The field of image restoration began primarily with the efforts of scientists involved in the space programs of both the United States and the former Soviet Union in the 1950s and early 1960s. These programs were responsible for producing many incredible images of the Earth and our solar system that, at that time, were unimaginable. Such images held untold scientific benefits which only became clear in the ensuing years as the race for the moon began to consume more and more of our scientific efforts and budgets. However, the images obtained from the various planetary missions of the time, such as the Ranger, Lunar Orbiter, and Mariner missions, were subject to many photographic degradations. These were a result of substandard imaging environments, the vibration in machinery and the spinning and tumbling of the spacecraft. Pictures from the later manned space missions were also blurred due to the inability of the astronaut to steady himself in a gravitationless environment while taking photographs. The degradation of images was no small problem, considering the enormous expense required to obtain such pictures in the first place. The loss of information due to image degradation could be devastating. For example, the 22 pictures produced during the Mariner IV flight to Mars in 1964 were later estimated to cost almost $10 million just in terms of the number of bits transmitted alone [83]. Any degradations reduced the scientific value of these images considerably and clearly cost the space agencies money.

This was probably the first instance in the engineering community where the extreme need for the ability to retrieve meaningful information from degraded images was encountered. As a result, it was not long before some of the most common algorithms from one-dimensional signal processing and estimation theory found their way into the realm of what is today known as “digital image restoration.”

The goal of this article is to introduce digital image restoration to the reader who is just beginning in this field, and to provide a review and analysis for the reader who may already be well-versed in image restoration. The perspective on the topic offered here is one that comes primarily from work done in the field of signal processing. Thus, many of the techniques and works cited here relate to classical signal processing approaches to estimation theory, filtering, and numerical analysis. In particular, the emphasis here is placed primarily on digital image restoration algorithms that grow out of an area known as “regularized least squares” methods. It should be noted, however, that digital image restoration is a very broad field, as we will discuss, and thus contains many other successful approaches that have been developed from different perspectives, such as optics, astronomy, and medical imaging, just to name a few.

In the process of reviewing this topic, we hope to address a number of very important issues in this field that are not typically discussed in the technical literature. The nature of these issues may be accurately summed up in these open questions to the image restoration research community: “Where have we been?”, “Where are we now?”, and “Where are we going?” Although these may seem questions too large to tackle in this forum, they are ones that warrant discussion now because of the relative maturity of the image restoration field. One indicator of this maturity is that reported improvements over tried-and-true algorithms in recent years might be...
considered quite small. Because of this, we would be well served now to take a step back and try to understand the contributions of the past and the needs of the future in order to best take advantage of the wealth of experience and knowledge in the area of digital image restoration.

Applications of Digital Image Restoration

The first encounters with digital image restoration in the engineering community were in the area of astronomical imaging, as previously mentioned. Ground-based imaging systems were subject to blurring due to the rapidly changing index of refraction of the atmosphere. Extraterrestrial observations of the Earth and the planets were degraded by motion blur as a result of slow camera shutter speeds relative to rapid spacecraft motion. Images obtained were often subject to noise of one form or another. For example, the astronomical imaging degradation problem is often characterized by Poisson noise, which is signal-dependent and has its roots in the photon-counting statistics involved with low light sources. Another type of noise found in other digital imaging applications is Gaussian noise, which often arises from the electronic components in the imaging system and broadcast transmission effects.

Not surprisingly, astronomical imaging is still one of the primary applications of digital image restoration today. Not only is it still necessary to restore various pictures obtained from spacecraft such as the space shuttle, but the well-publicized problems with the initial Hubble Space Telescope (HST) main mirror imperfections [87, 125] have provided an inordinate amount of material for the restoration community over the last few years. For example, Fig. 1 shows an HST picture of Saturn using the original Wide Field Planetary Camera (WFPC-I). The severe blurring was removed in Fig. 2 through digital restoration [42].

In the area of medical imaging, image restoration has certainly played a very important part. Restoration has been used for filtering of Poisson distributed film grain noise in chest X-rays, mammograms, and digital angiographic images [12, 32, 113], and for the removal of additive noise in Magnetic Resonance Imaging (MRI) [13, 88, 114]. Another emerging application of image restoration in medicine is in the area of quantitative autoradiography (QAR). In this field, images are obtained by exposing X-ray-sensitive film to a radioactive specimen. QAR is performed on post-mortem studies, and provides a higher resolution than techniques such as positron emission tomography (PET), X-ray computed tomography (CAT), and MRI, but still needs to be improved in resolution in order to study drug diffusion and cellular uptake in the brain. This can be accomplished through digital image restoration techniques [30]. Figure 3 shows a medical example of digital image restoration applied to an autoradiographic image of Cr-51 microspheres that are 10 microns in diameter. Figure 3(a) is the original image, and Fig. 3(b) is the restored image. The plot in Fig. 3(c) shows a line profile through the images demonstrating the improvement obtained through restoration. Here, an iterative restora-
tion algorithm that was formulated to consider the signal-dependent nature of film grain noise was used [30]. The full-width half maximum (FWHM) resolution of the microspheres was improved by 60% from about 259 microns to 103 microns.

Image restoration has also received some notoriety in the media, and particularly in the movies of the last decade. Ten years ago, the climax of the 1987 film "No Way Out," starring Kevin Costner, was based on the digital restoration of a blurry Polaroid negative image. The 1991 movie "JFK" made substantial use of a version of the famous Zapruder 8mm film of the assassination of President Kennedy, which has been enhanced and restored many times over the years. Similar restoration ideas showed up in the Michael Crichton book and subsequent 1993 film "Rising Sun," where researchers were needed to help restore the shadowy picture of a murderer from a surveillance videotape. Although some of these fictional uses of restoration were far-fetched, it is no surprise that digital image restoration has been used in law enforcement and forensic science for a number of years. For example, one of the most frequent needs for image restoration arises when viewing poor-quality security videotapes. In addition, the restoration of blurry photographs of license plates and crime scenes are often needed when such photographs can provide the only link for solving a crime. Such use of restoration is becoming more and more prevalent in our society. In fact, images restored in our laboratory were recently presented and accepted into evidence in a court of law for the first time by Dr. W. R. Oliver of the Office of the Armed Forces Medical Examiner [89]. Clearly, law enforcement agencies all over the world have made, and continue to make use of digital image restoration ideas in many forms.

Another application of this field which is especially important to our popular culture is the use of digital techniques to restore aging and deteriorated films. The idea of motion picture restoration is probably most often associated with the digital techniques used not only to eliminate scratches and dust from old movies, but also to colorize black-and-white films. For the purposes of this article, only a small subset of the vast amount of work being done in this area can be classified under the category of image restoration. Much of this work belongs to the field of computer graphics and enhancement. Nonetheless, some very important work has been done recently in the area of digital restoration of films. Some of the most interesting has been accomplished on animated films, such as the recent digital restoration of the film "Snow White and the Seven Dwarfs" by Walt Disney, which originally premiered in 1937 [22]. Though not restoring for blur degradation, the process used to correct for the cell dust, scratch and color fading problems with this original film could be classified as a form of spatially adaptive image restoration. There has been significant work in the area of restoration of image sequences in general as well, as discussed in [9, 10].

Perhaps the most exciting and expanding area of application for digital image restoration is that in the field of image and video coding. As techniques are developed to improve coding efficiency, and reduce the bit rates of coded images, artifacts such as blocking become quite a problem. Blocking artifacts are a result of the coarse quantization of transform coefficients used in typical image and video compression techniques. Usually, a discrete cosine transform (DCT) will be applied to prediction errors on blocks of 8 x 8 pixels. Intensity transitions between these blocks become more and more apparent when the high-frequency data is eliminated due to heavy quantization. Already, much has been accomplished to model these types of artifacts, and develop ways of restoring coded images as a post-processing step to be performed after decompression [70, 91, 102, 90, 129, 130]. In particular, very low bit rate coding applications such as mobile video communications impose bandwidth restrictions that require high compression. An example showing a still JPEG compressed image from a mobile video sequence at a compression ratio of 28:1 is shown in Fig. 4(a). Using a process based on mean field annealing and Markov Random Fields [91], a post-processed (restored) image is seen in Fig. 4(b). This image has most of the blocking artifacts removed, while still maintaining the important edges around the face in the picture. This idea of trading off smoothness and sharpness of an image in a spatially adaptive way forms the basis of regularization theory which is applied to the solution of the ill-posed restoration problem [118].

Digital image restoration is being used in many other applications as well. Just to name a few, restoration has been used to restore blurry X-ray images of aircraft wings to improve federal aviation inspection procedures [61]. It is used for restoring the motion induced effects present in still composite frames (produced by the superposition of two temporally spaced fields of a video image [77]), and, more generally, for restoring uniformly blurred television pictures [71]. Printing applications often require the use of restoration to ensure that halftone reproductions of continuous images are of high quality. In addition, restoration can improve the quality of continuous images generated from halftone images [34]. Digital restoration is also used to restore images of electronic piece parts taken in assembly-line manufacturing environments. Many defense-oriented applications require restoration, such as that of guided missiles, which may obtain distorted images due to the effects of pressure differences around a camera mounted on the missile. All in all, it is clear
that there is a very real and important place for image restoration technology today. Our task at hand now is to evaluate what types of applications may arise in the future and demand further innovation in this field. As a means of achieving this task, however, it is best to first understand the accomplishments of the past.

**Where Have We Been?**

A useful place to start is with a comprehensive definition of what digital image restoration is, and what it is not. Given such a definition, it will be easier to address the development of the various signal processing algorithms used for restoration, and to study how they affect the current trends in research.

Digital image restoration is a field of engineering that studies methods used to recover an original scene from degraded observations. It is an area that has been explored extensively in the signal processing, astronomical, and optics communities for some time. Many of the algorithms used in this area have their roots in well-developed areas of mathematics, such as estimation theory, the solution of ill-posed inverse problems, linear algebra and numerical analysis. Techniques used for image restoration are oriented toward modeling the degradations, usually blur and noise, and applying an inverse procedure to obtain an approximation of the original scene.

Image restoration is distinct from image enhancement techniques, which are designed to manipulate an image in order to produce results more pleasing to an observer, without making use of any particular degradation models. Image reconstruction techniques are also generally treated separately from restoration techniques, since they operate on a set of image projections and not on a full image. Restoration and reconstruction techniques do share the same objective, however, which is that of recovering the original image, and they end up solving the same mathematical problem, which is that of finding a solution to a set of linear or nonlinear equations. Some excellent treatment and review of different restoration and recovery techniques from a signal processing perspective can be found in these books and articles: [2, 8, 46, 48, 67, 109]. Much of the review material discussed here can be found with further detail in these references.

Developing techniques to perform the image restoration task requires the use of models not only for the degradations, but also for the images themselves. It will be valuable to study how some such models were used in the early applications and solutions in this field. Here, we will concern ourselves only with approaches based on digital techniques, although there have been significant efforts to restore degraded images through strictly optical and photographic means. There are a number of different ways in which to classify the many approaches to digital image restoration. One useful classification based on Deterministic and Stochastic approaches was given in Chapter 1 of [48]. In the second subsection below, we classify the well-known approaches to regularized least-squares restoration from the viewpoint of implementation technique, using the major categories: Direct, Iterative, and Recursive.

**Sources of Image Degradation**

In digital image processing, the general, discrete model for a linear degradation caused by blurring and additive noise can be given by the following superposition summation,

\[
y(i,j) = \sum_{k=1}^{M} \sum_{l=1}^{N} h(i,j; k, l)f(k, l) + n(i,j),
\]

where \( f(i,j) \) represents an original \( M \times N \) image, and \( y(i,j) \) is the degraded image which is acquired by the imaging system. In this formulation, \( n(i,j) \) represents an additive noise introduced by the system, and is usually taken to be a zero mean Gaussian distributed white noise term. In this article, we deal only with additive Gaussian noise, as it effectively models the noise in many different imaging scenarios. Many methods not detailed in this article utilize signal-dependent noise and lead to non-linear approaches to image restoration (see, for example, [62]).

In Equation (1), \( h(i,j;m,n) \) represents the two-dimensional point spread function (PSF) of the imaging system, which, in general, can be spatially varying. The difficulty in solving the restoration problem with a spatially varying blur commonly motivates the use of a stationary model for the blur. This leads to the following expression for the degradation system,

\[
y(i,j) = \sum_{k=1}^{M} \sum_{l=1}^{N} h(i-k, j-l)f(k, l) + n(i,j) = h(i,j) \ast \ast f(i,j) + n(i,j)
\]

where \( \ast \ast \) indicates two-dimensional convolution. The use of linear techniques for solving the restoration problem is facilitated by using this shift-invariant model. Models that utilize space-variant degradations are also common, but lead to more complex solutions.

An important aspect of image processing that deserves some mention here is that of the treatment of borders. The blurring process described by Equation (2) is linear. However, we often approximate this linear convolution by circular convolution, for mathematical reasons discussed later. This involves treating the image as one period from a two-dimensional periodic signal. The borders of an image are also often treated as symmetric extensions of the image, or as repeated instances of the edge pixel values. Such approaches seek to minimize the distortion at the borders caused by filtering algorithms which must perform deconvolution over the entire image. When implementing image restoration algorithms, it is very important to consider how the borders of the image are treated, as different approaches can result in very different restored images [2].

The following analytical models are frequently used in Equation (2) to represent the shift-invariant image degrada-
tion operator \([44, 67]\). The first two are encountered in the application of astronomical imaging mentioned before.

- **Motion Blur**: Represents the 1-D uniform local averaging of neighboring pixels, a common result of camera panning or fast object motion, shown here for horizontal motion,

\[
h(i) = \begin{cases} 
\frac{1}{L} & \text{if } \frac{-L}{2} \leq i \leq \frac{L}{2} \\
0 & \text{otherwise},
\end{cases}
\]  \(3\)

- **Atmospheric Turbulence Blur**: Common in remote sensing and aerial imaging, the blur due to long-term exposure though the atmosphere can be modeled by a Gaussian PSF,

\[
h(i, j) = K \exp\left(-\frac{i^2 + j^2}{2\sigma^2}\right),
\]  \(4\)

where \(K\) is a normalizing constant ensuring that the blur is of unit volume, and \(\sigma^2\) is the variance that determines the severity of the blur.

Photographic defocusing is also a problem in many different imaging situations. This type of blurring is primarily due to effects at the camera aperture that result in the spreading of a point of incoming light across a circle of confusion. A complete model of the camera’s focusing system depends on many parameters. These parameters include the focal length, the camera aperture size and shape, the distance between object and camera, the wavelength of the incoming light, and the effects due to diffraction \([7, 29]\). Accurate knowledge of all of these parameters is not frequently available after a picture has been taken. When the blur due to poor focusing is large, however, the following uniform models have been used as approximations of the PSF.

- **Uniform Out-of-Focus Blur**: This models the simple defocusing found in a variety of imaging systems as a uniform intensity distribution within a circular disk,

\[
h(i, j) = \begin{cases} 
\frac{1}{\pi R^2} & \text{if } \sqrt{i^2 + j^2} \leq R \\
0 & \text{otherwise},
\end{cases}
\]  \(5\)

- **Uniform 2-D Blur**: This is a more severe form of degradation that approximates an out-of-focus blur, and is used in many research simulations. This is the model for the blur used in the examples throughout this article,

\[
h(i, j) = \begin{cases} 
\frac{1}{L} & \text{if } \frac{L}{2} \leq i, j \leq \frac{L}{2} \\
0 & \text{otherwise},
\end{cases}
\]  \(6\)

where \(L\) is assumed to be an odd integer.

Usually, all blur-degraded images exhibit similar characteristics, namely a lowpass smoothing of the original image, attenuating the edge information which is very important for human visual perception \([86]\). In the process of trying to invert Equation \((1)\) to obtain an estimate of \(f(i,j)\), different artifacts may be introduced as a result of the characteristics of each blur operator. This issue will be discussed later. First, we will review some of the early or “classical” ways to perform the required inversion. As a tool for demonstrating these techniques, we can utilize an example of a synthetically blurred image which is often used for comparing results in the research literature. This image is referred to as the “Cam-eraman” image, and is seen in Fig. 5(a). Figure 5(b) shows the effects of a 7x7 uniform 2-D blur, at 40dB BSNR (Blurred Signal-to-Noise Ratio).

![Image](image_url)

**Some Classical Image Restoration Techniques**

In this section, we review a some of the many common approaches to image restoration that utilize minimum mean square error as an optimization criterion. The image degradation process is often represented in terms of a matrix-vector formulation of Equation \((1)\). This is given by

\[
y = Hf + n,
\]  \(9\)

where \(y, f,\) and \(n\) are the observed, original, and noise images, ordered lexicographically by stacking either the rows or the columns of each image into a vector. Assuming that the original image is of support \(N \times N\), then these vectors have support \(N^2\) x 1, and \(H\) represents the \(N^2 \times N^2\) superposition blur operator. When utilizing the stationary model of Equation \((2)\), \(H\) becomes a block-Toepplitz matrix representing the linear convolution operator \(h(i,j)\). Toeplitz, and block-Toeplitz matrices have special “banded” properties which make their use desirable for representing linear shift-invariant operators (see \([2]\) for further explanation of these matrices). By padding \(y\) and \(f\) appropriately with zeros so that the results of linear and circular convolution are the same, \(H\) becomes a block circulant matrix. This special matrix structure has the form

\[
H = \begin{bmatrix}
H(0) & H(N-1) & \cdots & H(1) \\
H(1) & H(0) & \cdots & H(2) \\
\vdots & \vdots & \ddots & \vdots \\
H(N-1) & H(N-2) & \cdots & H(0)
\end{bmatrix},
\]  \(10\)

where each sub-matrix \(H(i)\) is itself a circulant matrix.
BSNR

In most image restoration studies, the degradation modeled by blurring and additive noise is referred to in terms of a metric called the Blurred Signal-to-Noise Ratio (BSNR). This figure is defined in terms of the additive noise variance, \( \sigma_n^2 \), according to

\[
\text{BSNR} = 10 \cdot \log_{10} \left( \frac{1}{MN} \sum_{i,j} \left[ g(i,j) - \tilde{g}(i,j) \right]^2 \right) / \sigma_n^2
\]

for an \( M \times N \) image, where \( g(i,j) = y(i,j) - n(i,j) \) in Equation (1), and \( \tilde{g}(m,n) = E[x] \), which represents the expected value, or the mean, of \( g \).

For the purpose of objectively testing the performance of image restoration algorithms, the Improvement in SNR (ISNR) is often used. This metric is given by

\[
\text{ISNR} = 10 \cdot \log_{10} \left( \frac{\sum_{i,j} [f(i,j) - y(i,j)]^2}{\sum_{i,j} [f(i,j) - \hat{f}(i,j)]^2} \right)
\]

where \( f(i,j) \) and \( y(i,j) \) are the original and degraded intensity components, respectively, and \( \hat{f}(i,j) \) is the corresponding restored intensity field. Obviously, this metric can only be used for simulation cases when the original image is available. While mean squared error (MSE) metrics such as ISNR do not always reflect the perceptual properties of the human visual system, they serve to provide an objective standard by which to compare different techniques. However, in all cases presented here, it is important to consider the behavior of the various algorithms from the viewpoint of ringing and noise amplification, which can be a key indicator of improvement in quality for subjective comparisons of restoration algorithms.

\[
H(i) = \begin{bmatrix}
    h(i,0) & h(i,1) & \cdots & h(i,N-1) \\
    h(i,1) & h(i,2) & \cdots & h(i,0) \\
    \vdots & \vdots & \ddots & \vdots \\
    h(i,N-1) & h(i,0) & \cdots & h(i,1) 
\end{bmatrix}
\]

Notice that each block-row of \( H \) and each row of \( H(i) \) is a circular shift of the prior block-row or row, respectively. Representing \( H \) with a block circulant matrix could be further justified by using the result that the asymptotic distribution of the eigenvalues of a block Toeplitz and a block circulant matrix are the same [31]. The use of the block circulant approximation is important because it leads to desirable discrete frequency domain properties that can be used in solving Equation (9) [2]. These properties lead to efficient computation of inverse matrices, as discussed in the next section.

**Inverse Filtering.** Classical direct approaches to solving Equation (9) have dealt with finding an estimate \( \hat{f} \) which minimizes the norm

\[
\|y - H\hat{f}\|^2.
\]

thus providing a least squares fit to the data. This leads directly to the generalized inverse filter, which is given by the solution to

\[
(H^TH)\hat{f} = H^Ty.
\]

The critical issue that arises in this approach is that of noise amplification. This is due to the fact that the spectral properties of the noise are not taken into account. In order to examine this, consider the case when \( H \) (and, therefore, \( H^T \)) is block circulant, as described above. Such matrices can be diagonalized with the use of the 2-D Discrete Fourier Transform (DFT) [36]. This is because the eigenvalues of a block circulant matrix are the 2-D discrete Fourier coefficients of the impulse response of the degradation system which is used in uniquely defining \( H \), and the eigenvectors are the complex exponential basis functions of this transform. In matrix form, this relationship can be expressed by

\[
H = WHW^{-1}
\]

where \( H \) is a diagonal matrix comprising the 2-D DFT coefficients of \( h(i,j) \), and \( W^{-1} \) is a matrix containing the components of the complex exponential basis functions of the 2-D DFT. Pre-multiplication of both sides of Equation (14) by \( W^{-1} \), and post-multiplication by \( W \), or in this case, taking the 2-D DFT of the first row of \( H \), with the elements stacked into an \( N \times N \) image, gives the diagonal elements of \( H \).

Using this diagonalization approach, the matrix inverse problem of Equation (13) can be solved as a set of \( N^2 \) scalar problems. That is, using the DFT properties of block circulant matrices, and pre-multiplying both sides of Equation (13) by \( W^{-1} \), the solution can be written in the discrete frequency domain as

\[
\hat{F}(l) = \frac{\mathcal{H}^*(l)Y(l)}{\mathcal{H}(l)^2}.
\]

where \( \hat{F}(l) \), \( \mathcal{H}(l) \), and \( Y(l) \) denote the DFT of the restored image, \( \hat{f}(i,j) \), the PSF, \( h(i,j) \), and the observed image, \( y(i,j) \), as a function of the 2-D discrete frequency index \( l \), where \( l = (k_1,k_2) \) for \( k_1, k_2 = 0, \ldots, N-1 \), for an \( N \times N \) point DFT, and * denotes complex conjugate. Clearly, for frequencies at which \( \mathcal{H}(l) \) becomes very small, division by it results in amplification of the noise. Assuming that the degradation is lowpass, the small values of \( \mathcal{H}(l) \) are found at high frequen-
cies, where the noise is dominant over the image. For the frequencies where \( \mathcal{H}(l) \) is exactly zero (with respect to the accuracy of the particular computing environment being used) \( \hat{f}(l) \) is also equal to zero (which is the minimum norm least squares solution). Figure 6 shows the effects of a generalized inverse filter applied to the degraded image in Fig. 5(b). The restored image has been truncated to lie within the range of [0-255], however, the actual dynamic range of this image is much larger due to amplified noise. Clearly, this is not an acceptable restoration approach in this case.

In mathematical terms, the inverse problem represented in Equation (9) is ill-posed, if described in continuous infinite-dimensional space [85, 118]. In this case, the observation equation becomes a Fredholm integral equation of the first kind. The ill-posed nature of this problem implies that small bounded deviations in the data may lead to unbounded deviations in the solution. With respect to the discretized problem of Equation (9), the ill-posedness of the continuous problem results in the matrix \( H \) being ill-conditioned. It is interesting to note that the finer the discretization of the problem, the more ill-conditioned \( H \) becomes. Regularization theory is often used to solve ill-posed or ill-conditioned problems. The purpose of regularization is to provide an analysis of an ill-posed problem through the analysis of an associated well-posed problem, whose solution will yield meaningful answers and approximations to the ill-posed problem [48]. Techniques to accomplish this span a vast array of mathematical and heuristic concepts. In this section, some of the classical direct, iterative, and recursive approaches to this problem are reviewed.

**Direct Regularized Restoration Approaches.** Solving Equation (9) in a regularized fashion can lead to direct restoration approaches when considering either a stochastic or a deterministic model for the original image, \( f \). In both cases, the model represents prior information about the solution which can be used to make the problem well-posed.

**Stochastic Regularization.** Stochastic regularization can lead to the choice of a linear filtering approach that computes the estimate, \( \hat{f} \), according to

\[
\hat{f} = R^{-1} (Hf + R) = \frac{1}{\mathcal{H}(l) \mathcal{H}^*(l)} \frac{\mathcal{H}(l) \mathcal{H}^*(l) Y(l)}{S_g(l) \mathcal{H}^*(l) + S_n(l)}
\]

where \( S_g(l) \) and \( S_n(l) \) represent the power spectra of the original image and the noise, respectively.

Having these power spectra represents significant prior knowledge for the implementation of this filter. In most cases, however, the noise variance is known, or can be estimated from a flat region of the observed image [2]. In addition, it is possible to estimate \( S_g(l) \) in a number of different ways. The most common of these is to use the power spectrum of the observed image, \( S_y(l) \), as an estimate of \( S_g(l) \). Fig. 7 shows an example of a Wiener filter restoration of Fig. 5(b), using a periodogram estimate of the power spectrum computed from \( y \). The periodogram estimate of the power spectrum is simply defined according to [78]

\[
S_y(l) = \frac{1}{N^2} \left| Y(l) Y^*(l) \right|^2.
\]

**Deterministic Regularization.** The use of deterministic prior information about the original image can also be used for regularizing the restoration problem. For example, constrained least squares (CLS) restoration can be formulated by choosing an \( \hat{f} \) to minimize the Lagrangian

\[
\min \left[ \| y - Hf \|^2 + \alpha \| Gf \|^2 \right],
\]

where the term \( Gf \) generally represents a high pass filtered version of the image \( \hat{f} \). This is essentially a smoothness constraint which suggests that most images are relatively flat with limited high-frequency activity, and thus it is appropriate to minimize the amount of high-pass energy in the restored image. Use of the \( C \) operator provides an alternative
way to reduce the effects of the small singular values of $H$, occurring at high frequencies, while leaving the larger ones unchanged. One typical choice for $C$ is the 2-D Laplacian operator [37], given by

$$
C = \begin{bmatrix}
0.00 & -0.25 & 0.00 \\
-0.25 & -1.00 & -0.25 \\
0.00 & -0.25 & 0.00
\end{bmatrix}.
$$

(21)

In Equation (20), $\alpha$ represents the Lagrange multiplier, commonly referred to as the regularization parameter, which controls the tradeoff between fidelity to the data (as expressed by the term $\|y - H f\|^2$) and smoothness of the solution (as expressed by $\|C f\|^2$).

The minimization in Equation (20) leads to an equation of the form

$$
\hat{f} = \left(H^T H + \alpha C^T C\right)^{-1} H^T y.
$$

(22)

This also may be solved directly in the discrete frequency domain when block-circulant assumptions are used. The critical issue in the application of Equation (22) is the choice of $\alpha$. This problem has been investigated in a number of studies (see, for example, [25], and the references therein) and optimal techniques exist for finding an $\alpha$ given varying amounts of prior information about the noise and the signal. One way to use Equation (22), and choose $\alpha$ based on prior knowledge, follows a set theoretic approach [49, 46, 51]. With this method, a restored image is defined by an image which lies in the intersection of the two ellipsoids defined by

$$
Q_{hy} = \{f \mid \|y - H f\|^2 \leq \epsilon^2\},
$$

(23)

and

$$
Q_f = \{f \mid \|C f\|^2 \leq E^2\}.
$$

(24)

The equation of the center of one of the ellipsoids which bounds the intersection of $Q_{hy}$ and $Q_f$ is given by Equation (22) with $\alpha = (\epsilon/E)^2$. The same solution may be obtained with the Miller regularization approach [81].

Precise knowledge of both bounds $\epsilon^2$ and $E^2$ may not always be available. Several ways to estimate these bounds iteratively, based on the partially restored image at each step of the iteration, have been presented in [42, 41, 52, 43]. If the noise and signal variances are known or can be estimated, one choice is $\alpha = 1/\text{BSNR}$ [49, 46, 51]. Figure 8 shows an example of a CLS restoration applied to Fig. 5(b), using $\alpha = 1/\text{BSNR}$.

As an illustration of the behavior of the restored image in relation to the regularization parameter, Fig. 9(a) shows several direct CLS restorations at different choices for $\alpha$. Notice that with larger values of $\alpha$, and thus more regularization, the restored image tends to have more ringing, and yet with smaller values of $\alpha$, the restored image tends to have more amplified noise effects. This is seen in the corresponding error images shown in Fig. 9(b). The optimal solution, in the MSE sense, lies somewhere in the middle of the two extremes.

A more objective presentation of this idea can be seen in Fig. 10, where the variance, bias, and MSE of the direct CLS restoration filter are plotted as a function of $\alpha$. The variance and bias have been computed here in the frequency domain, as in [25], according to

$$
\text{Var}\left[\hat{f}(\alpha)\right] = \sigma_n^2 \sum_{i=1}^{N^2} \frac{|\mathcal{H}_i|^2}{\left(|\mathcal{H}_i|^2 + \alpha|c_i|^2\right)^2},
$$

(25)

and

$$
\text{Bias}\left[\hat{f}(\alpha)\right] = \sigma_n^2 \sum_{i=1}^{N^2} \frac{|F_i|^2 \alpha^2 |c_i|^4}{\left(|F_i|^2 + \alpha|c_i|^2\right)^2},
$$

(26)

where $\mathcal{H}_i$, $c_i$, and $F_i$ represent the 2-D discrete frequency components of the blur, Laplacian constraint operator, and original image, respectively, and $\sigma_n^2$ represents the variance of the additive noise. The bias of this estimator is a monotonically increasing function of $\alpha$, while the variance is a monotonically decreasing function of $\alpha$. Notice that the minimum MSE is encountered close to the intersection of these two curves, which is the point having equal bias and variance. Thus, one measure of objectively defining a good regularization parameter is to choose that $\alpha$ which gives the best compromise between these two types of errors. The properties of the bias and variance will be discussed in more detail in a later section.

While direct approaches solved in the frequency domain are among the most simple ways to restore noisy-blurred images, they are subject to a number of restrictions, most importantly the assumption that the image is globally stationary, and that a fair amount of prior information exists.

**Iterative Approaches.** Iterative image restoration algorithms have been investigated in some detail over the last decades (see, for example, [44, 46, 106], and the references therein, and [8, 49, 50, 51, 67]). The primary advantages of iterative techniques are that there is no need to explicitly implement the inverse of an operator and that the process may be monitored as it progresses. In addition, the effects of noise may be controlled with certain constraints, spatial adaptivity may be introduced, and parameters determining the solution

8. Result of Fig. 5(b) restored by a Constrained Least Squares filter, ISNR = 2.0 dB.
can be updated as the iteration proceeds [41-43]. Iterative algorithms are very well suited to restoring images suffering from a variety of degradations, such as linear, nonlinear, spatially varying, or spatially invariant blurs, and signal-dependent noise, because of the flexible framework provided by each approach, (see, for example, the work in [121, 122]).

One of the most basic of deterministic iterative techniques considers solving

$$ (H^T H + \alpha C^T C) f = H^T y $$

with the method of successive approximations [51]. This leads to the following iteration for $f$,

$$ f_{k+1} = f_k + \beta [H^T y - (H^T H + \alpha C^T C) f_k] $$

This iteration converges if

$$ 0 < \beta < \frac{2}{\lambda_{\text{max}}} $$

where $\lambda_{\text{max}}$ is the largest eigenvalue of the matrix $(H^T H + \alpha C^T C)$ [44]. This iteration is often referred to as the iterative CLS or Tikhonov-Miller method, depending on the way the regularization parameter $\alpha$ is computed. If the matrices $H$ and $C$ are block circulant, the iteration in Equation (28) can be implemented in the discrete frequency domain.

The termination criterion most frequently used compares the normalized change in energy at each iteration to a threshold such as
\[ \frac{\|f_{k+1} - f_k\|^2}{\|f_k\|^2} \leq 10^{-6}. \] (30)

Using this termination criterion, and a value of \( \beta = \left( \frac{2}{|r_{\text{max}}|} - 0.1 \right) \), Fig. 11 shows the result of the iteration of Equation (28) applied to Fig. 5(b), at convergence (after 321 iterations), using a value of \( \alpha = 0.0001 \), which is equal to \( \frac{1}{\text{BSNR}} \) in this case.

One of the benefits of an iterative approach such as this is that additional regularization may be obtained by terminating the iteration before convergence. The result shows some of the effects of stopping the iteration before it has converged to the solution obtained by the direct method. Early termination may be accomplished by monitoring the visual quality of the restored image at each iteration [44, 67]. Because the solution in the early iterations tends to have less amplified noise than at convergence, a "better" solution, based on subjective visual quality, may be obtained very simply by terminating the process. The choice of the regularization parameter, \( \alpha \), is still an issue with this approach, and it may be computed in a direct or iterative manner.

There are other deterministic techniques which can be used to perform iterative restoration as well. For example, the deterministic set-based approach described above can be generalized to form an iterative method called projections onto convex sets, or POCS, in which any number of prior constraints on a solution can be imposed as long as the constraint sets are closed convex [18]. Many techniques have used the POCS-based approach to perform iterative restoration with success in the past [93, 108, 110, 120, 130]. POCS has also been used in a recent stochastic based image recovery technique very successfully [107].

Other stochastic approaches also lead to iterative image restoration techniques. In particular, a large amount of research has been accomplished in the area of maximum likelihood solutions to this inverse problem. Some of this work has included various formulations of the Expectation-Maximization (EM) algorithm for image and blur identification and restoration [68, 66]. Such techniques are particularly effective when the blur operator and the signal and noise power spectra are unknown, as the identification process may be built into the iteration explicitly. When the degradation operator is known, but the signal and noise power spectra are not, an iterative Wiener filter results [35, 48] (Chapter 6), [68].

**Recursive Approaches.** Recursive filtering operations are beneficial because they also permit spatial adaptivity to be easily incorporated into the filter model [79]. In addition, they usually require less memory for storage than direct or iterative methods when reduced order models are used. The recursive equivalent of the Wiener filter is the discrete Kalman filter. This is a recursive filter based on an autoregressive (AR) parameterization of the prior statistical knowledge of \( f \). In the state space representation, consider that the global state vector for an \( M \times M \) image model, at pixel position \( (m,n) \), may be represented as

\[
\begin{align*}
& f(m, n) = \{ f(m, n), f(m, n - 1), \ldots, \\
& f(m - 1, n), f(m - 1, n - 1), \\
& \ldots, f(m - M + 1, n - M + 1) \}^T.
\end{align*}
\] (31)

The corresponding image model is then defined by

\[
\hat{f}(m, n) = A \hat{f}(m, n - 1) + w(m, n),
\] (32)

where \( w(m, n) = [1, 0, 0, \ldots, 0]^T w(m, n), \) \( w(m, n) \sim N(0, \sigma) \) (indicating that \( w(m,n) \) is Gaussian distributed with zero mean, and standard deviation \( \sigma \)), and \( A \) is an \( M \times M \) prediction matrix.

The Kalman filtering equations are reliant on the covariance of the error in the prediction governed by this model. They are also dependent upon an update based on the innovation term contributed by the new observation at each point in the recursion. Using notation defined thus far, the prediction and update terms for the Kalman filter are simply

**Prediction:**

\[
\hat{f}^+(m, n) = A \hat{f}(m, n - 1)
\] (33)

**Update:**

\[
P^+(m, n) = A P(m, n - 1) A^T + R_{\text{ww}}
\] (34)

\[
P(m, n) = [I - K(m, n) H] P^+(m, n)
\] (36)

\[
K(m, n) = P^+(m, n) H^T [H P^+(m, n) H^T + R_{\text{ww}}]^{-1}
\] (37)

where

\[
P^+(m, n) = E \left\{ (f(m, n) - \hat{f}^+(m, n))(f(m, n) - \hat{f}^+(m, n))^T \right\}
\] (38)

and

\[
P(m, n) = E \left\{ (f(m, n) - \hat{f}(m, n))(f(m, n) - \hat{f}(m, n))^T \right\}
\] (39)

In this model the observation noise and the model noise are modeled as zero mean Gaussian processes, with \( R_{\text{ww}} = E\{w w^T\} \) and \( R_{\text{ww}} = E\{w w^T\} \). Here, the observation state, \( y(m, n) \), is defined in a similar manner to \( f(m, n) \) in Equation (31).

Figure 12 shows the region of support of the global state vector extending into the past, which is updated at pixel \( (m,n) \).
in Equation (35). This figure also shows the pixels used in predicting the value of the current pixel. This model uses a non-symmetric half plane (NSHP) region of support to arrive at the predicted value of the current pixel. Reducing and keeping constant the support of the global state vector results in a suboptimal Kalman filter, the reduced update Kalman filter (RUKF) [127, 128]. The computation of this Kalman filter, however, becomes very efficient without sacrificing performance, due to the local extent of the correlation exhibited by most images. While the RUKF assumes that the gain is zero outside the local state, the reduced order model Kalman filter (ROMKF) [3] is based on the reduction of the dimension of the state vector which requires a modification to the state-space equations. Thus, the RUKF is a suboptimal Kalman filter based on the original 2-D model, whereas the ROMKF is the optimal Kalman filter based on modified state-space equations. Using an NSHP model estimated from a prototype image, a RUKF filter restoration of Fig. 5(b) is shown in Fig. 13. Notice that this restoration exhibits considerable ringing. This is primarily caused by the use of a constant prediction model here, which is invalid at the many edges in the scene. Moreover, this restoration illustrates the sensitivity of the Kalman filter to the parameters used in this prediction-update approach. Because of this, adaptive models are best used in the recursive filtering framework provided by Kalman filtering.

Image Restoration Artifacts

Some of the most common image restoration artifacts stem from global restrictions placed on the restored image by linear shift invariant (LSI) techniques. The two most prevalent artifacts are ringing around the edges in a restored image, and filtered noise causing false texturing in the flat regions of the image. The ringing artifacts are a function of the blur operator and the restoration filter, and are dependent on the image. On the other hand, the magnified noise effects are a property of the bandpass nature of the restoration filter, and are independent of the image. Assuming that the blur operator is lowpass with a sharp frequency cut-off (as found in motion blur and out-of-focus blurs), there will be zeros in the transfer function of this operator. Regularization essentially provides a means to reduce the effects of these zeros, which become poles in the restoration filter. Although this leads to reduced noise amplification in the neighborhood of the zero frequencies, it also results in the increase of an image-dependent component in the restoration error. To see this clearly, consider an LSI restoration filter, $G$. The restored image may be written as

$$
\hat{f} = Gy = GHf + Gn
$$

$$
= f + (GH - I)f + G_{n}
$$

$$
= f + e_{e} + e_{n}.
$$

(40)

So, the restored image comprises the original image, and two error terms, $e_{e} = (GH - I)f$, and $e_{n} = G_{n}$. Considering this problem in terms of its discrete frequency domain counterpart, at the spatial frequencies in the original image corresponding to zero frequencies in $H$, $(GH - I)$ will equal -1 and $e_{e}$ will be stimulated, giving rise to periodic ringing artifacts [44, 117].

Ringing is very evident in Figs. 7, 11, and 13. These results are highly regularized, a characteristic that tends to increase the image-dependent component of the error, since $GH$ deviates considerably from the identity. The tradeoff between $e_{e}$ and $e_{n}$ is the crucial issue in regularized image restoration, and has been analyzed in a number of different studies [65, 117]. One recent analysis of this problem treated the manifestation of these errors as the bias and variance of a CLS estimator, as shown in Fig. 10, and demonstrated the effects of the regularization parameter in terms of these image-dependent and image-independent error terms [25]. As the value of $\alpha$ becomes very small, the problem becomes underregularized, and the noise magnification component of the error dominates the solution. This is the case in Fig. 8, where the loose bound choice of the regularization parameter leads to a sharp but somewhat noisy result. The restrictions of LSI restoration filters naturally lead to the choice of spatially adaptive approaches to reduce some of the most common restoration artifacts. These approaches generally require some relaxation of the stationarity assumptions made.

So, in answer to the question, “Where have we been?” it is clear that we have been many places. We have learned how to accurately model the types of degradation found in images due to the environment used in obtaining them. We have applied many of the most fundamental techniques from estimation theory and numerical analysis to obtaining solutions for this problem. We have encountered a variety of linear filtering approaches along with iterative and recursive approaches to the solution. We have also studied the image restoration problem from the perspective of an appropriate
model for the image, from both a stochastic and deterministic viewpoint. Additionally, we, as a community, have implemented in fast software solutions all of the fundamental approaches to spatially invariant linear restoration.

Much progress has been made in the last five or ten years toward treating the restoration problem in a more nonstationary way. That is, the added benefits of spatially variant restoration have been explored from many angles, producing even better results than those that can be obtained with the techniques discussed in this section. We will briefly discuss some of the most important of these advances in next section.

Where Are We Now?

The algorithms described in the previous section represent the foundation of the approaches to the restoration problem today. They are successful approaches, and they have been applied to many different image restoration problems. Today, we find that a variety of new techniques are being investigated that attempt to improve upon these approaches, and use them as their fundamental basis. A good deal of the newer research is motivated by the desire to find ways to reduce the artifacts mentioned in the previous section. Additionally, special applications have also driven the development of new image-restoration approaches. In this section, we briefly discuss and cite some of these approaches.

Most of the algorithms in the realm of classical image restoration deal with some global assumptions about the behavior of an image, whether from the stochastic viewpoint of stationarity or a deterministic view of smoothness. Newer successful techniques which address the problem with such assumptions include the use of robust functionals [133] and total least squares [80]. However, spatially adaptive or nonstationary approaches have also been developed to alleviate some of the problems associated with such rigid global restrictions.

Spatially Adaptive Approaches

Spatially adaptive algorithms frequently incorporate the properties of the human visual system [1, 50, 51, 60, 97]. One such property is the reduced sensitivity to noise in regions with high spatial activity, such as edges. Because the visual system is also sensitive to sharp changes in an image, it is not desirable to smooth over the edges when performing restoration. Therefore, the application of a different restoration filter at each spatial location is desirable. Such filters should vary between the inverse filter (applied at sharp edges) and an over-regularized filter (applied in the flat regions). With a set theoretic restoration filter obtained from Equations (23) and (24), the spatial adaptivity can be introduced with the use of weighted norms [44, 45, 65]. The weight matrices are chosen to represent the masking and visibility functions[1] in [44, 50, 45]. An iterative algorithm is generally employed in the case of spatially adaptive restoration. The weight matrices can then be kept fixed, or be adapted at each iteration step based on the partially restored image [53].

Recursive methods can accommodate spatial adaptivity when using a stochastic model of the image, by changing the parameters of the model at the edges. For example, in [128] a local decision process was proposed to switch between the AR models that capture the correlation of the edges present at different spatial locations. In [39], an AR model was used that was driven by white noise with a space-variant variance to model the residual image. The multiple model approach can also lead to a reduction in ringing artifacts around the edges in a restored image, as in [117].

Maximum a posteriori (MAP) probability methods have also been proposed for nonstationary image restoration. These methods utilize space-variant density functions as prior knowledge to capture the nonstationarity of the original image. In [15, 26, 40], doubly stochastic Markov random fields were used as prior densities, and stochastic relaxation was used to minimize a nonconvex objective functional in each method.

Color Image Restoration

The problem of color image restoration presents a unique difficulty in that the multiple color channels are related. Thus, cross-channel correlations need to be exploited in order to achieve optimal restoration results. A number of approaches have been used to handle not only the color multichannel image restoration problem, but also other inherent multichannel problems such as image sequence restoration. Many such approaches can be found in [23, 24, 57, 92, 119, 135].

Neural Networks

Another interesting emerging approach in the image restoration realm is the use of neural networks [21, 94, 112, 131, 134]. Neural networks are especially well-suited to the image restoration task because they can effectively adapt to the local nature of the problem. They may also be used to realize well-known algorithms without the need for extensive underlying assumptions about the distribution of the parameters being estimated. They may also be used to estimate the regularization parameter in the CLS approach, and can be developed to alternate between learning and restoration cycles. Finally, neural processing techniques have recently led to efficient VLSI architectures for image restoration due to their highly parallel nature [55, 56, 69, 104].

Astronomical Image Restoration

Some approaches finding extensive application to the astronomical imaging problem described earlier are maximum entropy-based methods [124], hierarchical Bayesian approaches [82], modified Richardson-Lucy iterative approaches [126, 101, 72], and more recently preconditioned conjugate gradient methods [84]. These iterative approaches are well-suited to dealing with the unique problems of this application, in particular the signal-dependent nature of the noise. Also particularly effective at restoring astronomical
data, and maintaining photometric integrity, are a new class of general regularized iterative constrained least-squares techniques that evaluate the regularization parameter as a function of iteration in both the spatial and frequency domains [54]. For a useful review of current techniques in astronomical image restoration, see [87].

Wavelets

The use of wavelets for the task of image restoration and enhancement is a relatively new but rapidly emerging concept since their appearance in the image processing literature [74]. Although there has long been the view that a nonstationary approach may improve results substantially over those using stationary assumptions, the idea of multiresolution has not been a prevalent one. Instead, past adaptive restoration techniques, for example [46, 51, 117, 128], have examined the problem in the spatial domain, using various local measures to describe the type of activity near a pixel. However, a number of researchers are now beginning to analyze enhancement problems [17, 19, 75, 76, 111, 116], and restoration and recovery problems [11, 14, 59, 103, 115, 116] from the multiresolution/subband perspective.

One such example of the use of wavelets for restoration can be found in [4], where a new matrix formulation of a wavelet-based subband decomposition was presented. This formulation allows for the computation of the decomposition of both the signal and the convolution operator in the wavelet domain. This permits the conversion of any linear single channel space-invariant filtering problem into a multichannel one. In particular, this approach can be used to restore a single channel image with multichannel image restoration routine, like the approach to color image restoration followed in [16, 24, 38].

Also utilizing the wavelet concept, in [5], a new spatially adaptive restoration approach which uses a multiresolution Kalman smoothing filter [6, 73] was discussed. This filter was applied directly to the wavelet coefficients of a noisy image ordered onto quadtree structures, as seen in Fig. 14. In order to obtain the noisy wavelet coefficients, the observed image is first pre-filtered with a constrained least-squares filter having a small regularization parameter. This leaves the restored image under-regularized, and thus noisy. Subsequent wavelet domain filtering is applied in a spatially adaptive manner to remove this noise. A coupled approach is described in [5] to jointly optimize the pre-filtering with the wavelet-domain Kalman filtering.

The proposed method has the benefit that the majority of the regularization, or noise suppression, of the restoration is accomplished by the efficient multiscale filtering of wavelet detail coefficients ordered on quadtrees. Not only does this lead to potential parallel implementation schemes, but it permits adaptivity to the local edge information in the image. Figure 15 shows the result of such an approach applied to Fig. 5(b). This wavelet-based approach provides a particularly useful method for image restoration when the preservation of edges in the scene is of importance. Because the local adaptivity is based explicitly on the values of the wavelet detail coefficients, it is also efficient to implement and requires minimal added complexity over the initial deconvolution filter.

Many of the emerging identification and restoration techniques discussed here deal with the nonstationary nature of images, so they tend to be more complex than the classical methods presented previously. This is a natural result of the effort to find ways to improve on standard approaches. Fortunately, the computational resources available to today's researchers are generally sufficient to handle algorithms of much greater complexity than those previously used.

Where Are We Going?

It is clear that a vast array of knowledge is available to those who wish to restore degraded images. The question remains, however, as to what will become of the field known as digital image restoration. The ideas expressed in this section are merely a suggestion of what might be some avenues of exploration. There is certainly a great deal of knowledge to be obtained by continued research on the application of theoretical estimation techniques to this problem, but it seems appropriate to also consider the factors that will drive the continued research in this field. We can use these factors to lend insight to those seeking out new research projects in the area.

As we move into the next phase of research in this field, it will be useful to understand those areas in which the model of the degradation and restoration process can be improved. In Fig. 16, the various components of the most common degradation/restoration model are shown. Each of the elements of this model have been discussed in this article. It is evident from this figure and previous discussion here, that the amount of prior knowledge plays a major part in achieving the best restorations. Fig. 17 breaks down the three key areas that benefit from prior knowledge in the identification, estimation, and restoration processes. These are knowledge of the degradation, knowledge about the original image, and
Blur Identification

Most classical techniques assume that the convolution operator representing the blur is known a priori. This, however, is almost never the case in practical imaging situations. Thus, the area of blur identification, or blind deconvolution, is a very important subset of image restoration in which much important work has been performed (see [63, 64] for a recent review of this topic in this magazine).

The EM algorithm, for example, is an important tool in image restoration and blur identification that is used to solve for a maximum likelihood estimate of the parameters under consideration. As mentioned in an earlier section, the EM algorithm is one technique that has been successfully used to simultaneously identify the blur and restore an image [66, 68].

There are also several other emerging techniques that have been used to address this identification problem. The first of these is the generalized cross-validation approach [99]. This method considers a parameterized estimate of the blur based on the minimization of a restoration residual over the image and blur parameter space. This estimate excludes a single observed data point at a time. It is sometimes referred to as the “leave one out” method, as the accuracy of the estimates are analyzed on the basis of a prediction of the data left out at each step. An added benefit is that a Gaussian noise assumption is not required with cross-validation. The cross-validation approach has also been used for the estimation of the regularization parameter in [25, 98].

Another approach to the blur identification problem is based on the idea of residual spectral matching [105]. This consists of selecting the best PSF estimate from a pool of candidates that provides the best match between the restoration residual power spectrum and its expected value, assuming that the selected PSF is the correct one. This approach uses a Wiener filter as the restoration filter, and thus requires knowledge of the noise variance and the original image power spectrum. The residual spectral matching method is essentially a statistical analysis approach to the blur identification problem. Other work that treats the PSF as a stochastic signal has been examined using a maximum likelihood estimate of the blur obtained from differently blurred versions of an image with measurement errors in the PSF [123]. Another technique was developed to eliminate the use of prior knowledge of the extent of the blur [96]. This method considers modeling the blur in continuous spatial coordinates, permitting the likelihood function to be minimized with respect to the extent of the blur.

In a number of applications, more than one blurred version of an original image may be available. Improved restoration results can be obtained in such cases when the blurs are known, as reported in [58, 27, 47], but, more recently, this has been the case when the blurs are unknown [28, 33]. Finally, a new approach has recently been developed that jointly solves the blur identification and image restoration problem for a single image by minimizing a cost function based on a restoration error measure, a regularization term for the image, and a regularization term for the blur [132].

16. Image degradation model, and restoration process.

knowledge about the noise. Researchers are now attempting to improve the models used in identification and restoration by incorporating better prior knowledge into the problem. For example, it has been shown that digital restoration may fail when incomplete system models are used [95]. One example of an incomplete system model is that which excludes the image formation, or image gathering, process [20]. In moving from the continuous to the discrete domain, the sampling and reconstruction procedures may introduce enough distortion in there own right to influence a poor digital restoration. By addressing the continuous to discrete step directly in the system/restoration model, the degradation is better modeled, and improvements can be made over the incomplete system approach. Though more complex than classical approaches, techniques of the future will likely make use of more complete models to provide better image restorations. However, these complete models will be tied in a critical way to the application at hand.

From the viewpoint of applications, the future of image restoration depends upon the needs of a variety of video technologies that will require processing of digital images for a number of reasons. We already mentioned a number of applications in today’s world that make ready use of image restoration techniques. Some of the most interesting of these
lie in the area of consumer products. Consider, for example, hand-held video cameras that provide options such as digital focusing and adjustment for camera jerkiness. These digital enhancements to a widely accepted consumer product illustrate the importance of applying some of the well-developed approaches from the restoration and enhancement field in a useful manner. Many more digital consumer products that utilize digital images and video are clearly on the way today. Items such as personal communication systems (PCS) that utilize video may provide a great deal of opportunity for the application of restoration ideas to coded images. It should be important to obtain a perspective on the future of image restoration which considers the driving importance of consumer applications.

In the same way, a second area of application that will continue to draw on the expertise of image restoration experts is medical imaging. Already, a great deal has been done to investigate the application of restoration techniques to the problem of degraded image quality in this area. The technology that advances the field of medicine is critically linked to the world of image processing. One prominent piece of evidence that illustrates this is the emergence of "tele-medicine," which considers diagnosis from imaging over long distances through various forms of telecommunications. It is apparent that such remote diagnoses will rely heavily on the image quality at reception, and we may find that restoration has an important role to play in this and other similar applications in the medical realm today.

Of those emerging techniques that seem to hold promise for further research, blur identification seems to be one area that requires a good deal of research. The most critical fault of most of the algorithms discussed in this review is that they assume knowledge of the PSF of the imaging system. Although certain applications have special circumstances that permit good approximations of the PSF (such as those with blurry star field images), most real situations have difficult, spatially varying blurs. This makes studies into the blur identification problem all the more important to the future of this field. It may have been assumed 10 years ago that it would not be possible to take a blurry picture because of the great promise of digital image restoration. But the restoration community’s lack of persistent attention to the identification problem has made this assumption false. While blur identification is clearly the most difficult problem within this area, it is one that will hopefully provide considerable material for future researchers to come.

Spatially varying restoration methods such as those that use wavelets and Markov Random Fields also hold potential for further algorithmic research. Many results have been shown thus far that really preserve the edge-like nature of restored images using these techniques [100]. Again, this is an area that may still provide fertile ground for continued restoration research.

As mentioned in the beginning of this article, image restoration has matured considerably over the last several decades. The maturity of the field requires us to take an objective look at where restoration will go, and what contributions remain to be made in order to most effectively utilize the existing knowledge in this area. The discussion here has given some hints as to what directions we might take, but only offers one perspective on the substantial work in the field. Researchers coming from a tradition outside of signal processing, such as astronomical and medical imaging, certainly have additional and valuable viewpoints on this topic that will also help to guide the direction of future research in the field. Nonetheless, there remains a good deal of doubt about where the next major advances will come. Clearly, there is a real opportunity for interested researchers to begin shaping the next ten years, and the next century of digital image restoration right now.

Acknowledgment

The authors wish to thank Prof. Moon Gi Kang of the University of Minnesota, Duluth, and Dr. John Goyette of XiTec Inc. for their valuable contributions to this manuscript. In addition, the authors thank the anonymous reviewers of this article for their useful suggestions and diverse perspectives on the field of digital image restoration.

Mark R. Banham is a staff engineer with Digital Technology Research Laboratory, Chicago Corporate Research Laboratories, Motorola, Schaumburg, IL 60196-1078 (banham@csrl.motorola.com). Aggelos K. Katsaggelos is a professor with the Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL 60208-3118 (aggk@ece.nwu.edu).

References


