[4] H. I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-37, no. 6, June 1989.
[5] T. A. C. M. Claasen and W. F. G. Mecklenbräuker, "The Wigner distribution—a tool for time-frequency signal analysis, Part II," *Philips J. Res.*, vol. 35, nos. 4 and 5, 1980.
[6] T. A. C. M. Claasen and W. F. G. Mecklenbräuker, "The Wigner distribution—a tool for time-frequency signal analysis, Part III," *Philips J. Res.*, vol. 35, no. 6, 1980.
[7] L. Cohen, "Generalized phase-space distribution functions," *J. Math. Phys.*, vol. 7, no. 5, May 1966.
[8] L. Cohen and C. A. Pickover, "A comparison of joint time-frequency distributions for speech signals," in *Proc. ICASSP*, 1986.
[9] L. Cohen, "Time-frequency distributions—a review," *Proc. IEEE*, vol. 77, no. 7, July 1989.
[10] A. H. Nuttall, "Alias-free Wigner distribution function and complex ambiguity function for discrete-time samples," Naval Underwater Sys. Center, Tech. Rep. 8533, Apr. 1989.
[11] A. H. Nuttall, private communication.
[12] A. H. Nuttall, "Alias-free smoothed Wigner distribution function for discrete-time samples," *Naval Underwater Syst. Center, Tech. Rep. 8785*, Oct. 1990.
[13] M. Sun, C. C. Li, L. N. Sekhar, and R. J. Sclabassi, "Efficient computation of the discrete pseudo-Wigner distribution," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 11, Nov. 1989.
[14] Y. Zhao, L. E. Atlas, and R. J. Marks, II, "The use of cone-shaped kernels for generalized time-frequency representations of nonstationary signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 7, July 1990.

Maximum Likelihood Blur Identification and Image Restoration Using the EM Algorithm

A. K. Katsaggelos and K. T. Lay

***Abstract*—This correspondence describes an algorithm for the identification of the blur and the restoration of a noisy blurred image. The original image and the additive noise are modeled as zero-mean Gaussian random processes. Their covariance matrices are unknown parameters. The blurring process is specified by its point spread function, which is also unknown. Maximum likelihood estimation is used to find these unknown parameters. In turn, the EM algorithm is exploited in computing the maximum likelihood estimates. In applying the EM algorithm, the original image is part of the complete data; its estimate is computed in the E-step of the EM iterations. Explicit iterative expressions are derived for the estimation. Experimental results on simulated and photographically blurred images are shown.**

I. Introduction

Images are recorded to portray useful information. Unfortunately, a recorded image will almost certainly be a noisy-blurred version of an original image due to the imperfections of physical imaging systems [1]. It is desired that the original image be recovered from the observed image. In most practical situations, the first step in restoring an image is to identify the kind of degradation the image has suffered. Modeling the blurred image as the output of a two-dimensional (2D) linear space-invariant (LSI) system, the blur identification problem becomes the estimation of the unknown point-spread function (PSF) of the system.

The earliest work on blur identification concentrated on identifying PSF's that have zeros only on the unit bicircle [1]. In more recent work [2]–[4], the image and blur model identification problem is specified as 2D autoregressive moving-average (ARMA) parameter estimation. For the application of recursive blur identification methods, the blur PSF need to represent a causal minimum phase system [2]. If the assumption is not satisfied, recursive methods will identify the spectrally minimum phase equivalent of the nonminimum phase [4]. In this correspondence, the image and noise are modeled as multivariate Gaussian processes. Then maximum likelihood (ML) estimates of the parameters which characterize the Gaussian processes are sought. The expectation-maximization (EM) algorithm [5] is exploited in finding these parameters [6]–[8]. Simple iterative expressions are obtained in the discrete frequency domain. An algorithm similar to ours was independently derived in [9], [10]. However, the algorithm in [9], [10] is derived in the spatial domain and an autoregressive (AR) model is used for the image; furthermore, it assumes that the support of the PSF is known in advance.

II. Problem Formulation

In this work, we assume that the original image, denoted by $x(i, j)$, $i = 0, \cdots, N-1, j = 0, \cdots, N-1$, or by the $N^2 \times 1$ vector x in lexicographically ordered form [1], is modeled as a realization of a zero-mean Gaussian process with covariance Λ_X. The observed image $y(i, j)$, $i = 0, \cdots, N-1, j = 0, \cdots, N-1$, is modeled as the output of a 2D LSI system with PSF $\{d(p, q)\}$, which is corrupted by additive zero-mean white Gaussian noise $v(i, j)$, with covariance matrix Λ_V. That is, $y(i, j)$ is expressed as

$$y(i,j) = \sum_{(p,q)\in S_D} d(p, q)\, x(i-p, j-q) + v(i,j) \quad (1)$$

where S_D is the finite support region of the distortion filter. In matrix/vector form, (1) can be rewritten as [1]

$$y = Dx + v. \quad (2)$$

With the standard assumption that x and v are uncorrelated, the observed image y is Gaussian with probability density function (pdf) equal to

$$f_Y(y;\phi) = \left|2\pi(D\Lambda_X D^H + \Lambda_V)\right|^{-1/2} \cdot \exp\left\{\frac{-1}{2} y^H (D\Lambda_X D^H + \Lambda_V)^{-1} y\right\} \quad (3)$$

where ^{H} denotes the Hermitian (i.e., conjugate transpose) of a matrix or a vector, and $|\cdot|$ denotes the determinant of a matrix. The pdf of y is specified by the set of the unknown parameters $\phi = \{D, \Lambda_X, \Lambda_V\}$. A ML approach is followed in estimating ϕ. The ML estimation of the parameter set is the determination of the value of ϕ such that the logarithm of the likelihood function (i.e., $\log f_Y(y; \phi)$) is maximized. By studying the likelihood function it is clear that if no structure is imposed on the matrices D, Λ_X, and Λ_V, the number of unknowns involved is very large, thus the ML identification problem becomes unmeaningful and unmanageable. Furthermore, the estimation of $\{d(p, q)\}$ is not unique, because only second-order statistics of the blurred image are used, which do not contain information about the phase of the blur. This is an equivalent requirement to the minimum phase requirement for the blur, imposed by recursive identification techniques. In order to obtain a unique solution to the ML image and blur identification and restoration problem, additional information about the unknown parameters is incorporated into the solution process.

Manuscript received July 31, 1989; revised May 14, 1990. This work was supported in part by the NSF under Grant MIP-8614217.

The authors are with the Department of Electrical Engineering and Computer Science, Northwestern University, McCormick School of Engineering and Applied Science, The Technological Institute, Evanston IL 60208.

IEEE Log Number 9041619.

A. Constraints on the Unknown Parameters

The structure we are imposing on Λ_X and Λ_V results from the commonly used assumptions in the field of image restoration [1]. First we assume that the additive noise v is white, with variance σ_V^2, that is,

$$\Lambda_V = \sigma_V^2 I \tag{4}$$

where I is the identity matrix. Further, we assume that the random process x is stationary which results in Λ_X being a block-Toeplitz matrix [1]. A block-Toeplitz matrix is asymptotically equivalent to a block-circulant matrix as the dimension of the matrix becomes large [11]. For average size images, the dimensions of Λ_X are large indeed; therefore, the block-circulant approximation is a valid one. Associated with Λ_X is the 2D sequences $\{l_X(p, q)\}$. The matrix D in (2) is also block Toeplitz; therefore, it is approximated by a block circulant matrix. Block circulant matrices can be diagonalized with a transformation matrix W constructed from discrete Fourier kernels [1]. More specifically, it holds that

$$\Lambda_X = W Q_X W^{-1} \tag{5}$$

$$D = W Q_D W^{-1}. \tag{6}$$

All matrices W, Q_X, and Q_D are of size $N^2 \times N^2$. Q_X and Q_D are diagonal matrices with elements, the raster scanned 2D DFT values of the 2D sequences of $\{l_X(p, q)\}$ and $\{d(p, q)\}$, denoted, respectively, by $S_X(m, n)$ and $\Delta(m, n)$. Due to the particular form of W^{-1}, the product $W^{-1}\beta$, where β is an $N^2 \times 1$ vector obtained by stacking an $N \times N$ 2D sequence $\beta(i, j)$, is the stacked 2D DFT of $\beta(i, j)$.

As mentioned previously, the estimate of $\{d(p, q)\}$ is not unique since its Fourier phase is not determined. Nonuniqueness can in general be avoided by enforcing the solution to satisfy a set of constraints. Most PSF's of particular interest can be assumed to be symmetric, i.e., $d(p, q) = d(-p, -q)$. In this case the phase of the DFT of $\{d(p, q)\}$ is zero or $\pm\pi$. Unfortunately, uniqueness of the ML solution is not always established by the symmetry assumption, due primarily to the phase ambiguity. Therefore, additional constraints are required. Such constraints are the following: i) The blurring mechanism preserves energy [1], which results in

$$\sum_{(i,j)\in S_D} d(i, j) = 1. \tag{7}$$

ii) The PSF coefficients are nonnegative and iii) the support S_D is finite. A procedure for determining the support S_D and for reinforcing this last constraint is described in Section IV.

III. The EM Iterations for Estimating ϕ

Since the direct maximization of $f_Y(y; \phi)$ is complicated due to its high nonlinearity, the EM algorithm is applied to estimate ϕ. In applying the EM algorithm, the observation y represents the incomplete data and a set of complete data z has to be chosen properly. The term "incomplete data" implies the existence of two sample spaces Ω_Y and Ω_Z and a many-to-one mapping from Ω_Z and Ω_Y. In the E-step of the EM algorithm, the conditional expectation of log $f_Z(z; \phi)$ conditioned upon the observed data y and the current estimate of the relevant parameters is computed, where $f_Z(z; \phi)$ is the pdf of the complete data; in the M-step, this expectation is maximized. In a compact form, the EM algorithm can be expressed as the alternate computation of the following two equations:

$$\phi^{(p+1)} = \arg\left\{\max_{\{\phi\}} Q(\phi; \phi^{(p)})\right\} \tag{8}$$

and

$$\begin{aligned} Q(\phi; \phi^{(p)}) &= E[\log f_Z(z; \phi) \mid y; \phi^{(p)}] \\ &= \int_{\Omega_{Z|y}} f_{Z|Y}(z \mid y; \phi^{(p)}) \log f_Z(z; \phi)\, dz \end{aligned} \tag{9}$$

where $\phi^{(p)}$ is the estimate of ϕ at the pth iteration step and the integration is over all possible values of z that may produce the observed result y, denoted by $\Omega_{Z|y}$.

We choose $z = [x^H y^H]^H$ as the complete data (other choices of complete data are discussed in [7], [12], [13]). In the linear Gaussian case, the EM algorithm reduces to the minimization of [8], [14]

$$\begin{aligned} F(\phi; \phi^{(p)}) = {} & \log|\Lambda_X| + \log|\Lambda_V| \\ & + \mathrm{tr}\left\{(\Lambda_X^{-1} + D^H\Lambda_V^{-1}D)\Lambda_{X|y}^{(p)}\right\} \\ & + \mu_{X|y}^{(p)H}(\Lambda_X^{-1} + D^H\Lambda_V^{-1}D)\mu_{X|y}^{(p)} \\ & - 2y^H\Lambda_V^{-1}D\mu_{X|y}^{(p)} + y^H\Lambda_V^{-1}y \end{aligned} \tag{10}$$

where $\mu_{X|y}^{(p)}$ and $\Lambda_{X|y}^{(p)}$ are the conditional mean and conditional covariance of x given y and $\phi^{(p)}$, respectively. Note that $\mu_{X|y}^{(p)}$ represents the restored image since it is the minimum mean-square error estimate of the original image. Because of the block-circulant structure imposed on D and Λ_X, (10) can be expressed in the frequency domain as [6]–[8]

$$\begin{aligned} F(\phi; \phi^{(p)}) = {} & N^2 \log \sigma_V^2 + \frac{1}{\sigma_V^2}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}\Bigg\{|\Delta(m, n)|^2 \\ & \cdot\left(S_{X|y}^{(p)}(m, n) + \frac{1}{N^2}\left|M_{X|y}^{(p)}(m, n)\right|^2\right) \\ & + \frac{1}{N^2}\Big(|Y(m, n)|^2 - 2\,\mathrm{Re}\,[Y^*(m, n) \\ & \cdot \Delta(m, n)\, M_{X|y}^{(p)}(m, n)]\Big)\Bigg\} \\ & + \sum_{m=0}^{N-1}\sum_{n=0}^{N-1}\Bigg\{\log S_X(m, n) + \frac{1}{S_X(m, n)} \\ & \cdot\left(S_{X|y}^{(p)}(m, n) + \frac{1}{N^2}\left|M_{X|y}^{(p)}(m, n)\right|^2\right)\Bigg\} \end{aligned} \tag{11}$$

where

$$M_{X|y}^{(p)}(m, n) = \frac{\Delta^{(p)*}(m, n)\, S_X^{(p)}(m, n)}{\left|\Delta^{(p)}(m, n)\right|^2 S_X^{(p)}(m, n) + \sigma_V^{2(p)}}\, Y(m, n), \tag{12}$$

$$S_{X|y}^{(p)}(m, n) = \frac{S_X^{(p)}(m, n)\, \sigma_V^{2(p)}}{\left|\Delta^{(p)}(m, n)\right|^2 S_X^{(p)}(m, n) + \sigma_V^{2(p)}}. \tag{13}$$

In (11), $Y(m, n)$ and $M_X^{(p)}(m, n)$ are respectively, the 2D DFT's of the observed image $y(i, j)$ and the unstacked vector $\mu_{X|Y}^{(p)}$ into an $N \times N$ array, * denotes complex conjugation and Re $[\eta]$ and $|\eta|$ denote, respectively, the real part and the magnitude of the complex number η. Taking the partial derivatives of $F(\phi; \phi^{(p)})$ with respect to $S_X(m, n)$ and $\Delta(m, n)$ and setting them equal to zero, we obtain the solutions that minimize $F(\phi; \phi^{(p)})$, which represent $S_X^{(p+1)}(m, n)$ and $\Delta^{(p+1)}(m, n)$, respectively. They are

$$S_X^{(p+1)}(m, n) = S_{X|y}^{(p)}(m, n) + \frac{1}{N^2}\left|M_{X|y}^{(p)}(m, n)\right|^2 \tag{14}$$

$$\Delta^{(p+1)}(m, n) = \frac{1}{N^2}\,\frac{Y(m, n)\, M_{X|y}^{(p)*}(m, n)}{S_{X|y}^{(p)}(m, n) + \frac{1}{N^2}\left|M_{X|y}^{(p)}(m, n)\right|^2} \tag{15}$$

Substituting (15) into (11) and then minimizing $F(\phi; \phi^{(p)})$ with respect to σ_V^2, we obtain

$$\sigma_V^{2(p+1)} = \frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \Big\{ \left|\Delta^{(p+1)}(m, n)\right|^2 \Big(S_{X|y}^{(p)}(m, n) + \frac{1}{N^2} \left|M_{X|y}^{(p)}(m, n)\right|^2 \Big) + \frac{1}{N^2} \Big(\left|Y(m, n)\right|^2 - 2\,\mathrm{Re}\left[Y^*(m, n)\, \Delta^{(p+1)}(m, n)\, M_{X|y}^{(p)}(m, n)\right] \Big) \Big\}. \tag{16}$$

Equation (12) shows that the restored image (i.e., $M_{X|y}^{(p)}$) is the output of a Wiener filter, based on the available estimate of ϕ, with the observed image as input.

IV. Experimental Results

Experimental results with simulated and photographically blurred images are described in this section. Equations (14)–(16) were used for estimating the unknown parameters. Convergence thresholds were set for $\sigma_V^{2(p)}$, $S_X^{2(p)}$, and $\Delta^{(p)}$ for terminating the EM iterations. More specially, the iterations were terminated when $\xi_v^{(p)} = |v^{(p)} - v^{(p-1)}|/|v^{(p-1)}| < 10^{-4}$ for all $v \in \{\sigma_V^2, S_X, \Delta\}$. Equation (12) was used to compute the restored image. An arbitrary value (preferably somewhat larger than the true value of σ_V^2) was assigned to $\sigma_V^{2(0)}$, the initial value of the estimate of σ_V^2. The Blackman–Tukey algorithm [15] was used to compute S_Y, an estimate of the power spectral density of y, which in turn was used as $S_X^{(0)}$, the initial value of the estimate of S_X. The 2D impulse was used as $d^{(0)}$, the initial estimate of the PSF. The constraints mentioned in Section II-A were applied at each iteration step. Since the EM algorithm converges to a stationary point of the likelihood function, it is reasonable to conjecture that if the initial guess for the PSF is sufficiently good, then the algorithm has a better chance to converge to the true values. This conjecture was used in reinforcing the finite support constraint for the PSF. That is, no knowledge is assumed *a priori*, about S_D in (1); instead, after convergence of the EM iteration was reached, the estimate of $\{d(i, j)\}$ was truncated, normalized, and used as $\{d^{(0)}(i, j)\}$ in restarting a new iteration cycle (i.e., another run of the EM iterations until convergence). In truncating the PSF, the rule was applied according to which, starting with $d(0, 0)$ pairs of adjacent values of the estimates of $\{d(i, j)\}$ were compared; if they were different by at least an order of magnitude, the smaller one was disregarded, therefore determining the support S_D.

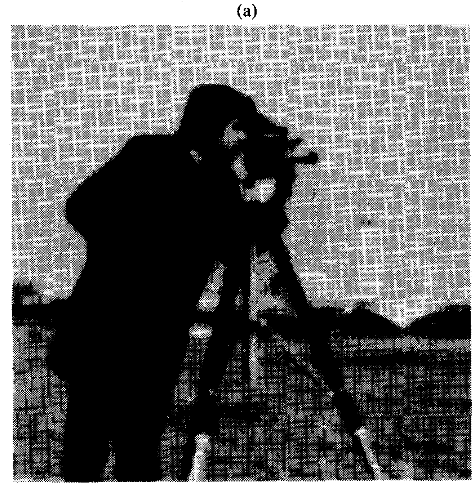
In the following the number of iteration cycles is denoted by I_c while the number of iterations at the ith iteration cycle is denoted by $I_p(i)$. In comparing the blur identification result to the true blur, the following figure of merit was used:

$$\epsilon = \frac{\|d - \hat{d}\|}{\|d\|} = \frac{\sqrt{\sum_{(i,j)\in S_D \cup S_{\hat{D}}} \left|d(i, j) - \hat{d}(i, j)\right|^2}}{\sqrt{\sum_{(i,j)\in S_D} \left|d(i, j)\right|^2}} \tag{17}$$

where $\{d(i, j)\}$ and $\{\hat{d}(i, j)\}$ denote the true and the estimated PSF's, and S_D and $S_{\hat{D}}$ their respective supports.

Fig. 1(a) shows a noisy-blurred picture. The PSF is 2D Gaussian, with values shown in Table I, and the signal-to-noise ratio (SNR) is equal to 50 dB. The restored image is shown in Fig. 1(b) and the values of the estimated blur with the number of iterations run are shown in Table II. The values of $\{d(i, j)\}$ not shown in Table II (i.e., outside the 5 × 5 support region) were very small and thus truncated. A general observation with this approach is that increased sharpness in the restored image is traded with noise amplification, as we increase the number of iteration cycles. This effect can be seen in Fig. 2. Fig. 2(a) shows the degraded image with SNR = 30 dB and blur PSF as shown in Table I. The restored

(a)



(b)

Fig. 1. (a) Noisy blurred image; 2D Gaussian blur, SNR = 50 dB. (b) Restored image of Fig. 1(a).

TABLE I
Values of the 2D Gaussian PSF

$d(i, j)$	j: −2	−1	0	1	2
i: −2	.0030	.0133	.0219	.0133	.0030
−1	.0133	.0596	.0983	.0596	.0133
0	.0219	.0983	.1621	.0983	.0219
1	.0133	.0596	.0983	.0596	.0133
2	.0030	.0133	.0219	.0133	.0030

TABLE II
Estimated PSF for Fig. 1(a) with $d^{(0)}$ a 2D Impulse (ϵ = 0.2426)

$\hat{d}(i, j)$	j: −2	−1	0	1	2
i: −2	.0056	.0171	.0262	.0171	.0056
−1	.0171	.0786	.1205	.0786	.0171
0	.0262	.1205	.1971	.1205	.0262
1	.0171	.0786	.1205	.0786	.0171
2	.0056	.0171	.0262	.0171	.0056
$I_c = 1$; $I_p(1) = 18$					

Fig. 2. (a) Noisy blurred image: 2D Gaussian blur, SNR = 30 dB. Restored images of Fig. 2(a): (b) one iteration cycle; (c) four iteration cycles.

image at the end of one iteration cycle is shown in Fig. 2(b). Fig. 2(c) shows another restoration, after four iteration cycles. The estimated values of the $\{d(i, j)\}$ are shown in Table III. A photo-

TABLE III
ESTIMATED PSF FOR FIG. 2(a) ($\epsilon = 0.6104$)

$\hat{d}(i, j)$	j: -2	-1	0	1	2
i: -2	$-.0023$	.0193	.0268	.0193	$-.0023$
-1	.0158	.0942	.0883	.0942	.0158
0	.0136	.0844	.3199	.0844	.0136
1	.0158	.0942	.0883	.0942	.0158
2	$-.0023$	.0193	.0268	.0193	$-.0023$

$I_c = 4$; $I_p(1) = 18$, $I_p(2) = 16$
$I_p(3) = 16$, $I_p(4) = 13$

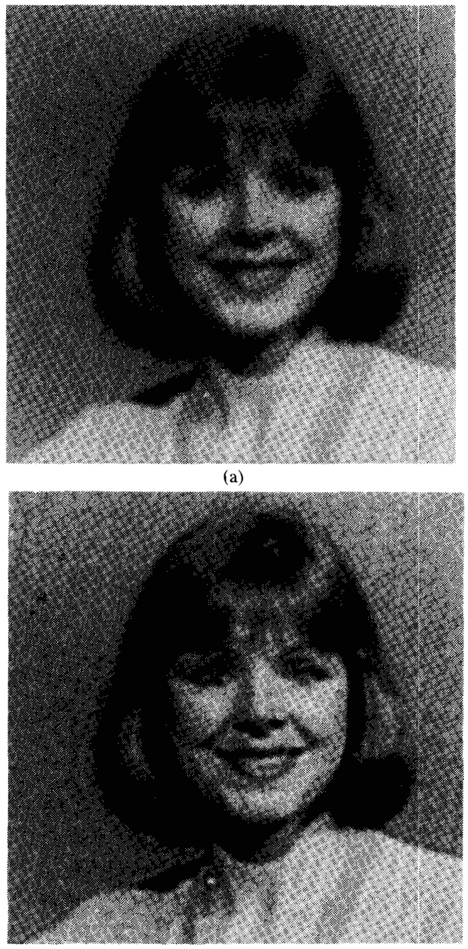


Fig. 3. (a) A photographically blurred image. (b) Restored image of Fig. 3(a).

graphically blurred image (courtesy of Kodak Research Laboratories) is shown next in Fig. 3(a). The restored image is shown in Fig. 3(b). As in previous experiments, sharpness in the restored image has been traded with noise amplification.

Based on our experiments we have made the following observations: a) Rerunning the EM iterations trades sharpness in the restored image with noise amplification. (b) Proper incorporation of constraints improves the image restoration and parameter identification results significantly. (c) The estimated σ_V^2 was usually smaller than the actual one. This resulted in a sharper but a more noisy restored image than the Wiener filter restoration with the true values for d, σ_V^2, and S_X. (d) The restoration results are sensitive to variations in the values of the PSF, while they are quite insensitive to variations in the values of σ_V^2, as was also reported in [2].

V. Conclusions

In this correspondence, we have proposed an iterative algorithm for the identification of the blur and image parameters and the restoration of a noisy blurred image. The algorithm obtains the maximum likelihood estimates of the unknown parameters with the use of the EM algorithm. One of the advantages of the proposed algorithm is that no knowledge about the type of the distortion or its support region needs to be incorporated into the iteration. The finite support of the blur PSF is reinforced as explained in Section IV. The use of an AR image model can be considered as a special case of our algorithm [8]. However, by estimating S_X, issues relating to the minimum phase requirement and the determination of the support of the AR model are avoided.

References

[1] H. C. Andrews and B. R. Hunt, *Digital Image Restoration.* Englewood Cliffs, NJ: Prentice-Hall, 1977.

[2] A. M. Tekalp, H. Kaufman, and J. W. Woods, "Identification of image and blur parameters for the restoration of noncausal blurs," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 963-972, Aug. 1986.

[3] J. Biemond, F. G. van der Putten, and J. W. Woods, "Identification and restoration of images with symmetric noncausal blurs," *IEEE Trans. Circuit Syst.*, vol. CAS-35, pp. 385-392, Apr. 1988.

[4] A. M. Tekalp and H. Kaufman, "On statistical identification of a class of linear space-invariant image blurs using nonminimum-phase ARMA models," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, pp. 1360-1363, Aug. 1988.

[5] A. D. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc.*, vol. B39, pp. 1-37, 1977.

[6] K. T. Lay and A. K. Katsaggelos, "Simultaneous identification and restoration of images using maximum likelihood estimation and the EM algorithm," in *Proc. 26th Annu. Allerton Conf. Comm., Contr., Computing* (Urbana-Champaign, IL), Sept. 1988, pp. 661-662.

[7] A. K. Katsaggelos and K. T. Lay, "Simultaneous identification and restoration of images using maximum likelihood estimation," in *Proc. 1989 Int. Conf. Contr., Application* (Israel), Apr. 1989, pp. TP-5-4.1-TP-5-4.5.

[8] A. K. Katsaggelos and K. T. Lay, "Simultaneous blur identification and image restoration using the EM algorithm," *Proc. SPIE Int. Soc. Opt. Eng.*, pp. 1474-1485, Nov. 1989.

[9] R. L. Lagendijk, D. L. Angwin, H. Kaufman, and J. Biemond, "Recursive and iterative methods for image identification and restoration," in *Proc. 1988 EUSIPCO-88* (Grenoble, France), Sept. 1988, pp. 235-238.

[10] R. L. Lagendijk, J. Biemond, and D. E. Boekee, "Blur identification using the expectation-maximization algorithm," in *Proc. Int. Conf. Acoust., Speech, Signal Processing* (Glasgow, Scotland), May 1989, pp. 1397-1400.

[11] R. M. Gray, "On the asymptotic eigenvalue distribution of Toeplitz matrices," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 725-730, Nov. 1985.

[12] K. T. Lay and A. K. Katsaggelos, "Image identification and restoration based on the expectation-maximization algorithm," *Opt. Eng.*, vol. 29, no. 5, pp. 436-445, May 1990.

[13] A. K. Katsaggelos and K. T. Lay, "Identification and restoration of images using the expectation-maximization algorithm," in *Digital Image Restoration*, A. K. Katsaggelos, Ed. Springer, 1991.

[14] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, pp. 477-489, Apr. 1988.

[15] S. M. Kay, *Modern Spectral Estimation—Theory and Application.* Englewood Cliffs, NJ: Prentice-Hall, 1988.

On the Use of Pitch Predictors with High Temporal Resolution

Peter Kroon and Bishnu S. Atal

***Abstract*—The use of pitch predictors in linear predictive coding systems is an efficient way to represent periodicity in the speech signal. Typically, the predictor is described by one parameter representing the delay in samples, and 1 to 3 predictor coefficients. Multiple predictor coefficients can provide interpolation for periodicities that are not a multiple of the sampling interval and allow for a frequency-dependent gain. In this correspondence, we describe a first-order pitch predictor whose delay is specified as an integer number of samples plus a fraction of a sample at the current sampling frequency. This realization has a better performance than conventional multiple coefficient predictors and leads to more efficient coding of the predictor parameters.**

Introduction

Periodicity in speech signals can be efficiently modeled with a slowly time-varying linear prediction filter [1], [2]. This predictor, often referred to as pitch predictor, is used in many speech coding systems such as adaptive predictive coders and more recently in multipulse and CELP coders. These latter applications have revived interest in properties of the pitch predictors (e.g., [3]). This correspondence introduces a new representation for a pitch predictor, which leads to better predictor performance and more efficient parameter quantization over conventional pitch predictors.

The general form of an odd-order pitch predictor with delay M and predictor coefficients $b(k)$ is given by

$$P(z) = 1 - \sum_{k=-(p-1)/2}^{(p-1)/2} b(k) z^{-(M+k)}, \qquad p = 1, 3, \cdots. \quad (1)$$

The predictor parameters M and $b(k)$ are determined either from the speech or the LPC residual signal, such that the squared error between the original signal and its predicted value is minimized. The delay M is found by doing an exhaustive search over its allowed range. For a given value of M, the coefficients can be found by solving a set of simultaneous linear equations in the unknowns $b(k)$ [1], [2]. The value of M is the equivalent in number of samples of a delay in the range from 2 to 20 ms. For periodic signals, this delay would correspond to a pitch period (or possibly an integral number of pitch periods). The delay would be random for nonperiodic signals. Typically, 1 to 3 predictor coefficients are used, and their values are adapted in time at rates varying from 50 to 200 times/s.

A good measure of predictor performance is the prediction gain [4]. Its value depends on many factors, such as how frequently the predictor parameters are updated, the predictor order, and the amount of periodicity in the input signal. For sampled signals the prediction gain also depends on the sampling frequency f_s. Increasing the sampling frequency increases the average prediction gain (see Appendix). Typical values for the prediction gain for periodic segments are in the range of 6 to 15 dB.

Higher order predictors result in a higher prediction gain, but more bits are needed to encode the additional coefficients (2 to 3 b/coefficient [2]). Multiple coefficients provide interpolation between the samples, if the pitch delay does not correspond to an integer number of samples. Moreover, multiple coefficients allow representation of a frequency-dependent gain factor which is useful

Manuscript received August 27, 1989; revised April 30, 1990.

IEEE Log Number 9041621.

The authors are with the Speech Research Department, AT&T Bell Laboratories, Murray Hill, NJ 07074.

1053-587X/91/0300-0733$01.00