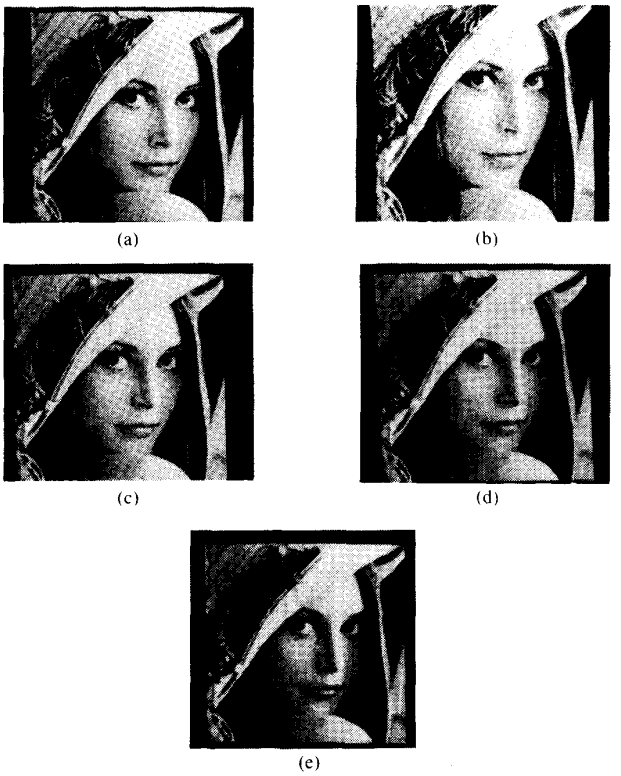
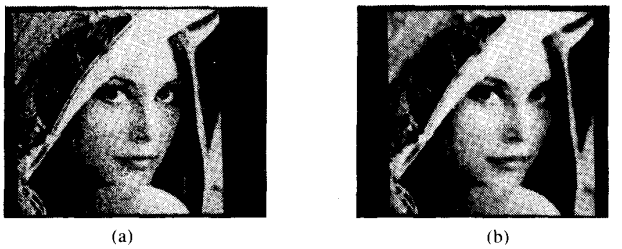


(a) (b) (c) (d) (e)

Fig. 3. (a) Original image F. (b) Original image degraded by additive noise with $\sigma_\epsilon^2 = 225$. (c) Degraded image filtered with LV_∞. (d) Degraded image filtered with LV_4. (e) Degraded image filtered with the median filter.



(a) (b)

Fig. 4. (a) Image F degraded with multiplicative noise with $\sigma_\epsilon = 0.3$. (b) Noise removed by LV_4.

RMSE for the LV_4 and the LV_∞ filtered images [Fig. 3(d) and (c)] are 8.21 and 9.92, respectively. The RMSE for the LV_6 (not shown) is 7.94.

VI. Conclusions

We have presented a new perspective into unifying many current image smoothing techniques by introducing the concepts of modeling error and confidence measures. We have shown that the median, the sample absolute gradient, and the sample variance may be used to obtain effective confidence measures. We have presented a family of filters based on the local sample variance which gives a tradeoff between noise removal and edge preservation.

References

[1] Y. H. Lee and S. A. Kassam, "Generalized median filtering and related nonlinear filtering techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 672-683, June 1985.

[2] A. C. Bovik, T. S. Huang, and D. C. Munson, Jr., "A generalization of median filtering using linear combinations of order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 1342-1350, Dec. 1983.

[3] J. S. Lee, "Digital image smoothing and the sigma filter," *Comput. Vision Image Processing*, vol. 24, pp. 255-269, 1983.

[4] C. Rey and R. K. Ward, "An adaptive minimum-variance predictor-smoother algorithm for image processing," in *Proc. 22nd Annu. Allerton Conf. Commun., Contr., Comput.*, Univ. Illinois, Urbana-Champaign, Oct. 1984, pp. 122-131.

[5] D. Wang, A. Vagnucci, and D. Li, "Image enhancement by gradient inverse smoothing scheme," *Comput. Graphics Image Processing*, vol. 15, pp. 167-187, 1981.

[6] M. Nagao and T. Matsuyama, "Edge preserving smoothing," *Comput. Graphics Image Processing*, vol. 9, pp. 394-407, 1979.

[7] J. S. Lee, "Digital image enhancement and noise filtering by use of local statistics," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-2, pp. 165-168, Mar. 1980.

[8] ——, "Refined filtering of image noise using local statistics," *Comput. Graphics Image Processing*, vol. 15, pp. 380-389, 1981.

[9] ——, "Speckle analysis and smoothing of synthetic aperture radar image," *Comput. Graphics Image Processing*, vol. 17, pp. 24-32, 1981.

[10] W. K. Pratt, *Digital Picture Processing*. New York: Wiley, 1978.

[11] A. Rosenfeld and A. C. Kak, *Digital Picture Processing, Vol. 1*, 2nd ed. New York: Academic, 1982.

[12] G. L. Anderson and A. N. Netravali, "Image restoration on a subjective criterion," *IEEE Trans. Syst., Man, Cybern.*, Dec. 1976.

[13] I. Pitas and A. N. Venetsanopoulos, "Nonlinear mean filters in image processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 573-584, June 1986.

[14] R. L. Stevenson and G. R. Arce, "Morphological filters: Statistics and further syntactical properties," *IEEE Trans. Circuits Syst.*, vol. CAS-34, Nov. 1987.

[15] A. C. Bovik, T. S. Huang, and D. C. Munson, "Edge-sensitive image restoration using order-constrained least squares methods," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 1253-1263, Oct. 1985.

[16] Y. H. Lee and A. T. Fam, "An edge gradient enhancing adaptive order statistic filter," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 680-695, May 1987.

[17] Y. H. Lee, Y. Lee, and A. T. Fam, "Successive partioning and selection filters," in *Proc. 24th Annu. Allerton Conf.*, Oct. 1986.

[18] P. D. Wendt, E. J. Coyle, and N. C. Gallagher, "Stack filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 898-891, Aug. 1986.

[19] G. R. Arce, N. C. Gallagher, and T. A. Nodes, "Median filter: Theory and applications," in *Advances in Computer Vision and Image Processing*, vol. 2, T. S. Huang, Ed. Greenwich, CT: JAI Press, 1986.

Interpolating Autoregressive Processes: A Bound on the Restoration Error

RAYMOND N. J. VELDHUIS

***Abstract*—This correspondence presents an upper bound for the restoration error variance of a sample restoration method for autoregressive processes which was presented in [1]. This bound is valid for the case that one sample is unknown in a realization of an autoregressive process of arbitrary, finite order.**

I. Introduction

In [1], a method was presented that can be used to restore groups of unknown samples in a realization of an autoregressive process.

Manuscript received April 11, 1988; revised December 1, 1988.

The author is with Philips Research Laboratories, 5600 JA Eindhoven, The Netherlands.

IEEE Log Number 8929360.

It was observed that the restoration results were better if the signal spectrum was peaky. In this correspondence, an upper bound for the restoration error variance is derived that is indeed lower if the autoregressive process has poles close to the unit circle of the complex plane. This situation corresponds to a peaky signal spectrum. The upper bound is derived for the case that one sample is unknown in a realization of an autoregressive process of arbitrary, finite order. The main results of [1] are briefly summarized in Section II and the bound is derived in Section III.

II. Sample Restoration

It is assumed that s_k, $k = 0, \cdots, N-1$ is a segment of realization of a stationary autoregressive process s_k, $k = -\infty, \cdots, +\infty$.[1] This means that there exist a finite positive integer p, called the *order of prediction*; real numbers $a_0, a_1, \cdots, a_p$, $a_0 = 1$, called the *prediction coefficients*; and a zero-mean white noise process e_k, $k = -\infty, \cdots, +\infty$, called the *excitation noise*, with variance σ_e^2, such that

$$\sum_{l=0}^{p} a_l s_{k-l} = e_k, \quad k = -\infty, \cdots, +\infty. \tag{1}$$

The spectrum $S(\theta)$ of the autoregressive process of (1) is given by

$$S(\theta) = \frac{\sigma_e^2}{\left|\sum_{l=0}^{p} a_l \exp(-j\theta l)\right|^2} = \frac{\sigma_e^2}{\sum_{l=-p}^{p} b_l \exp(-j\theta l)} \tag{2}$$

where

$$b_l = \sum_{k=0}^{p} a_k a_{k+l} \tag{3}$$

with the convention that $a_k = 0$ for $k > p$ and $k < 0$. An autoregressive process can be regarded as the output of an all-pole filter with a transfer function

$$\frac{1}{A(e^{j\theta})} = \frac{1}{\sum_{l=0}^{p} a_l \exp(-j\theta l)}$$

excited with zero-mean white noise with variance σ_e^2. The zeros of the polynomial

$$A(z) = \sum_{l=0}^{p} a_l z^{-l}, \qquad z \in \cdots \tag{4}$$

are the poles of the filter. They must be within the unit circle of the complex plane. Note that $-a_p$ is the product of the zeros and that the spectrum is more peaky if the zeros are close to the unit circle.

The samples $s_{t(i)}$, $i = 1, \cdots, m$ with $p \le t(1) < \cdots < t(m) \le N - p - 1$ are unknown. In [1], estimates for the unknown samples are obtained by first estimating the prediction coefficients for which

$$\sum_{k=p}^{N-1} \left(\sum_{l=0}^{p} a_l s_{k-l} \right)^2 \tag{5}$$

is minimal, substituting zeros for the unknown samples, and then minimizing the same expression as a function of the unknown samples, substituting the previously obtained estimates for the prediction coefficients. As a result, the estimates for the unknown samples, arranged in the vector $\hat{\mathbf{x}} = [\hat{s}_{t(1)}, \cdots, \hat{s}_{t(m)}]^T$, are obtained as the solutions of the system

$$\mathbf{B}\hat{\mathbf{x}} = -\mathbf{z} \tag{6}$$

with

$$b_{i,j} = b_{t(j)-t(i)}, \qquad i, j = 1, \cdots, m \tag{7}$$

[1]In this correspondence, stochastic variables are boldface.

and

$$z_i = \sum_{k=-p}^{p} b_k v_{k-t(i)}, \qquad i = 1, \cdots, m. \tag{8}$$

Here, b_k, $k = -p, \cdots, p$ is defined in (3) and v_k is obtained from s_k by substituting zeros for the unknown samples. This procedure can be made iterative by again estimating prediction coefficients for which (5) is minimal, substituting the previously obtained estimates for the prediction coefficients, and so on.

From now on, it is assumed that the prediction coefficients are known. As in [1], the expected relative restoration error per sample

$$\mathcal{E}\left\{ \frac{\sum_{i=1}^{m} (\hat{s}_{t(i)} - s_{t(i)})^2}{mR(0)} \right\}$$

is used as an indication of the performance of the restoration method. It is shown in [1] that the expected relative restoration error per sample is given by

$$\mathcal{E}\left\{ \frac{\sum_{i=1}^{m} (\hat{s}_{t(i)} - s_{t(i)})^2}{mR(0)} \right\} = \frac{\sigma_e^2 \operatorname{trace}(\mathbf{B}^{-1})}{mR(0)}. \tag{9}$$

Here $R(k)$ is the kth autocorrelation coefficient of the stochastic signal s_i. In the case of one unknown sample, this comes down to

$$\frac{\sigma_e^2}{b_0 R(0)}. \tag{10}$$

In the following section, an upper bound for (10) is derived which is low for signals with a peaky spectrum.

III. The Restoration Error

In this section, the following upper bound

$$\frac{\sigma_e^2}{b_0 R(0)} \le \frac{1 - a_p^2}{1 + a_p^2} = \frac{1 - \left(\prod_{i=1}^{p} \rho_i\right)^2}{1 + \left(\prod_{i=1}^{p} \rho_i\right)^2} \tag{11}$$

for the relative restoration error variance is derived. In (11), ρ_i denotes the modulus of the ith zero of (4). If the signal spectrum is more peaky, the ρ_i are closer to one and the right-hand side of (11) becomes small.

Note that $a_p \neq 0$ because the autoregressive process is of order p. For $b_0 = 1 + a_1^2 + \cdots + a_p^2$ (3), it follows straightforwardly that

$$b_0 \ge 1 + a_p^2. \tag{12}$$

A similar expression for $R(0)/\sigma_e^2$ is more difficult to obtain. It comes, in fact, as a byproduct of the Levinson–Durbin [2] algorithm. On multiplying both sides of (1) by s_{k-m}, $m = 0, \cdots, p$ and by taking the expected value, one obtains after some manipulations the so-called *Yule–Walker* equations [2]:

$$\mathbf{R}^{(p)}\mathbf{a} = -\mathbf{r}^{(p)}. \tag{13}$$

Here, $\mathbf{R}^{(p)}$ is the $p \times p$ autocorrelation matrix, defined by

$$r_{i,j}^{(p)} = R(i-j), \qquad i, j = 1, \cdots, p,$$

$\mathbf{r}^{(p)}$ is a vector of autocorrelation coefficients, defined by

$$\mathbf{r}^{(p)} = [R(1), \cdots, R(p)]^T,$$

and $\mathbf{a}$ is the vector of prediction coefficients, defined by

$$\mathbf{a} = [a_1, \cdots, a_p]^T.$$

The Levinson–Durbin algorithm is an iterative procedure to solve (13) for $\mathbf{a}$ in p iteration steps [2]. In the ith iteration step, the system

$$\mathbf{R}^{(i)}\mathbf{a}^{(i)} = -\mathbf{r}^{(i)}$$

is solved for $\mathbf{a}^{(i)}$ by using the $\mathbf{a}^{(i-1)}$ already obtained. Here, $\mathbf{R}^{(i)}$ and $\mathbf{r}^{(i)}$ are defined as above, with p replaced by i. The Levinson–Durbin algorithm also computes σ_e^2 in p iteration steps. Starting from $(\sigma_e^2)^{(0)} = R(0)$, in the ith iteration step, $(\sigma_e^2)^{(i)}$ is computed by

$$(\sigma_e^2)^{(i)} = \left(1 - (a_i^{(i)})^2\right)(\sigma_e^2)^{(i-1)}. \tag{14}$$

For a polynomial $A(z)$ (4) of order p, with all its zeros inside the unit circle of the complex plane, one has that $|a_i^{(i)}| < 1$, $i = 1, \cdots, p$. The results of the Levinson–Durbin algorithm are sequences $1, \mathbf{a}^{(1)}, \cdots, \mathbf{a}^{(p)}$ and $R(0), (\sigma_e^2)^{(1)}, \cdots, (\sigma_e^2)^{(p)}$, with $\mathbf{a}^{(p)} = \mathbf{a}$ and $(\sigma_e^2)^{(p)} = \sigma_e^2$. By repeatedly using (14), one obtains

$$\frac{R(0)}{\sigma_e^2} = \frac{1}{\left(1 - (a_1^{(1)})^2\right)\left(1 - (a_2^{(2)})^2\right) \cdots \left(1 - (a_p^{(p)})^2\right)}.$$

Because $a_p^{(p)} = a_p \neq 0$ and $|a_i^{(i)}| < 1$, this gives

$$\frac{R(0)}{\sigma_e^2} \geq \frac{1}{1 - a_p^2}. \tag{15}$$

The polynomial $A(z)$ in (4) can also be written as

$$A(z) = (1 - \alpha_1 z^{-1})(1 - \alpha_2 z^{-1}) \cdots (1 - \alpha_p z^{-1})$$

where the α_i, $i = 1, \cdots, p$ are the complex zeros. For a_p, one has that $a_p = -\alpha_1 \alpha_2 \cdots \alpha_p$. Because $A(z)$ has real coefficients, the zeros occur in conjugated pairs, and if $\alpha_i = \rho_i \exp(j\phi_i)$, then

$$a_p = \rho_1 \rho_2 \cdots \rho_p. \tag{16}$$

Finally, combining (12), (15), and (16) gives the upper bound (11). Of course, (15) can also be substituted into (9), which results in an upper bound for the case that there are more unknown samples. It seems, because of the factor $1 - a_p^2$, that this bound is also low if the poles are close to the unit circle, but this cannot be proved because there are no specific results on the behavior of trace $(\mathbf{B}^{-1})$.

The polynomial

$$1 - a_p z^{-p}$$

has zeros

$$\alpha_k = \sqrt[p]{|a_p|}\, e^{j(2\pi/p)k}, \qquad k = 1, \cdots, p$$

if $a_p > 0$ or

$$\alpha_k = \sqrt[p]{|a_p|}\, e^{j((2\pi/p)k + (\pi/p))}, \qquad k = 1, \cdots, p$$

if $a_p < 0$. It is obvious that for this polynomial, (12) holds with the equality sign. It follows from the Levinson–Durbin algorithm that $a_i^{(i)} = 0$, $i = 1, \cdots, p-1$ if and only if $a_i = 0$, $i = 1, \cdots, p-1$; therefore, (15) also holds with the equality sign. Consequently, (11) holds with the equality sign.

Acknowledgment

The author wishes to thank his colleagues A. van den Enden and G. Janssen for spending their time discussing the results presented here, and especially P. Delsarte for providing an alternative "Levinson-free" proof of (15).

References

[1] A. J. E. M. Janssen, R. N. J. Veldhuis, and L. B. Vries, "Adaptive interpolation of discrete-time signals that can be modeled as autoregressive processes," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 317–330, Apr. 1986.

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

A Class of Mixed Transversal and Ladder Adaptive Filters with Pure Order Updates

XIAO-HU YU AND ZHEN-YA HE

***Abstract*—The long-term numerical stability of the recursive least-squares (RLS) adaptive filters is a critical problem which has received extensive attention in recent years. This correspondence presents a new class of exact RLS adaptive filter structures with pure order updates—named mixed transversal and ladder (MITAL)—in which the propagation of numerical errors along with time is limited to at most M sampling periods, and the long-term numerical stability can be thus guaranteed. The new adaptive structures can be treated either as transversal forms or as ladder forms to meet requirements of different applications. The outstanding numerical performances of the new adaptive filters are illustrated by computer simulations.**

I. Introduction

The long-term numerical stability of the recursive least-squares (RLS) adaptive filters is a critical problem which has received extensive attention in recent years. It has been analytically observed by Ljung and Ljung [1] that the finite wordlength errors in the fast transversal filter (FTF) algorithms [2] are propagated in exponentially increasing fashion, making the algorithms diverge suddenly after a large number of iterations. It also has been found by Fabre [3] that in the case of exponential window with the forgetting factor strictly less than 1, instability inevitably occurs in any of the FTF-like algorithms. Some rescue procedures were proposed to avoid the disastrous effects of this divergence [2], [5], [6], but unfortunately these ways usually lead to a "blind period" or degradation of the performances, which undoubtedly prevents wide-scale use of the fast algorithms.

Starting with Botto [7], several researchers [8], [9] very recently have made attempts in the direction of stabilizing the FTF-like RLS algorithms. The common ground for their efforts is that the fast-RLS algorithms can be modeled as a nonlinear time-varying system, with which the numerical stability problem is transformed into stabilization of the system. However, it is still an open issue to find a generally satisfactory solution, because the stability analysis of the nonlinear time-varying system is a very complicated problem, which depends not only upon characteristics of the input signals but also upon the order of the filter being used.

Owing to the order-recursive characteristic, the adaptive ladder forms exhibit better numerical stability [10]. However, the possibility still exists for numerical error accumulation, due to the computation of the reflection coefficients that has to be performed time recursively [11]. Another limitation of the conventional ladder forms is that they cannot directly provide the Wiener solution or transversal filter coefficients which are just desired for many practical applications.

In this correspondence, we present a new class of RLS algorithms—named "mixed transversal and ladder (MITAL)"—in which the sequential LS solution is obtained by two steps. First, two correlation vectors of the observed data sequences are iterated time recursively that carry the essential information for the estimation problem [15], and second, based on the current values of the two correlators, the desired solution is determined by a purely

Manuscript received December 27, 1987; revised December 12, 1988. This work was supported in part by the Ph.D. Student Research Foundation of the China Education Commission.

The authors are with the Department of Radio Engineering, Nanjing Institute of Technology, Nanjing, Jiang-Su Province, People's Republic of China.

IEEE Log Number 8929358.

0096-3518/89/0900-1464$01.00 © 1989 IEEE