Invited Paper

TUTORIAL REVIEW OF RECENT DEVELOPMENTS IN DIGITAL IMAGE RESTORATION

M. Ibrahim Sezan and A. Murat Tekalp*

Imaging Information Systems Research Laboratories, Eastman Kodak Company
Rochester, NY 14650-1816 USA

*Department of Electrical Engineering, University of Rochester
Rochester, NY 14627 USA

ABSTRACT

In this review, we consider the three fundamental aspects of image restoration: (i) modeling, (ii) model identification methods, and (iii) restoration methods. Modeling refers to determining a model of the relationship between the ideal image and the observed degraded image, as well as modeling the ideal image itself, on the basis of *a priori* information. Model parameters are determined by various identification methods. Restoration algorithms are discussed in two categories: general algorithms and specialized algorithms. We also furnish a brief discussion of present and future research directions.

1. INTRODUCTION

Image restoration is the problem of finding an estimate of an ideal image from its blurred and noisy rendition. An ideal image refers to the image that would be recorded by an imaging system in the absence of any blur and noise. Image restoration finds applications in a wide variety of fields such as photojournalism, forensic analysis, consumer and commercial photography, space imaging, and medical imaging.

It is well-known that image restoration is an ill-posed inverse problem. That is, a unique solution may not exist, and/or solution(s) may be extremely sensitive to noise. To alleviate the effects of ill-posedness, it is necessary to have some *a priori* information about the ideal image as well as the blur and the imaging system. *A priori* information is used in modeling and algorithm development.

Information about the nature of blur (e.g., linear and space-variant) is used to determine the blur formation model. *A priori* knowledge of the input/output and noise characteristics of the imaging sensor and other system components is used in determining the image recording and noise model. The blur formation and the image recording and noise models determine the overall observation model that describes the input/output relationship of the imaging system.

In image modeling, the ideal image can be modeled, for instance on the basis of an *a priori* Markovian assumption. In algorithm development, *a priori* information is used in defining constraints on the image estimate, and/or in defining a criterion or a quantitative description of the image estimate. At this stage, a more precise definition of the image restoration problem can be given as: Given a noisy and blurred image, find an estimate of the ideal image using *a priori* information about the blur, the imaging system and the ideal image. Research in digital image restoration has started in late 1960's. Detailed discussions of the earlier developments in image restoration can be found in a number of survey articles and books (see Ref. 1 and the references therein). The book by Andrews and Hunt, [2] published in 1977, is the only book available that specializes in the field. It provides an excellent discussion of the developments that have been made till that time. A recent survey article by Sezan and Tekalp [1] emphasizes the advances that have been made in the last decade.

We consider three important aspects of the problem (i) modeling, (ii) model identification methods, and (iii) restoration methods. In modeling, we discuss the observation, and image models. We then provide an overview of the methods for image and blur model identification. Restoration methods are discussed in two categories: (i) general methods, and (ii) specialized methods. In the first category we discuss the fundamental methods which are general in nature. Specialized methods are extensions or applications of the fundamental methods to specific restoration problems. Images illustrating the results obtained by some of the methods discussed here are going to be shown during our oral presentation.

2. MODELING

2.1. Observation model

In digital image restoration, the blurred image is modeled as

$$\boldsymbol{b} = \boldsymbol{D}\boldsymbol{s} \tag{1}$$

where $\boldsymbol{b}$ and $\boldsymbol{s}$ denote the vectors representing the lexicographical ordering of the blurred and the ideal image, respectively. The matrix $\boldsymbol{D}$ is the blur operator. That is, blur formation is modeled by a linear system.

The blur may be space-invariant or space-variant. In the case of space-invariant blurs, $\boldsymbol{D}$ represents a convolution, and (1) can be expressed as

$$b(i,j) = \sum_{(k,\ell)\in\mathcal{S}_d} d(k,\ell)\, s(i-k, j-\ell), \qquad (i,j)\in\mathcal{S}_b \tag{2}$$

where $\mathcal{S}_b$, d and $\mathcal{S}_d$ denote the blurred image support, blurring point spread function (PSF) and its support, respectively. In the case of a space-variant blur, $\boldsymbol{D}$ represents a superposition, and the blurred image can be expressed as

$$b(i,j) = \sum_{(k,\ell)\in\mathcal{S}_d(i,j)} d(i,j;k,\ell)\, s(k,\ell), \qquad (i,j)\in\mathcal{S}_b \tag{3}$$

where $d(i,j;k,l)$ and $\mathcal{S}_d(i,j)$ denote the PSF and its support, respectively, at pixel location (i,j).

If the type of the blur is known *a priori*, it may be possible to derive a parametric model of the PSF. This is the case for blurs due to atmospheric turbulence, relative motion between the scene and the imaging system, and defocused lenses.[2,3] Models of the PSF of defocused lenses have been derived using the principles of both geometrical and physical optics.[3]

The observed image is modeled by

$$\boldsymbol{r} = f(\,\boldsymbol{b}\,) + \boldsymbol{v} \tag{4}$$

where the memoryless point mapping $f(\cdot)$ models the response of the imaging sensor to the blurred image distribution which is in the light intensity domain. The mapping $f(\cdot)$ is determined from the input/output characteristic of the imaging sensor (e.g., $d - \log E$ curve of film). The sensor mapping is nonlinear in the case of images recorded on photographic film or paper. However, in the case of images recorded by solid state sensors, the sensor mapping is linear. The observation noise $\boldsymbol{v}$ accounts for both sensor noise and noise introduced by other components of the imaging system such as the quantizer and amplifiers. The sensor noise is in general signal dependent, resulting in a signal-dependent observation noise.

In the literature, observation noise is almost always approximated by a zero-mean Gaussian random field that is additive and uncorrelated to the image signal for the purpose of mathematical tractability. It should also be noted that the effects of a nonlinear sensor mapping has often been ignored in algorithm development. The sensor nonlinearity, which is of importance for film-based systems, has been explicitly accounted for in the restoration methods discussed in Sec. 4.2.3.

2.2. Image model

Images can be represented as either 2-D deterministic sequences or random fields, which can be expressed in a compact vector notation using a lexicographic ordering. In the deterministic case, they are assumed to be members of an appropriate Hilbert space (e.g., a Euclidean space with the usual inner product and norm). In the random case, images are assumed to have a certain probability distribution function (pdf). Furthermore, each realization of the random field can be considered as a member of a well-defined Hilbert space.[4]

In the random case, models have been developed for the pdf of the ideal image in the context of maximum *a posteriori* (MAP) image restoration. For example, Hunt [5], and then Trussell and Hunt [6] have proposed a Gaussian distribution with space-varying mean and stationary covariance as a model for the pdf of the image. Geman and Geman [7] have used a Gibbs distribution to model the pdf of the image.

On the other hand, if the image is assumed to be a realization of a Gauss-Markov random process, then it can be statistically modeled through an autoregressive (AR) difference equation[8]

$$s(i,j) \;=\; \sum_{(k,\ell)\in\mathcal{S}_c} c(k,\ell)\, s(i-k,j-l) \;+\; w(i,j) \tag{5}$$

where $\mathcal{S}_c$ is the image model support (one-dimensional (1-D), or two-dimensional (2-D); causal or noncausal where causality is defined on the basis of raster scanning of images), and $c(k,l)$ for $(k,\ell)\in\mathcal{S}_c$ denote the model coefficients. The quantity $w(i,j)$ denotes the modeling error which is Gaussian distributed. The model coefficients are determined such that the modeling error is minimum in the mean square sense. [8] Note that if the image is nonstationary, the model coefficients are space-variant. In that case, the model coefficients in (5) are replaced by $c(i,j;k,\ell)$. Such models have been discussed in the literature.[1]

3. IDENTIFICATION

In this section, we discuss methods for identifying the parameters that are involved in the image and observation models for subsequent use in the restoration algorithms. Methods for blur identification are discussed first. We then discuss methods that identify both the blur and the image model simultaneously. Finally, methods for identifying parameters, in various restoration algorithms, which are pertinent to the ideal image and observation noise are discussed.

3.1. Blur identification

Blur identification, in general, refers to identification of the PSF coefficients defined in (2) or (3). A parametric model for the PSF of certain types of blur (e.g., motion or defocus) can be derived using the principles of optics if the imaging system is completely known. This is an attractive alternative in blur identification, especially when the number of the model parameters is less than the number of the PSF coefficients.

The methods proposed by Gennery [9] and Canon [10] make the following two assumptions: (i) the blurring PSF has zero crossings in the frequency domain and it can be completely characterized by the location of these zero-crossings, and (ii) the location of zero-crossings can be determined from the Fourier transform [9] or power cepstrum [10] of the observed image. These methods are very simple to use and they have been successfully applied in many cases.[9–12] It is indeed true that the PSF models for motion and focus blurs do have zeros in the frequency domain, and they can be uniquely identified by the location of these zero-crossings.[3] On the other hand, blurring PSFs that do not have zero crossings in the frequency domain (e.g., Gaussian PSF modeling atmospheric turbulence) cannot be identified by these techniques. Furthermore, the identification of the zero-crossings from the observed image may be quite difficult due to the presence of observation noise. To overcome the effects of noise, Erdem and Tekalp [12] have proposed using the bispectrum of the observed image to identify the frequency-domain zero-crossings of the PSF. This approach is motivated by the fact that the bispectrum computed from the observed image is not affected, in principle, by the observation noise.

3.2. Simultaneous image and blur model identification

Modern techniques, which are based on parametric modeling of both the image process and the blurring PSF are capable of identifying both models simultaneously. These computationally more expensive techniques have been shown to be useful, when the simple methods mentioned above do fail. In these methods, the image is modeled as a causal, AR Gauss-Markov random field as in (5), and the blurred image is modeled as in (2). The parameters of the image model are the AR model coefficients and the variance of the modeling error. In general, the support of the image model and the extent of the PSF are either assumed to be known, or can be estimated.

If the image model (5) is incorporated into the observation equation (4) (with linear sensor mapping), the observed image can be expressed as a noisy observation of an autoregressive moving average (ARMA) field which represents the blurred image. Thus, image and blur identification becomes an ARMA parameter identification problem where the AR part corresponds to the image model and the moving average (MA) part corresponds to the blur model. If the image model can be identified *a priori*, for instance from a prototype image (see the next section), the problem of blur identification becomes an MA model identification problem. In other words, the blurred image can be expressed as the output of an all-zero system driven by a Gaussian noise process. The zeros of the system transfer function correspond to the zeros of the z-transform of the PSF. Thus, the identification problem becomes equivalent to finding the zeros of the PSF in the z-domain. In this sense, the MA approach can be viewed as an extension of the classical frequency-domain approach to the complex z-domain.

The ARMA model identification problem can be solved by a least squares (LS) approach [13] ignoring the presence of noise. The stability requirements of the resulting recursive equations allows the identification of only the magnitude of the blur frequency response (BFR)[13]. If the PSF is symmetric and noncausal, the phase of the BFR is either zero or $\pm\pi$. In that case, the phase information can be retrieved from the zero crossings of the BFR. The 2-D ARMA model identification problem can be decomposed into a set of parallel 1-D ARMA model identification problems, where 1-D least squares identification methods can be implemented in parallel.[14]

The main idea of the maximum likelihood (ML) approach to image/blur identification is to find the parameter values (including, in principle, the observation noise variance) which have most likely resulted in the observed image. The parameter values are found by maximizing a likelihood functional. The ML approaches can be classified into three categories: (i) gradient-based optimization based approach, (ii) the expectation maximization (E-M) algorithm based approach, and (iii) prediction-error based approach. The first approach[15] involves the maximization of the likelihood function using a steepest descent type algorithm. The major drawbacks of this approach are its slow speed of convergence and computational complexity.[15]

The ML approaches discussed by Lagendijk et al.[15], and Lay and Katsaggelos [16] are based on solving the ML problem iteratively by the expectation maximization (EM) algorithm.[17] One important advantage of using the EM algorithm is that it involves the solution of linear equations, whereas the original ML problem is a nonlinear optimization problem.[17] The EM algorithm can be formulated in either spatial [15] or frequency domain.[16] Briefly stated, the EM algorithm is a general approach for maximizing a likelihood index (or posterior distribution) when some of the data are "missing" in some sense, and observation of the missing data would have greatly simplified the estimation of the parameters. In the E-step of the algorithm, the conditional expectation of the log-likelihood of "complete" data, given the observed "incomplete" data and an initial estimate of the parameters, is determined. In the next step (the M step), the parameter values that maximize this conditional expectation are found. Since the dependence of this conditional expectation on the parameter set is identical to the dependence of the log-likelihood of the complete data on the parameter set, the complete data set should be defined such that its log-likelihood can easily be maximized (hence the simplification of the problem in the case of observing the missing data).

As shown in Refs. 15 and 16, the choice of incomplete data set as $\{r\}$ and the complete data set as $\{s\ r\}$ yields simultaneous identification and restoration, and satisfies the simplicity requirement. In this case, the linear minimum mean square error estimate of the actual image, i.e., a restoration (see Section 4.1.1) is determined in the E-step of the algorithm, whereas image and blur identification is performed in the M-step. Interesting consequences of other choices of complete/incomplete data sets are discussed by Lay and Katsaggelos.[16]

It should be noted that the EM algorithm has problems similar to those of the ARMA identification methods. Namely, the phase of the BFR may not be uniquely determined. (This is due to the fact that the ML formulation does not involve the phase of the BFR.[15,16]) Fortunately, constraining the PSF to be symmetric with nonnegative coefficients almost always alleviates this uniqueness problem.[15,16] As in the case of other ML implementations, the EM algorithm may converge to a local extremum, which may result in undesirable parameter estimates. It is experimentally shown that the convergence point is dependent on the initialization of the algorithm. In order to decrease the number of parameters, and to improve the initializations, Lagendijk et al. [18] proposed a hierarchical implementation of the EM algorithm, where identification is performed at different image resolution levels and the results of one level is used in initializing the identification in the next resolution level. We should also mention that the ML approach, like the other methods discussed, require, in general, the *a priori* knowledge of the image model support and the extent of the PSF (if not parametrized otherwise). A possible strategy here is to exaggerate the size of these supports (at the expense of increasing the number of parameters) and to estimate the actual support by discarding the estimated coefficients with insignificant values.[15] The frequency-domain EM formulation given by Lay and Katsaggelos,[16] does not require the *a priori* knowledge of the image model support because the image is modeled by a zero-mean Gaussian random field and thus characterized by its covariance matrix.

In Angwin's approach [19] - the prediction-error approach - the likelihood function is expressed recursively in prediction error form, where the prediction errors are obtained by a Kalman filter. Optimization of the likelihood index is performed via a gradient-based search in the parameter space. At each iteration, the evaluation of the likelihood index requires the implementation of the Kalman filter (resulting in an intermediate restoration). In order to alleviate the computational burden of this approach, Angwin proposed the use of reduced order model Kalman filtering (ROMKF). [19] In this approach, the observation noise variance and the variance of the image modeling error are estimated *a priori*. Therefore, ML is applied to estimate image model parameters and PSF coefficients only. Angwin has proposed a parametric modeling of the PSF, which effectively decreases the number of PSF parameters, and, thus, possibly the number of local extrema.

In the above, the image and blur models are assumed to be space-invariant. There are several approaches which assume that the image model varies slowly over the image (i.e., locally space-invariant), and make use of space-invariant identification techniques within a moving local window.[19] Some other approaches use a decision mechanism to choose from a predetermined finite set of image models at each image point.[20,21]

3.3. Image parameter identification

In the previous section, we have considered the identification of AR image models from blurred images. Here we consider methods to identify other parametric models, that are used in various restoration algorithms, which are pertinent to the ideal image.

An estimate of the power spectrum of the ideal image can be determined from the AR image parameters that can be estimated using a least squares approach. [22] On the other hand, non-parametric spectral estimation is also possible through the application of periodogram-based methods to a prototype image.[22] Power spectrum estimate is used in Wiener image restoration or in defining constraints on the solution in set-theoretic methods. [23–25]

In the context of maximum *a posteriori* (MAP) methods [5–7] (Section 4.1.1), the prior distribution is often modeled by a parametric pdf, such as a Gaussian [5,6] or a Gibbsian [7] pdf. A standard method for estimating these parameters do not exist. They can be estimated from the observed image, or a prototype image in a number of ways. [5-7]

3.4. Noise parameter identification

As we have mentioned earlier, almost all practical implementations of the restoration algorithms assume that the observation noise is a zero-mean, white Gaussian process that is uncorrelated to the image signal. In this case, the noise field is completely characterized by its variance, which is commonly estimated by the sample variance computed over a low-contrast local region of the observed image. As we are going to see in the following section, noise variance plays an important role in defining the constraints used in some of the restoration methods.

4. RESTORATION

In the past, image restoration methods have been classified in a number of different ways such as linear vs. nonlinear methods, iterative vs. noniterative methods, deterministic vs. stochastic methods, etc. Here, we classify image restoration methods into two major categories: (i) general methods, and (ii) specialized methods. The first class refers to fundamental methods which are general in nature. Specialized methods refer to extensions or applications of the fundamental methods to specific restoration problems. Each category is further classified within itself.

4.1. General methods

We classify general methods according to how a solution to the restoration problem is defined. A solution, i.e., an estimate of the ideal image, can be defined on the basis of (i) an optimality criterion, (ii) an optimality criterion subject to constraints, or (iii) a set of constraints which define the feasible solutions.

4.1.1. Criterion-based methods

The minimum-norm least squares method: A least squares solution minimizes the norm of the residual $\rho(\boldsymbol{x}) \doteq \boldsymbol{D}\boldsymbol{x} - \boldsymbol{r}$, or equivalently solves the normal equation $\boldsymbol{D}^*\boldsymbol{D}\boldsymbol{x} = \boldsymbol{D}^*\boldsymbol{r}$. (Note that this criterion effectively ignores the presence of noise.) The least squares solution with minimum norm (energy) is called the pseudo-inverse (or generalized) solution (PIS).

The PIS can be determined using a singular value decomposition (SVD) analysis [2] (which reduces to the well-known frequency domain pseudo-inverse filtering when the blur is convolutional and the blur operator can be represented by a circulant matrix), or using an iterative algorithm based on iterative minimization of $\rho(\boldsymbol{x})$ or iterative solution of the normal equation. The noise sensitivity of PIS can be reduced by truncating the singular value expansion, or by terminating the iterations in the case of iterative algorithms, at the expense of reduced resolution. Truncation or termination strategies are not well-defined unless additional information is available. Sullivan and Katsaggelos, [26] show that a termination rule for iterative algorithms can be defined if the second order statistics of the noise and image processes can be estimated.

The linear minimum mean square error method: The linear minimum mean square error (LMMSE) method finds the linear estimate of the ideal image for which the mean square error between the estimate and the ideal image is minimum. The estimate is obtained on the basis of *a priori* second order statistical information about the image and noise processes. In the case of stationary processes and space-invariant blurs, the LMMSE estimator takes the form of the Wiener filter.[2]

A Kalman filter determines the causal LMMSE estimate recursively. It is based on a state-space representation of the imaging system.[27] The image is assumed to be a causal Gauss-Markov random field represented by a finite order difference equation as in (5). In the first step of Kalman filtering, a prediction of the state variables is formed on the basis of the previous state of the system. In the second step, the predictions are updated on the basis of the observed image data to form the estimate for the present state of the system.

Early applications of Kalman filtering to the restoration of blurred and noisy images have been restricted to one-dimensional (1-D) Kalman filters (processing each line independently) using 1-D image models. [28] Woods and Radewan,[27] have formulated a two-dimensional (2-D) reduced-update scalar Kalman filter (i.e., filtering a pixel at a time) using a 2-D AR image model with non-symmetric half plane (NSHP) causal support. In the reduced update Kalman filter (RUKF), the update procedure is limited to only those state variables in a neighborhood of the pixel currently being processed.[27] The main assumption here is that a pixel is insignificantly correlated with pixels outside a certain neighborhood about itself. The performance degradation due to the reduced update has been found to be negligible.[27] On the other hand, RUKF results in significant computational savings.

Murphy and Silverman[29] have formulated a vector Kalman filter based on a 2-D semicausal, Gauss-Markov image model, where the filtering is performed one line at a time. Biemond et al.[30] propose 2-D hybrid Kalman filtering using a semicausal image model, and derive parallel Kalman filters operating on the columns of the image after its rows are decorrelated by a Fourier transformation. This approach is based on a vector-matrix representations of the image and the observation models and then approximating the resulting Toeplitz matrices by circulant matrices.

A recent attempt to reduce both the computational and storage requirements of Kalman filtering using 2-D image models is the formulation of the reduced order model Kalman filtering (ROMKF). [19] In ROMKF the state vector is truncated to a size which is on the order of the image model support about the pixel currently being processed. (ROMKF is a scalar filter and uses a NSHP image model as the RUKF.) Angwin reports experimental results indicating that the performance of ROMKF and RUKF are similar.[19]

In all of the above cases, the state vector was one-dimensional. In other words, the recursion of the Kalman filter is in one direction only (horizontal or vertical). Several researchers (e.g., Ref. 31) have proposed using higher-dimensional state-space models to reduce the effective size of the state vector. The complexity of these higher-dimensional state-space model based formulations, however, limits the practical use of these approaches.

<u>The maximum *a posteriori* probability method:</u> In maximum *a posteriori* probability (MAP) restoration, the criterion is to find the estimate which maximizes the probability density of the ideal image conditioned on the observation (*a posteriori* probability). Using the Bayes rule, the posterior probability can be expressed in terms of probability densities of the observed image conditioned on the ideal image, the ideal image and the observed image. The first two densities are determined from the *a priori* knowledge of the probability distributions of the actual image and the noise processes. (The probability density of the observed image does not affect the maximization of the posterior density.)

Hunt [5] has shown that under Gaussian assumptions for both the image and noise processes, the MAP solution can be obtained either by using the Picard iteration or by directly maximizing the posterior density using a gradient ascent algorithm. Later, Trussell and Hunt [6] have proposed the use of a modified form of the Picard iteration (which significantly increases the convergence rate) and suggested the use of the variance of the residual signal in determining the convergence. Geman and Geman [7] propose using a Gibbs distribution as the probability distribution of the actual image. The maximization of the posterior distribution in this case becomes an extremely complex problem. Geman and Geman use simulated annealing to maximize the posterior distribution.

It should be emphasized that the MAP approach differs from the methods discussed above in that it is capable of accounting for the nonlinear sensor characteristics. (It can easily be verified that in the case of a linear sensor and under Gaussian assumptions, the MAP estimate reduces to the LMMSE estimate.)

4.1.2. Constrained-optimal methods

These methods optimize an optimality criterion subject to constraints on the solution. The constraints and the criterion reflect the *a priori* information about the ideal image.

The constrained least squares method: A well-known constrained-optimal method is the constrained least squares (CLS) algorithm which is based on Tikhonov regularization.[32] In the CLS algorithm, the estimate of the image is defined to be that member of the set $C_\sigma \doteq \{\boldsymbol{x} : \ ||\boldsymbol{D}\boldsymbol{x} - \boldsymbol{r}||^2 \leq \sigma^2\}$ which minimizes a stabilizing functional, usually of the form $J(\boldsymbol{x}) \doteq ||\boldsymbol{A}\boldsymbol{x}||^2$. The operator $\boldsymbol{A}$ is chosen such that the minimization of J implicitly incorporates some *a priori* information about the ideal image (e.g., smoothness) into the problem. The defining parameter of the set, σ, is related to the observation noise variance. A frequency domain implementation of the CLS algorithm in the case of space-invariant blurs has been given by Hunt.[33] Recently, neural network structures implementing constrained least squares image restoration have been proposed.[34]

The Miller-regularization method: In Miller regularization, the estimate is defined to be that member of the intersection of the sets C_σ and $C_\epsilon \doteq \{\boldsymbol{x} : \ ||\boldsymbol{A}\boldsymbol{x}||^2 \leq \epsilon^2\}$ which minimizes the functional $J(\boldsymbol{x}) = ||\boldsymbol{D}\boldsymbol{x} - \boldsymbol{r}||^2 + (\sigma/\epsilon)^2||\boldsymbol{A}\boldsymbol{x}||^2$. [32] The operator $\boldsymbol{A}$ is chosen such that restricting the solution to C_ϵ reflects some kind of *a priori* information about the ideal image. Further the parameters σ and ϵ should be such that the set intersection is nonempty.[32] It can be shown that the Miller estimate uniquely solves an Euler equation whose parameter is (σ/ϵ), (the regularization parameter) which is assumed to be known *a priori.* [32] The regularization parameter can be determined from an estimate of the signal-to-noise ratio. [35] A recent work by Reeves and Mersereau [36] suggests using the cross-validation principle to determine the regularization parameter. An iterative algorithm to determine the Miller estimate has been derived and applied to image restoration.[35] This iterative algorithm is formulated in a weighted Hilbert space for ringing reduction by Lagendijk et al. [37] Lagendijk et al. have proposed augmenting this algorithm with projection operators projecting onto convex constraint sets. The augmented algorithm can be viewed as a hybrid Miller-projections onto convex sets (POCS) algorithm.

The maximum entropy method: The maximum entropy (ME) approach maximizes the entropy of the image data subject to the constraint that the resulting solution should reproduce the observation exactly [38] or within a tolerable uncertainty defined by the statistics of the observation noise.[39] Note that maximizing the entropy reflects the smoothness property of the actual image. (In the absence of constraints, the entropy is highest for a constant-valued image). One important aspect of the ME approach is that the nonnegativity constraint is implicitly imposed on the solution because the entropy is defined in terms of the natural logarithms of the image pixels. A number of ME approaches have been discussed in the literature. These approaches vary in the way they implement the ME principle. A common feature of all these approaches is their computational complexity. Without claiming completeness, we refer to several representative papers.

Frieden is the first to apply the ME principle to restoration.[38] In his formulation the sum of the entropy of the image and the noise process is maximized subject to the observation equation, and subject to a predetermined value of the sum of the pixel values. The formulation proposed by Gull and Daniell [39] can be viewed as another form of Tikhonov regularization; the entropy of the image (analogous to the stabilizing functional) is maximized subject to the the constraint that the the sample variance of the residual signal should be equal to the noise variance within a certain confidence factor. The optimization problem is solved using an ascent algorithm. Trussell [40] shows that in the case of a prior distribution defined in terms of the image entropy, the MAP solution is identical to the solution obtained by this ME formulation. A more complex but more powerful ME formulation has been given by Burch et al.[41]

4.1.3. Constraint-based methods

In constraint based methods, any image that satisfies the *a priori* constraints (i.e., a feasible image) is defined to be an estimate of the ideal image.

Set-theoretic methods: In set-theoretic methods, given an *a priori* constraint on the ideal image, a set is defined such that its members are consistent with the constraint, and the ideal image is a member of the set (henceforth the constraint sets). Thus, the ideal image is a member of the intersection of (any finite number) constraint sets. In set-theoretic methods, an estimate of the ideal image is defined to be any member of the intersection set, i.e., a feasible solution satisfying the constraints.

Set-theoretic methods vary according to the mathematical properties of the constraint sets. In the method of projections onto convex sets (POCS), the constraint sets are closed convex. A feasible solution is found by performing successive projections onto the constraint sets starting from an initial point and then repeating the process iteratively. The iterations converge to a feasible solution in the intersection set (see Chaps. 2 and 11 in Ref. 42). It should be noted that the convergence point is affected by the choice of the initialization. First applications of POCS have been to spectral extrapolation and image reconstruction from projections. [42] Trussell and Civanlar [23] are the first to apply POCS to image restoration. Chen et al.[43] have investigated the application of POCS to blind deblurring of images. Recent applications of POCS are concerned with specific image restoration problems, and they are discussed in Section 4.2. We refer the reader to Refs. 23,25,42-44 for the examples of convex constraint sets that are of importance in image restoration problems. Several extensions of the POCS algorithm to nonintersecting sets [45] and nonconvex sets (Chap. 8 in Ref. 42) have been discussed in the literature. Leahy and Goutis[46] have suggested optimizing a convex criterion functional, such as minimum energy or cross entropy, over the intersection of two convex constraint sets.

In the method of fuzzy sets (FS), the constraints are defined using fuzzy sets.[24] More precisely, the constraints are reflected in the membership functions defining the fuzzy sets. In this case, a feasible solution is defined as one which has a high grade of membership (e.g., above a certain threshold) in the intersection set.

Constrained iterative methods All iterative constraint-based algorithms can be expressed in the form $\boldsymbol{C}\boldsymbol{x}_k = \boldsymbol{x}_{k+1}$. The operator $\boldsymbol{C}$ (linear or nonlinear) is defined on the basis of image constraints such that if an image, say $\boldsymbol{y}$, satisfies these constraints, then $\boldsymbol{C}\boldsymbol{y} = \boldsymbol{y}$ (i.e., $\boldsymbol{y}$ is a fixed point of the operator $\boldsymbol{C}$). Therefore, a fixed point of the constraint operator can be defined as the estimate of the ideal image. A variety of iterative algorithms have been developed on these principles. We refer the reader to the papers by Shafer et al.[47]

It should be noted that POCS could also be included in this category. In POCS, the operator $\boldsymbol{C}$ is the composition of the projection operators projecting onto the convex constraint sets. The convergence proof of POCS can be given on the basis of the fact that $\boldsymbol{C}$ is a certain type of nonexpansive operator called the "reasonable wanderer".[42]

4.2. Specialized algorithms

4.2.1. Restoration of images degraded by space-variant blurs

In principle, all fundamental methods apply to the restoration of images degraded by space-variant blurs. However, implementations of these methods may be computationally formidable due to large matrix operations. Therefore, several specialized methods are developed to attack the space-variant restoration problem. These methods can be classified into three different categories: (i) sectioning; (ii) coordinate transformation; and (iii) direct approaches.

The main assumption in the sectioning methods is that the blur is approximately space-invariant within local regions of the image. Therefore, the entire image can be restored by applying well-known space-invariant techniques to the local image regions.[48] A major drawback of sectioning methods is the generation of artifacts at the region boundaries.

Robbins et al.,[49] and then Sawchuck [50] show that for certain kinds of space-varying blurs (e.g., in the case of relative motion between the object and the imaging system that takes place on trajectories which are not parallel to each other), a coordinate transformation (CTR) can be applied to the observed image so that the blur in the transformed coordinates becomes space-invariant. Therefore, the transformed image can be restored by a space-invariant filter and then transformed back to obtain the final restored image. It should be noted that the statistical properties of the image and noise processes are affected by the coordinate transformation. In particular, the stationarity in the original spatial coordinates is not preserved in the transform coordinate system. The results reported by Robbins et al.[49] and Sawchuck [50] are obtained by inverse filtering, and thus this statistical issue is of no concern. In general, the CTR method is applicable to only a limited class of space-variant blurs.

The lack of generality of the CTR motivates the direct approaches. For instance, Tekalp and Pavlovic [51] derive the RUKF for space variant blurs and show by examples that it is superior to CTR when a coordinate transformation can be found. Recently, Angwin [19] has derived the ROMKF using a space-variant model for the blurring PSF.

4.2.2. Restoration of images degraded by random blurs

Random blur occurs when the PSF fluctuates randomly, for instance due to atmospheric turbulance or random relative motion between the subject and the image recording device. In this case, blur identification problem becomes the identification of the statistical properties of the blur.

A naive approach to the problem of random blur is to use the expected value of the PSF process in one of the the restoration algorithms discussed above. The resulting restoration, however, may be quite unsatisfactory.[52] Slepian [53] has derived the LMMSE estimate in the case of random blur. The superposition approach, where the estimate of the actual image is expressed as a weighted superposition of the shifted versions of the degraded image, has been applied to the problem by Ward and Saleh.[54] This approach assumes that the blur process is uncorrelated and its mean is known *a priori.* A recent approach is proposed by Combettes and Trussell in the context of POCS, where the fluctuations in the PSF are reflected in the bounds defining the constraint sets.[52] Another recent approach is by Quan and Ward where a Wiener-type filter is derived.[55]

4.2.3. Restoration of images recorded by nonlinear sensors

Except for the MAP approach discussed in Section 4.1.2, none of the algorithms discussed above are equipped to handle sensor nonlinearity. A direct but naive approach to restoration when $f(\cdot)$ represents a pointwise nonlinearity is expanding the observation model into its Taylor series about the mean of the observed image and obtaining an approximate restoration on the basis of the linearized model.[2] However, the results do not show significant improvement over those obtained by ignoring the sensor nonlinearity.

The MAP methods discussed in Section 4.1.1 are capable of taking the sensor nonlinearity into account. An alternative approach, which is relatively less demanding computationally, has been recently developed for the case of photographic film.[11]. The observed density domain image data is back-transformed to the exposure domain. The observation equation in the exposure domain exhibits a convolutional relationship between the actual and the blurred image with multiplicative noise. Pavlovic and Tekalp [11] derive a LMMSE deconvolution filter in the exposure domain in the presence of multiplicative noise under certain assumptions. Their results suggest that accounting for the sensor nonlinearity may dramatically improve restoration results.

4.2.4. Image restoration with artifact reduction

It has been shown by Lagendijk et al.,[37] and Tekalp and Sezan [56] that linear space-invariant restoration algorithms introduce artifacts to the restored image. These artifacts can be classified as [56] (i) boundary-truncation; (ii) filtered-noise; (iii) filter-deviation; and (iv) PSF-error artifacts.

Boundary-truncation artifacts are due to errors in the convolutional model of the blurred image at the boundaries.[56] Filtered-noise artifacts are due to filtering of the noise field by the restoration filter. The deviation of the restoration filter from the exact inverse filter (e.g., in the case of a regularized filter) gives rise to filter-deviation artifacts. PSF-error artifacts are due to errors made in modeling and/or identifying the blurring PSF. Boundary-truncation artifacts can be reduced by preprocessing the boundaries of the observed image prior to restoration.[56] Reduction of other types of artifacts, on the other hand, lead to more complex, possibly nonlinear and space-variant, restoration algorithms.

Lagendijk et al.[37] have suggested trading ringing (a special case of filter-deviation artifacts) with the filtered-noise artifacts, spatially at each pixel depending on whether the pixel belongs to an edge or a uniform region. Towards this end, they have formulated the iterative Miller algorithm in a weighted Hilbert space so that its deviation from inverse filtering is small at the edges, at the expense of filtered noise artifacts. Tekalp et al.[57] proposed an edge adaptive Kalman filter to suppress the ringing artifacts. A MAP-based decision logic is employed to select from a predetermined set of image models at each pixel. The main idea of this approach is that the better the image model fits the image, the better the reduction of ringing artifacts that can be achieved. Sezan and Tekalp [44] have developed a POCS algorithm that forces the restoration to be equal to the Wiener restoration at a certain set of frequencies. Outside this set of frequencies, over which the sources of artifacts are significant, the Wiener solution is discarded and the solution is determined such that the resulting restoration is consistent with regional smoothness constraints. Recently, Sezan and Trussell [25] has developed constraints based on prototype images for set-theoretic image restoration with artifact reduction.

4.2.5. Restoration of multispectral images

A direct approach to multispectral image restoration is to ignore the correlations among different spectral bands and restore each band independently using one of the algorithms discussed above. This so-called "independent-channel" approach has been reasonably successful. Hunt and Kubler [58] show that Karhunen-Loeve (KL) transformation decorrelates the image channels, and thus the independent-channel approach is applicable. Unfortunately, the use of the KL transform may be impractical because it is image dependent. In the case of color images, as a suboptimum but easy-to-use alternative, Hunt and Kubler [58] suggest the use of the National Television Systems Committee (NTSC) YIQ transformation which is image independent. One should realize, however, that uncorrelated noise becomes correlated to the image signal under the YIQ transformation, because it is not an orthogonal transformation. Thus, the use of restoration filters, such as the Wiener filter, that assumes uncorrelated noise is not theoretically founded.

Recent efforts in multispectral image restoration are concentrated on making use of the correlations between the spectral bands of images.[59,60] Galatsanos and Chin [59] derives a Wiener filter that makes use of *a priori* cross power spectrum between color channels. Their derivation is based on approximating Toeplitz matrices with circular ones. Sezan and Trussell [60] show that the sensitivity of the multispectral Wiener filtering approach on cross power spectral estimates can be quite high. Tekalp et al.[61] and Angwin [19] use a multispectral AR image model to restore multispectral images using the RUKF and ROMKF approaches, respectively. In Ref. 60, Sezan and Trussell develop a framework for set-theoretic restoration of multispectral images, and propose a number of multispectral image constraints.

5. PRESENT AND FUTURE RESEARCH

At present, the following are the limiting factors in the applications of digital image restoration: (i) lack of fast and reliable blur identification methods, (ii) lack of general and efficient algorithms for the identification of space-variant blurs, and restoration of images degraded by such blurs, and (iii) presence of artifacts in restored images. Present and future research in digital image restoration is expected to focus on image and blur model identification, restoration of images blurred by space-variant blurs and reduction of restoration artifacts. Developing specialized algorithms for restoring image sequences is also an important future research topic due to the growing use of video systems and surveillance cameras.

6. REFERENCES

1. M. I. Sezan and A. M. Tekalp, "Survey of recent developments in digital image restoration," *Opt. Eng*, Vol. 29(5), 393-404 (1990).
2. H. C. Andrews and B. R. Hunt, *Digital Image Restoration*, Prentice Hall, New Jersey (1977).
3. H. C. Lee, "A review of image-blur models in a photographic system using the principles of optics," *Opt. Eng.* Vol. 29(5), 405-421 (1990).
4. J. N. Franklin, "Well-posed stochastic extensions of ill-posed linear problems," *J. Math. Anal. Appl.*, 31, 682-716 (1970).
5. B. R. Hunt, "Bayesian methods in nonlinear digital image restoration," *IEEE Trans. Comput.* C-26(3), 219-229 (1977).
6. H. J. Trussell and B. R. Hunt, "Improved methods of maximum a posteriori restoration," *IEEE Trans. Comput.* C-27(1), 57-62 (1979).
7. S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.* PAMI-6(6), 721-741 (1984).
8. A. K. Jain, "Advances in mathematical models for image processing," *Proc. IEEE* 69(5), 502-528 (1981).
9. D. B. Gennery, "Determination of optical transfer function by inspection of frequency-domain plot," *J. Opt. Soc. Amer.* 63(12), 1571-1577 (1973).
10. M. Cannon, "Blind deconvolution of spatially invariant image blurs with phase," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-24(1), 58-63
11. G. Pavlovic and A. M. Tekalp, "Restoration in the presence of multiplicative noise with application to scanned photographic images," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, 1913-1916 (1990).
12. A. T. Erdem and A. M. Tekalp, "Blur identification using bispectrum," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, 1961-1964 (1990).
13. A. M. Tekalp and H. Kaufman, "On statistical identification of a class of linear space-invariant image blur models using nonminimum phase ARMA models," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-36(8), 1360-1363 (1988).
14. 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-23(3), 385-394 (1988).
15. R. L. Lagendijk, A. M. Tekalp and J. Biemond, "Maximum likelihood image and blur identification: A unifying approach," *Opt. Eng.* Vol. 29(5), 422-435 (1990).
16. K. T. Lay and A. K. Katsaggelos, "Blur identification and image restoration based on the EM algorithm," *Opt. Eng.* Vol. 29(5), 436-445 (1990).
17. A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data," *J. Royal Statist. Soc. B.* 39, 1-38 (1977).
18. R. L. Lagendijk, J. Biemond, and D. E. Boekee, "Hierarchical blur identification," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, 1889-1892 (1990).
19. D. L. Angwin, "Adaptive image restoration using reduced order model based Kalman filters," Ph. D. dissertation, RPI (September 1989).
20. S. A. Rajala and R. P. DeFigueiredo, "Adaptive nonlinear image restoration by a modified Kalman filtering approach," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-29(5), 1033-1042 (1981).
21. A. T. Erdem and A. M. Tekalp, "Decision-directed adaptive image restoration using multiple image and blur models," in *Proc. Int. Conf. Contr. Appl.* (1989).
22. S. M. Kay, *Modern Spectral Estimation: Theory and Application*, Prentice Hall, New Jersey (1987).

23. H. J. Trussell and M. R. Civanlar, "Feasible solution in signal restoration," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-32(4), pp. 201-212 (1984).
24. M. R. Civanlar and H. J. Trussell, "Digital signal restoration using fuzzy sets," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-34(4), 919-936 (1986).
25. M. I. Sezan and H. J. Trussell, "Prototype image-based constraints for set-theoretic image restoration," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, 1897-1900 (1990).
26. B. J. Sullivan and A. K. Katsaggelos, "A new termination rule for linear iterative image restoration algorithms," *Opt. Eng.* Vol. 29(5), 471-477 (1990).
27. J. W. Woods and V. K. Ingle, "Kalman filtering in two-dimensions-further results," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-29, 188-197 (1981).
28. A. O. Aboutalib, M. S. Murphy, and L. M. Silverman, "Digital restoration of images degraded by general motion blurs, *IEEE Trans. Automat. Contr.* AC-22, 294-302 (1977).
29. M. S. Murphy and L. M. Silverman, "Image model representation and line-by-line recursive restoration," *IEEE Trans. Automat. Contr.*, AC-23, 809-816 (1978)
30. J. Biemond, J. Rieske and J. J. Gerbrands, "A fast Kalman filter for images degraded by both blur and noise," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-31(5), 1248-1256 (1983).
31. Z. Wu, "Multidimensional state-space model Kalman filtering with application to image restoration," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-33(6), 1576-1592, (1985).
32. M. Bertero, C. De Mol, and G. A. Viano, 'The stability of inverse problems,' in *Inverse Scattering Problems in Optics*, H. P. Baltes, ed., Springer Verlag, New York (1980).
33. B. R. Hunt, "The application of constrained least squares estimation to image restoration," *IEEE Trans. Comput.* C-22(9), 805-812 (1973).
34. S. J. Yeh, H. Stark and M. I. Sezan, "Hopfield-type neural networks: their set-theoretic formulations as associative memories, classifiers, and their application to image restoration," in *Digital Image Restoration*, A. Katsaggelos, ed., Springer Verlag, Berlin (forthcoming in 1990).
35. A. K. Katsaggelos, 'Iterative image restoration algorithms," *Opt. Eng.* 28(7), 735-748 (1989).
36. S. J. Reeves and R. M. Mersereau, "Optimal estimation of the regularization parameter and stabilizing functional for regularized image restoration," *Opt. Eng.* Vol. 29(5), 446-454 (1990).
37. R. L. Lagendijk, J. Biemond, and D. E. Boekee, "Regularized iterative image restoration with ringing reduction," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-36(12), pp. 1874-1888 (1988).
38. B. R. Frieden, "Restoring with maximum likelihood and maximum entropy," *J. Opt. Soc. Am.* 62(4), 511-518 (1972).
39. S. F. Gull and G. J. Daniell, "Image reconstruction from incomplete and noisy data," *Nature* 272, 686-690 (1978).
40. H. J. Trussell, "The relationship between image restoration by the maximum *a posteriori* method and a maximum entropy method," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-28(1), 114-117 (1980).
41. S. F. Burch, S. F. Gull and J. Skilling, "Image restoration by a powerful maximum entropy method," *Comp. Vision, Graph., Image Proc.* 23, 113-128 (1983).
42. H. Stark, ed., *Image Recovery: Theory and Application*, Academic, Florida (1987).
43. C. T. Chen, M. I. Sezan, and A. M. Tekalp, "Effects of constraints, initialization, and finite-word length in blind deblurring of images by convex projections," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, 1201-1204 (1987).
44. M. I. Sezan and A. M. Tekalp, "Adaptive image restoration with artifact suppression using the theory of convex projections," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-38(1) (1990).
45. D. C. Youla and V. Velasco, "Extensions of a result on the synthesis of signals in the presence of inconsistent constraints," *IEEE Trans. Circuits Syst.* CAS-33(4), 465-467 (1986).

46. R. M. Leahy and C. E. Goutis, "An optimal technique for constraint-based image restoration and reconstruction," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-34(6), 1629-1642 (1986).
47. R. W. Schafer, R. M. Mersereau and M. A. Richards, "Constrained iterative restoration algorithms," *Proc. IEEE* 69(4), 432-450 (1981).
48. H. J. Trussell and B. R. Hunt, "Image restoration of space-variant blurs by sectioned methods," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-26(6) 608-609 (1978).
49. G. M. Robbins and T. S. Huang, "Inverse filtering for linear shift-variant imaging systems," *Proc. IEEE* 60(7) (1972).
50. A. A. Sawchuck, "Space-variant image restoration by coordinate transformations," *J. Opt. Soc. Am.* 64(2), 138-144 (1974).
51. A. M. Tekalp and G. Pavlovic, "Space-variant and color image restoration using Kalman filtering," in *Proc. Int. Symp. Circ. Syst.* (1989).
52. P. L. Combettes and H. J. Trussell, "Methods for digital restoration of signals degraded by a stochastic impulse response," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-37(3), 393-401 (1989).
53. D. Slepian, "Linear least squares filtering of distorted images," *J. Opt. Soc. Amer.* 57(7), 918-922 (1967).
54. R. K. Ward and B. E. A. Saleh, "Deblurring random blur," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-35(10), 1494-1498 (1987).
55. L. Quan and R. K. Ward, "Restoration of randomly blurred images by the Wiener filter," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-37(4), 589-592 (1989).
56. A. M. Tekalp and M. I. Sezan, "Quantitative analysis of artifacts in linear space-invariant image restoration," 1(2) *Multidimensional Systems and Signal Processing* (1990).
57. A. M. Tekalp, H. Kaufman and J. Woods, "Edge-adaptive Kalman filtering for image restoration with ringing suppression," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-37(6), 892-899, (1989).
58. B. R. Hunt and O. Kubler, "Karhunen-Loeve multispectral image restoration, Part I: Theory,' *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-32(6), 592-599 (1984).
59. N. P. Galatsanos and R. T. Chin, "Digital restoration of multi-channel images," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-37(3), 415-421 (1989).
60. M. I. Sezan and H. J. Trussell, "Use of a priori knowledge in multispectral image restoration," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, 1429-1432 (1989).
61. A. M. Tekalp and G. Pavlovic, "Multichannel image modeling and Kalman filtering for multispectral image restoration," *Signal Process.* (1990).