Image identification and restoration based on the expectation-maximization algorithm

K. T. Lay
Aggelos K. Katsaggelos, MEMBER SPIE
Northwestern University
Department of Electrical Engineering
and Computer Science
The Technological Institute
Evanston, Illinois 60208

Abstract. In this paper, the problem of identifying the image and blur parameters and restoring a noisy blurred image is addressed. Specifying the blurring process by its point spread function (PSF), the blur identification problem is formulated as the maximum likelihood estimation (MLE) of the PSF. Modeling the original image and the additive noise as zero-mean Gaussian processes, the MLE of their covariance matrices is also computed. An iterative approach, called the EM (expectation-maximization) algorithm, is used to find the maximum likelihood estimates of the relevant unknown parameters. In applying the EM algorithm, the original image is chosen to be part of the complete data; its estimate is computed in the E-step of the EM iterations and represents the restored image. Two algorithms for identification/restoration, based on two different choices of complete data, are derived and compared. Simultaneous blur identification and restoration is performed by the first algorithm, while the identification of the blur results from a separate minimization in the second algorithm. Experiments with simulated and photographically blurred images are shown.

Subject terms: image restoration; maximum likelihood estimation; blur identification; expectation-maximization algorithm.

Optical Engineering 29(5), 436-445 (May 1990).

CONTENTS

1. INTRODUCTION

Due to the imperfections of physical imaging systems, a recorded image will almost certainly be a degraded version of an original image. It is desired that an estimate of the original image is recovered from the observed image. To obtain such an estimate, the identification of the blur (and the image parameters, depending on the restoration approach used) is required. The identification of the blur is thought of as a harder problem than the identification of the image parameters, due to the higher sensitivity of the restoration results to errors in the blur parameters than to errors in the image parameters. Most of the work on

Invited Paper RR-104 received Oct. 23, 1989; revised manuscript received Jan. 9, 1989; accepted for publication Jan. 10, 1990.

restoration in the literature was done under the assumption that the blur is exactly known.[1,2] However, this is hardly the actual case since usually we do not have enough knowledge about the mechanism of the distortion process. Therefore, the degradation that an image has suffered has to be estimated from the noisy blurred image. By modeling the blurring process by a linear space-invariant (LSI) system, the blur identification problem becomes the estimation of the unknown point spread function (PSF) of the LSI system.

The earliest work on blur identification concentrated on PSFs the Fourier transforms of which have zeros on the unit bi-circle.[1] The identification problem then was to search for these zeros in the frequency domain. A drawback of such an approach is that it is very difficult to find the location of the zeros due primarily to the presence of noise. In more recent work,[3,4] the image and blur model identification problem is specified as the estimation of 2-D autoregressive moving-average (ARMA) parameters. Tekalp et al.[3] computed the maximum likelihood (ML) estimates of these ARMA parameters by first decomposing the PSF into four causal quarterplane convolutional factors and then identifying each of these factors recursively. Biemond et al.[4] showed that the 2-D ARMA identification can be done in parallel, where each of the parallel channels requires the identification of a one-dimensional complex ARMA process. Lagendijk et al.[5] and Katsaggelos et al.[6] formulated the blur identification and image restoration problem as a constrained ML problem. The resulting nonlinear minimization problem was solved by an iterative gradient-type method. A priori knowledge about the blur and the image model was incorporated into the iterations. One advantage of iterative techniques is that they are free from the causality restriction and offer the possibility of incorporating prior knowledge about the original blur and image model into the identification and restoration procedure.

In this paper, two iterative algorithms for image and blur identification and image restoration are derived based on the expectation-maximization (EM) algorithm.[7] More specifically, the image and noise are modeled as multivariate Gaussian processes. Then the method of maximum likelihood estimation is applied to find the parameters that characterize the Gaussian processes. Direct maximization of the likelihood function is difficult because it is highly nonlinear with respect to these parameters. The EM algorithm is exploited to find these parameters.[8–10] In the application of the EM algorithm, a set of complete data has to be chosen properly. In this paper, two different choices of complete data are investigated. The first choice results in the estimation of all unknown parameters, while the second choice is not capable of estimating the blur parameters, which need to be estimated with the use of another optimization technique. Due to the assumption that the image is represented by a stationary process and that the blurring process is LSI, the image covariance and distortion are represented by block-circulant matrices. In this case, the problem under consideration is solved in the frequency domain and simple iterative expressions are obtained. Many effective estimation algorithms in various signal processing applications other than the image identification/restoration have been derived based on the EM algorithm.[11–14] An algorithm similar to the first algorithm to be presented in this paper was independently derived by Lagendijk et al.[15] However, the algorithm in Ref. 15 is derived in the spatial domain and an autoregressive (AR) model is used for the image; furthermore, it assumes the a priori knowledge of the support of the PSF. A review of other maximum likelihood image and blur identification algorithms is given in Ref. 16.

The image and blur models are presented in Sec. 2 and in Sec. 3 the blur identification is formulated as a MLE problem. In Sec. 4, two identification/restoration algorithms are derived based on two different choices of complete data. Experiments with simulated and photographically blurred images are shown in Sec. 5.

2. IMAGE AND BLUR MODELS

In this work we assume that the original image, denoted by $x(i,j)$, $i = 0,\ldots,N-1$, $j=0,\ldots,N-1$, or by the $N^2\times 1$ vector x in lexicographically ordered form,[1] is modeled as a realization of a random process that is Gaussian with zero mean. Its probability density function (pdf) is equal to

$$f_X(x) = |2\pi\Lambda_X|^{-1/2}\exp\left(\frac{-1}{2}x^H\Lambda_X^{-1}x\right) , \tag{1}$$

where Λ_X is the covariance matrix of x, the superscript H denotes the Hermitian (i.e., conjugate transpose) of a matrix and a vector, and $|\cdot|$ denotes the determinant of a matrix. A special case of this representation is when x is modeled as a 2-D AR process, with causal support, driven by a zero-mean homogeneous Gaussian white noise process.[17]

The observed image $y(i,j)$, $i=0,\ldots,N-1$, $j=0,\ldots,N-1$, is modeled as the output of a two-dimensional LSI system with PSF $\{d(p,q)\}$, which is corrupted by additive zero-mean Gaussian noise $v(i,j)$, with covariance matrix Λ_V, which is uncorrelated with $x(i,j)$. That is, $y(i,j)$ is expressed as

$$y(i,j) = \sum_{(p,q)\in S_D} d(p,q)x(i-p,j-q) + v(i,j) , \tag{2}$$

where S_D is the finite support region of the distortion filter. The support region of the distortion filter can have any shape. In other words, there are no constraints imposed on S_D, like the constraint, for example, that $\{d(p,q)\}$ represents a causal minimum phase system, which is imposed by recursive identification methods. In matrix/vector form, Eq. (2) can be rewritten as

$$y = Dx + v . \tag{3}$$

Equation (2) can be rewritten in the continuous frequency domain according to the convolution theorem.[18] Since the discrete Fourier transform (DFT) will be used in implementing convolution, we assume that Eq. (2) represents circular convolution [2-D sequences can be padded with zeros in such a way that the result of linear convolution equals that of circular convolution,[18] or the observed image can be preprocessed around its boundaries so that Eq. (2) is consistent with the circulant convolution of $\{d(p,q)\}$ with $\{x(p,q)\}$[19]]. Then the matrix D in Eq. (3) is assumed to be block circulant.[1]

3. ML PARAMETER IDENTIFICATION

The assumed image and blur models are specified in terms of the unknown deterministic parameters $\phi = \{\Lambda_X,\Lambda_V,D\}$. The image identification and restoration problem at hand is then the identification of ϕ and the estimation of an image as close as possible to the original one, given the observed image y. A ML approach is followed next in carrying out the task.

3.1. Formulation

Since x and v are uncorrelated, the observed image y is also Gaussian with pdf equal to

$$f_Y(y) = |2\pi(D\Lambda_X D^H + \Lambda_V)|^{-1/2} \exp\left[\frac{-1}{2} y^T (D\Lambda_X D^H + \Lambda_V)^{-1} y\right] , \tag{4}$$

where the inverse of the matrix $(D\Lambda_X D^H)$ is assumed to be defined. To emphasize that $f_Y(y)$ is parameterized, we rewrite it as $f_Y(y;\phi)$. The ML estimation of the parameter set ϕ is the determination of ϕ_{ML} that maximizes the logarithm of the likelihood function $f_Y(y;\phi)$. That is,

$$\phi_{ML} = \arg\{\max_{\{\phi\}} f_Y(y;\phi)\} = \arg\{\max_{\{\phi\}} \log f_Y(y;\phi)\} . \tag{5}$$

Taking the logarithm of Eq. (4) and disregarding constant additive and multiplicative terms, the maximization of the log-likelihood function becomes the minimization of the function $L(\phi)$, given by

$$L(\phi) = \log|D\Lambda_X D^H + \Lambda_V| + y^T (D\Lambda_X D^H + \Lambda_V)^{-1} y , \tag{6}$$

where it is assumed that $\log|D\Lambda_X D^H + \Lambda_V|$ is defined. By studying the function $L(\phi)$ 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. With so many unknowns and only one observation (i.e., y), the ML identification problem becomes unmanageable. Furthermore, the estimate of $\{d(p,q)\}$ is not unique because the ML approach to image and blur identification uses only second order statistics of the blurred image, since all pdfs are assumed to be Gaussian. More specifically, the second order statistics of the blurred image do not contain information about the phase of the blur, and therefore the phase of the ML solution for the blur is in general undetermined. To obtain a unique solution to the ML image and blur identification and restoration problem, additional information about the unknown parameters needs to be incorporated into the solution process.

3.2. 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,

$$\Lambda_V = \sigma_V^2 I , \tag{7}$$

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.[20] 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 2-D sequences $\{l_X(p,q)\}$. The matrix D in Eq. (3) was also assumed to be block-circulant. Block-circulant matrices can be diagonalized with a transformation matrix W constructed from discrete Fourier kernels.[21,22] More specifically, it holds that

$$\Lambda_X = W Q_X W^{-1} , \tag{8}$$

$$D = W Q_D W^{-1} . \tag{9}$$

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 that are the raster scanned 2-D DFT values of the 2-D 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 $N \times N$ 2-D sequence $\beta(i,j)$, is the stacked 2-D DFT of $\beta(i,j)$.

As mentioned in the previous section, the estimate of $\{d(p,q)\}$ is not unique due to no determination of its Fourier phase. Non-uniqueness can in general be avoided by enforcing the solution to satisfy a set of constraints. Most PSFs 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 π. 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

$$\sum_{(i,j)\in S_D} d(i,j) = 1 . \tag{10}$$

(ii) The PSF coefficients are nonnegative. (iii) The support S_D is finite. Similar constraints are commonly used in image restoration and blur identification.[16] A procedure for determining the support S_D and for reinforcing this last constraint is described in Sec. 5.

4. ML PARAMETER IDENTIFICATION VIA THE EM ALGORITHM

The likelihood function $L(\phi)$ in Eq. (6) is nonlinear with respect to ϕ. Therefore, analytic solutions for ϕ cannot be found in general and numerical optimization techniques need to be employed. Gradient based techniques were developed in Ref. 3. In this work, an iterative approach, called the expectation-maximization algorithm, is applied for obtaining an estimate of ϕ. The algorithm is outlined in the next section for the linear Gaussian case.

4.1. The EM algorithm in the linear Gaussian case

The EM algorithm, first proposed by Dempster et al.,[7] is a general iterative approach for solving ML estimation problems. As the name implies, the EM algorithm consists of two steps—the expectation step (E-step) and the maximization step (M-step). 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 to Ω_Y. The observed data y is a realization from Ω_Y. The corresponding z in Ω_Z is not observed directly but only indirectly through y. The choice of the complete data may not be unique. 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 compact form, the EM algorithm can be expressed as the alternate computation of the following two equations:

$$\phi^{(p+1)} = \arg\{\max_{\{\phi\}} Q(\phi;\phi^{(p)})\} \tag{11}$$

and

$$Q(\phi;\phi^{(p)}) = E[\log f_Z(z;\phi)|y;\phi^{(p)}] = \int_{\Omega_{Z|y}} f_{Z|Y}(z|y;\phi^{(p)}) \log f_Z(z;\phi)\, dz \ , \tag{12}$$

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 $\Omega_{Z|y}$. By using the EM algorithm, $f_Y(y;\phi)$ converges to a stationary point if $Q(\phi;\phi^{(p)})$ is continuous in both ϕ and $\phi^{(p)}$.[7,23] The maximization in Eq. (11) ensures that $f_Y(y;\phi)$ is monotonically nondecreasing. The EM algorithm is an attractive alternative to the direct maximization of $f_Y(y;\phi)$ if the solution to Eq. (11) can be found relatively easily. However, since the EM algorithm does not necessarily drive the likelihood function to converge to the global maximum of $f_Y(y;\phi)$, the choice of the starting point $\phi^{(0)}$ can affect the estimation significantly.

Let us now denote the relation between observed and complete data by

$$y = Hz \ , \tag{13}$$

where H represents a linear noninvertible transformation. Since it is assumed that x and y are zero-mean Gaussian processes, z is also zero-mean Gaussian with pdf equal to

$$f_Z(z) = |2\pi\Lambda_Z|^{-1/2}\exp\left(\frac{-1}{2} z^H \Lambda_Z^{-1} z\right) \ . \tag{14}$$

The conditional pdf of z given y is also Gaussian with conditional mean and conditional covariance respectively equal to[24]

$$\mu_{Z|y} = E[z|y] = \Lambda_{ZY}\Lambda_Y^{-1} y = \Lambda_Z H^H (H\Lambda_Z H^H)^{-1} y \tag{15}$$

and

$$\Lambda_{Z|y} = E[(z-\mu_{Z|y})(z-\mu_{Z|y})^H|y] = \Lambda_Z - \Lambda_{ZY}\Lambda_Y^{-1}\Lambda_{YZ} = \Lambda_Z - \Lambda_Z H^H (H\Lambda_Z H^H)^{-1} H\Lambda_Z \ , \tag{16}$$

where $\Lambda_Y = H\Lambda_Z H^H$, $\Lambda_{YZ} = H\Lambda_Z$, and $\Lambda_{ZY} = \Lambda_Z H^H$. As is shown in Ref. 11, the EM algorithm in the linear Gaussian case [i.e., Eq. (13)] reduces to the minimization of

$$F(\phi;\phi^{(p)}) = \log|\Lambda_Z| + \mathrm{tr}(\Lambda_Z^{-1} C_{Z|y}^{(p)}) \ , \tag{17}$$

where $\mathrm{tr}(A)$ denotes the trace of a matrix A, $C_{Z|y}^{(p)} = E[zz^H|y;\phi^{(p)}] = \Lambda_{Z|y}^{(p)} + \mu_{Z|y}^{(p)}\mu_{Z|y}^{(p)H}$, $\mu_{Z|y}^{(p)} = E[z|y;\phi^{(p)}]$, and $\Lambda_{Z|y}^{(p)} = E[(z-\mu_{Z|y})(z-\mu_{Z|y})^H|y;\phi^{(p)}]$. Substituting the expression for $C_{Z|y}^{(p)}$ into Eq. (17), we obtain

$$F(\phi;\phi^{(p)}) = \log|\Lambda_Z| + \mathrm{tr}(\Lambda_Z^{-1}\Lambda_{Z|y}^{(p)}) + \mu_{Z|y}^{(p)H}\Lambda_Z^{-1}\mu_{Z|y}^{(p)} \ . \tag{18}$$

That is, the E-step of the algorithm is the computation of $F(\phi;\phi^{(p)})$ in Eq. (18) with the use of the expressions in Eqs. (15) and (16), while the M-step of the algorithm is described by the following equation:

$$\phi^{(p+1)} = \arg\{\min_{\{\phi\}} F(\phi;\phi^{(p)})\} \ . \tag{19}$$

Note that $\mu_{Z|y}$ has to be computed in the E-step based on the most updated estimate of the unknown parameter set, i.e. $\phi^{(p)}$. The computation of $\mu_{Z|y}$ provides us with the restored image since the conditional mean of x based on the observed image y (denoted by $\mu_{X|y}$) is part of $\mu_{Z|y}$ and represents the minimum mean square error estimate of the original image.

4.2. Choices of complete data

An important step in the application of the EM algorithm is the specification of the mapping H in Eq. (13). Clearly, Eq. (3) can be rewritten as

$$y = [\,0 \quad I\,]\begin{bmatrix} x \\ y \end{bmatrix} = [\,D \quad I\,]\begin{bmatrix} x \\ v \end{bmatrix} = [\,I \quad I\,]\begin{bmatrix} Dx \\ v \end{bmatrix} \ , \tag{20}$$

where 0 and I represent the $N^2 \times N^2$ zero and identity matrices, respectively. According to Eq. (20), there are three candidates for representing the complete data, namely, $\{x,y\}$, $\{x,v\}$, and $\{Dx,v\}$. However, only the choice of $\{x,y\}$ as the complete data results in the simultaneous identification of the blur and the image parameters and the restoration of the image. The choice of $\{x,v\}$ as the complete data results in the identification of Λ_X and Λ_V only and the restoration of the image, while D needs to be identified in a separate minimization step. That is, the maximization of the likelihood function is carried out in two separate steps. Finally, the choice of $\{Dx,v\}$ as the complete data results in the identification of Λ_X and Λ_V and an estimate of Dx only; that is, noise-smoothing is performed instead of restoration. Although only the first choice fully justifies the term "complete data," the first two choices of complete data are considered in the following since a restored image is obtained in both cases. The resulting algorithms are compared experimentally in Sec. 5.

4.2.1. {x,y} *as the complete data*

Choosing the original and observed images as the complete data, the expressions $H = [0\ I]$ and $z = [x^H\ y^H]^H$ are obtained according to Eqs. (13) and (20). The covariance matrix of z takes the form

$$\Lambda_Z = E[zz^H] = \begin{bmatrix} \Lambda_X & \Lambda_X D^H \\ D\Lambda_X & D\Lambda_X D^H + \Lambda_V \end{bmatrix} \ , \tag{21}$$

and its inverse is equal to[25]

$$\Lambda_Z^{-1} = \begin{bmatrix} \Lambda_X^{-1} + D^H\Lambda_V^{-1}D & -D^H\Lambda_V^{-1} \\ -\Lambda_V^{-1}D & \Lambda_V^{-1} \end{bmatrix} \ . \tag{22}$$

Substituting Eqs. (21) and (22) into Eqs. (15), (16), and (17), Eq. (18) can be rewritten as

$$F(\phi;\phi^{(p)}) = \log|\Lambda_X| + \log|\Lambda_V| + \mathrm{tr}[(\Lambda_X^{-1} + D^H\Lambda_V^{-1}D)\Lambda_{X|y}^{(p)}] + \mu_{X|y}^{(p)H}(\Lambda_X^{-1} + D^H\Lambda_V^{-1}D)\mu_{X|y}^{(p)} - 2y^H\Lambda_V^{-1}D\mu_{X|y}^{(p)} + y^H\Lambda_V^{-1}y \ , \tag{23}$$

where

$$\mu_{x|y}^{(p)} = \Lambda_x^{(p)} D^{(p)H} (D^{(p)} \Lambda_x^{(p)} D^{(p)H} + \Lambda_v^{(p)})^{-1} y \ , \tag{24}$$

$$\Lambda_{x|y}^{(p)} = \Lambda_x^{(p)} - \Lambda_x^{(p)} D^{(p)H} (D^{(p)} \Lambda_x^{(p)} D^{(p)H} + \Lambda_v^{(p)})^{-1} D^{(p)} \Lambda_x^{(p)} \ . \tag{25}$$

Due to the constraints on the unknown parameters described in Sec. 3.2, Eq. (23) can be written in the discrete frequency domain as follows:

$$\begin{aligned} F(\phi;\phi^{(p)}) = & \\ & N^2 \log \sigma_v^2 + \frac{1}{\sigma_v^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \left(|\Delta(m,n)|^2 \left[S_{x|y}^{(p)}(m,n) + \frac{1}{N^2} |M_{x|y}^{(p)}(m,n)|^2 \right] \right. \\ & \left. + \frac{1}{N^2} \{ |Y(m,n)|^2 - 2\mathrm{Re}[Y^*(m,n)\Delta(m,n) M_{x|y}^{(p)}(m,n)] \} \right) \\ & + \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \left\{ \log S_x(m,n) \right. \\ & \left. + \frac{1}{S_x(m,n)} \left[S_{x|y}^{(p)}(m,n) + \frac{1}{N^2} |M_{x|y}^{(p)}(m,n)|^2 \right] \right\} \ , \end{aligned} \tag{26}$$

where

$$M_{x|y}^{(p)}(m,n) = \frac{\Delta^{(p)*}(m,n) S_x^{(p)}(m,n)}{|\Delta^{(p)}(m,n)|^2 S_x^{(p)}(m,n) + \sigma_v^{2(p)}} Y(m,n) \ , \tag{27}$$

$$S_{x|y}^{(p)}(m,n) = \frac{S_x^{(p)}(m,n) \sigma_v^{2(p)}}{|\Delta^{(p)}(m,n)|^2 S_x^{(p)}(m,n) + \sigma_v^{2(p)}} \ . \tag{28}$$

In Eq. (26), $Y(m,n)$ is the 2-D DFT of the observed image $y(i,j)$ and $M_{x|y}^{(p)}(m,n)$ is the 2-D DFT of the unstacked vector $\mu_{x|y}^{(p)}$ into an $N \times N$ array. 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 $F(\phi;\phi^{(p)})$ and represent $S_x^{(p+1)}(m,n)$ and $\Delta^{(p+1)}(m,n)$. They are

$$S_x^{(p+1)}(m,n) = S_{x|y}^{(p)}(m,n) + \frac{1}{N^2} |M_{x|y}^{(p)}(m,n)|^2 \ , \tag{29}$$

$$\Delta^{(p+1)}(m,n) = \frac{1}{N^2} \frac{Y(m,n) M_{x|y}^{(p)*}(m,n)}{S_{x|y}^{(p)}(m,n) + (1/N^2) |M_{x|y}^{(p)}(m,n)|^2} \ , \tag{30}$$

where $M_{x|y}^{(p)}$ and $S_{x|y}^{(p)}$ are computed by Eqs. (27) and (28). Substituting Eq. (30) into Eq. (26) and then minimizing $F(\phi;\phi^{(p)})$ with respect to σ_v^2, we obtain

$$\begin{aligned} \sigma_v^{2(p+1)} = & \\ & \frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \left(|\Delta^{(p+1)}(m,n)|^2 \left[S_{x|y}^{(p)}(m,n) + \frac{1}{N^2} |M_{x|y}^{(p)}(m,n)|^2 \right] \right. \\ & \left. + \frac{1}{N^2} \{ |Y(m,n)|^2 - 2\mathrm{Re}[Y^*(m,n)\Delta^{(p+1)}(m,n) M_{x|y}^{(p)}(m,n)] \} \right) \ . \end{aligned} \tag{31}$$

Equation (27) 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.

4.2.2. {x,v} *as the complete data*

Choosing $\{x,v\}$ as the complete data, we obtain $H = [D \ I]$ and $z = [x^H \ v^H]^H$. Because x and v are uncorrelated, the covariance matrix of z is equal to

$$\Lambda_z = \begin{bmatrix} \Lambda_x & 0 \\ 0 & \Lambda_v \end{bmatrix} \ , \tag{32}$$

and its inverse is equal to

$$\Lambda_z^{-1} = \begin{bmatrix} \Lambda_x^{-1} & 0 \\ 0 & \Lambda_v^{-1} \end{bmatrix} \ . \tag{33}$$

Substituting Eqs. (32) and (33) into Eqs. (15), (16), and (17) and expressing $F(\phi;\phi^{(p)})$ in the frequency domain, as was done in the previous section, Eq. (18) can be rewritten as

$$\begin{aligned} F(\phi;\phi^{(p)}) = & \\ & N^2 \log \sigma_v^2 + \frac{1}{\sigma_v^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \left\{ \left[S_{v|y}^{(p)}(m,n) + \frac{1}{N^2} |M_{v|y}^{(p)}(m,n)|^2 \right] \right\} \\ & + \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \left\{ \log S_x(m,n) \right. \\ & \left. + \frac{1}{S_x(m,n)} \left[S_{x|y}^{(p)}(m,n) + \frac{1}{N^2} |M_{x|y}^{(p)}(m,n)|^2 \right] \right\} \ , \end{aligned} \tag{34}$$

where $S_{x|y}^{(p)}(m,n)$, $M_{x|y}^{(p)}(m,n)$, $S_{v|y}^{(p)}(m,n)$, and $M_{v|y}^{(p)}(m,n)$ are the 2-D DFT values of the conditional means and conditional covariances of x and v given y, respectively. $M_{x|y}^{(p)}(m,n)$ and $S_{x|y}^{(p)}(m,n)$ are computed by Eqs. (27) and (28), respectively, while $M_{v|y}^{(p)}(m,n)$ and $S_{v|y}^{(p)}(m,n)$ are computed by the following two equations:

$$M_{v|y}^{(p)}(m,n) = \frac{\sigma_v^{2(p)}}{|\Delta^{(p)}(m,n)|^2 S_x^{(p)}(m,n) + \sigma_v^{2(p)}} Y(m,n) \ , \tag{35}$$

$$S_{v|y}^{(p)}(m,n) = \frac{|\Delta^{(p)}(m,n)|^2 S_x^{(p)}(m,n) \sigma_v^{2(p)}}{|\Delta^{(p)}(m,n)|^2 S_x^{(p)}(m,n) + \sigma_v^{2(p)}} \ . \tag{36}$$

Note that the estimation of D [or equivalently, $\Delta(m,n)$] cannot be obtained by minimizing Eq. (34) because $\Delta(m,n)$ does not appear in this equation. To estimate D, a separate minimization must be carried out. For convenience of exposure, let us divide the set of unknown parameters ϕ into two sets ϕ_1 and ϕ_2, where $\phi_1 = \{\Lambda_v, \Lambda_x\}$ and $\phi_2 = \{D\}$. Then the functional to be minimized in Eq. (34) should be specified as $F(\phi_1;\phi^{(p)})$. Setting the partial derivatives of $F(\phi_1;\phi^{(p)})$ with respect to σ_v^2 and $S_x(m,n)$ equal to zero, the solutions that minimize $F(\phi_1;\phi^{(p)})$ are equal to

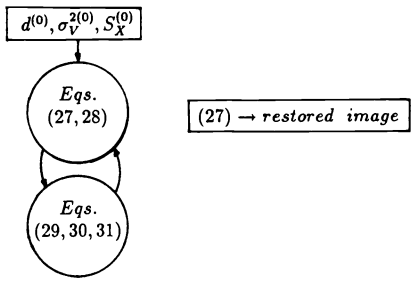


Fig. 1. The CD-xy algorithm.

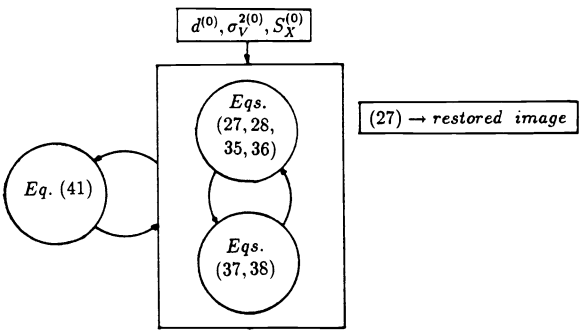


Fig. 2. The CD-xv algorithm.

$$S_X^{(p+1)}(m,n) = S_{X|y}^{(p)}(m,n) + \frac{1}{N^2}|M_{X|y}^{(p)}(m,n)|^2 , \tag{37}$$

$$\sigma_V^{2(p+1)} = \frac{1}{N^2}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}\left[S_{V|y}^{(p)}(m,n) + \frac{1}{N^2}|M_{V|y}^{(p)}(m,n)|^2\right] . \tag{38}$$

In evaluating Eqs. (37) and (38) at the $(p+1)$st iteration step, with the use of Eqs. (27), (28), (35), and (36), an estimate of $\Delta(m,n)$ at the pth step is required. Such an estimate can be obtained by minimizing

$$J(D;\phi_1^{(p)}) = \log|D\Lambda_X^{(p)}D^H + \Lambda_V^{(p)}| + y^H(D\Lambda_X^{(p)}D^H + \Lambda_V^{(p)})^{-1}y . \tag{39}$$

Expressed in the frequency domain, $J(D;\phi_1^{(p)})$ becomes

$$J(\Delta(m,n);\phi_1^{(p)}) = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1}\log[|\Delta(m,n)|^2 S_X^{(p)}(m,n) + \sigma_V^{2(p)}] + \frac{(1/N^2)|Y(m,n)|^2}{|\Delta(m,n)|^2 S_X^{(p)}(m,n) + \sigma_V^{2(p)}} . \tag{40}$$

By setting the partial derivative with respect to $\Delta(m,n)$ equal to zero, we obtain

$$|\Delta^{(p)}(m,n)|^2 = \begin{cases} \dfrac{(1/N^2)|Y(m,n)|^2 - \sigma_V^{2(p)}}{S_X^{(p)}(m,n)} , & \text{if } \dfrac{1}{N^2}|Y(m,n)|^2 > \sigma_V^{2(p)} , \\ 0, & \text{otherwise} . \end{cases} \tag{41}$$

From Eq. (41) we observe that only the magnitude of $\Delta^{(p)}(m,n)$ is available, as was mentioned earlier. A similar observation holds for Eq. (30), according to which the phase of $\Delta(m,n)$ is equal to the phase of $\Delta^{(0)}(m,n)$. Note that the estimation of ϕ is done through the alteration of Eqs. (37)-(38) and Eq. (41).

The effect of mixing the optimization procedure into the EM algorithm has not been completely analyzed theoretically. That is, the convergence properties of the EM algorithm do not necessarily hold, although both the applications of Eqs. (37)-(38) and Eq. (41) increase the likelihood function. However, based on the experimental results the algorithm derived in this section always converges to a stationary point; furthermore, the results are comparable to those obtained with the algorithm of Sec. 4.2.1.

The algorithms developed in Secs. 4.2.1 and 4.2.2 are shown in block diagram form in Figs. 1 and 2, respectively. For convenient referencing, we call the two cases of choices of the complete data CD-xy and CD-xv, respectively. Figure 1 summarizes the CD-xy algorithm, while Fig. 2 summarizes the CD-xv algorithm.

5. EXPERIMENTAL RESULTS

Experimental results with simulated and photographically blurred images are described in this section. Both algorithms proposed in the previous section were tested. Convergence thresholds were set for $\sigma_V^{2(p)}$, $S_X^{(p)}$, and $\Delta^{(p)}$ for terminating the EM iterations. That is, if $\xi_\upsilon^{(p)} = |\upsilon^{(p)} - \upsilon^{(p-1)}|/|\upsilon^{(p-1)}| < \delta$, $\upsilon \in \{\sigma_V^2, S_X, \Delta\}$, where $\delta = 10^{-4}$ was used, the iterations were terminated. Equation (27) was used to compute the restored image. An arbitrary value (preferably somewhat larger than the true value of σ_V^2) was chosen as $\sigma_V^{2(0)}$, the initial value of the estimate of σ_V^2. The Blackman-Tukey algorithm[26] was used to compute S_Y, an estimate of the power spectral density of y, which in turn was used as $S_X^{(0)}$, in both the CD-xy and CD-xv cases. We know that the EM algorithm has the drawback that it converges to a stationary point of the likelihood function. Therefore, it is reasonable to conjecture that if the initial guess for the PSF (i.e., $d^{(0)}$) is sufficiently good, then the algorithm has a better chance to converge to the true values. This conjecture was verified, by rerunning the CD-xy algorithm with different initial conditions $d^{(0)}$, such as a 2-D impulse, a decaying exponential function, and a 3*3 pillbox function, as is further explained later. In all cases, no knowledge was assumed about the support of the PSF, thus representing a major advantage of the proposed algorithm. The above conjecture was also used in reinforcing the finite support constraint for the PSF. That is, 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 that if two adjacent values of the estimates of $\{d(i,j)\}$ were different by at least an order of magnitude, the smaller one was disregarded. 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 various blur identification results, the following figure of merit was used:

$$\varepsilon = \frac{\|d - \hat{d}\|}{\|d\|} = \frac{\left[\sum_{(i,j)\in S_D \cup S_{\hat{D}}} |d(i,j) - \hat{d}(i,j)|^2\right]^{1/2}}{\left[\sum_{(i,j)\in S_D} |d(i,j)|^2\right]^{1/2}} , \tag{42}$$

where $\{d(i,j)\}$ and $\{\hat{d}(i,j)\}$ denote the original and the estimated PSFs with their respective region of support S_D and $S_{\hat{D}}$.

Fig. 3. Original cameraman image.

(a)

(b)

Fig. 4. (a) Noisy blurred image; 2-D Gaussian blur, SNR = 50 db. (b) Restored image of Fig. 4(a) by the CD-xy algorithm.

(a)

(b)

(c)

Fig. 5. (a) Noisy blurred image; 2-D Gaussian blur, SNR = 30 dB. Restored images of Fig. 5(a) by the CD-xy algorithm after (b) one iteration cycle and (c) four iteration cycles.

TABLE I. Values of the 2–D 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. 4(a) by the CD-xy algorithm. ε = 0.2426.

$d(i,j)$	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
$I_c = 1$; $I_p(1) = 18$					

TABLE III. Estimated PSF for Fig. 5(a) by the CD-xy algorithm. ε = 0.6104.

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

Figure 3 shows the original cameraman image. Figure 4(a) shows a degraded version of Fig. 3. The PSF is 2-D Gaussian, with values as shown in Table I, and signal-to-noise ratio (SNR) equal to 50 dB. The restored image by applying the CD-xy algorithm is shown in Fig. 4(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 at least an order of magnitude smaller than the smallest value appearing in Table II and thus were 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 Figs. 5 and 6. Figure 5(a) shows the degraded cameraman image with SNR = 30 dB and blur PSF as shown in Table I. The restored image at the end of one iteration cycle is shown in Fig. 5(b). Figure 5(c) shows another restoration, after four iteration cycles. The estimated values of the $\{d(i,j)\}$ are shown in Table III. Figure 6(a) shows the degraded cameraman image with SNR = 50 dB and blur PSF the truncated 1-D Gaussian function shown in Table IV. The restored images after one and three iteration cycles are shown respectively in Figs. 6(b) and

(a) (b) (c)

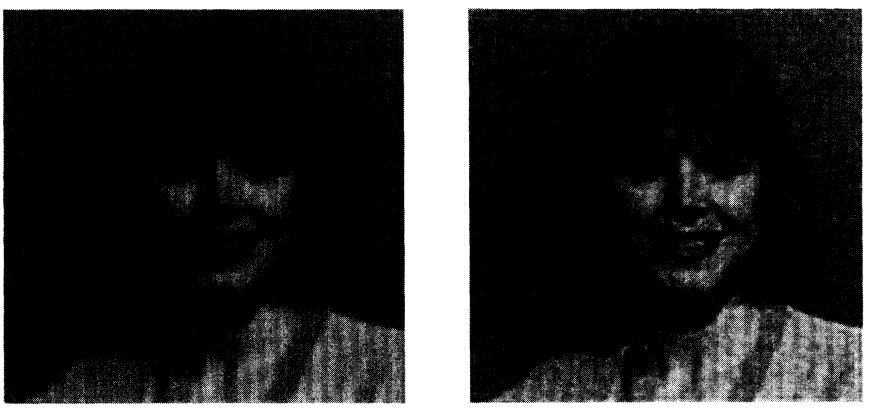
Fig. 6. (a) Noisy blurred image; 1-D Gaussian blur, SNR = 50 dB. Restored images of Fig. 6(a) by the CD-xy algorithm after (b) one iteration cycle and (c) three iteration cycles.

TABLE IV. Values of the 1-D Gaussian PSF.

$d(i,j)$	j: -4	-3	-2	-1	0	1	2	3	4
i: 0	.0052	.0298	.1039	.2199	.2824	.2199	.1039	.0298	.0052

TABLE V. Estimated PSF for Fig. 6(a) by the CD-xy algorithm. ε = 0.1489.

$d(i,j)$	j: -4	-3	-2	-1	0	1	2	3	4
i: -1	-.0079	-.0067	.0028	.0184	.0276	.0184	.0028	-.0067	-.0079
0	.0108	.0211	.0881	.2080	.2725	.2080	.0881	.0211	.0108
1	-.0079	-.0067	.0028	.0184	.0276	.0184	.0028	-.0067	-.0079
$I_c = 3$; $I_p(1) = 18$, $I_p(2) = 16$, $I_p(3) = 9$									



(a) (b)

Fig. 7. (a) Photographically blurred image. (b) Restored image of Fig. 7(a) by the CD-xy algorithm.

TABLE VI. Estimated PSF for Fig. 4(a) by the CD-xv algorithm. ε = 0.5751.

$d(i,j)$	j: -1	0	1
i: -1	.0705	.1491	.0705
0	.1377	.2853	.1377
1	.0705	.1491	.0705
$I_t = 4$; $I_e(1) = 12$, $I_e(2) = 13$, $I_e(3) = 10$, $I_e(4) = 11$			

TABLE VII. Estimated PSF for Fig. 6(a) by the CD-xv algorithm. ε = 0.2082.

$d(i,j)$	j: -2	-1	0	1	2
i: 0	.0801	.2455	.3487	.2455	.0801
$I_t = 6$; $I_e(1) = 23$, $I_e(2) = 14$, $I_e(3) = 17$, $I_e(4) = 12$, $I_e(5) = 12$, $I_e(6) = 9$					

Fig. 8. Restored image of Fig. 4(a) by the CD-xv algorithm.

Fig. 9. Restored image of Fig. 6(a) by the CD-xv algorithm.

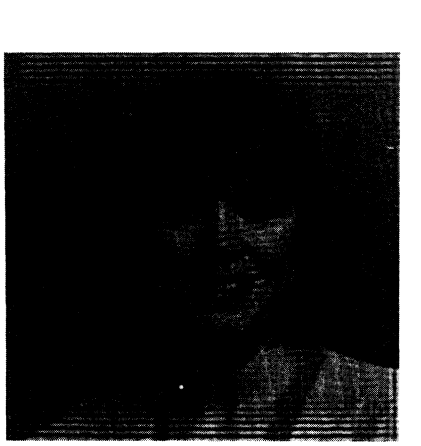


Fig. 10. Restored image of Fig. 7(a) by the CD-xv algorithm.

Fig. 11. Restored image of Fig. 4(a) by the CD-xy algorithm and $d^{(0)}$ given by Eq. (43).

TABLE VIII. Estimated PSF for Fig. 4(a) by the CD-xy algorithm and $d^{(0)}$ given by Eq. (43). $\varepsilon = 0.1801$.

$d(i,j)$	j: -2	-1	0	1	2
i: -2	.0167	.0201	.0202	.0201	.0167
-1	.0182	.0580	.0871	.0580	.0182
0	.0163	.0837	.1324	.0837	.0163
1	.0182	.0580	.0871	.0580	.0182
2	.0167	.0201	.0202	.0201	.0167
$I_c = 3$; $I_p(1) = 12$, $I_p(2) = 8$, $I_p(3) = 8$					

TABLE IX. Estimated PSF for Fig. 4(a) by the CD-xy algorithm and $d^{(0)}$ a 3*3 pillbox. $\varepsilon = 0.2347$.

$d(i,j)$	j: -2	-1	0	1	2
i: -2	.0203	.0269	.0254	.0269	.0103
-1	.0228	.0786	.1132	.0786	.0228
0	.0186	.1099	.1686	.1099	.0228
1	.0228	.0786	.1132	.0786	.0228
2	.0203	.0269	.0254	.0269	.0103
$I_c = 4$; $I_p(1) = 28$, $I_p(2) = 15$, $I_p(3) = 16$, $I_p(4) = 10$					

Fig. 12. Restored image of Fig. 4(a) by the CD-xy algorithm and $d^{(0)}$ a 3*3 pillbox.

6(c). The estimated values of $\{d(i,j)\}$ corresponding to Fig. 6(c) are shown in Table V. As is clear from these values, the estimated PSF can be considered to be 1-D, although no information about its type and its support region was incorporated into the algorithm. A photographically blurred image (courtesy of Kodak Research Laboratories) is shown next in Fig. 7(a). The CD-xy algorithm was applied to it; the restored image is shown in Fig. 7(b). As in previous experiments, sharpness in the restored image has been traded with noise amplification.

The CD-xv algorithm was tested next. The restoration of Fig. 4(a) is shown in Fig. 8 and the values of the estimated blur in Table VI. As was done previously, the values of $\{d(i,j)\}$ not shown in the table were very small and thus were truncated. The restoration of Fig. 6(a) is shown in Fig. 9 and the estimated blur in Table VII. Finally, the restoration of Fig. 7(a) is shown in Fig. 10. In Tables VI and VII the quantities specifying the number of iterations are I_t and $I_e(j)$, where I_t is the number of alternations of Eqs. (37)-(38) and Eq. (41) and $I_e(j)$ is the number of the EM iterations [i.e., Eqs. (37)-(38)] at the jth alternation of Eqs. (37)-(38) and Eq. (41).

In all previous experiments, the 2-D impulse was used as the initial guess for the PSF. In verifying the effect of the initial guess on the solutions, as discussed above, the CD-xy algorithm was applied to Fig. 4(a) with the initial choice

$$d^{(0)}(i,j) = 2^{-(i+j)}/6.25, \quad -2 \leq i,j \leq 2 , \tag{43}$$

which better resembles the 2-D Gaussian blur than the 2-D impulse function. The restored image is shown in Fig. 11 and the estimated blur is shown in Table VIII. Comparing the restored images in Figs. 4(b) and 11 and the estimated blurs in Tables II and VIII, we see that with a better starting point, the EM algorithm generates a better result, based on the visual quality of the images and the parameter ε. We also experimented with the choice of a 3*3 pillbox as $d^{(0)}$. The restored image is shown in Fig. 12 and the estimated blur is shown in Table IX. The estimated values of d in Table IX are somewhat better than those in Table II since the 3*3 pillbox is closer in the mean square sense to the actual blur than the 2-D impulse function.

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 estimate of 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 $\{d(i,j)\}$, σ_v^2, and S_X. (d) The restoration results are very sensitive to variations in the values of the PSF, while they are quite insensitive to variations in the values of σ_v^2 and S_X, as was also reported in Ref. 3. (e) Based on the quality of the restored images and the accuracy of the estimated PSF, it seems that the performance of the algorithm with $\{x,y\}$ as the complete data is better than the performance of the algorithm with $\{x,v\}$ as the complete data. This is primarily due to the fact that when $\{x,v\}$ is used as the complete data, decomposition of the unknown parameters into two sets (i.e., $\{\Lambda_X,\Lambda_V\}$ and $\{D\}$) is unavoidable.

6. DISCUSSION AND CONCLUSIONS

In this paper, we have proposed two iterative algorithms, based on two different choices of complete data, for the identification of the blur and image parameters and the restoration of a noisy blurred image. The algorithms obtain the maximum likelihood estimates of the unknown parameters with the use of the EM algorithm. Explicit expressions of the EM iterations are obtained in the discrete frequency domain. The restored image is computed in the E-step 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. In most of our experiments the iteration was initialized with a 2-D impulse, although other initial guesses of the PSF were tested for comparison purposes. Another advantage of the EM iterations is that a priori knowledge about the blur and image can be incorporated into the identification/restoration process. The experimental results corresponding to the two choices of complete data were compared. It turns out that the choice of $\{x,y\}$ as the complete data seems to result in a more effective algorithm.

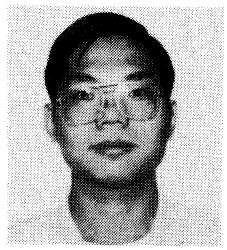
The parametric statistical model used for the original image is quite powerful and has found many applications in the area of image processing.[1] Clearly, the main reason for its use, in cases where the assumed multivariate Gaussian behavior may not be borne out by the data or the physical reality of the problem, is its mathematical tractability. Both assumptions of zero mean and stationarity for the original image result in an image model being able to describe only the gross or global properties of a class of images. Therefore, the development of efficient algorithms for identification and restoration based on an image model with nonstationary mean and/or covariance is currently under investigation. Such algorithms are expected to provide better restoration results at an additional computational expense. On the other hand, a simplified image model may prove to be equally efficient, based on the observation, which is supported by experimental evidence, that the restoration results are not very sensitive to variations of the image model parameters. Such a simplification is obtained by using, for example, an autoregressive model for the image field. If the process driving the AR model is Gaussian, then the same results presented in this work can be obtained, where the image covariance matrix is described in terms of the AR coefficients. The iterations presented in this paper can also be written in the spatial domain, where it may be easier to incorporate information about the blur into the iteration, as was described in Refs. 5 and 6. However, one of the questions to face in this case is the a priori determination of the extent of the support of the PSF and the number of equations to be used in solving for the unknown blur and model parameters.

7. ACKNOWLEDGMENT

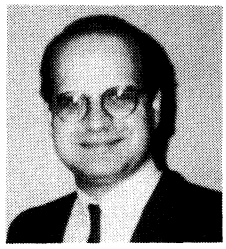
This material is based on work supported in part by the National Science Foundation under grant No. MIP-8614217.

8. REFERENCES

1. H. C. Andrews and B. R. Hunt, *Digital Image Restoration*, Prentice-Hall, Englewood Cliffs, N.J. (1977).
2. A. Rosenfeld and A. C. Kak, *Digital Picture Processing*, Vol. 1, Academic, New York (1982).
3. 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 Proc. ASSP-34, 963–972 (1986).
4. J. Biemond, F. G. van der Putten, and J. W. Woods, "Identification and restoration of images with symmetric noncausal blurs," IEEE Trans. Circ. Syst. CAS-35, 385–392 (1988).
5. R. L. Lagendijk, A. K. Katsaggelos, and J. Biemond, "Iterative identification and restoration," in *Proc. IEEE Int. Conf. on Acoust., Speech, and Signal Processing*, pp. 992–995 (1988).
6. A. K. Katsaggelos, R. L. Lagendijk, and J. Biemond, "Constrained iterative identification and restoration of images," in *Proc. Fourth European Signal Processing Conf. EUSIPCO '88*, pp. 1585–1588, North Holland (1988).
7. A. D. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," J. Roy. Stat. Soc. B39, 1–37 (1977).
8. K. T. Lay and A. K. Katsaggelos, "Simultaneous identification and restoration of images using maximum likelihood estimation and the EM algorithm," in *26th Annual Allerton Conf. on Communication, Control and Computing*, pp. 661–662 (1988).
9. K. T. Lay and A. K. Katsaggelos, "Maximum likelihood image identification and restoration based on the EM algorithm," in *Proc. Conf. on Information Science and Systems*, pp. 656–662 (1989).
10. A. K. Katsaggelos and K. T. Lay, "Simultaneous identification and restoration of images using maximum likelihood estimation," in *Proc. IEEE 1989 Int. Conf. on Control, and Applications*, pp. TP-5-4.1–TP-5-4.5 (1989).
11. M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," IEEE Trans. Acoust. Speech Signal Proc. ASSP-36, 477–489 (1988).
12. M. Segal and E. Weinstein, "The cascade EM algorithm," Proc. IEEE 76, 1388–1390 (1988).
13. M. I. Miller and D. L. Snyder, "The role of likelihood and entropy in incomplete-data problems: applications to estimating point-process intensities and Toeplitz constrained covariances," Proc. IEEE 75, 892–907 (1987).
14. B. R. Musicus and J. S. Lim, "Maximum likelihood parameter estimation of noisy data," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, pp. 224–227 (1979).
15. R. L. Lagendijk, D. L. Angwin, H. Kaufman, and J. Biemond, "Recursive and iterative methods for image identification and restoration," in *Proc. Fourth European Signal Processing Conf. EUSIPCO '88*, pp. 235–238, North Holland (1988).
16. R. L. Lagendijk, A. M. Tekalp, and J. Biemond, "Maximum likelihood image and blur identification: a unifying approach," Opt. Eng. 29(5), (1990).
17. A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, Englewood Cliffs, N.J. (1989).
18. D. E. Dudgeon and R. M. Mersereau, *Multidimensional Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, N.J. (1984).
19. J. W. Woods, J. Biemond, and A. M. Tekalp, "Boundary value problems in image restoration," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, pp. 692–695 (1985).
20. R. M. Gray, "On the asymptotic eigenvalue distribution of Toeplitz matrices," IEEE Trans. Info. Theory IT-18, 725–730 (1985).
21. R. C. Gonzalez and P. Wintz, *Digital Image Processing*, second edition, Addison-Wesley, Reading, Mass. (1987).
22. B. R. Hunt, "The application of constrained least squares estimation to image restoration by digital computer," IEEE Trans. Comput. C-22, 805–812 (1973).
23. C. F. Wu, "On the convergence properties of the EM algorithm," Ann. Statis. 11, 95–103 (1983).
24. J. M. Mendel, *Lessons in Digital Estimation Theory*, Prentice-Hall, Englewood Cliffs, N.J. (1987).
25. T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, N.J. (1980).
26. S. M. Kay, *Modern Spectral Estimation—Theory & Application*, Prentice-Hall, Englewood Cliffs, N.J. (1988).



Kuen-Tsair Lay was born in Taiwan in 1961. He received the BSEE and MSEE degrees from National Taiwan University, Taipei, in 1983 and 1987, respectively. From 1984 to 1985, he was an instructor in the Chinese Military Academy, Kaohsiung, Taiwan. Since 1987 he has been a research assistant in the Department of Electrical Engineering and Computer Science of the Technological Institute, Northwestern University, where he is engaged in research on image restoration and blur identification. His current research interests include multidimensional digital signal processing, modeling and identification of imaging systems, and estimation of parameters formulated as mathematical optimization.



Aggelos K. Katsaggelos was born in Arnea, Greece, in 1956. He received the Diploma degree in electrical and mechanical engineering from the Aristotelian University of Thessaloniki, Greece, in 1979 and the MS and Ph.D. degrees, both in electrical engineering, from the Georgia Institute of Technology, Atlanta, in 1981 and 1985, respectively. From 1980 to 1985 he was a research assistant in the Digital Signal Processing Laboratory of the Electrical Engineering School at Georgia Tech, where he was engaged in research on image restoration. He is currently an assistant professor in the Department of Electrical Engineering and Computer Science at Northwestern University. During the 1986–1987 academic year he was an assistant professor at Polytechnic University, Department of Electrical Engineering and Computer Science, Brooklyn, N.Y. His current research interests include signal and image processing, processing of moving images, computational vision, and VLSI implementation of signal processing algorithms. Dr. Katsaggelos is a member of the Associate Staff, Department of Medicine, at Evanston Hospital. He is also a member of IEEE, SPIE, the IEEE-CAS Technical Committee on Visual Signal Processing and Communications, the Technical Chamber of Commerce of Greece, and Sigma Xi. He is the editor of the book *Digital Image Restoration* (Springer-Verlag, New York).