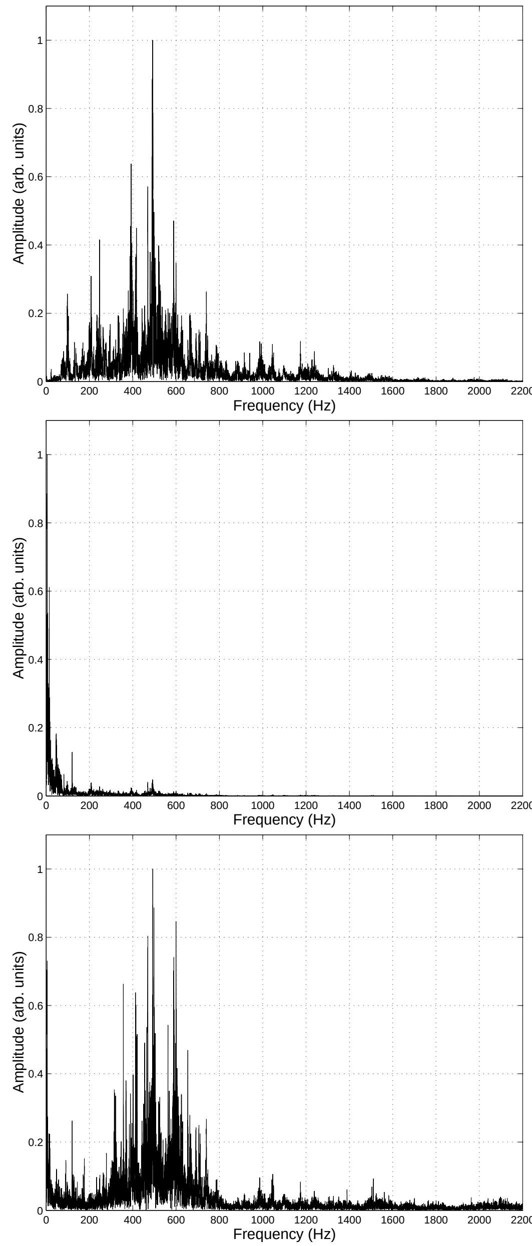

Figure 4: Top: Frequency spectrum of the music segment. Center: The frequency spectrum of the sound wave after it has been played through the computer audio speaker and recorded on the computer microphone. Bottom: The frequency spectrum of the waveform after it was processed using the SeDDaRA technique.

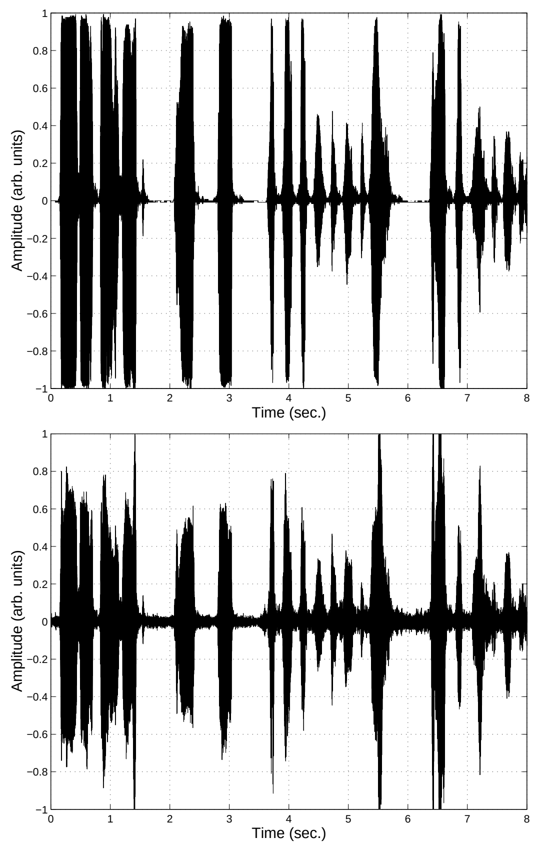

Figure 5: Top: Sound clip from a speech by John F. Kennedy Bottom: Restoration of the Clip using SeDDaRA.

3.3 Deconvolution with unknown impulse function

The method was then applied to an audio recording with undefined frequency degradation. A potential application of this technique is to restore audio information that was not recorded under optimum settings. To test this application and blind deconvolution in general, SeDDaRA was applied to a sound clip from a speech by President John F. Kennedy. [17] The reference data was a voice of a man recorded recently with significantly better sound quality. Ideally, one would restore a degraded recording of a certain person's voice but using a good quality recording of that person's voice as the reference data, if available. This would best preserve the frequency qualities of the person's voice.

The original and restored recordings are shown in Fig. 5. The frequency distributions of the reference data, the original recording, and the restored version are shown in Fig. 6.

From Fig. 6, the reference data (top) has an increased response in the 0 to 500 Hz range and decreased response for frequencies above 800 Hz. After processing, the 200 to 500 Hz region has been comparatively amplified while frequencies above 800 Hz have been attenuated. These qualities are evident in the restored waveform Fig. 5 (bottom) producing a perceptively better reproduction of the voice. However, the reference waveform does have diminished response near 600 Hz, which may suggest the reference sound wave was not as good as approximation as needed.

As with any blind deconvolution method, some model of the transfer function or non-degraded signal must be estimated to perform the operation. For SeDDaRA, the reference signal shapes the frequency response of the restored waveform. Thus, the more attention given to the choice of reference signal, the better the restoration will be.

4 CONCLUSION

A blind deconvolution process has been presented that compares the frequency response of a degraded audio-frequency signal to a good-quality signal with the desired frequency response. A transfer function is derived from the two signals and used to restore the degraded signal. The qualitative experiments presented here verify that the algorithm can be applied to acoustic waveforms to restore the frequency characteristics of the signal. The success of the restoration is dependent on the choice of an appropriate reference signal.

The SeDDaRA process has several unique characteristics that are not found in current signal processing algorithms. At the core of the process, this method extracts a reasonably good approximation for the degradation of a signal in a comparatively short amount of time, provided the degradation is invariant across the data set. This algorithm is easy to implement, and can be inserted into existing signal processing packages without much difficulty. As demonstrated, the method works well on a wide variety of signal types, including images, and acoustic waveforms. This is accomplished without direct information about the type or extent of aberration.

Potential commercial applications include research-quality signal processing, restoration of degraded or non-optimum audio signals, and potentially real-time processing of digital signals, such as those in cell phones. SeDDaRA may also find application in recording studios and home sound systems to counteract effects created by room acoustics, and enhance the quality of the reproduction.

References

[1] J.N. Caron, N.M. Namazi, and C.J. Rollins, "Noniterative blind data restoration by use of an extracted filter function," *Applied Optics*, **41**(32), p. 6884 (2002).

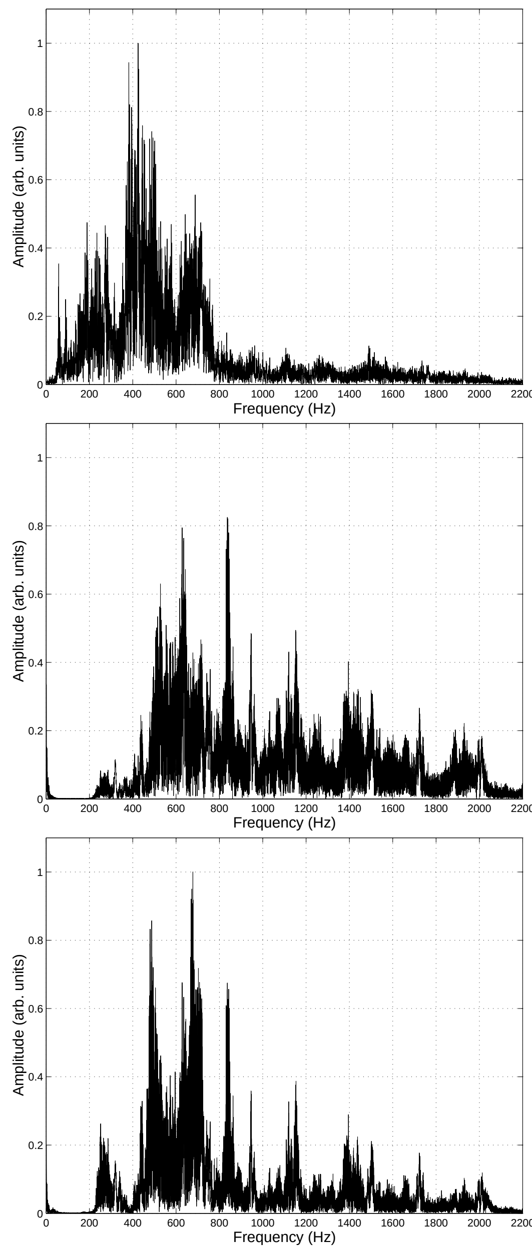

Figure 6: Top: Reference frequency spectrum of a man's voice. Center: The frequency spectrum of the sound wave taken from a speech by JFK. Bottom: The frequency spectrum of the restored signal.

[2] J.N. Caron, N.M. Namazi, R.L. Lucke, C.J. Rollins, and P.R. Lynn, Jr., "Blind data restoration with an extracted filter function," *Optics Letters* **26**(15), p. 1164 (2001).

[3] J.N. Caron, U.S. Patent pending, anticipated acceptance date: August, 2004.

[4] J.G. Proakis and D.G. Manolakis, *Introduction to Signal Processing*, Macmillan Publishing Company, New York, NY (1988).

[5] P.M. Clarkson and H. Stark, eds., *Signal Processing Methods for Audio, Images, and Telecommunications*, Academic Press, San Diego, CA (1995).

[6] Ayers, and Dainty, "Iterative Blind Deconvolution Method and Its Applications," *Optics Letters*, **13**, p. 547 (1988).

[7] J.-Y. Lee and A.K. Nandi, "Extraction of impacting signals using blind deconvolution," *Journal of Sound and Vibration*, **232**(5), p. 945 (2000).

[8] A.K. Nandi, D. Mämpel, and B. Roscher, "Blind deconvolution of ultrasonic signals in nondestructive testing applications," *IEEE Transactions on Signal Processing*, **45**(5), p. 1382 (1997).

[9] S. Barraza-Felix and B.R. Frieden, "Regularization of the image division approach to blind deconvolution," *Applied Optics*, **38**(11), p. 2232 (1999).

[10] K. Torkkola, "Blind separation for audio signals - are we there yet," *Proceedings of the 1st International Workshop on Independent Component Analysis and Blind Signal Separation*, Aussois, France, p. 239 (1999).

[11] M. Joho, H. Mathis, and G. S. Moschytz, "An FFT-based algorithm for multichannel blind deconvolution," *IEEE International Symposium on Circuits and Systems*, Orlando, FL, p. III-203 (1999).

[12] N. Wiener, "The Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications," Wiley, New York (1949).

[13] C.W. Helstrom, "Image restoration by the method of least-squares," *Journal of the Optical Society of America*, **57**, p. 297 (1967).

[14] D. Slepian, "Linear least-squares filtering of distorted images," *Journal of the Optical Society of America*, **57**, p. 918 (1998).

[15] O. Kirkeby, P.A. Nelson, H. Hamada, and F. Orduna-Bustamante, "Fast deconvolution of multichannels using regularization," *IEEE Transactions on Speech and Audio Processing*, **6**(2), p. 189 (1998).

[16] *United Sounds of Bassoon*, Milan Turkovi, Daniele Damiano, Clelia Goldings, and Richard Galler, performers, KOCH International GmbH (1994).

[17] From the Inaugural Address by President John F. Kennedy Washington, D.C. January 20, 1961. Source: John Fitzgerald Kennedy Library, Columbia Point, Boston, MA, USA.