Direct Line Spectral Frequency Adaptation in Second Order Cascade Sections

A. A. (Louis) Beex[1] and Gaguk Zakaria[1, 2]
[1]Systems Group – DSP Research Laboratory
The Bradley Department of Electrical and Computer Engineering,
VIRGINIA TECH, Blacksburg, VA 24061-0111 USA
[2]Hughes Network Systems, Germantown, MD 20876 USA

Abstract

We propose the adaptive cascade recursive least squares with subsection adaptation (CRLS-SA) algorithm for direct estimation of line spectral frequencies (LSF) for use in speech coding. The least squares based CRLS-SA algorithm has low computational complexity since each section is adapted independently from other sections. Using CRLS-SA, the LSF can be obtained directly from the coefficients of each section of the CRLS-SA structure. Two CRLS-SA approaches can be used for estimating LSF: use CRLS-SA as an AR inverse filter and obtain the LSF by converting the coefficients of each second order filtes using a look-up table or, alternatively, construct a cascade of second order LSF polynomials P(z) and Q(z), and adapt each using CRLS-SA. The results show that the LSF estimates from CRLS-SA are competitive with the results obtained by the popular autocorrelation method followed by approximation using Chebyshev polynomials.

1. Introduction

In CELP based speech coding, the linear prediction (LP) coefficients are usually converted into *line spectral pairs* (LSP) [5], or *line spectral frequencies* (LSF). The LSF are closely related to the formant frequencies, so that the LSF encode LP speech spectra more efficiently than LP coefficients [7]. The LSF are usually approximated using Chebyshev polynomials.

Using a cascade of second order adaptive filters to estimate the LSF is straightforward because the LSF can be obtained directly from the coefficients of each section with very little additional computation. There have been some attempts to compute the LSF using a cascade adaptive filter, such as using an LMS adaptive filter [1, 2, 11], or a recursive least squares adaptive filter [8]. However, LMS is very slow to adapt, especially for speech signals, and recursive least squares (RLS) is computationally very expensive. Also, the use of unstable gradient filters [2], extensive filtering to find gradients [11], or non-trivial extensions to orders higher than four [1], are less desirable features.

CRLS-SA is a least-squares based cascade adaptive filter, where each section is adapted independently [12]. Since each section is a second order filter, the computational effort required by CRLS-SA is approximately 8*4*N/2, where N is the order of the filter. Also, for short data records, CRLS-SA gives a result closer to the known signal compared to the widely used autocorrelation method [12].

We propose two methods to use CRLS-SA for estimating LSF. First, we obtain LSF by directly converting the LP coefficients for each section of the cascade. Second, when each section of the cascade structure uses the symmetric and anti-symmetric polynomials, $P(z)$ and $Q(z)$ respectively, the LSF coefficients of the latter can be adapted directly.

2. Review of CRLS-SA

The cascade recursive least squares with subsection adaptation (CRLS-SA) algorithm represents a least squares based cascade adaptive filter. The weighted-least-squares cost function to be minimized is:

$$\varepsilon_n = \sum_{i=1}^{n} \lambda^{n-i} e_n^2 \tag{1}$$

where

$$e_n = \left\{ \prod_{k=1}^{M} A_k(z) \right\} y_n \tag{2}$$

is the output of the LP filter of order N, and

$$A_k(z) = 1 - a_{n,k,1} z^{-1} - a_{n,k,2} z^{-2} \tag{3}$$

is a second order LP filter, y_n is the input signal, and n indicates the time index. The CRLS-SA is shown in

0-7803-5700-0/99/$10.00©1999 IEEE

Figure 1, where each section is represented by (3), and M is $N/2$.

In order to adapt the coefficients of each section, we need to compute the gradients, defined as:

$$\varphi_{n,k,i} = \frac{e_n}{a_{n,k,i}} \quad ; \text{i=1,2.} \tag{4}$$

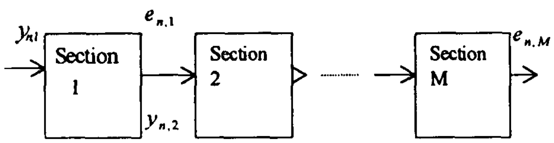

Fig. 1: Cascade Adaptive Filter.

It can be shown that $\varphi_{n,k,1}$ can be computed efficiently by passing the output e_{n-1} to the filter $A_k^{-1}(z)$ [12], i.e.

$$\varphi_{n,k,1} = z^{-1}\frac{e_n}{A_k(z)} \tag{5a}$$

$$\varphi'_{n,k,2} = \varphi_{n-1,k,1} \tag{5b}$$

Since each section represents a different pole pair, the gradients of different sections generally occupy different frequency bands. Hence, the gradients for different sections are pretty much decoupled and each section can be adapted independently. The CRLS-SA algorithm is summarized in the Appendix.

3. Line Spectral Frequencies For A Cascade Of Second Order Filters

A linear prediction (LP) filter of order N can be denoted as:

$$A(z) = 1 - \sum_{i=1}^{N} a_i z^{-i} \tag{6}$$

where the a_i's are the LP coefficients. $A(z)$ can be written as a cascade of several second order LP filters:

$$A(z) = \prod_{k=1}^{N/2} A_k(z) \tag{7a}$$

where

$$A_k(z) = 1 - a_{k,1}z^{-1} - a_{k,2}z^{-2} \tag{7b}$$

From $A_k(z)$, we can form the respectively symmetric and anti-symmetric second order filters, $P_k(z)$ and $Q_k(z)$:

$$\begin{aligned} P_k(z) &= A_k(z) + z^{-3}A_k\left(z^{-1}\right) \\ &= 1 - \left(a_{k,1} + a_{k,2}\right)z^{-1} - \left(a_{k,1} + a_{k,2}\right)z^{-2} + z^{-3} \end{aligned} \tag{8}$$

$$\begin{aligned} Q_k(z) &= A_k(z) - z^{-3}A_k\left(z^{-1}\right) \\ &= 1 - \left(a_{k,1} - a_{k,2}\right)z^{-1} - \left(a_{k,2} - a_{k,1}\right)z^{-2} - z^{-3} \end{aligned} \tag{9}$$

$P_k(z)$ and $Q_k(z)$ have one root at $z=-1$ and $z=1$ respectively, so that $P_k(z)$ and $Q_k(z)$ can be written as:

$$P_k(z) = \left(1 + z^{-1}\right)\widetilde{P}_k(z) \tag{10a}$$

$$Q_k(z) = \left(1 - z^{-1}\right)\widetilde{Q}_k(z) \tag{10b}$$

where

$$\widetilde{P}_k(z) = 1 + p_k z^{-1} + z^{-2} \tag{10c}$$

$$\widetilde{Q}_k(z) = 1 + q_k z^{-1} + z^{-2} \tag{10d}$$

Factoring $(1+z^{-1})$ out of $P_k(z)$ and $(1-z^{-1})$ out of $Q_k(z)$, we can show that

$$p_k = -\left(1 + a_{k,1} + a_{k,2}\right) \tag{11a}$$

$$q_k = \left(1 - a_{k,1} + a_{k,2}\right) \tag{11b}$$

We know that a second order filter with complex conjugate roots can be represented by

$$B(z) = 1 - 2r\cos\theta z^{-1} + r^2 z^{-2} \tag{12}$$

where θ relates to the frequency and r to the radius of the root pair. If $r=1$, $B(z)$ has its roots on the unit circle. So, from p_k and q_k we can get the frequencies of $\widetilde{P}_k(z)$ and $\widetilde{Q}_k(z)$ as

$$\theta_{p,k} = \cos^{-1}\left(-p_k/2\right) \tag{13a}$$

$$\theta_{q,k} = \cos^{-1}\left(-q_k/2\right) \tag{13b}$$

By using table look-up, the LSF $\theta_{p,k}$ and $\theta_{q,k}$ can be obtained directly from p_k and q_k respectively without additional computational load. No Chebychev polynomial approximation is needed due to the second order section approach. For $A_k(z)$ to be minimum phase, the LSF must alternate on the unit circle, which for the second order sections implies

$$0 < \theta_{p,k} < \theta_{q,k} < \pi \; ; \; k = 1,\cdots,M \tag{13c}$$

where M equals $N/2$ [2].

4. Direct LSF Adaptation

The cascade second order symmetric and anti-symmetric filter is as shown in Figure 1, where each section has the structure indicated in Figure 2.

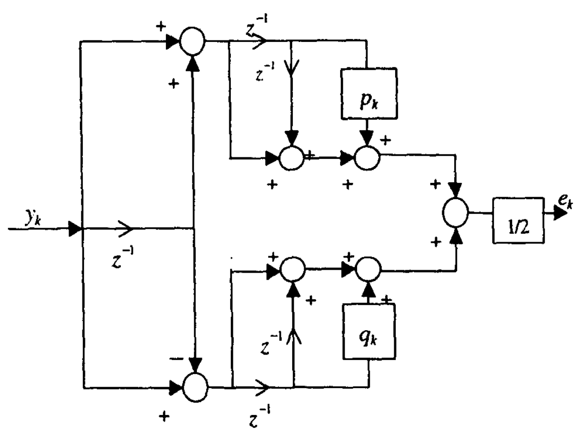

Fig. 2: Direct LSF Adaptation Section $A_k(z)$.

The structure in Figure 2 implements (10a-10d). We can thus represent $A(z)$ as follows:

$$A(z)=\prod_{k=1}^{M}\left\{\frac{1}{2}\left[\left(1+z^{-1}\right)\tilde{P}_k(z)+\left(1-z^{-1}\right)\tilde{Q}_k(z)\right]\right\} \quad (14)$$

We can show that the gradients $\varphi_{n,p,k}$ and $\varphi_{n,q,k}$ for adapting $\tilde{P}_k(z)$ and $\tilde{Q}_k(z)$ respectively are

$$\varphi_{n,p,k}=-\left(1+z^{-1}\right)z^{-1}\frac{e_{n,M}}{2A_k(z)} \quad (15a)$$

$$\varphi_{n,q,k}=-\left(1-z^{-1}\right)z^{-1}\frac{e_{n,M}}{2A_k(z)} \quad (15b)$$

We can use the CRLS-SA algorithm in the Appendix by using (15a) and (15b) instead of (A1a) and (A1b) respectively, and defining:

$$\underline{\varphi}_{n,k}=[\varphi_{n,p,k} \quad \varphi_{n,q,k}] \quad (15a)$$

instead of as in (A1c), and using

$$\underline{\hat{a}}_{n,k}=\left[p_{n,k} \quad q_{n,k}\right] \quad (15b)$$

instead of (A5).

Furthermore, the second order LSF parameters can be used directly for synthesis as follows.

$$Y_k(z)=\frac{E_k(z)}{A_k(z)}$$

$$=\frac{E_k(z)}{\frac{1}{2}\{P_k(z)+Q_k(z)\}}$$

$$=\frac{2E_k(z)}{P_k(z)+Q_k(z)}$$

$$=\frac{2E_k(z)}{(1+z^{-1})\tilde{P}_k(z)+(1-z^{-1})\tilde{Q}_k(z)} \quad (16)$$

Multiplying out we get:

$$\begin{aligned}2E_k(z)&=Y_k(z)\left[\left(1+z^{-1}\right)\tilde{P}_k(z)+\left(1-z^{-1}\right)\tilde{Q}_k(z)\right]\\&=Y_k(z)\left\{\tilde{P}_k(z)+\tilde{Q}_k(z)\right\}+z^{-1}Y_k\left\{\tilde{P}_k(z)-\tilde{Q}_k(z)\right\}\end{aligned} \quad (17)$$

Replacing $\tilde{P}_k(z)$ using (10c) and $\tilde{Q}_k(z)$ using (10d) yields

$$Y_k(z)\left\{2+(p_k+q_k)z^{-1}+2z^{-2}\right\}=2E_k(z)-z^{-2}Y_k(p_k-q_k)$$

An inverse z-transform then yields the desired relation:

$$y_{n,k}=\frac{1}{2}\left\{2e_{n,k}-y_{n-1,k}(p_k+q_k)-y_{n-2,k}(2+p_k-q_k)\right\} \quad (18)$$

which is reflected in the diagram of Figure 3.

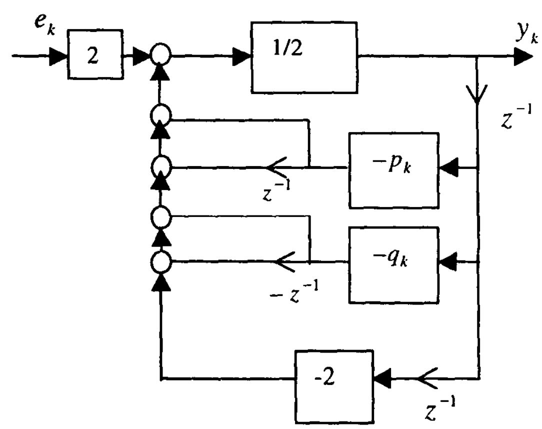

Fig. 3: Synthesis Using Cascade LSF.

5. Performance Results

To evaluate the performance of LSF estimation using CRLS-SA, 100 speech-like signal realizations are generated. The spectrum of the underlying tenth order AR process is shown in Figure 4.

The performance of the LSF estimator is evaluated by measuring the Itakura distance [3] between the estimated

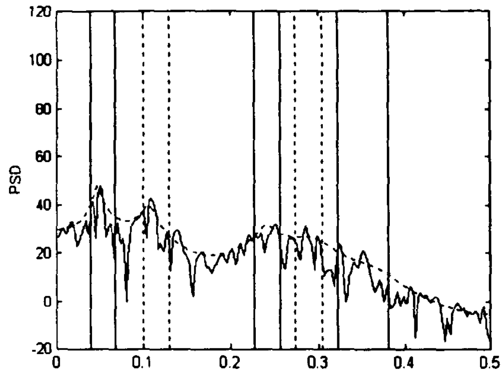

Fig. 4: LSF and Spectra of Speech and the Underlying AR Process.

AR process, obtained by converting the LSF back to an AR process, and the actual AR process. The length of the data records is 240 samples. For comparison, we also compute the LSF using the autocorrelation method, where from the autocorrelation estimates the polynomials $P(z)$ and $Q(z)$ are formed. Next the roots of $P(z)$ and $Q(z)$ are computed.

Typical results for LSF generated with the Direct LSF CRLS-SA algorithm of this paper are shown in Table 1.

Table 1. Statistical Performance Comparison

Method		Direct LSF	Autocorrelation
LSF root	Actual	Mean/Std.Dev.	Mean/ Std.Dev.
P1	0.0445	0.0394/0.0067	0.0550/0.0110
Q1	0.0671	0.0679/0.0089	0.0685/0.0127
P2	0.1026	0.0991/0.0052	0.1050/0.0079
Q2	0.1241	0.1296/0.0072	0.1253/0.0058
P3	0.2276	0.2273/0.0058	0.2299/0.0069
Q3	0.2505	0.2569/0.0061	0.2496/0.0099
P4	0.2711	0.2737/0.0055	0.2856/0.0049
Q4	0.3088	0.3038/0.0049	0.2860/0.0169
P5	0.3219	0.3219/0.0080	0.3350/0.0217
Q5	0.3763	0.3807/0.0065	0.3612/0.0288

From the results in Table 1, based on 100 realizations, we see that both methods produce consistent results in terms of mean and standard deviation. Note however that Direct LSF CRLS-SA yields less bias and less variance in terms of the corresponding LSF root estimates. Recall too that for the autocorrelation method, the roots of $P(z)$ and $Q(z)$ are obtained using Chebyshev approximation [5, 7]. By contrast, for Direct LSF CRLS-SA the line spectral frequencies themselves have been adapted so that these can be transmitted without further computation.

The histograms of the Itakura distance, associated with the AR process estimated by CRLS-SA and with the autocorrelation method, are shown in Figures 5 and 6 respectively.

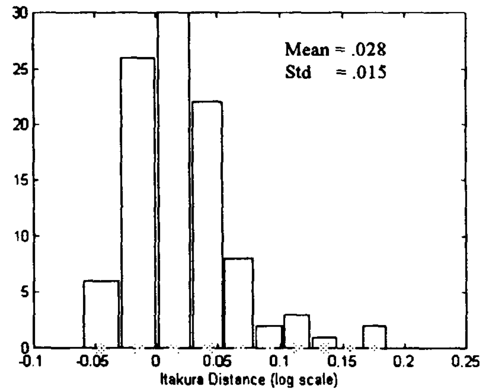

Fig. 5: Histogram of the Itakura Distance Produced by CRLS-SA.

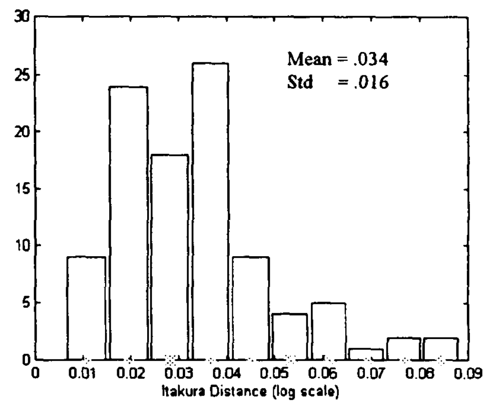

Fig. 6: Histogram of the Itakura Distance Produced by Autocorrelation Method.

We see in Figures 5 and 6 that the mean of the Itakura distance for the AR process obtained by CRLS-SA is smaller than that obtained from the autocorrelation method. Also, the histogram in Figure 5 shows negative and positive values indicating the possibility of unbiased estimation. In Figure 6 on the other hand, the Itakura distance is exclusively positive, indicative of the biased estimates produced by the autocorrelation method. Note that also the variance of the Itakura distance for the AR process obtained by CRLS-SA is smaller than that obtained from the autocorrelation method. A possible explanation for the better CRLS-SA estimator is that, for our relatively short data record, the autocorrelation method has not yet been able to estimate the underlying AR process as well as CRLS-SA [12].

6. Conclusion

The CRLS-SA algorithm was proposed for directly finding the line spectral frequencies (LSF) for the cascaded second-order-sections corresponding to a higher order AR process. This turns out to be a good alternative to using the popular autocorrelation method, which entails additional computation for conversion to LSF. Two different approaches can be used. The first is to use CRLS-SA as a cascade of second order AR inverse filters and to obtain the LSF by simply converting the coefficients of each section using a look-up table. The second approach directly adapts the LSF coefficients by constructing a cascade of second order section structures based on the LSF polynomials $P(z)$ and $Q(z)$. Simulation with, spectrally, speech-like signals shows that our direct adaptation method produces results that are statistically competitive with those from the popular autocorrelation approach.

7. References

[1] B. M. G. Cheetham, *Adaptive LSP Filter*, Electronics Letters, vol. 23, no. 2, pp. 89-90, January 1987.

[2] P. C. Ching and K.C. Ho, *An Efficient Adaptive LSP Method for Speech Encoding*, IEEE Region 10 Conference on Computer and Communication Systems, Hongkong, September 1990.

[3] J. R. Deller Jr., J. G. Proakis, J. H. L. Hansen, Discrete-time Processing of Speech Signals, New York, Macmillan, 1993.

[4] S. Haykin, Adaptive Filter Theory, Englewood Cliffs, Prentice Hall, 1991.

[5] ITU-T, Coding of Speech at 8 Kbps Using Conjugate-Structure Algebraic-Code-Excited Linear-Predictive (CS-ACELP) Coding, Draft Recommendation G.729, June 1995.

[6] L. B. Jackson and S. L. Wood, *Linear Prediction in cascade form*, *IEEE Transactions on ASSP*, vol. ASSP-26, pp. 519-528, Dec. 1978.

[7] P. Kabal and R. P. Ramachandran, *The Computation of Line Spectral Frequencies Using Chebyshev Polynomials*, IEEE Transactions on ASSP, vol.ASSP-34, no.6, December 1986.

[8] J.M.T. Romano, M. Belanger, L.C. Coradine, Least Squares Adaptive Filter in Cascade Form for Line Pair Spectrum Modeling, Signal Processing V: Theories and Applications, L.Toores, E. Masgrau, and M.A. Lagunas (eds.), Elsevier Science Publishers B.V., 1990.

[9] M. R. Schroeder and B. S. Atal, *Code-excited Linear Prediction (CELP): High-quality speech at very low bit rates*, ICASP, pp. 25.1.1-25.1.4, 1985.

[10] K. Swaminathan et. al, *A Robust Low Rate Voice Coding for Wireless Communications*, IEEE Workshop on Speech Coding, pp. 75-76, September 1997.

[11] J. Taihui, Z. Youwei, and G. Junying, *A New Algorithm for Adaptive Line Spectral Pair Filter and Its Convergence Behavior*, ICCT'96, p.2, vol. 1151, pp. 436-40, Beijing PRC, 5-7 May 1996.

[12] G. Zakaria and A. A. (Louis) Beex, *Cascade Recursive Least Squares with Subsection Adaptation for AR parameter Estimation*, ICASSP 98, Seattle, pp. 953-956, 12-15 May 1998.

APPENDIX

$$\varphi_{n,k,1} = \frac{e_{n,N/2}}{A_k(z)} \tag{A1a}$$

$$\varphi_{n,k,2} = \varphi_{n-1,k,1} \tag{A1b}$$

$$\underline{\varphi}_{n,k} = \begin{bmatrix}\varphi_{n,k,1} & \varphi_{n,k,2}\end{bmatrix} \tag{A1c}$$

$$\underline{\kappa}_{n,k} = \frac{P_{n-1,k}\,\underline{\varphi}_{n,k}}{\lambda + \underline{\varphi}_{n,k}^{H} P_{n-1,k}\,\underline{\varphi}_{n,k}} \tag{A2}$$

$$P_{n,k} = \lambda^{-1} P_{n-1,k} - \lambda^{-1}\underline{\kappa}_{n,k}\,\underline{\varphi}_{n,k}^{H} P_{n-1,k} \tag{A3}$$

$$e_{n,k} = y_{n,k+1} = d_{n,k} - \underline{\hat{a}}_{n-1,k}^{H}\,\underline{y}_{n,k} \tag{A4}$$

$$\underline{\hat{a}}_{n,k} = \underline{\hat{a}}_{n-1,k} + \underline{\kappa}_{n,k} e_{n,k} \tag{A5}$$