ISSN No:-2456-2165

Spectral Analysis in the Presence of Noise Using Linear Prediction Coding and Homomorphic Processing

[1]Salem A. Alturky, [2]Abdullah A. Alnaim,
Saudi Aramco

Abstract:- It is claimed that the spectral analysis schemes, which are based on an explicit model of the speech signal (homomorphic and linear predictive analysis), are sensitive to errors in the model. In particular, speech is usually recorded in the presence of some type of background noise. In this research paper, the changes in the spectral envelope caused by the presence of additive white noise will be explored.

Keywords:- *Fourier analysis, Homomorphic analysis, Linear Prediction Coding.*

I. INTRODUCTION

Linear prediction based speech (LPC) analysis is known to be sensitive to the presence of additive noise. For low pitch frequencies, the LPC analysis largely ignores the pitch related fine structure of the speech spectrum. The sensitivity of the LPC parameters increases rapidly as a function of the pitch frequency. This sensitivity can be directly traced to themean-squared prediction error criterion used in the LPC analysis. In this experiment, we present a noise-compensated method for LPC analysis, which ensures good spectral matching between the original speech spectrum and some noisy speech spectrums.

II. PROCEDURE

A. Speech file and Gaussian random noise generation

Using a Gaussian random noise generator in MATLAB, you will be able to produce a white noise sequence. Three cases will be explored where the signal-to-noise ratio is 0 dB, 10 dB, and 20 dB for a low-pitch speech file (100 Hz) that will be generated by applying a train of unit samples to a cascade of second-order IIR filters. The parameters of this speech file are described in table 1.

F1 (Hz)	BW1 (Hz)	F2 (Hz)	BW2 (Hz)	F3 (Hz)	BW3 (Hz)	Pitch (Hz)
270	40	2290	70	3010	170	100

Table 1 - Parameters of a synthetic speech file representing the low-pitch vowel /i/

To do this, a random noise file will be generated of length equal to the original speech file and estimate its variance. Then, the random noise file samples should be scaled so that the quantity $10*\log_{10}$ [variance-of –speech-file/variance-of-noise] achieves the desired dB values. Finally, the scaled random noise have to be added to the speech file to obtain the noisy speech files.

B. Linear Predictive Analysis

In this section, a portion of the speech signal will be extracted by applying a 51.2 msec hanning window starting at 20 msec of the speech file. Then, Fourier analysis will be used to determine the spectra of these records and plot their magnitude. After that, linear predictive analysis(Durbin-Levinson Recursion Method) of these records will be used to obtain the spectral envelope information as well asplotting them using model orders 4, 6 and 14.

C. Homomorphic Analysis

In this section, homomorphic analysis will be used on these data records to obtain spectral envelope information by extracting the low-time (low-quefrency) part of the cepstrum. It is recommended to use at least a 1024 point FFT for the homomorphic analysis. Since the pitch period is known, it will be easy to know how many samples the low-quefrency window should be. Finally, you plot the smoothed spectral magnitude obtained by homomorphic methods.

III. RESULTS AND DISCUSSION

A. Speech file and Gaussian random noise generated

The all-pole model for the linear prediction is most widely used and implemented when applying linear predictive analysis. The equations derived from the all-pole model approach are relatively straight-forward to solve. In this study, we chose to design an all-pole digital filter using the given information presented in table 1. After that, we can apply a train of unit samples to the digital filter to generate our synthetic test speech file for our experiment. On figure 1, you can see a plot of the magnitude of the generated all-pole digital filter representing the low-pitch vowel /i/. The plot shows the desired three formants location, which exactly matches the original formants location based on the given information presented in table 1.

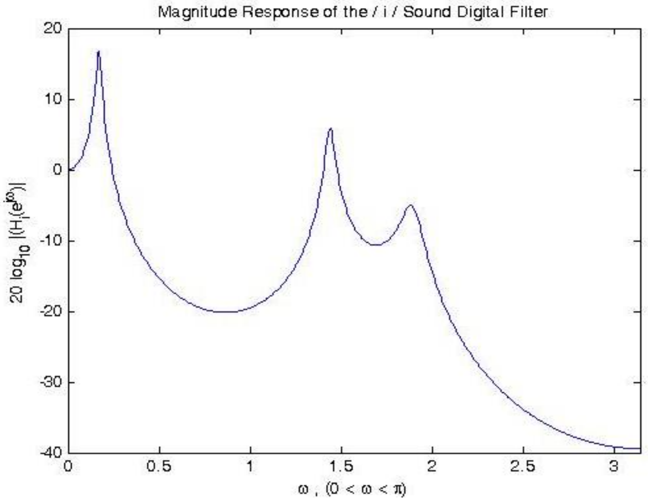


Figure 1 - Magnitude of the All-pole Digital Filter Representing the Low-Pitch Vowel /i/

Three cases of signal-to-noise ratio: 0 dB, 10 dB, and 14 dB are applied to the low-pitch speech file. You can see the resulting noisy speech files on figure 2. From figure 2, you can see that as the signal-to-noise ratio increases, more energy is presented on signals. The generated noisy signals will be used to examine the sensitivity of the linear prediction coding (LPC) method as well as the homomorphic method.

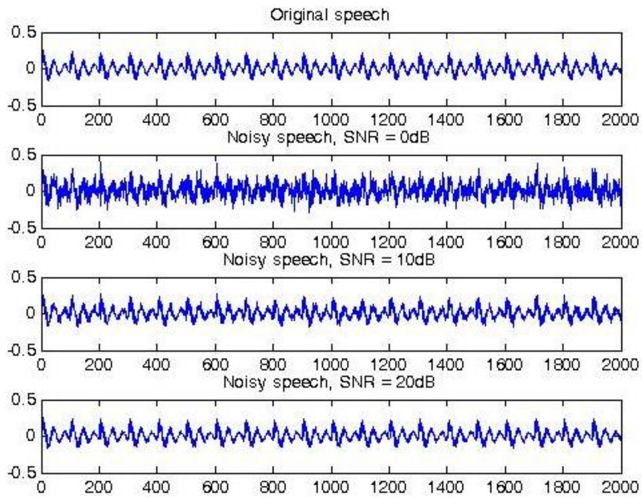


Figure 2 - Noisy Speech Files

B. Linear Predictive Analysis

The prediction of current sample as a linear combination of past P samples forms the basis of linear prediction analysis where P is the order of prediction. The primary objective of linear predictive analysis is to compute the linear predictive coefficients, which minimize the prediction error e(n). Linear predictive analysis separates the given short-term sequence of speech into its slowly varying vocal tract component represented by linear predictive filter H(z), and fast varying excitation component given by the linear predictive residual e(n). As the linear predictive spectrum provides the vocal tract characteristics, the vocal tract resonances (formants) can be estimated from the linear predictive spectrum. Various formant locations can be obtained by picking the peaks from the magnitude linear predictive spectrum.

The Levinson–Durbin algorithm is a recursive order-update method for calculation of linear predictor coefficients. The Durbin algorithm starts with a predictor of order zero for which $E^{(0)}=r_{xx}(0)$. The algorithm then computes the coefficients of a predictor of order i, using the coefficients of a predictor of order i-1. In the process of solving for the coefficients of a predictor of order P, the solutions for the predictor coefficients of all orders less than P are also obtained. The autocorrelation method of linear prediction minimizes the error signal over all time.

From the linear predictive analysis, you can notice how the spectra vary with the different model orders. As the signal-to-noise ratio increases as well as the model order, the more energy is presented on the noisy signals and the result would be a good estimation of spectra, because you can easily detect more spectrums with more energy. From figures 3, 4, and 5, you can observe that the 20 dB noisy signal with the model order 14 has the best spectral matching of the original speech spectrum.

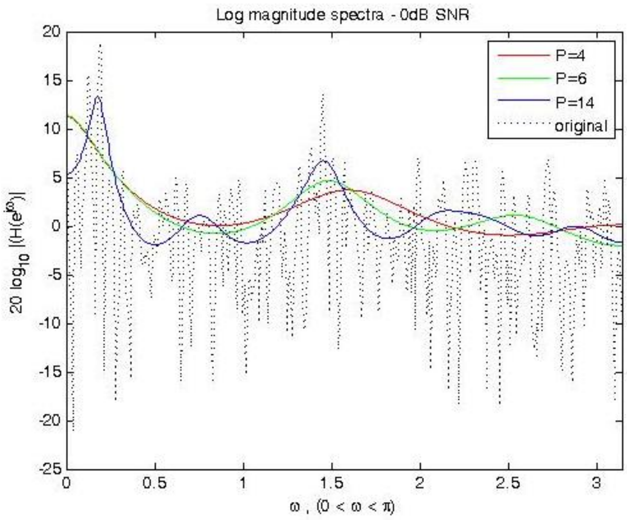


Figure 3 -Spectral Envelope Information for the 0 dB Noisy Speech File Using Model Orders 4, 6, and 14

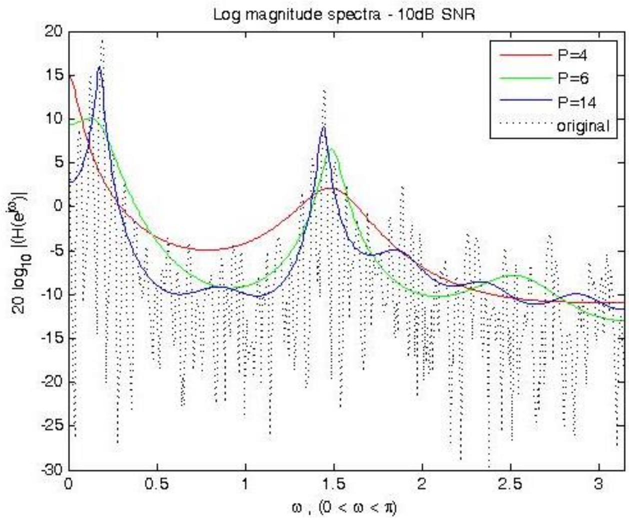


Figure 4 - Spectral Envelope Information for the 10 dB Noisy Speech File Using Model Orders 4, 6, and 14

ISSN No:-2456-2165

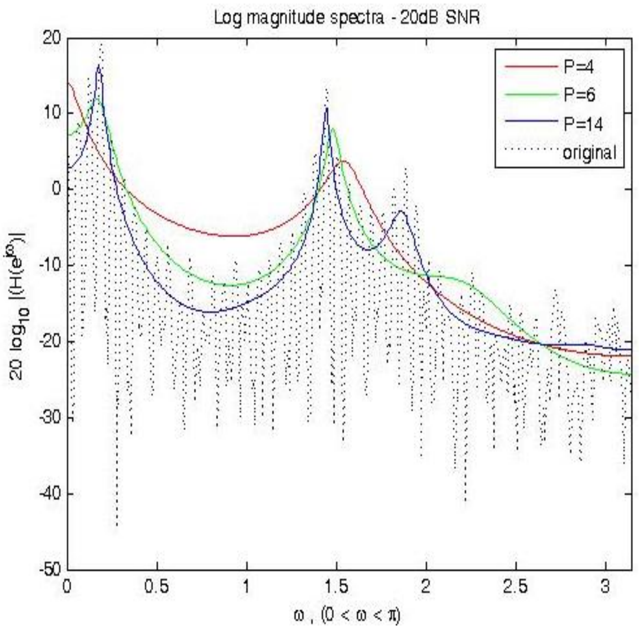


Figure 5 - Spectral Envelope Information for the 14 dB Noisy Speech File Using Model Orders 4, 6, and 14

C. Homomorphic Analysis

Homomorphic filtering was developed as a general method of separating signals, which have been non-additively combined. Homomorphic filtering can also be used simultaneously to separate the vocal tract impulse response and the excitation. Cepstral methods may be used to estimate pitch and voicing as well. Presence of a peak in the high-quefrency cepstrum indicates voicing. Homomorphic passes the desired signal unaltered, while removing the undesired signal. For linear systems, this is analogous to additive noise removal. In figure 6, you can see a plot of the cepstrum of the noisy speech files considering the three signal-to-noise ratio cases. For the case where the signal-to-noise ratio is equal to 20 dB, the plot shows a good estimation for the formants of the original speech file (the peaks of the noisy signals represents the formants). One can conclude that as the signal-to-noise ratio increases, the more accurate estimation of the formants.

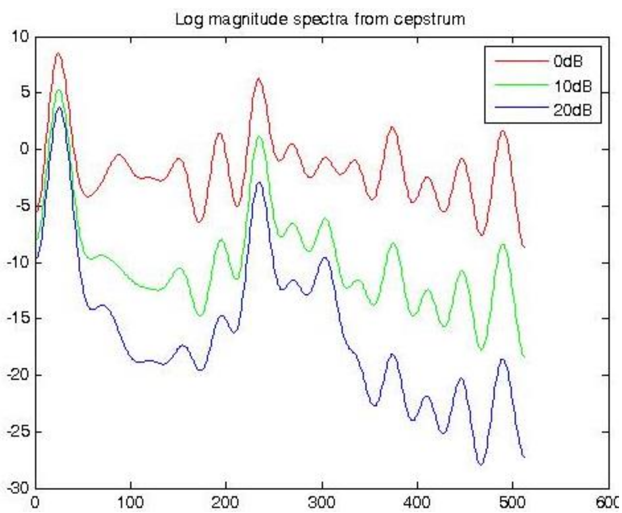


Figure 6 - Cepstrum Plot for the 0, 10, and 20 dB Noisy Speech Files

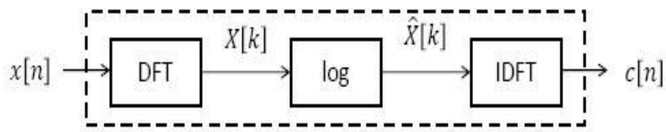
IV. APPLICATIONS OF LINEAR PREDICTIVE CODING

Linear prediction modeling is used in a diverse area of applications, such as data forecasting, speech coding, video coding, speech recognition, model-based spectral analysis, model-based interpolation, signal restoration, and impulse/step event detection. Linear prediction models are extensively used in speech processing, in low bit-rate speech coders, speech enhancement and speech recognition. Here in this experiment, we used the Levinson-Durbin Recursion algorithm to estimate the spectrums of three noisy signals, and then compare them with the original spectrum of the synthetic speech file. This experiment related to the area of speech coding, speech recognition, and model-based spectral analysis. We can use this technique to predict the original formants (resonances frequency that represent the vocal tract) of a noisy speech signal.

V. CONCLUSION

As we have discussed in this research paper, Linear prediction analysis is known to be sensitive to the presence of additive noise. From the linear predictive analysis, you can notice how the spectra vary with the different model orders. We have concluded that as the signal-to-noise ratio increases as well as the model order, the more energy is presented on the noisy signals and the result would be a good estimation of spectrums, because we can easily detect more spectrums with more energy (more SNR). In addition, using the Homomorphic analysis, we have shown that the Cepstral method may be used to estimate pitch and voicing. For example, presence of a peak in the high-quefrency cepstrum indicates voicing, which gives the spectrums. This experiment shows that both methods, the Linear Prediction analysis and the homomorphic analysis, give an accurate estimation of the spectrums with the presence of an additive noise. You can refer to the appendix section in this paper to look for the process of accomplishing both techniques.

APPENDIX



Homomorphic Speech Processing [2]

$$
\begin{aligned}
& for\ i = 1, 2, \ldots, p \\
& \quad E_0 = r(0) \\
& \quad k_i = \left(r(i) - \sum_{j=1}^{i-1} a_{i-1}(j) r(i-j) \right) / E_{i-1} \\
& \quad for\ \mathrm{j} = 1, 2, \ldots, i-1 \\
& \qquad a_i(i) = k_i \\
& \qquad a_i(j) = a_{i-1}(j) - k_i a_{i-1}(i-j) \\
& \quad E_i = (1 - k_i^2) E_{i-1}
\end{aligned}
$$

Levinson Durbin Algorithm [2]

REFERENCES

[1]. D. G. Childers. Speech Processing and Synthesis Toolboxes. John Wiley & Sons, Inc., New York, 2000.

[2]. Lawrence Rabiner and Ronald Schafer. 2010. Theory and Applications of Digital Speech Processing (1st ed.). Prentice Hall Press, Upper Saddle River, NJ, USA.

[3]. S. Chandra and Lin Wen. Experimental comparison between stationary and nonstationary formulations of linear prediction applied to voiced speech analysis. Acoustics, Speech, and Signal Processing [see also IEEE Transactions on Signal Processing], IEEE Transactions on, 22(6):403–415, 1974.

[4]. Singhal, Sharad. "On Reducing Pitch Sensitivity of LPC Parameters." The Journal of the Acoustical Society of America 72.S1 (1982): S78. Print.

[5]. Trabelsi, A.; Boyer, F.R.; Savaria, Y.; Boukadoum, M.; "Improving LPC analysis of speech in additive noise," Circuits and Systems, 2007. NEWCAS 2007. IEEE Northeast Workshop on , vol., no., pp.93-96, 5-8 Aug.2007. doi: 10.1109/NEWCAS.2007.4487956

[6]. Vaseghi, Saeed V. Advanced Digital Signal Processing and Noise Reduction. Chichester: John Wiley, 2000. Print.

[7]. J.R. Deller Jr., J.H.L. Hansen, J.G. Proakis, Discrete-Time Processing of Speech Signals, IEEE Press, New York, 2000.

[8]. G. Kopec, A. Oppenheim and J. Tribolet, "Speech analysis homomorphic prediction," in IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 25, no. 1, pp. 40-49, February 1977, doi: 10.1109/TASSP.1977.1162909.

[9]. Paulikas, Sarunas & Karpavičius, R.. (2007). Application of Linear Prediction Coefficients Interpolation in Speech Signal Coding. Electronics and Electrical Engineering. 80. 39-42.

[10]. Rahman, M. Shahidur & Shimamura, Tetsuya. (2005). Linear prediction using homomorphic deconvolution in the autocorrelation domain. Proceedings - IEEE International Symposium on Circuits and Systems. 2855 - 2858 Vol. 3. 10.1109/ISCAS.2005.1465222.