REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE May 1996	3. REPORT TYPE AND DATES COVERED Professional Paper
4. TITLE AND SUBTITLE EFFICIENT CALCULATION OF SPECTRAL TILT FROM VARIOUS LPC PARAMETERS		5. FUNDING NUMBERS PR: CE77 PE: 0603013N WU: DN302104
6. AUTHOR(S) V. Goncharoff, E. VonColln, and R. Morris		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Command, Control and Ocean Surveillance Center (NCCOSC) RDT&E Division San Diego, CA 92152–5001		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Naval Engineering Logistics Office P. O. Box 2304, Eads Street Arlington, VA 22201		10. SPONSORING/MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES		
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.		12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

A coarse measure of a discrete signal's power distribution vs. frequency is its "spectral tilt," which we define as the slope of least-squares linear fit to the log power spectrum ($0 \leq \omega \leq \pi$ rads/sec). One application of calculating spectral tilt is to help discriminate between voiced and unvoiced/silent time segments while processing sampled speech signals. When linear predictive coding (LPC) of speech is used to model its short-time spectrum, it is computationally more efficient to calculate spectral tilt directly from the LPC model parameters instead of from the log power spectrum. In this paper we present methods for calculating spectral tilt from cepstral coefficients, pole values, and polynomial coefficients. These methods are all based on calculating spectral tilt as a weighted sum of cepstral coefficients.

19960523 208

~~Published in *Proceedings of the IASTED International Conference on Signal and Image Processing*, pp. 60–63, November 1995.~~

14. SUBJECT TERMS Mission Area: Speech Technology Spectral Tilt Linear Predictive Coding (LPC) Polynomial Coefficients Cepstral Coefficients Pole Values			15. NUMBER OF PAGES 16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT SAME AS REPORT

DTIC QUALITY INSPECTED 1

21a. NAME OF RESPONSIBLE INDIVIDUAL	21b. TELEPHONE *(include Area Code)*	21c. OFFICE SYMBOL
E. VonColln	(619) 553–3655	Code 44213

Efficient Calculation of Spectral Tilt from Various LPC Parameters

Vladimir Goncharoff*, Eric VonColln and Robert Morris

*University of Illinois at Chicago
College of Engineering, EECS Dept., Mail Code 154
851 S. Morgan St., Chicago, IL 60607-7053
Internet: goncharo@eecs.uic.edu

NCCOSC
RDT&E Division, Code 44213
53245 Patterson Rd., San Diego, CA 92152
Internet: voncolln@cod.nosc.mil

Abstract

A coarse measure of a discrete signal's power distribution vs. frequency is its "spectral tilt", which we define as the slope of least-squares linear fit to the log power spectrum ($0 \le \omega \le \pi$ rads/sec). One application of calculating spectral tilt is to help discriminate between voiced and unvoiced/silent time segments while processing sampled speech signals. When linear predictive coding (LPC) of speech is used to model its short-time spectrum, it is computationally more efficient to calculate spectral tilt directly from the LPC model parameters instead of from the log power spectrum. In this paper we present methods for calculating spectral tilt from cepstral coefficients, pole values, and polynomial coefficients. These methods are all based on calculating spectral tilt as a weighted sum of cepstral coefficients.

Introduction

Calculation of linear spectral tilt has been of interest to scientists in various disciplines [1-2], as it provides a coarse measure of a signal's power distribution vs. frequency. Given z-domain transfer function $H(z)$ we define spectral tilt to be m_{opt}, the slope of line $y(\omega) = m_{opt}\omega + b_{opt}$ that best fits the log power spectrum $\ln|H(e^{j\omega})|^2$ over the range $0 \le \omega \le \pi$, by the method of least-squares. That is, when error E is defined as

$$E = \int_0^{\pi} \left\{ m\omega + b - \ln\left|H\left(e^{j\omega}\right)\right|^2 \right\}^2 d\omega ,$$

E is minimum when $b = b_{opt}$ and $m = m_{opt}$. In the continuous-frequency case, calculating spectral tilt using this formula is difficult except for the simplest $|H(e^{j\omega})|^2$. An approximation to the spectral tilt is usually found from M samples of $\ln|H(e^{j\omega})|^2$, taken uniformly over the range $0 \le \omega \le \pi$. Estimating spectral tilt in this manner falls within the general class of least-squares polynomial interpolation [3], and estimation error is reduced to zero as $M \to \infty$. When $H(z)$ is in the form of an all-pole model,

$$H(z) = \frac{G}{\left(1 - p_1 z^{-1}\right)\left(1 - p_2 z^{-1}\right)\cdots\left(1 - p_N z^{-1}\right)} ,$$

it is advantageous to find m_{opt} from the N pole values directly, without first having to calculate samples of $\ln|H(e^{j\omega})|^2$. This not only reduces the number of calculations required, but also provides an exact measure of spectral tilt of the LPC log power spectrum. The rest of this paper deals with estimating m_{opt} from various LPC filter parameters, such as the pole values $\{p_1, p_2, \ldots, p_N\}$.

Background

Let us represent the spectral tilt of $S(\omega) = \ln|H(e^{j\omega})|^2$ using functional notation: $m = Tilt\{S(\omega)\}$. The underlying principle to our method is the linearity property of least-squares spectral tilt calculation.

That is, if $S(\omega) = a_1 S_1(\omega) + a_2 S_2(\omega)$, then

$$Tilt\{S(\omega)\} = Tilt\{a_1 S_1(\omega) + a_2 S_2(\omega)\}$$
$$= a_1 Tilt\{S_1(\omega)\} + a_2 Tilt\{S_2(\omega)\}.$$

Thus when $S(\omega) = \Sigma S_k(\omega)$, spectral tilt may be found as the sum of tilts from each term $S_k(\omega)$:

$$m = \sum_k Tilt\{S_k(\omega)\} = \sum_k m_k$$

To prove that the linearity property holds in this case, consider the following: when optimally fitting samples of $S(\omega)$ $(0 \le \omega \le \pi)$ with a first-order polynomial by the method of least squares, b_{opt} and m_{opt} may be found by solving the following system of normal equations $Ax = y$:

$$\begin{bmatrix} M & \sum_{k=0}^{M-1} \omega_k \\ \sum_{k=0}^{M-1} \omega_k & \sum_{k=0}^{M-1} \omega_k^2 \end{bmatrix} \begin{bmatrix} b_{opt} \\ m_{opt} \end{bmatrix} = \begin{bmatrix} \sum_{k=0}^{M-1} S(\omega_k) \\ \sum_{k=0}^{M-1} \omega_k S(\omega_k) \end{bmatrix}$$

Here the polynomial being defined is $y(\omega) = m_{opt}\omega + b_{opt}$, and frequency values $\omega_k = k\pi/(M-1)$ $(k = 0, 1, \cdots, M-1)$ are M uniformly-spaced samples of ω over the interval $0 \le \omega \le \pi$. Vector y may be expressed as $y = V^T s$, where:

$$s = \begin{bmatrix} S(\omega_0) \\ S(\omega_1) \\ S(\omega_2) \\ \vdots \\ S(\omega_{M-1}) \end{bmatrix} \qquad V = \begin{bmatrix} 1 & \omega_0 \\ 1 & \omega_1 \\ 1 & \omega_2 \\ \vdots & \vdots \\ 1 & \omega_{M-1} \end{bmatrix}$$

Vector x may be found after a single matrix multiplication: $x = A^{-1}V^T s = Ds$ (A^{-1} will always exist for unique sampled values of ω). Note that matrix D is independent of $S(\omega)$.

Since matrix multiplication is a linear operation, replacing s with a sum of terms yields:

$$x = D\sum_{i=1}^{N} s_i = \sum_{i=1}^{N} Ds_i = \sum_{i=1}^{N} x_i$$
$$= \begin{bmatrix} b_1 + b_2 + \cdots + b_N \\ m_1 + m_2 + \cdots + m_N \end{bmatrix} = \begin{bmatrix} b_{opt} \\ m_{opt} \end{bmatrix}$$

Thus we have proven that for polynomial fitting of sampled points by the method of least-squares, the sum of polynomial coefficients calculated to optimally fit each additive component of s equals the polynomial coefficients calculated to optimally fit s. The true spectral tilt of $S(\omega)$ is approached when the number of sample points $M \to \infty$, so this property is also valid in the continuous frequency domain.

Spectral Tilt from Cepstral Coefficients

The log power spectrum $S(\omega)$ may be expressed as a summation of cosine functions, weighted by the cepstral coefficients [4]:

$$S(\omega) = \ln\left|H(e^{j\omega})\right|^2 = 2\ln\left|H(e^{j\omega})\right|$$
$$= 2\sum_{k=1}^{\infty} c_k \cos(k\omega) = \sum_{k=1}^{\infty} S_k(\omega)$$

Thus $S(\omega)$ is the sum of infinitely many cosine terms. Using the linearity property discussed above, we may express spectral tilt in terms of the cepstral coefficients:

$$m = \sum_{k=1}^{\infty} m_k = \sum_{k=1}^{\infty} Tilt\{2c_k \cos(k\omega)\}$$
$$= 2\sum_{k=1}^{\infty} c_k Tilt\{\cos(k\omega)\} = \frac{-48}{\pi^3} \sum_{k=1,3,5,\cdots}^{\infty} \frac{c_k}{k^2}$$

Because the even-indexed cepstral coefficients are the weights for cosine waves having an integer number of periods in the range $0 \le \omega \le \pi$ rads/sec, there is zero contribution to spectral tilt from these terms.

Spectral Tilt from Linear Predictor Poles

The log power spectrum of an N-th order LPC filter may also be expressed in terms of its N pole values and gain constant G:

$$S(\omega) = \ln\left|H\left(e^{j\omega}\right)\right|^2 = \ln\left|G\prod_{k=1}^{N}\left(1 - p_k e^{-j\omega}\right)^{-1}\right|^2$$

$$= \ln G^2 - 2\sum_{k=1}^{N}\ln\left|1 - p_k e^{-j\omega}\right|$$

In this form $S(\omega)$ is the sum of $N+1$ terms. The spectral tilt of the term $\ln G^2$ equals zero, leaving the following expression for spectral tilt of $S(\omega)$ in terms of its N pole values:

$$m = Tilt\{S(\omega)\} = -2\sum_{k=1}^{N} Tilt\left\{\ln\left|1 - p_k e^{-j\omega}\right|\right\}$$

We were not able to solve the integral and express m in terms of the pole values $\{p_1, p_2, \ldots, p_N\}$ directly from this equation. An alternate solution method exists, however: to first calculate cepstral coefficients from the pole values, then solve for spectral tilt by the formula derived in the previous section. The cepstral coefficients corresponding to the N-th order all-pole log power spectrum may be found as follows [5]:

$$c_k = \frac{1}{k}\sum_{n=1}^{N}\left(p_n\right)^k$$

Combining this result with the previous expression for spectral tilt in terms of the cepstral coefficients, we obtain the following infinite summation:

$$m = \frac{-48}{\pi^3}\sum_{k=1,3,5,\cdots}^{\infty}\left\{\frac{1}{k^3}\sum_{n=1}^{N}\left(p_n\right)^k\right\}$$

Fortunately, due to the p^k / k^3 terms, the solution for spectral tilt m converges rapidly with increasing k (typically, only fifteen or so terms yield a solution to full IEEE double floating-point precision). Another LPC pole-based method is to pre-calculate spectral tilt for various combinations of pole magnitude and angle, then store the results in a look-up table. We obtained good quantization levels by uniform sampling of θ vs. $\ln(1-r)$, where $p = re^{j\theta}$.

Spectral Tilt from Linear Predictor Coefficients

The log power spectrum of an N-th order LPC filter may also be expressed in terms of its linear predictor coefficients and gain constant G:

$$S(\omega) = \ln\left|H\left(e^{j\omega}\right)\right|^2 = \ln G^2 - \ln\left|1 - \sum_{k=1}^{N}\alpha_k e^{-j\omega k}\right|^2$$

In this form $S(\omega)$ may not be expressed as a sum of terms, where each term depends on a coefficient α_k. The approach we take is to take advantage of a recursive solution for the cepstral coefficients from the LPC coefficients [6]:

$$c_k = \begin{cases} \alpha_k + \frac{1}{k}\sum_{n=1}^{k-1} n c_n \alpha_{k-n}, & 1 \le k \le N; \\ \frac{1}{k}\sum_{n=k-N}^{k-1} n c_n \alpha_{k-n}, & k > N. \end{cases}$$

The cepstral coefficients are then used to calculate spectral tilt from the formula previously derived. In most cases this method will be more computationally efficient than factoring the LPC polynomial to obtain the pole values, and solving for spectral tilt using the pole equations.

Experimental Results

We have applied the equations derived above to measuring the spectral tilt of sampled speech data: a 30 msec hanning window was stepped across the speech samples with 0.1 msec steps, the windowed data zero-padded to 1024 points prior to calculating its FFT.

The MATLAB "polyfit" routine was used to estimate the spectral tilt of 513 samples of $\ln|H(e^{j\omega})|^2$, spaced uniformly between 0 and 5 KHz. A 180 ms section of the speech waveform used for this experiment is plotted as Figure 1(a). The corresponding spectral tilt, calculated directly from $\ln|H(e^{j\omega})|^2$, is plotted as the solid line in Figure 1(b). (The slope value is expressed in dB/kHz.) This plot shows interesting spectral slope features corresponding to silence, onset and end of voiced speech, sibilant and plosive sounds.

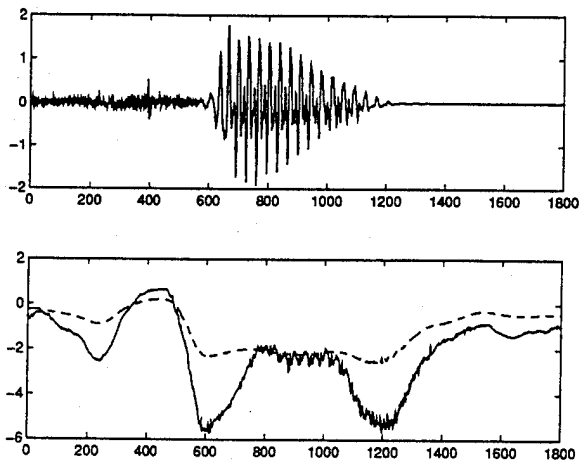

Figure 1(a). A 180 ms section of female speech selected for analysis.

Figure 1(b). Solid line: exact spectral tilt, calculated from the log-magnitude spectrum. Dashed line: spectral tilt estimated from the poles of the corresponding LPC-14 spectral model.

Next we calculated LPC-14 coefficients corresponding to each short-time spectrum (autocorrelation method), factored the denominator polynomial to obtain pole values for each frame, and applied the formula presented above that calculated LPC spectral tilt from the pole values. The result is plotted as the dashed line in Figure 1(b). We also found that lower LPC orders provide reasonable estimates of spectral tilt, as shown in Fig. 2.

Conclusion

From the equations presented here one may calculate spectral tilt directly from various LPC model parameters, avoiding the necessity and computational expense of first calculating the log power spectrum. We rely on the linearity property of least-squares polynomial data fitting to simplify our expressions. The results may be applied to efficiently estimate the short-time spectral tilt of speech signals to help discriminate between various speech waveform components.

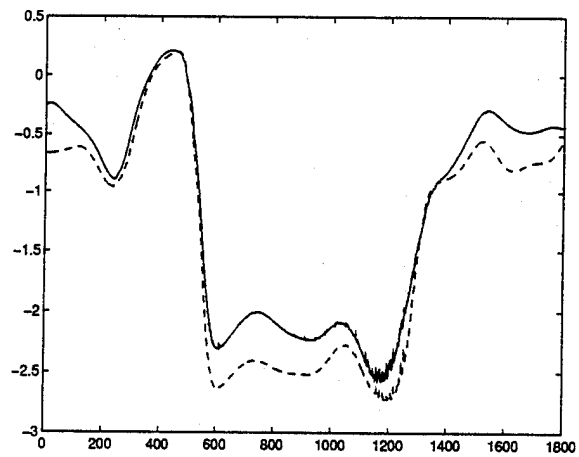

Figure 2. Solid line: the LPC-14 spectral tilt from Figure 1(b). Dashed line: LPC-2 spectral tilt for comparison.

[1] Y. Qi and B. Winberg, "Spectral slope of vowels produced by tracheoesophageal speakers." *Journal of Speech and Hearing Research*, 34:243, 1991.

[2] R. Trawinski, J. Domyslawska, and R. Ciurylo, "Method of accurately measuring the spectral tilt in the Ebert spectrograph." *Applied Optics*, 32:4828, 1993.

[3] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, Numerical Recipes: the art of scientific computing (FORTRAN version). Cambridge University Press, Cambridge, 1990, ISDN 0-521-38330-7.

[4] J. Deller, Jr., J. Proakis, and J. Hansen, Discrete-Time Processing of Speech Signals. Macmillan Publishing Co., New York, 1993.

[5] K. Assaleh, "Supplementary orthogonal cepstral features." *Proc. IEEE Int. Conf. Acoust, Speech and Signal Process. 1995 (Detroit)*, 1:413.

[6] M. Schroeder, "Direct (nonrecursive) relations between cepstrum and predictor coefficients." *Trans. Acoust., Speech and Signal Process.*, 29:297, April 1981.