All-Pole Modeling Technique Based on Weighted Sum of LSP Polynomials

Tom Bäckström, *Member, IEEE*, and Paavo Alku

***Abstract*—This study presents a new technique called weighted-sum line spectrum pair (WLSP) where an all-pole filter is defined by using a sum of weighted line spectrum pair polynomials. WLSP yields a stable all-pole filter of order m, whose autocorrelation function coincides with that of the input signal between indices 0 and $m-1$. By sacrificing the exact matching at index m, WLSP models the autocorrelation of the input signal at the indices above m more accurately than conventional linear prediction (LP). Experiments with vowels show that, in comparison to the conventional LP, WLSP yields all-pole spectra that model formants with an increased dynamic range between formant peaks and spectral valleys.**

***Index Terms*—All-pole modeling, line spectrum pair, linear prediction.**

I. Introduction

LINEAR PREDICTION (LP) is among the most widely used methods of speech processing. It has an established role, epecially in speech compression, and it is used in various low bit rate speech coders, e.g., [1]. Even though different variations of LP have been developed, the most widely used is the autocorrelation method of linear prediction. In this method, an optimal predictor (a finite-impulse response of order m) is determined by minimizing the square of the prediction error, the residual, over an infinite time interval [2].

The popularity of the conventional autocorrelation method of LP is based largely on the fact that it yields a stable all-pole model for the speech spectrum, which is accurate enough for most applications when presented by few parameters. The ability of LP to model the speech spectrum is explained by the autocorrelation function of the all-pole filter, which matches exactly the autocorrelation of the input signal between indices 0 and m, when the prediction order equals m [3]. It should be noted, however, that in determining the conventional LP predictor of order m, no information is used from the autocorrelation function of the input signal beyond the time index m.

The line spectrum pair (LSP) decomposition is one of the methods developed for presenting the LP information [4]. In this technique, the predictor computed by linear prediction is split into a symmetric and an antisymmetric polynomial. It has been proved that the roots of these two polynomials, the LSPs, are located on the unit circle and interlaced, if the original LP predictor is minimum phase [5]. Furthermore, it has been shown that LSPs behave well when interpolated [6]. Due to these properties, the LSP decomposition is widely used currently in quantization of LP information, e.g., [1].

In this study, we present a new linear predictive algorithm called weighted-sum line spectrum pair (WLSP), which yields an all-pole filter of order m to model the speech spectrum. In contrast to the conventional autocorrelation method of LP, the proposed algorithm also takes advantage of the autocorrelation of the input signal beyond time index m in order to obtain a more accurate all-pole model for the speech spectrum. WLSP exploits the LSP decomposition in a manner different from that typically used in speech coding; the LSP decomposition is not computed in order to quantize the LP information but rather as a computational tool, which defines stable all-pole filters with the proposed autocorrelation matching property.

Manuscript received December 3, 2001; revised October 10, 2002. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. P. C. Ching.

The authors are with the Helsinki University of Technology (HUT), Laboratory of Acoustics and Audio Signal Processing, FIN-02015 HUT, Espoo, Finland (e-mail:tom.backstrom@hut.fi).

Digital Object Identifier 10.1109/LSP.2003.811635

II. Background

The conventional LP predictor of order m for a signal $x(n)$ as given by [2] is

$$A_m(z) = 1 + \sum_{i=1}^{m} a_i z^{-i}. \quad (1)$$

The coefficient vector $\mathbf{a} = \langle a_i \rangle_{i=0,\dots,m}$, where $a_0 = 1$, can be solved from the normal equations $\mathbf{Ra} = \sigma^2[1, 0, \dots, 0]^T$. The autocorrelation matrix $\mathbf{R}$ is defined as the expected value of correlation $\mathbf{xx}^T$, i.e., $\mathbf{R} = E[\mathbf{xx}^T]$, where signal $\mathbf{x}$ is assumed to be wide-sense stationary and the residual energy is $\sigma^2 = \mathbf{a}^T\mathbf{Ra}$.

The symmetric and antisymmetric LSP polynomials are defined as $P(z) = A_m(z) + z^{-m-1}A_m(z^{-1})$ and $Q(z) = A_m(z) - z^{-m-1}A_m(z^{-1})$, respectively [5]. By defining a zero-extended vector $\hat{\mathbf{a}} = [\mathbf{a}^T, 0]^T$ (where $\mathbf{a}$ is the LP coefficient vector), the LSP polynomials can be defined equivalently in matrix form as

$$\mathbf{p} = \hat{\mathbf{a}} + \hat{\mathbf{a}}_R \text{ and } \mathbf{q} = \hat{\mathbf{a}} - \hat{\mathbf{a}}_R \quad (2)$$

where subscript R denotes reversal of rows.

III. Methods

A. Weighted Sum of the LSP Polynomials

The proposed method is based on predictor polynomial $D(z, \lambda)$ defined as

$$D(z, \lambda) = \lambda P(z) + (1 - \lambda)Q(z) \quad (3)$$

i.e., $D(z,\lambda)$ is a weighted sum, with weighting parameter λ, of the LSP polynomials $P(z)$ and $Q(z)$ in the polynomial space. It is worth noting that with $\lambda = 1/2$ (3) becomes $D(z,1/2) = 1/2[P(z)+Q(z)] = A_m(z)$. This is the equation according to which the LP predictor is reconstructed from the LSP polynomials [5].

Mathematical properties of $D(z,\lambda)$ are described in this section by focusing on three issues: the energy of the residual, autocorrelation of $D^{-1}(z,\lambda)$ and root tracks of $D(z,\lambda)$.

B. Residual of $D(z,\lambda)$

For the zero extended LP coefficient vector $\hat{\mathbf{a}}$ we have $\mathbf{R}\hat{\mathbf{a}} = [\sigma^2, 0, \ldots, 0, \gamma]^T$, where $\gamma = \sum_{i=0}^{m} a_i R(m-i+1)$. The symmetric LSP polynomial thus yields

$$\mathbf{Rp} = \mathbf{R}(\hat{\mathbf{a}} + \hat{\mathbf{a}}_R) = \begin{bmatrix} \sigma^2+\gamma \\ 0 \\ \vdots \\ 0 \\ \sigma^2+\gamma \end{bmatrix} \tag{4}$$

and the antisymmetric LSP polynomial yields similarly

$$\mathbf{Rq} = \mathbf{R}(\hat{\mathbf{a}} - \hat{\mathbf{a}}_R) = \left[\sigma^2-\gamma, 0, \ldots, 0, \gamma-\sigma^2\right]^T. \tag{5}$$

Polynomial $D(z,\lambda)$ can be written in vector form as $\mathbf{d}(\lambda) = \lambda\mathbf{p} + (1-\lambda)\mathbf{q}$ and we obtain

$$\mathbf{Rd}(\lambda) = \mathbf{R}\left[\lambda\mathbf{p} + (1-\lambda)\mathbf{q}\right] = \begin{bmatrix} \sigma^2+(2\lambda-1)\gamma \\ 0 \\ \vdots \\ 0 \\ (2\lambda-1)\sigma^2+\gamma \end{bmatrix}. \tag{6}$$

From this equation, we can write the following expression for the residual energy:

$$\begin{aligned} E[e^2(n)] &= E\left[\mathbf{d}^T(\lambda)\mathbf{x}\mathbf{x}^T\mathbf{d}(\lambda)\right] = \mathbf{d}(\lambda)^T\mathbf{R}\mathbf{d}(\lambda) \\ &= \sigma^2 + 2(2\lambda-1)\gamma + (2\lambda-1)^2\sigma^2. \end{aligned} \tag{7}$$

Vector $\mathbf{d}(\lambda)$ corresponds to the LP predictor of order $m+1$, if λ is chosen such that the last row of right hand side vector in (6) becomes equal to zero, i.e., $\lambda = -\gamma/2\sigma^2 + 1/2$. This is, in fact, one iteration step of the Levinson–Durbin recursion [7], which we will call the Levinson–Durbin solution.

Incrementing the predictor order in LP decreases the energy of the residual. Thus, the residual energy for the model order m should be greater than or equal to the residual energy for the model order $m+1$ [(7)], i.e., $|\sigma^2| \geq |\sigma^2 + 2(2\lambda-1)\gamma + (2\lambda-1)^2\sigma^2|$. Substituting the Levinson–Durbin solution for γ yields $1 \geq 4|(1-\lambda)\lambda|$, which holds true when λ is chosen such that $\lambda \in [1-\sqrt{2}/2), 1+\sqrt{2}/2]$. That is, if λ is chosen within these limits, the residual energy given by $D(z,\lambda)$ (of order $m+1$) is smaller than the residual energy given by the LP predictor of order m.

C. Autocorrelation of $D^{-1}(z,\lambda)$

Given the autocorrelation $R(i)$ $(0 \leq i \leq m)$ of an input signal, each possible value of $R(m+1)$ (possible values are such that matrix $\mathbf{R}$ is invertible) has a corresponding value of λ. That is, if λ is chosen according to the Levinson–Durbin recursion and γ is calculated with a given value of $R(m+1)$, then $D(z,\lambda)$ is the LP predictor of order $m+1$ and all autocorrelation values will be exactly fitted up to $m+1$. If a different value is chosen for λ, the model then matches some other autocorrelation $\widehat{R}(i)$, $0 \leq i \leq m+1$, where $R(i) = \widehat{R}(i)$ for $0 \leq i \leq m$. In other words, $D^{-1}(z,\lambda)$ always matches autocorrelations of the input signal in range $0 \leq i \leq m$.

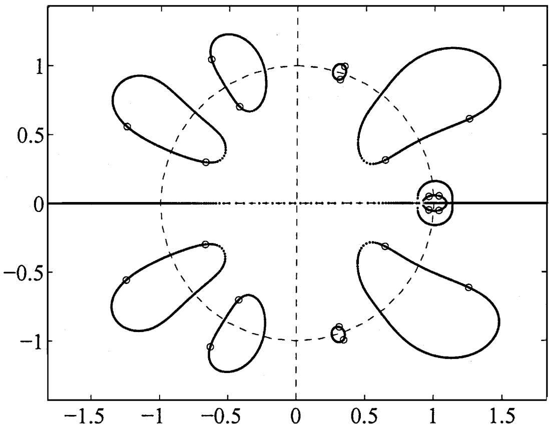

Fig. 1. Root tracks of $D(z,\lambda)$ as a function of $\lambda \in \mathbb{R}$, an example with $m = 10$. Small circles inside the unit circle correspond to the roots of the LP polynomial $A_m(z)(\lambda = 1/2)$, and small circles outside the unit circle correspond to their mirror image partners (i.e., roots of $A_m(z^{-1})$, $\lambda = \pm\infty$). The LSP polynomials roots are at the intersections of the unit circle and the $D(z,\lambda)$ root tracks ($\lambda = 0$ or $\lambda = 1$).

D. Root Tracks of $D(z,\lambda)$

Polynomial $D(z,\lambda)$ is a linear combination, and thus a continuous transformation in the orthogonal space spanned by $P(z)$ and $Q(z)$. We can readily prove that the roots of polynomial $D(z,\lambda)$ follow continuous tracks in the Z-plane as functions of λ. An example with $m = 10$ of the polynomial root space of $D(z,\lambda)$ as a function of $\lambda \in \mathbb{R}$ is shown in Fig. 1.

Apart from being continuous, the root tracks of $D(z,\lambda)$ are also closed paths, i.e., their ending points at $\lambda = \pm\infty$ coincide. In fact, when λ goes to infinity (either positive or negative), then roots of $D(z,\lambda)$ will become roots of $A_m(z^{-1})$ (i.e., the mirror image partners of the roots of $A_m(z)$), i.e., $\lim_{\lambda\to\pm\infty} D(z,\lambda)/(1+\lambda z^{-1}) = 2z^{-m}A_m(z^{-1})$.

Since the zeros of the LSP polynomials $P(z)$ and $Q(z)$ are interlaced on the unit circle, their sum $A(z) = (1/2)[P(z) + Q(z)]$ will be minimum-phase [5]. Scaling of $P(z)$ and $Q(z)$ with positive coefficients does not alter their zeros and, consequently, polynomial $D(z,\lambda)$ will also be minimum-phase in the open interval $\lambda \in (0,1)$. Furthermore, setting $\lambda = 1/2$ yields the original LP polynomial $A_{m-1}(z)$ of order $m-1$ and $\lambda = -(\gamma/2\sigma^2) + (1/2)$ yields the LP polynomial $A_m(z)$ of order m.

IV. Algorithm

The weighted sum of the LSP polynomials (3), together with the autocorrelation properties of $D^{-1}(z,\lambda)$ (described in Section III-C), serves as a basis for the proposed WLSP filters. Given an input signal $x(n)$, defining a stable WLSP all-pole filter of order m comprises the following stages. (The time index of the autocorrelation function is denoted by i.)

1) Calculate LP polynomial $A_{m-1}(z)$ for $x(n)$ using conventional linear prediction with the autocorrelation criterion. Notice that LP analysis of order m yields LSP polynomials of order $m+1$ [5]. Therefore, in order to generate an mth order WLSP filter, the orders of $P(z)$ and $Q(z)$ in (3) have to be equal to m, which corresponds to defining an LP predictor of order $m-1$.
2) Construct the LSP polynomials $P(z)$ and $Q(z)$ from $A_{m-1}(z)$ as defined in [5].
3) Define, using (3), an all-pole filter $D^{-1}(z,\lambda)$ of order m with an additional parameter λ. Optimize the all-pole filter by searching for the value of λ that minimizes the absolute error between the (normalized) autocorrelations of $x(n)$ and $D^{-1}(z,\lambda)$ for $m \leq i \leq m+L$. Notice from Section III-C that autocorrelations of $x(n)$ and $D^{-1}(z,\lambda)$ are equal for $0 \leq i < z$, when the order of $D^{-1}(z,\lambda)$ equals m.

Note that in Step 3), parameter λ has to be in the open interval $\lambda \in (0,1)$ to guarantee stability. Further, in our current experiments, parameter λ was optimized iteratively with a straightforward brute-force approach. Refinement of the optimization technique is left for further studies.

V. RESULTS

The proposed all-pole modeling technique was compared to conventional LP analysis by analyzing the spectra of seven Finnish vowels. The speech sounds were produced by six speakers (two female, four male). The two linear predictive analyses were computed by using a predictor order $m = 10$, a 20-ms Hamming window, autocorrelation window of length $L = 20$ (samples) and a sampling frequency of 8 kHz.

The comparison of the behaviors of LP and WLSP in modeling of formants was done by using the following procedure. Given an all-pole spectrum (in decibels), the formant peak was defined as a local maximum of the spectrum. The spectral valley was then defined as a local minimum of the spectrum following this peak. The level difference of these two spectral components, denoted by L_{diff}, was computed to characterize the dynamics of the all-pole spectrum in the vicinity of the corresponding formant. Finally, the L_{diff} given by LP was subtracted from the L_{diff} yielded by WLSP for all the formants extracted. This difference $\Delta L = L_{\text{diff,WLSP}} - L_{\text{diff,LP}}$ is positive when WLSP models the formant with larger dynamics than LP. The value of ΔL is shown in Fig. 2 averaged over all extracted formants and all subjects. It should be noted that ΔL is positive for all vowels. Fig. 3 shows the fast Fourier transform spectrum of a male vowel /a/ together with the all-pole spectra given by LP and WLSP.

To be able to compute all-pole models of speech with increased spectral dynamics between formant peaks and valleys (i.e., increasing ΔL) is beneficial, for example, in speech coding. Namely, using all-pole filters with large spectral dynamics results in improved quality of the decoded speech, because the noise level that originates from quantization of the residual is attenuated in formant valleys.

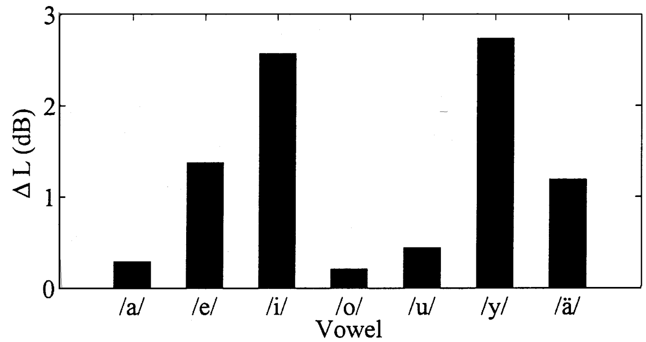

Fig. 2. Level difference (L_{diff}) was computed between a formant peak and the following spectral valley. Difference of this measure computed for LP and WLSP ($\Delta L = L_{\text{diff,WLSP}} - L_{\text{diff,LP}}$) with $m = 10$ and $L = 20$ is shown for seven Finnish vowels.

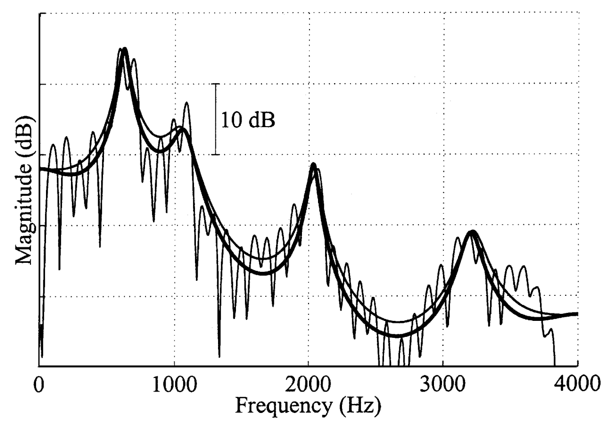

Fig. 3. Fast Fourier transform spectrum of a male vowel /a/ together with all-pole spectra $(m = 10)$ of LP (thin line) and WLSP ($L = 20$, thick line).

The two all-pole modeling techniques were also compared in terms of the normalized residual energies, i.e., the energy of the residual divided by the energy of the input signal. The results showed that the conventional LP analysis yielded, on average, 0.17 dB smaller values of the normalized residual energy averaged over all subjects and vowels. This degradation of performance of WLSP compared to the conventional LP analysis in terms of the normalized residual energy is explained by the fact (see Section III-C) that the autocorrelation function of the all-pole filter given by the conventional LP matches precisely the mth autocorrelation term of the input signal, while this is not achieved by WLSP.

VI. CONCLUSION

In this letter, we have presented a new all-pole modeling technique called weighted-sum line spectrum pair (WLSP), based on the weighted sum of the LSP polynomials. With this technique, we define an all-pole filter of order m, whose autocorrelation exactly matches the $m-1$ first autocorrelation values of the input signal, but the exact match at index m is sacrificed in order to achieve improved matching in the upper autocorrelation range. A mathematical treatise demonstrates that the proposed method is well behaved, and stability of the all-pole filter is guaranteed.

Experiments show that WLSP yields, in comparison to conventional LP, spectral models with larger level differences between formant peaks and spectral valleys. Consequently, the presented methods could be applied in coding and enhancement of speech, by suppressing quantization noise in spectral valleys.

REFERENCES

[1] ETSI, *Digital Cellular Telecommunications System (Phase 2); Enhanced Full Rate (Efr) Speech Transcoding, Gsm 06.60 Version 4.0.0*: ETSI, Aug. 1997.

[2] J. Makhoul, "Linear prediction: a tutorial review," *Proc. IEEE*, vol. 63, pp. 561–580, Apr. 1975.

[3] M. R. Schroeder, "Linear prediction, extremal entropy and prior information in speech signal analysis and synthesis," *Speech Commun.*, vol. 1, no. 1, pp. 9–20, 1982.

[4] F. Itakura, "Line spectrum representation of linear predictive coefficients of speech signals," *J. Acoust. Soc. Amer.*, vol. 57, no. 1, p. 35, 1975.

[5] F. K. Soong and B.-H. Juang, "Line spectrum pair (LSP) and speech data compression," in *Proc. ICASSP*, vol. 1, San Diego, CA, Mar. 1984, pp. 1.10.1–1.10.4.

[6] K. K. Paliwal and W. B. Kleijn, "Quantization of LPC parameters," in *Speech Coding and Synthesis*, W. B. Kleijn and K. K. Paliwal, Eds. Elsevier, Amsterdam, The Netherlands, 1995, pp. 433–466.

[7] P. Delsarte and Y. V. Genin, "The split Levinson algorithm," *IEEE Trans. Acoust. Speech Signal Processing*, vol. ASSP-34, pp. 470–478, June 1986.