

HOMOMORPHIC PREDICTION[†]

Alan Oppenheim, Gary Kopec^{*} and José Tribolet^{**}

Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science
Research Laboratory of Electronics

Abstract

Two commonly used signal analysis techniques are linear prediction and homomorphic filtering. Each has particular advantages and limitations. This paper considers several ways of combining these methods to capitalize on the advantages of both. The resulting techniques, referred to collectively as homomorphic prediction, are potentially useful for pole-zero modelling and inverse filtering of mixed-phase signals. Results on seismic data and speech are presented.

Introduction

Two commonly used signal analysis techniques are homomorphic filtering^{1,2} and linear prediction.^{3,4,5} Each of these has particular advantages and limitations.

Linear prediction is directed primarily at modelling a signal as the response of an all-pole system. Its chief advantage over other identification methods is that for signals well matched to the model it provides an accurate representation with a small number of easily-calculated parameters. However, in situations where spectral zeros are important linear prediction is less satisfactory. Furthermore, it assumes that the signal is either minimum phase or maximum phase but not mixed phase.

Homomorphic filtering was developed as a general method of separating signals which have been nonadditively combined. Unlike linear prediction, it is not a parametric technique and does not presuppose a specific model. Therefore, it is effective on a wide class of signals, including those which are mixed phase and those characterized by both poles and zeros. However, the absence of an underlying model also means that homomorphic analysis does not exploit as much structure in a signal as

does linear prediction. Thus, it may be far less efficient than an appropriate parametric technique when dealing with highly structured data. It appears possible to combine these two techniques in a number of ways into new methods of analysis which embody the advantages of both.⁶ These combined methods will be referred to collectively as homomorphic prediction. In this paper we consider two specific forms of homomorphic prediction. In one of these, the resulting analysis technique is useful for pole-zero modelling of signals. The other is directed at utilizing the advantages of linear prediction for inverse filtering of mixed-phase signals.

The basic strategy for combining linear prediction with cepstral analysis is to use homomorphic processing to transform a general signal into one or more other signals whose structures are consistent with the assumptions of linear prediction. In this way the generality of homomorphic analysis is combined with the efficiency of linear prediction. In the next section we briefly review some of the properties of homomorphic analysis that suggest this approach. We then discuss several specific ways of combining the two techniques.

Homomorphic Signal Processing

Homomorphic signal processing is based on the transformation of a signal $x(n)$ as depicted in Figure 1. Letting $\hat{X}(z)$ and $X(z)$ denote the z -transforms of $\hat{x}(n)$ and $x(n)$, the system $D_*[\cdot]$ is defined by the relation

$$\hat{X}(z) = \log X(z) \quad (1)$$

where the complex logarithm of $X(z)$ is appropriately defined. The system $L[\cdot]$ is a linear system and the system $D_*^{-1}[\cdot]$ is the inverse of $D_*[\cdot]$. The signal $\hat{x}(n)$ is commonly referred to as the complex cepstrum. There are a number of properties of the complex cepstrum that are

particularly useful, and these have been discussed elsewhere^{1,2}. Of particular interest for this paper are the following:

(1) With $x(n)$ expressed as the convolution of its minimum phase and maximum phase components, denoted as $x_{\min}(n)$ and $x_{\max}(n)$ respectively,

$$\hat{x}(n) = \hat{x}_{\min}(n) + \hat{x}_{\max}(n)$$

Furthermore, $\hat{x}_{\min}(n)$ is zero for $n < 0$ and $\hat{x}_{\max}(n)$ is zero for $n > 0$. Thus, the complex cepstrum provides a means for factoring a signal into its minimum phase and maximum phase components. Specifically, by choosing the linear system in Figure 1 such that

$$\hat{y}(n) = \begin{cases} 0 & n < 0 \\ \frac{1}{2} \hat{x}(0) & n = 0 \\ \hat{x}(n) & n > 0 \end{cases}$$

the output $y(n)$ will be equal to $x_{\min}(n)$. Alternatively, choosing the linear system such that

$$\hat{y}(n) = \begin{cases} \hat{x}(n) & n < 0 \\ \frac{1}{2} \hat{x}(0) & n = 0 \\ 0 & n > 0 \end{cases}$$

will result in an output $y(n)$ equal to $x_{\max}(n)$.

(2) From the above it follows that a mixed phase signal can be converted to a minimum phase signal with the same spectral magnitude. This is accomplished by choosing the linear system such that

$$\hat{y}(n) = \begin{cases} 0 & n < 0 \\ \hat{x}(0) & n = 0 \\ \hat{x}(n) + \hat{x}(-n) & n > 0 \end{cases} \quad (2)$$

The resulting output $y(n)$ will then be a minimum phase signal with the same spectral magnitude as $x(n)$.

Homomorphic Prediction

Given the above properties of the complex cepstrum we now outline several ways of using homomorphic filtering to prepare a signal for analysis by linear prediction. As discussed previously, one limitation of linear prediction is its restriction to minimum or maximum phase signals. However, by exploiting

property (1), a mixed phase signal can be factored into its minimum and maximum phase components. These can then be analyzed separately using linear prediction.

A second limitation of linear prediction is its inability to locate spectral zeros. A number of methods have been proposed^{7,8}. These require, however, that the data have the proper time registration and no linear phase component. Furthermore, in some cases, such as speech analysis, the basic pulse to be modelled is convolved with an impulse train. While all-pole modelling using linear prediction can successfully be applied to the convolved signal, the available methods for pole-zero analysis require that the signal first be deconvolved, i.e. that the basic pulse be estimated. Homomorphic filtering can be used to obtain a minimum phase approximation to the basic pulse. A minimum phase signal by definition has no linear phase component and hence is properly aligned. For speech processing the homomorphic filtering can also be used simultaneously to separate the vocal tract impulse response and excitation⁹. Following the homomorphic filtering one of the available methods can be used to determine the minimum phase counterparts of the poles and zeros. Alternatively, the signal can first be factored into its minimum and maximum phase elements. Because neither of these has a linear phase component they are also synchronized properly.

Example I: Inverse Filter Design for Mixed Phase Signals

We shall illustrate now the effectiveness of homomorphic prediction in designing an inverse filter for an air gun signature. A block diagram of this technique is presented in Figure 2. The data was sampled at 1.22 msec and is plotted in Figure 3. Its complex cepstrum was computed (Figure 4) and from it the minimum phase and maximum phase components of the data were obtained (Figures 5 and 6). Linear prediction was applied to the minimum phase component yielding a causal predictive filter $h_{\min}(n)$ of length M_1 . Similarly, linear prediction was applied to the maximum phase component yielding an anti-causal predictive filter $h_{\max}(n)$ of length M_2 . Finally, the original data is processed through the two-sided inverse filter formed by cascading the two predictive filters defined above. The output of the inverse filter for $M_1=17$ and $M_2=2$ is shown in Figure 7.

Note the compression that was achieved with such a short filter, there is practically only one bubble pulse remaining and the peak ratio was substantially improved from 66% in the original data to 20% in the processed data.

Example II: Pole-Zero Modelling of Speech

Figure 8 is a block diagram of one approach to the pole-zero modelling of speech by homomorphic prediction. As explained earlier, the basic idea is to do pole-zero analysis on a minimum-phase estimate of the vocal tract impulse response, $\{v_{mp}(n)\}$, rather than on the speech waveform itself. Figure 8(a) shows how homomorphic filtering can be used to obtain $\{v_{mp}(n)\}$. When linear prediction (covariance method) is applied to this signal [Figure 8(b)] the roots of the resulting predictor polynomial, $\hat{H}_V(z)$, are estimates of the poles of $V_{mp}(z)$. In Figure 8(c) $\hat{H}_V(z)$ is used to construct the inverse of the LPC error sequence, $\{\tilde{e}^{-1}(n)\}$. If $V(z)$ is rational and its poles were identified correctly, the z-transform of the error signal is the numerator of $V(z)$. In that case the inverse error signal is all-pole and the singularities of $E^{-1}(z)$ correspond to the zeros of $V(z)$. This suggests that in general the analysis of $\{\tilde{e}^{-1}(n)\}$ by linear prediction [Figure 1(d)] will produce a polynomial, $\hat{H}_{e^{-1}}(z)$, such that $\hat{H}_{e^{-1}}(z)/\hat{H}_V(z)$ is a good rational approximation to $V(z)$ [Figure 8(e)]. Instead of first computing $\{\tilde{e}(n)\}$ and then inverting its spectrum to get the inverse error signal it is computationally more convenient to calculate the inverse vocal tract impulse response $\{v_{mp}^{-1}(n)\}$ and then filter it through $1/\hat{H}_V(z)$ to obtain $\{\tilde{e}^{-1}(n)\}$. This is shown in Figure 8(a) and Figure 8(c).

Figure 9 illustrates the above procedure on a sample of voiced speech. The analysis was performed on a 54 msec segment of intervocalic $|m|$ from the utterance "Say momo again". Spectra corresponding to various stages of the processing are shown. The alphabetic tags associated with the plots refer to the labeled points of the block diagram.

References

1. A. V. Oppenheim, R. W. Schafer and T. G. Stockham, "Nonlinear Filtering of Multiplied and Convolved Signals," Proceedings of the IEEE, vol. 56, no. 8, pp. 1264-1291, August 1968.
2. A. V. Oppenheim and R. W. Schafer, Digital Signal Processing, Chapter 10, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1975.
3. J. Makhoul, "Linear Prediction: A Tutorial Review," Proceedings IEEE Special Issue on Digital Signal Processing, vol. 63, no. 4, April 1975
4. L. Wood and S. Treitel, "Seismic Signal Processing," IEEE Proceedings Special Issue on Digital Signal Processing, vol. 63, no. 4, April 1975
5. B. S. Atal and S. L. Hanauer, "Speech Analysis and Synthesis by Linear Prediction of the Speech Wave," J. Acoustical Society of America, vol. 50, pp. 637-655, 1971.
6. A. V. Oppenheim, G. Kopec, J. Tribolet "Signal Analysis by Homomorphic Prediction," submitted to the IEEE Trans. Acoustics, Speech, and Signal Processing, April 1975.
7. J. L. Shanks, "Recursion Filters for Digital Processing," Geophysics, vol. 32, no. 1, pp. 33-51, February 1967.
8. T. Durkin, "The Fitting of Time-Series Models," Rev. Inst. Int. Statist., vol. 28, no. 3, pp. 233-243, 1960.
9. A. V. Oppenheim and R. W. Schafer, "Homomorphic Analysis of Speech," IEEE Trans. Audio and Electroacoustics vol. AU-16, no. 2, pp. 221-226, June 1968.

† This work was supported in part by the Advanced Research Projects Agency monitored by ONR under Contract N00014-75-C-0951 and in part by the National Science Foundation under Grant ENG71-02319-A02.

* Supported by Fannie and John Hertz Foundation.

** Supported by the Instituto de Alta Cultura, Portugal.

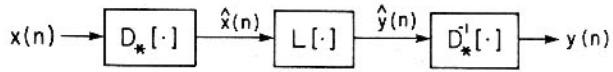


Figure 1: Canonic Homomorphic Processor

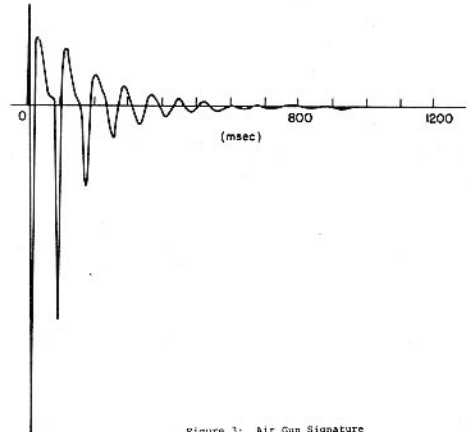


Figure 3: Air Gun Signature

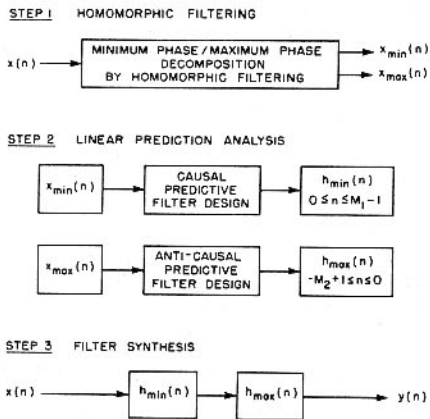


Figure 2: Inverse Filter Design by Homomorphic Prediction

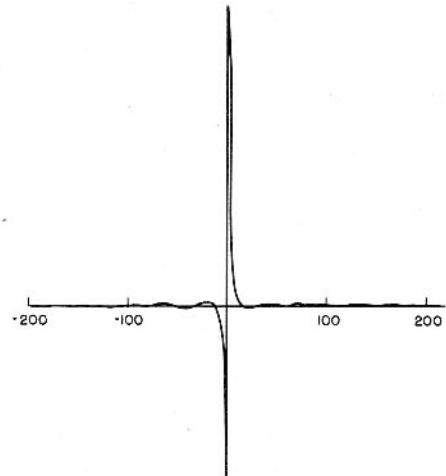


Figure 4: Complex Cepstrum of Air Gun Signature

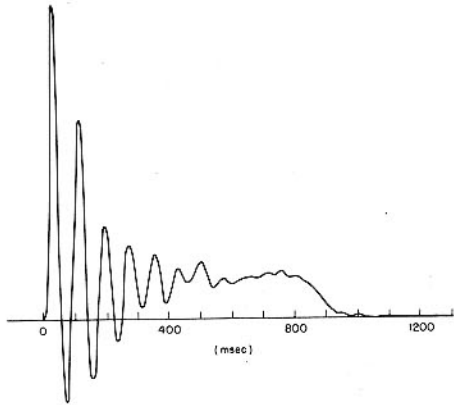


Figure 5: Minimum Phase Component of Air Gun Signature



Figure 7: Air Gun Signature After Inverse Filtering

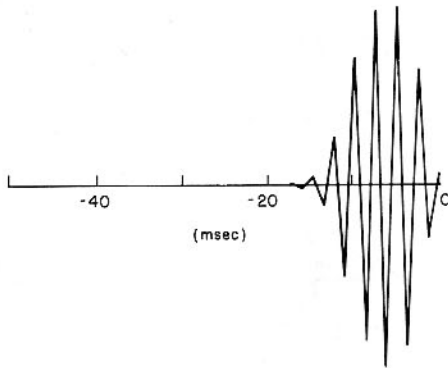


Figure 6: Maximum Phase Component of Air Gun Signature

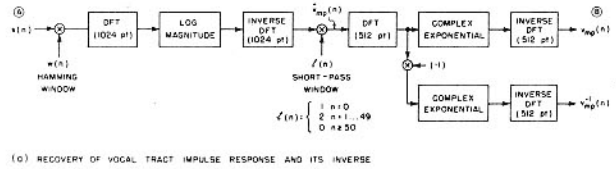


Figure 8: Pole-Zero Speech Analysis by Homomorphic Prediction

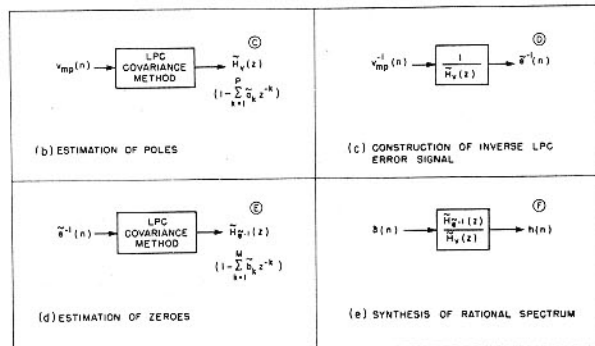
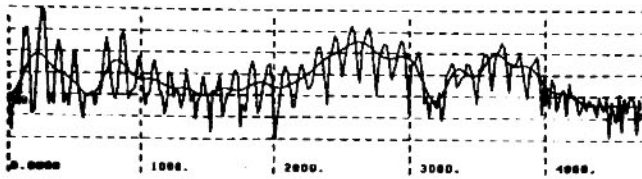
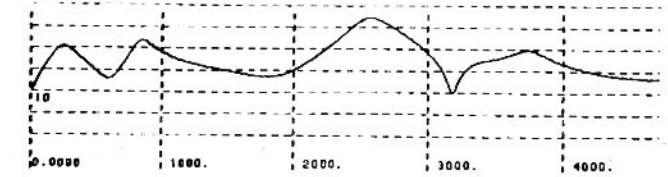


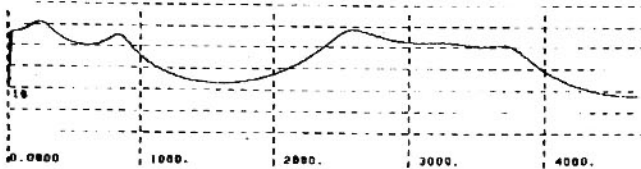
Figure 8: continued



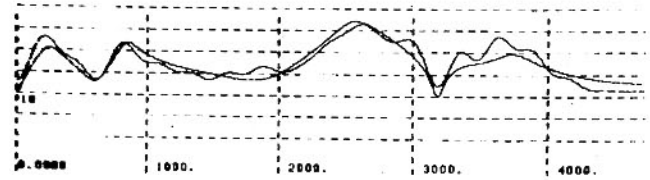
log magnitude spectra of: (A)- windowed speech segment
(B)- vocal tract impulse response



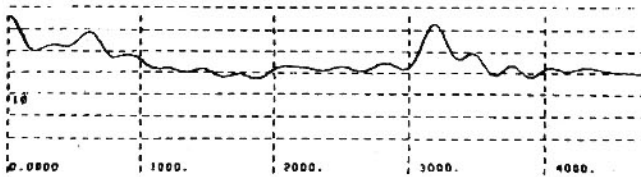
(F)- 12 pole/10 zero rational approximation to vocal tract response



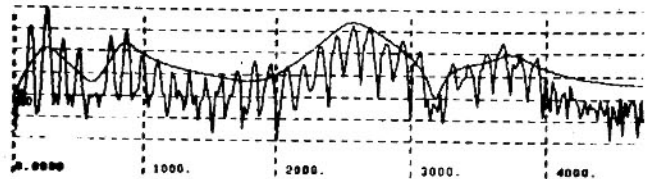
(C)- 12 pole LPC spectrum from (B); estimated vocal tract poles



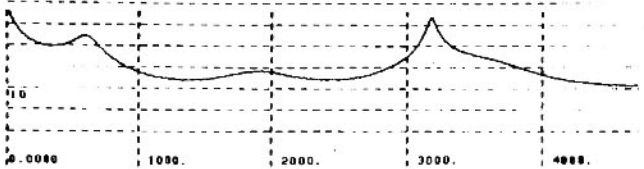
comparison of (F) and (B)



(D)- spectrum of inverse LPC error signal



comparison of (F) and (A)



(E)- 10 pole LPC spectrum from (D); estimated vocal tract zeros

Spectra associated with various steps in the pole-zero analysis of a segment of voiced speech by homomorphic prediction.

Alphabetic labels refer to points indicated in the block diagram.

Figure 9: continued

Figure 9