We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,800
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
1. Introduction

Image restoration or deconvolution of a blurred natural image is a mature research activity with a rich set of available techniques and algorithms, well-summarised in review articles, Banham & Katsaggelos (1997); Kundur & Hatzinakos (1996). Despite this history and volume of work, there is current research activity motivated by the desire to find yet superior methods to restore the ground truth image (GTI). Important performance metrics to assess the efficacy of restoration methods include: restoration accuracy, computational complexity and convergence speed. In this chapter we use these performance metrics in the development of restoration methods of greatest utility for real-world applications where complexity/speed is a major concern and the evaluation of image restoration needs to take into account the highly structured features of natural images and, to a lesser extent, the human visual system.

The scope of this work focusses on non-blind image restoration where the point spread function (PSF) of the blur convolutional kernel is known. Blind deconvolution is, by its nature, a more challenging problem, Haykin (1994); Kundur & Hatzinakos (1996). However with effective and efficient PSF estimation techniques, Fergus et al. (2006); Joshi et al. (2008); Krahmer et al. (2006); Nayar & Ben-Ezra (2004); Oliveira et al. (2007), the research trend has been to handling blind deconvolution in two steps, with PSF estimation as the first step and image estimation as the second step, Levin et al. (2009). This motivates us to focus on efficient algorithms for image restoration where the blur convolutional kernel is known.

In this chapter, we first analyze existing linear deterministic restoration models and develop a class of novel models with better performance. Then using regularization as the basis, we link linear deterministic and stochastic restoration models. By introducing a previously developed novel visual metric to image regularization analysis, we study the purported superior performance of stochastic prior models and demonstrate that those models are not superior to simpler linear deterministic prior models. In addition, we show that the high complexity “derivative likelihood” models under the maximum a posteriori (MAP) framework offer no significant advantage to a properly configured, efficient “normal likelihood” model.

2. Quadratic regularization in image restoration

2.1 Regularization

Image acquisition being an inverse problem can be modeled by a continuous model in an infinite dimensional space, which is categorized as a (linear) Fredholm integral equation of
the first kind, Demoment (1989). In the sense of Hardamard, Hadamard (1952), a solution
to a well-posed problem satisfies the conditions of existence, uniqueness and stability. As
Fredholm integral equations of the first kind do not meet the criteria for a well-posed problem,
image restoration belongs into the general class of problems which are classified as ill-posed
implies that, small bounded perturbations in the data may lead to unbounded deviations in
the solution, Phillips (1962).

For images defined on a discrete set, linear algebra can be used to find solutions for
ill-posed problems such as image restoration. One of the simplest methods to restore images
affected by a linear distortion is the use of the pseudo inverse, Albert (1972), for which the
solution fulfils the first two conditions (existence and uniqueness) of Hardamard’s well-posed
problem, Hadamard (1952), but fails in meeting the stability condition. This motivates or
leads to regularization as one of the most widely accepted and used techniques, in which
the solution fulfils all three conditions of a well-posed problem. The concept underlying
regularization is to find an acceptable solution from imperfect data, for which, the problem
should be stated more completely by including some extra or priori information, Miller (1970);

Regularization approaches to image restoration are classified broadly in two ways: stochastic
regularization which uses the knowledge of covariance matrices of the GTI and noise; and
deterministic regularization which deems that most natural images are relatively featureless
with limited high-frequency activity, Banham & Katsaggelos (1997). While stochastic
regularization has been used extensively in the past, with important contributions to the field
such as Wiener filter, Wiener (1942), recently, much emphasis has been on the use of derivative
filters with deterministic regularization, Fergus et al. (2006); Levin et al. (2007). Thus, our
contribution in this chapter relates to deterministic regularization and the term regularization,
henceforth, refers to deterministic regularization techniques.

Among many regularization techniques, Tikhonov, Tikhonov & Arsenin (1977) regularization
is one of the first and best-known techniques for stabilization. It was proposed in Tikhonov &
Arsenin (1977), that the solution for

\[ b = Kg + n, \]

where \( b \) is the measured data, \( g \) is the original data (ground truth), \( K \) is the distortion
operator or the transformation and \( n \) represents additive random noise, can be achieved
by constrained minimization of a functional \( \Phi(g) \), which is called the stabilizing functional.
Under the stabilizing functional approach, the image restoration problem is formulated as
determining an estimate \( \hat{g} \) of \( g \), which minimizes the functional \( \Phi(g) \) under the condition that
the estimate \( \hat{g} \) satisfies

\[ \| b - K\hat{g} \|^2 = \delta, \]

where \( \delta \) is a positive constant and \( \| \cdot \| \) denotes the Frobenius norm

\[ \| A \| = \sqrt{\sum_{i=1}^{m_1} \sum_{j=1}^{m_2} a_{ij}^2} \]
for some matrix \( A \) and \( a_{ij} \) is the \((i,j)\) entry. The constrained minimization problem in (2) can be solved by the method of Lagrange multipliers, which is to determine \( \hat{g} \), an estimate of ground truth \( g \), by minimizing the functional

\[
\| b - K \bar{g} \|^2 + \lambda \Phi(g),
\]

where \( \lambda > 0 \) is the Lagrange multiplier and is often called as the regularization parameter. As the regularization parameter, \( \lambda \) controls the tradeoff between the solution accuracy \( \| b - K \bar{g} \|^2 \) and its degree of regularity \( \Phi(g) \), choosing a proper value for \( \lambda \) is important in image restoration.

The first term in (4), named as data-fidelity term fits to the data, while stabilizing functional incorporates “believed” properties of the GTI. Generally the data-fidelity term is a standard fixed choice. In contrast, the richness and variety of image restoration techniques comes down to different choices of the regularization term, reflecting different implicit models. As the choice of the stabilizing functional can take a variety of forms, in this chapter, we selected two widely used model classes for our analysis: the fast quadratic stabilizing functions introduced in section 2.2, and Sparse and Laplacian prior methods in section 3.2. The latter model class can be developed from relating the stabilizing functional to the priori knowledge using a probabilistic viewpoint and is claimed to have better performance, Levin et al. (2007).

When the stabilizing functional \( \Phi(g) \) in (4) belongs to the class of nonnegative quadratic functionals, the minimization problem can be expressed as

\[
\hat{g} = \arg \min_g \| b - K \bar{g} \|^2 + \lambda \| D \bar{g} \|^2,
\]

where \( D \) is a bounded linear operator, Miller (1970) and is often called the regularization operator or stabilizing operator. It is shown in Hunt (1973), that the minimization problem in (5) can be formulated as a constrained least squares image restoration problem when the solution \( g \) satisfies the necessary and sufficient condition of

\[
(K^T K + \lambda (D^T D)) \hat{g} = K^T b.
\]

This leads to the closed form solution for (5) in the form

\[
\hat{g} = (K^T K + \lambda (D^T D))^{-1} K^T b.
\]

We extend, in a trivial way, the above formulation by considering \( D \) as the combination of \( R \) component regularization operators, in the form of

\[
D \bar{g} \triangleq (D_1^T, D_2^T, \ldots, D_R^T)^T \bar{g}.
\]

With the introduction of \( R \) Lagrange multipliers, the general form of (5) can be expressed as

\[
\hat{g} = \arg \min_g \| b - K \bar{g} \|^2 + \sum_{r=1}^{R} \lambda_r \| D_r \bar{g} \|^2,
\]

for which, the closed form solution is given by

\[
\hat{g} = (K K^T + \sum_{r=1}^{R} \lambda_r (D_r^T D_r))^{-1} K^T b.
\]
As the images are of limited support and when the corresponding hypothesis of uniformity on image edges can be made, the matrices $K$ and $D$ in (10) have a special structure and are called block circulant matrices, Hunt (1971). As circulant matrices can be diagonalized by the discrete Fourier transform, the minimization in (9) can be solved extremely quickly using the Fourier domain techniques, Hunt (1973).

### 2.2 Regularization operators as components in quadratic stabilizing functionals

The generality of the regularization operator allows the development of a class of linear operators and the minimization in (9) will be the source of many regularizing solutions for (1) depending on the choice of the regularization operator. This choice is usually based on the known details of the image formation process and plays an important role in the regularization.

The simplest regularization operator is when $D$ is an identity matrix, where $Dg = g$, and the regularized solution for this was referred as minimum norm restoration, Hunt & Andrews (1977). In general, $D$ often takes the form of a sparsifying operator such as a discrete approximation of a derivative operator. Through the experiments in Zhu & Mumford (1997), it was shown that even though the statistics of natural images vary from image to image, the histograms for the response of derivative filters are relatively consistent and scale invariant across the images. Taking these factors into consideration, in this section, we discuss a class of regularization operators based on the partial derivative operators (PDO), which could be used in the quadratic stabilizing functional.

#### 2.2.1 First order partial derivative operator

When first order partial derivative operators (FOPDO) are considered as the regularization operators, $Dg$ in (8) can be expressed as

$$Dg = \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} g,$$

where $\partial_x$ and $\partial_y$ are any discrete space, spatially invariant linear operators that emulate first order derivative in $x$ and $y$ directions, respectively, Levin et al. (2007). This type of regularization uses two component regularization operators.

#### 2.2.2 Second order partial derivative operator

Second order partial derivative operators (SOPDO), can be derived mainly in two forms.

1. **Isotropic SOPDO** – When the regularization operator takes the form

$$Dg = \begin{pmatrix} \partial_{xx} \\ \partial_{yy} \end{pmatrix} g,$$

it is called the isotropic SOPDO. Though the SOPDO defined above cannot be considered as a true isotropic differential operator, such as the continuous Laplacian operator, it gives the simplest possible isotropic operator with even-order derivatives, Leung & Lu (1995). Similar to FOPDO, $\partial_{xx}$ and $\partial_{yy}$ represent any discrete space, spatially invariant linear operators that emulate second order derivatives.
2. Non-isotropic SOPDO – The non-isotropic SOPDO is formed by
\[ \mathcal{D}g = \begin{pmatrix} \partial_{xx} \\ \partial_{xy} \\ \partial_{yy} \end{pmatrix} g. \]

As the edges and lines in images may occur in any direction, when the differential operator is isotropic it would give better results than a non-isotropic differential operator, Leung & Lu (1995).

\section*{2.2.3 Mixed partial derivative operator}

In general, considering only even-order derivatives, the use of directional derivatives in more than one dimensional can be expressed as
\[ \mathcal{D}g = (\partial_{s_1}^{p_1}, \partial_{s_2}^{p_2}, \ldots, \partial_{s_m}^{p_m}) g. \] (11)

where \( p \) is the order of the derivatives, \( m \) is the number of dimensions and \( s_1 \) to \( s_m \) represent the direction of the derivatives.

Using the above general model, we introduce a new regularization operator, with different combinations of higher order derivative operators. In this discussion, we limit the use of higher order derivative operators up to the second order, and the new PDO is called first and second order derivative operator (FSOPDO). With FSOPDO, \( \mathcal{D}g \) in (11) takes the form
\[ \mathcal{D}g = \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_{xx} \\ \partial_{yy} \end{pmatrix} g. \]

These quadratic regularization functionals are compared in a new perspective with the widely used prior models which are believed to result in better performance in section 3.5.

\section*{2.3 Non-blind image restoration through SOPDO}

\subsection*{2.3.1 Noisy image deconvolution}

Although most previous image restoration algorithms have considered FOPDO as the regularization model, Levin et al. (2007), we claim that SOPDO has better performance in terms of the difference between the ground truth and the estimated data on images which are susceptible to noise. Here, we deal only with additive Gaussian noise, as it effectively models the noise in many different imaging scenarios. In this section, we study in detail a few simulation results which are used to do comparison evaluations with other existing image restoration techniques.

We take non-isotropic SOPDO as the regularization operator for image restoration through least squares restoration as given in section 2.2. In order to compare the performance of the non-isotropic SOPDO prior model, we take two regularization models, FOPDO and a sparse stabilizing functional defined in Levin et al. (2007). Relating regularization to probability, the stabilizing functional in image restoration is also referred as prior model and

www.intechopen.com
Table 1. Comparison of stabilizing functional model.

<table>
<thead>
<tr>
<th>Smoothing weight</th>
<th>Non-isotropic SOPDO MSE $\times 10^{-4}$</th>
<th>Time in seconds</th>
<th>FOPDO MSE $\times 10^{-4}$</th>
<th>Time in seconds</th>
<th>Sparse Iterations</th>
<th>MSE $\times 10^{-4}$</th>
<th>Time in seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>4.978</td>
<td>0.43</td>
<td>5.168</td>
<td>0.44</td>
<td>10</td>
<td>8.050</td>
<td>27</td>
</tr>
<tr>
<td>0.05</td>
<td>4.646</td>
<td>0.42</td>
<td>4.730</td>
<td>0.44</td>
<td>10</td>
<td>6.899</td>
<td>27</td>
</tr>
<tr>
<td>0.01</td>
<td>6.043</td>
<td>0.42</td>
<td>7.561</td>
<td>0.44</td>
<td>10</td>
<td>4.289</td>
<td>27</td>
</tr>
</tbody>
</table>

for the quadratic regularization functionals, where the value of $R$ depends on the respective model such as $R = 2$ for FOPDO and $R = 3$ for non-isotropic SOPDO model.

2.3.2 Efficiency in deconvolution

As SOPDO can use frequency domain deconvolution techniques, it can be implemented highly efficiently than most of the recent non-blind deblurring techniques. The comparison was done with the Sparse deconvolution algorithm in Levin et al. (2007) named as “Levin Sparse deconvolution” and the non-blind deconvolution of, Shan et al. (2008) (distributed online) named as “Shan executable”. The results in Table 2 support the claim that...
### Table 2. Efficiency in non-blind image deconvolution

<table>
<thead>
<tr>
<th>Restoration Technique</th>
<th>Image size in pixels</th>
<th>Kernel size in pixels</th>
<th>Efficiency in seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levin Sparse deconv. (50 iterations)</td>
<td>484 × 752</td>
<td>19 × 27</td>
<td>576</td>
</tr>
<tr>
<td></td>
<td>910 × 903</td>
<td>99 × 99</td>
<td>556</td>
</tr>
<tr>
<td></td>
<td>1107 × 1694</td>
<td>99 × 99</td>
<td>1240</td>
</tr>
<tr>
<td></td>
<td>99 × 99</td>
<td>2429</td>
<td></td>
</tr>
<tr>
<td>Levin Sparse deconv. (10 iterations)</td>
<td>99 × 99</td>
<td>124</td>
<td></td>
</tr>
<tr>
<td>Shan executable</td>
<td>484 × 752</td>
<td>19 × 27</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>910 × 903</td>
<td>99 × 99</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>1107 × 1694</td>
<td>99 × 99</td>
<td>166</td>
</tr>
<tr>
<td>Non-isotropic SOPDO deconvolution</td>
<td>484 × 752</td>
<td>2.34</td>
<td></td>
</tr>
<tr>
<td></td>
<td>910 × 903</td>
<td>5.54</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1107 × 1694</td>
<td>13.17</td>
<td></td>
</tr>
</tbody>
</table>

### Table 3. Efficiency results on scaling non-isotropic SOPDO regularization

Further, we tested for the robustness of non-isotropic SOPDO regularization by using different sized images with varying sized kernels. The detailed results are shown in Table 3. All the images used for this experiment are color images, having separate rgb (red, green, blue) channels. The image and kernel sizes are given in pixels and the efficiency was measured in seconds. The results clearly show the robustness and the efficiency of the non-isotropic SOPDO regularization model with respect to different scales of image and kernel.

#### 2.3.3 Performance in deconvolution

Several computational experiments were carried out in order to compare non-isotropic SOPDO regularization with “Levin Sparse deconvolution” and “Shan executable”. The performance of these deconvolution techniques on a naturally blurred, highly textured image, given in Shan et al. (2008), are shown in Fig.1 and Fig.2. The blur kernel used in this experiment was retrieved through the blind deconvolution package of, Shan et al. (2008) distributed online. Closer visual inspection of the image results show that non-isotropic SOPDO technique best shows the tree branches and leaves while the other techniques have a blurring effect still remaining on the estimated result. This fact is further discussed and evidenced by evaluating the deconvolution in a new perspective in section 3.5.
3. Comparison of sparse prior models to quadratic regularization

3.1 Key issues

While the development of regularized solutions for ill-posed problems is widely discussed in the signal processing literature, recently by looking at the ill-posed image restoration problem from a probabilistic view point, some researchers claim that the Sparse prior model, Fergus et al. (2006); Levin et al. (2007) (discussed in more detail below in section 3.2.1) outperforms quadratic regularization models (discussed in section 2.2). The analytical study in this section addresses the following problems:

1. Are sparse prior models superior to quadratic regularization models?
2. What is the source of better performance of sparse prior models?
3. Are fast quadratic regularization models good enough for image restoration?

3.2 Regularization – Bayesian interpretation

Inverse problems such as image restoration are seen as probabilistic inference problems, where lack of information is compensated by assumptions. Therefore, it is not surprising, when the nature of the regularization detailed above is taken into consideration, to see that there is a close relationship between regularization and Bayesian estimation. Applying Bayes theorem to the image restoration problem in (1), for a known blur kernel, the posterior distribution can be written as

\[ p(g|b) \propto p(b|g) p(g), \]  

where \( p(b|g) \) represents the likelihood and \( p(g) \) represents the prior for the ground truth image. The estimation of the GTI based on posterior distribution can be classified in several ways. The minimum mean-square error estimate represents the mean of the posterior density, the MAP estimate stands for the mode of the posterior density while the maximum likelihood
Under the MAP technique, estimation of the GTI simplifies to
\[
\hat{g} = \arg \max_g p(g|b). \tag{15}
\]
Considering the non-blind image deconvolution process, we convert (15) to an energy minimization problem, where the energy is defined as
\[
E(g) \triangleq - \log \{ p(g|b) \}. \tag{16}
\]
Different likelihood and prior models on the ground truth have been applied for image restoration in literature. An analysis of existing prior models can be found in Mignotte (2006).
Considering the fact that for a given $g$, the variation in $b$ is due to the noise $n$, Hunt (1977), together with the above definitions, non-blind image restoration problem can be recast as seeking the unknown GTI, $g(i, j)$, that minimizes the functional

$$
\| b - K \hat{g} \|^2 + \sum_{r=1}^{R} \sum_{ij} \lambda_r \rho(D_r g(i, j)),
$$

(17)

where $D_r$ is the $r$th of $R$ linear operators, $i, j$ are pixel indices, $\lambda_r > 0$ are the regularization parameters, $\| \cdot \|$ stands for the Frobenius norm and $\rho(\cdot)$ is a scalar memoryless nonlinear mapping, generally taking the form

$$
\rho(z) \triangleq |z|^a
$$

(18)

for judicious choice of real parameter $a$ (not necessarily integer).

Many techniques belong to this class and differ only in: the set of linear operators $D_r$, $r = 1, 2, \ldots, R$, and the nonlinear mapping $\rho(z)$ (or choice of $a$). Numerous image restoration techniques have been developed under this framework from the early work, Geman & Geman (1984); Greig et al. (1989) to the most recent research, Fergus et al. (2006); Levin et al. (2007; 2009); Shan et al. (2008).

### 3.2.1 Sparse prior model

In recent literature, it is shown that, when derivative filters are applied to natural images, the filter outputs tend to be sparse, Olshausen et al. (1996); Simoncelli (1997). That is, the histogram of the derivative filtered image peaks at zero and falls off much faster than a Gaussian distribution. These heavy tailed natural image priors are used in a number of applications in image processing literature, such as denoising, Roth & Black (2005); Simoncelli (1999), reflection separation, Levin & Weiss (2007); Weiss (2001) and deconvolution, Levin (2007); Shan et al. (2008) in which, they are implemented in various ways such as student-$t$ distributions, Roth & Black (2005) and scale mixtures of Gaussian distributions, Fergus et al. (2006); Portilla et al. (2003).

In Levin et al. (2007), sparsity is incorporated by having $D_r$ as the derivative filters and $\alpha = 0.8$ in (18) as the prior term, which results in

$$
\| b - K \hat{g} \|^2 + \sum_{r=1}^{R} \sum_{ij} \lambda_r \| D_r \hat{g}(i,j) \|_1^{0.8}.
$$

(19)

This can be solved in the spatial domain using the Conjugate Gradient algorithm, Barrett et al. (1994).

### 3.2.2 Laplacian prior model

Although not as close as the Sparse prior to the natural image priors, Laplacian prior with $\alpha = 1$ in (18) is expected to result in a less smooth solution than the Gaussian prior. With the Laplacian prior, the optimization becomes

$$
\| b - K \hat{g} \|^2 + \sum_{r=1}^{R} \lambda_r \| D_r \hat{g} \|_1.
$$

(20)
Recently, much attention has been paid in solving $L^1$ norm regularization problems through compressed sensing. In Kim et al. (2007), an efficient method for optimizing a solution to a problem similar to (20) was discussed when $D_r$ are invertible.

### 3.2.3 Gaussian prior model

When $\alpha = 2$, minimization in (17) is called the Gaussian prior deconvolution in Levin et al. (2007) and is equivalent to the quadratic regularization problem in (9). Thus, in this chapter, we use the terms Gaussian prior and quadratic (specifically isotropic SOPDO) regularization interchangeably.

### 3.3 Image restoration evaluation

#### 3.3.1 Visual metric for evaluation

For all the restoration performance analysis and comparisons in this paper, we use a recently developed visual metric called SSIM (Structured SIMilarity) index, Wang et al. (2004), which has not been used for the comparison of prior models in the image restoration literature to date. The approach of SSIM is motivated by the highly structured characteristics of the natural image, where the strong neighborhood dependencies carry important information about the structures of the objects in the visual scene, Wang & Bovik (2009).

Assuming $x$ and $y$ are local image patches representing the same patch in the original and estimated images, the local SSIM index measures the similarities of three elements of the image patches: the similarity $\ell(x,y)$ of the local patch luminances (brightness values), the similarity $c(x,y)$ of the local patch contrasts, and the similarity $s(x,y)$ of the local patch structures. These local similarities are expressed using simple, easily computed statistics, and combined together to form local SSIM, $S(x,y)$, Wang & Bovik (2009).

$$
S(x,y) = \ell(x,y) \cdot c(x,y) \cdot s(x,y)
= \left( \frac{2\mu_x \mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \right) \cdot \left( \frac{2\sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \right) \cdot \left( \frac{\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3} \right),
$$

(21)

where $\mu_x$ and $\mu_y$ are the local sample means of $x$ and $y$, respectively, $\sigma_x$ and $\sigma_y$ are the local sample standard deviations of $x$ and $y$, respectively, and $\sigma_{xy}$ is the sample cross correlation of $x$ and $y$ after removing their means. The items $C_1$, $C_2$, and $C_3$ are small positive constants that stabilize each term, so that near zero sample means, variances or correlations do not lead to numerical instability.

Due to the fact that the underlying principle of SSIM is to extract the structural information which complies with the human visual system, SSIM maps are asserted to be a better signal fidelity measurement over MSE, Wang & Bovik (2009). In evaluating images through MSE, all image pixels are treated equally and content dependent variations in image fidelity are not accounted for. The two main indicators in SSIM evaluations, mean SSIM (MSSIM) and SSIM maps have values in the range from 0 to 1, where 1 indicates the best restoration. Although MSSIM and SSIM maps are generally used as visual fidelity metrics, we evaluate image restoration with the histogram of the SSIM map, as it provides an accurate view of the local restoration.
3.3.2 Image restoration models

Ignoring the presence of noise in image acquisition represented by (1), general image restoration could be represented by the model shown in Fig. 3, where \( \mathcal{L} \) represents the deblur process. The notation in (1) and the representation in Fig. 3 may be an over-simplification. From physical intuition, we could see that even though \( g \) is continuous by nature, image recording imposes limitations on the spatial extent of \( g \) and \( b \), leading to artifacts which impact on the final estimate of image restoration.

As illustrated in Fig. 4, we categorize these spatial artifacts in two ways. The “Model artifacts” are those, which are not present on naturally blurred images, but introduced in blur simulations as a result of sharp intensity transitions at the boundary of a finite image. Generating a blur image from a finite GTI causes unnatural blur distortions in the vicinity of the boundary of the image. Suppression of these “Model artifacts” could be accomplished by preprocessing the observed degraded image with techniques such as truncation and reducing the size of the blurred image. On the other hand, “Process artifacts” come along with the deblur process \( \mathcal{L} \) due to finite \( b \), which affect the performance of most deconvolution algorithms.

In order to show the effect of “Process artifacts”, we restore an image, originally, of size 255 × 255 pixels, but truncated in order to remove the “Model artifacts” introduced by a 13 × 13 pixels blur kernel, making the final image of size 242 × 242. The results of restoration with Sparse, Laplacian and Gaussian priors are shown in Fig. 5. In this experiment, deconvolution with Sparse and Laplacian priors were carried out using iterative re-weighted least squares (IRLS) method, Meer (2004), through the code available online, Levin et al. (2007), while the Gaussian prior is processed with both IRLS and fast Fourier techniques (FFT) separately. In our simulations, we processed IRLS for 150 iterations beyond which there were no further improvements. Analyzing the results of the performance of the Gaussian prior model with the FFT and IRLS techniques, we see that process artifacts are better handled by the IRLS technique than the FFT and this result is justified by the IRLS processing of Sparse and Laplacian prior models.

Both the “Model” and “Process” artifacts discussed above are not part of natural images, but are imposed artificially by the image modeling and processing techniques. Thus we claim that
3.4 Performance of quadratic regularization operators

In order to achieve the objective of studying the performance of different operators in the quadratic regularization as detailed in section 2.2, we carried out some simulations, where we avoided the effect of “Model artifacts” by taking a boundary strip off from the blurred image. In our evaluations, we used FOPDO, SOPDO and FSOPDO models to compare the performance. From this point onwards the term SOPDO refers to isotropic-SOPDO unless stated otherwise.

The simulations, for which the results are demonstrated in Fig. 6, are executed in the same environment as the simulation for Fig. 5, but with quadratic regularization models. We evaluated the performance of the regularization models under varying regularization parameter \( \lambda \) values as discussed in section 2.2. While the choice of parameters representing \( \lambda \) for FOPDO, SOPDO and FSOPDO are given in Table 4, the actual values for the respective parameters are given in Table 5.

While the overall SSIM values for few of the simulation results under varying \( \lambda \) values are shown in Table 5, the histogram distribution representing the first line of Table 5 is shown in Fig. 7. Overall, by analyzing these results, we claim that, in the presence of “Process artifacts”, a better performance could be achieved with FSOPDO over FOPDO and SOPDO.
3.5 Regularization model performance comparison

As shown earlier in section 2.2, we modeled the regularization of image restoration based on the quadratic regularization terms (sometimes called as the least squares regularization) and in section 3.2, we discussed the existing probabilistic models under a MAP framework.

In the next section we compare the performance of these quadratic regularization models by removing the “Process” and “Model” artifacts.
Fig. 8. Image restoration model for a naturally blurred image, where $K$ is the blur process, $L$ is the deblur process, $g$, $b$, $z$, $\hat{g}$ stand for GTI, blur image, deblurred image with artifacts, and the final estimated GTI respectively. The process $P_M(z)$ decouples “Model” and “Process” artifacts from the deblurred image.

These models form a method of regularization in image restoration. This section is devoted for the comparison of these models. The comparison in this section will guide us for making recommendations for the appropriate regularization technique and is discussed at the end of this section.

### 3.5.1 SSIM performance comparison

As the objective of our simulations is to evaluate the contribution of the regularization models towards image restoration, we use the restoration model shown in Fig. 8, where we decouple artifact effects from restoration by projecting the estimated image with

$$P_M(z)(i,j) = \begin{cases} z(i,j), & \text{if } i,j \in M \\ 0, & \text{otherwise} \end{cases}$$

(22)

where $M$ is a region without “Model” and “Process” artifacts.

To be consistent with the SSIM map region in Fig. 5, we take a large image of support $1024 \times 1024$ and project the final image to a $242 \times 242$ region within the inner region of the estimated image, which is least affected by the artifacts. The restoration was carried out with FFT processing of the Gaussian prior and IRLS processing of Sparse and Laplacian priors. The comparison of the performance of the priors is shown in Fig. 9. In it we note that Gaussian prior with FFT processing has performed as well as or better to the Sparse and Laplacian prior models.

As the literature claims that iterative algorithms such as conjugate gradient algorithms suppress noise and perform better in noisy blur image restoration, we simulated a noisy blurred restoration under the same conditions given for Fig. 9, but with different regularization parameter values, as more weight should now be given to the prior over data. The noise added was Gaussian with zero mean and 0.01 variance. The optimal results we obtained for varying $\lambda$ are shown in Fig. 10. With these results, we claim that Gaussian prior handles noisy images as better as the Sparse and Laplacian prior models.

Thus, these results pave a new path of thinking and we claim that quadratic regularization with SOPDO model, when appropriately configured and used in a realistic context, free from unnatural artifacts, is comparable to Sparse prior model in terms of image restoration performance under the SSIM criterion.

### 3.5.2 Efficiency comparison

As the optimization problem in least-squares regularization is convex and as the fast Fourier techniques could be applied for the computation, for an image of size $L \times L$ pixels, the restoration through least-squares regulation has a complexity of $O(L \log L)$ operations. In
contrast, when a Sparse prior is used, the optimization problem is no longer convex and cannot be minimized in closed form. Using the conjugate gradient method, Barrett et al. (1994), or IRLS method, the optimization can be solved in $O(L l_{\text{max}})$ where $l_{\text{max}}$ represent the maximum number of iterations.

A few simulation results on efficiency are shown in Table 6, where all the values are in seconds and represent the time taken for the restoration using each of the respective model. While the quadratic regularization deconvolution was carried out using Fourier domain techniques, the Sparse deconvolution was carried out using the IRLS method. Under the IRLS algorithm, it is experienced that in order to achieve an acceptable result, the number of iterations should be at least 50 and better results could be achieved when the number of iterations are above 100. From the results shown, it is evident that when the size of the image increases, the relative efficiency of the restoration through Sparse prior model becomes extremely low.
### Table 6. Efficiency of regularization operators. The times taken for restoration of grey-scale and colored images are given in seconds for each of the regularization operators.

<table>
<thead>
<tr>
<th>Image size in pixels</th>
<th>Time for Quadratic regularization with FFT</th>
<th>Time for Sparse prior in 50 iterations</th>
<th>Time for Sparse prior in 100 iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>grey-scale 255 × 255</td>
<td>0.08</td>
<td>30</td>
<td>60</td>
</tr>
<tr>
<td>grey-scale 1024 × 1024</td>
<td>1.1</td>
<td>303</td>
<td>596</td>
</tr>
<tr>
<td>colored 484 × 752</td>
<td>1.7</td>
<td>371</td>
<td>730</td>
</tr>
<tr>
<td>colored 1107 × 1694</td>
<td>9.0</td>
<td>2180</td>
<td>4292</td>
</tr>
</tbody>
</table>

**3.5.3 Regularization recommendations**

In addition to lower efficiency and not-superior performance, Sparse prior models lack in proper theoretical guidelines for selecting the best regularization parameter. In contrast, the quadratic regularization models can use well-established methods such as L-curve criterion, Hansen (1998) and the Generalized Cross Validation criterion, Hansen (1998) for choosing the value of $\lambda$. Difficulties in selecting the optimal converging point in non-convex minimization techniques such as IRLS also is an issue.

According to the theoretical and experimental details provided above, we propose that if we could decouple image restoration and “Process artifact” handling, then the use of quadratic regularization models will result in more efficient and effective image restoration in comparison to Sparse and Laplacian prior models. The decoupling of image restoration and “Process” artifact handling could be achieved through techniques such as tiling, Liu & Jia (2008), which enables the uses of the efficient least squares regularization.

Thus, coming back to our problem formulation in section 3.1, we claim that:

1. Sparse prior models are not superior to quadratic regularization models in terms of performance in image restoration.
2. In terms of efficiency, Sparse prior models are significantly inferior to quadratic regularization models.
3. The performance through Sparse prior model increases over quadratic regularization models when boundary effects are not addressed and processing artifacts are not compensated for.
4. Quadratic regularization models provide the best image restoration for large images in terms of efficiency and effectiveness while they provide a good enough solution for other images when the boundary artifacts are taken care of.

Analyzing the above items further, if the improvements of the Sparse prior model are in artifact handling, not in image restoration, we can pose the following questions.

“Do more complicated prior models such as Sparse, which are asserted be better matched to natural images, actually help image restoration in terms of restoring natural image features?”

“If those complicated prior models hold no significant advantage, is it worth the effort spend on them compared to simple and efficient prior models which restore closer or better than those prior models?”

www.intechopen.com
4. Likelihood model analysis

Different likelihood models in the prior model in (14) have been studied in various ways. The fact that most of these models are not justified with proper theoretical foundations encouraged us to analyze and understand the variations and the validity and accuracy of the (implicit) underlying assumptions, which could explain the different performances.

This investigation guides our development of a new scheme for the multiple image likelihood model described in section 4.1.2. The likelihood model analysis is carried out using this new model and the theoretical analysis is corroborated by the computational experiments detailed in section 4.3.

4.1 Likelihood models in image restoration

4.1.1 Likelihood model for a single image

In image restoration literature, the likelihood for a single image is defined by modeling the image noise \( n \) as a set of independently and identically distributed (i.i.d.) random variables following a Gaussian distribution for all pixels, which is given by

\[
p(b|g) = \prod_{\ell_1=1}^{L_1} \prod_{\ell_2=1}^{L_2} N(n(\ell_1, \ell_2)|0, \sigma^2),
\]

(23)

where \( N(\cdot | \mu, \sigma^2) \) denotes a Gaussian distribution with mean \( \mu \) and variance \( \sigma^2 \), while \( L_1 \) and \( L_2 \) represent the image support.

4.1.2 Likelihood model for multiple images

Based on the above likelihood model for a single image, we develop a new model for the likelihood of multiple images as detailed below.

Given a set of \( R \) degraded images of a common GTI \( g \), the posterior distribution for the GTI can be derived by extending (14), resulting in

\[
p(g|b_1, b_2, \ldots, b_R) \propto p(b_1, b_2, \ldots, b_R|g) p(g),
\]

(24)

where, generalizing (1),

\[
b_r = K_r g + n_r, \quad r = 1, 2, 3, \ldots, R
\]

(25)

and \( K_r \) are operators representing possibly different but known blurs, and \( n_r \) are noise images. Under the assumption that \( n_p \) is independent of \( n_q \) for all \( p \neq q \), the likelihood in (24) is

\[
p(b_1, b_2, \ldots, b_R|g) = \prod_{r=1}^{R} N(n_r).
\]

(26)

Thus, for a group of \( R \) images satisfying the noise independency condition in (26), the likelihood can be modeled as

\[
p(b_1, b_2, \ldots, b_R|g) = \prod_{r=1}^{R} \prod_{\ell_1=1}^{L_1} \prod_{\ell_2=1}^{L_2} N(n_r(\ell_1, \ell_2)|0, \sigma_r),
\]

(27)

where \( \sigma_r \) represent the standard deviation of the Gaussian distribution for \( n_r \). This new model for the likelihood of multiple images will be used for the analysis of likelihood models in the next section.
4.1.3 Likelihood models for analysis

Out of the various likelihood models introduced in the literature of image restoration, we consider two recent approaches in Levin et al. (2007) and, Shan et al. (2008) for our analysis.

In Levin et al. (2007), the single image likelihood conforms to (23) and is explicitly given by
\[ p(b|g) \propto e^{-\frac{1}{2\sigma^2} \|Kg-b\|^2}, \quad (28) \]

where \(\|\cdot\|\) stands for the Frobenius norm.

In Shan et al. (2008), the likelihood is defined with different orders of partial derivatives, denoted by operator \(\partial_r\), of a single degraded image. For ease of understanding, we represent their model in the form
\[ p(b|g) = \prod_{\partial_r \in \Theta} \prod_{\ell_1=1}^{L_1} \prod_{\ell_2=1}^{L_2} N(n(\ell_1, \ell_2)|0, \sigma) \cdot N(\partial_r n(\ell_1, \ell_2)|0, \sigma_{\omega(\partial_r)}), \quad r = 1, 2, 3, \ldots, R \quad (29) \]

where \(\Theta\) is a set of partial derivative operators given by
\[ \Theta \triangleq \{ \partial_1, \partial_2, \partial_3, \ldots, \partial_R \}. \quad (30) \]

For example, in Shan et al. (2008), the set \(\Theta\) has the elements \{\(\partial_x, \partial_y, \partial_{xx}, \partial_{xy}, \partial_{yy}\}\}, in which, \(\partial_x\) is the first order derivative in \(x\) direction and \(\partial_y\) is the first order derivative in \(y\) direction and similar interpretations hold for higher order derivatives.

Further, Shan et al. (2008) shows that the partial derivatives of \(n\) also follow Normal distributions with standard deviation values based on the order of the partial derivative operator. The standard deviations of the partial derivatives are specified in the form
\[ \sigma_{\omega(\partial_r)} = (\sqrt{2})^{\omega(\partial_r)} \sigma, \quad (31) \]

where \(\omega(\partial_r)\) represents the order of the partial derivative operator \(\partial_r\). Few example elements of the set \(\Theta\) in (30) with the respective standard deviation values are given in Table 7.

As there was no analysis presented behind using the higher order partial derivatives of noise in Shan et al. (2008) leading to (29), we provide an interpretation of formula, based on our new general likelihood model for a group of degraded images of a common ground truth \(g\) in (27).

Guided by the likelihood expression (27), we can define a virtual group of images for the likelihood model in (29) as
\[ b_r = \partial_r b, \quad r = 1, 2, 3, \ldots, R \quad (32) \]
and in order to align with model (25), define

\[ K_r g \triangleq \partial_r (g \ast k), \quad r = 1, 2, 3, \ldots, R \]  
\[ n_r \triangleq \partial_r n, \quad r = 1, 2, 3, \ldots, R \]  

(33)  
(34)

where \( \ast \) stands for the convolution operator and \( k \) is the blur kernel.

From this, we infer that the likelihood (29) implicitly assumes \( \partial_p n \) is independent of \( \partial_q n \) for all \( p \neq q \). Since all virtual images are derived from a single degraded image, we can infer this is a strong assumption made to simplify the likelihood expression. In principle, it should be possible to formulate a model without recourse to the derivative images which add limited new information. We corroborate this claim in section 4.3 with experiments.

4.2 Frequency domain deconvolution

In this section we approach image deconvolution with FOPDO regularization and with different likelihood models discussed above. For our analysis, we consider the likelihood models of (23) and (29) using terminology “normal likelihood” and “derivative likelihood” with the notation using subscripts “\( n \)” and “\( d \)”, respectively. With our experiments, we limit the set \( \Theta \) in (30), going up to second order partial derivative operators in (29) and we take elements of \( \Theta \) from the following values

\[ \Theta = \{ \partial_x, \partial_y, \partial_{xx}, \partial_{xy}, \partial_{yy} \} \].

(35)

4.2.1 Normal likelihood deconvolution

Under FOPDO regularization as detailed in section 2.2, the stabilizing functional \( \Phi(g) \) takes the form

\[ \Phi(g) \triangleq \| \partial_x g \|^2 + \| \partial_y g \|^2. \]  

(36)

Applying this stabilizing functional to the MAP framework detailed in section 4, the energy functional under “normal likelihood”, can be derived as

\[ E_n(g) = \| g \ast k - b \|^2 + \lambda \Phi(g). \]  

(37)

According to the convolution theorem, the convolution operation in the spatial domain becomes an element-wise product in the frequency domain making \( \mathcal{F}(g \ast k) = G \ast K \) where \( \mathcal{F}(-) \) stands for discrete Fourier transform, \( G \) for \( \mathcal{F}(g) \), \( K \) for \( \mathcal{F}(k) \) and “\( \ast \)” denotes element-wise product. Based on the above property, transforming (37) into frequency domain and applying Plancherels theorem, Bracewell & Kahn (1966), we derive the energy in the frequency domain for (37) as follows.

\[ \mathcal{F}(E_n(g)) = \| G \ast K - B \|^2 + \lambda \mathcal{F}(\Phi(g)), \]  

(38)

where

\[ \mathcal{F}(\Phi(g)) = \| \mathcal{F}(\partial_x) \ast G \|^2 + \| \mathcal{F}(\partial_y) \ast G \|^2, \]

\( B \) stands for \( \mathcal{F}(b) \) and given \( \partial_x \) takes the form of a (convolution) matrix, then \( \mathcal{F}(\partial_x) \) denotes its Fourier transform.
Minimizing the energy in (38) and solving for estimated $G$ denoted as $\hat{G}$ results in

$$\hat{G}_n = \frac{B * K}{K * K + \lambda \Psi},$$  \hspace{1cm} (39)$$

where

$$\Psi = \mathcal{F}(\partial_x) * \mathcal{F}(\partial_x) + \mathcal{F}(\partial_y) * \mathcal{F}(\partial_y),$$

and $\hat{G}_n$ is the Fourier transform of the estimated GTI under “normal likelihood”. $\bar{C}$ stands for the complex conjugate and the division is performed element-wise. The estimated ground truth image $\hat{g}_n$ can be derived by taking the inverse Fourier transform of $\hat{G}_n$. With the above derivations, it is evident that Fourier domain expression used to estimate the GTI is:

1. simple and leads to a closed form solution and
2. amenable to Fast Fourier Techniques leading to a highly efficient solution.

**4.2.2 Derivative likelihood deconvolution**

We now give an analogous derivation for the “derivative likelihood”. The energy functional in this case is derived similar to (37),

$$E_d(g) = \sum_{\partial r \in \Theta} \frac{1}{2} \left\| \partial r (g * k) - \partial r b \right\|^2 + \lambda \Phi(g).$$  \hspace{1cm} (40)$$

Transforming (40) into the frequency domain results in

$$\mathcal{F}(E_d(g)) = \sum_{\partial r \in \Theta} \frac{1}{2} \left( \left\| \mathcal{F}(\partial r) * G * K - \mathcal{F}(\partial r) * B \right\|^2 + \lambda \mathcal{F}(\Phi(g)) \right)^2,$$  \hspace{1cm} (41)$$

where, $\partial r$ is a matrix convolution operator representing a partial order derivative operator and $\mathcal{F}(\partial r)$ denotes its Fourier transformation.

By minimizing the energy (41), we compute the estimated $G$,

$$\hat{G}_d = \frac{B * K * \Omega}{K * K * \Omega + \lambda \Psi},$$  \hspace{1cm} (42)$$

where

$$\Omega = \sum_{\partial r \in \Theta} 1 \mathcal{F}(\partial r) = \mathcal{F}(\partial r).$$

By taking the inverse Fourier transforms of (42), we could get the estimated GTI, $\hat{g}_d$, under “derivative likelihood” model similar to “normal likelihood” model.
In order to come up with the most effective and efficient restoration algorithm, we investigate the contribution of each of the likelihood models for estimating the GTI: (39) corresponding to “normal likelihood” and (42) corresponding to “derivative likelihood”, respectively.

We used the same “Picasso” image which was used in Shan et al. (2008) for experiments using the likelihood model in (29). The ground truth images are estimated using the Fourier domain techniques, specifically applying (39) and (42) for the “Normal” and “Derivative” likelihood models respectively. The experiment results are given in Table 8. In order to eliminate the “Model” and “Process” artifacts as discussed in section 3.3.2, in all our simulations, the blurring was carried out under the assumption that the images are periodic.

The MSE values in the table are given as multiples of $10^{-4}$, while the value of $\lambda$ is given in multiples of $10^{-5}$. The values in bold in Table 8 refer to the optimal MSE values the respective likelihood model could reach for varying $\lambda$. As the results show clearly, the “normal likelihood” model has a better estimate for the GTI than the “derivative likelihood” model, we claim that applying “normal likelihood” in the image restoration algorithm results in a better restoration.

Our investigation was further extended to analyze whether higher order derivatives of noise contribute to the spatial randomness of noise as claimed in Shan et al. (2008). The noise maps given in Fig. 11 are computed for different values of $\lambda$ in (39) and (42). As per the results Fig. 11(c) and Fig. 11(d), when the effect of the prior becomes smaller (i.e., the weight on the data fitting term or the likelihood becomes larger), the noise estimate is more spatially random, but with the increase in the weight of the prior, the noise estimate becomes structured (signal dependant), see Fig. 11(e) and Fig. 11(f). We experienced these results regardless of the likelihood model we used. Based on the above results, we claim that using higher order partial derivatives in the likelihood model for non-blind deconvolution does not result in a better noise map estimation while the same noise map estimation can be achieved through the normal likelihood model with the appropriate Lagrange multiplier.

Hence, through the likelihood model analysis based on benchmark image, we conclude that higher order derivatives in the likelihood model are not required for better performance whereas applying single image likelihood model with appropriate regularization results in a more effective non-blind image restoration.

<table>
<thead>
<tr>
<th>$\lambda$ ($\times 10^{-5}$)</th>
<th>MSE Normal likelihood ($\times 10^{-4}$)</th>
<th>MSE Derivative likelihood ($\times 10^{-4}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>2.8448</td>
<td>4.4371</td>
</tr>
<tr>
<td>5</td>
<td>2.4148</td>
<td>4.0862</td>
</tr>
<tr>
<td>0.25</td>
<td>2.2292</td>
<td>4.1852</td>
</tr>
</tbody>
</table>

Table 8. Comparison of likelihood models

4.3 Likelihood model analysis
5. Contributions

In this chapter, we have contributed to regularization based image restoration techniques in the following:

1. We have developed a general class of quadratic regularization models based on partial derivative operators (PDO), section 2.2. Out of those models, we have shown that the Second Order Partial Derivative Operator (SOPDO) model performs better than First Order Partial Derivative Operator (FOPDO) model for images susceptible to noise, while the novel First and Second Order Partial Derivative Operator (FSOPDO) model performs better than both FOPDO and SOPDO models.
2. We have used the Structured Similarity index (SSIM) map, Mean SSIM (MSSIM) value and histograms of SSIM maps as novel visual metrics for comparison and evaluation of regularization models in image restoration, section 3.3.2.

3. We have critically evaluated Sparse and Laplacian prior models against Quadratic regularization models using the novel visual metrics discussed in section 3.5. By eliminating the effects of processing and modeling artifacts, not present when capturing actual blurred natural images, we have shown that Sparse and Laplacian derivative prior models, which are claimed to be consistent with natural images, do not significantly contribute in restoring natural image features and have significantly slower relative restoration performance.

4. Finally, we have analyzed and evaluated multiple derivative operator based restoration methods under MAP/ML framework with a novel model to represent the likelihood based on multiple images, section 4.1.2. By using this novel model, we demonstrate that complex higher order derivative likelihood models are not required for better performance in image restoration.

6. References


www.intechopen.com


This book represents a sample of recent contributions of researchers all around the world in the field of image restoration. The book consists of 15 chapters organized in three main sections (Theory, Applications, Interdisciplinarity). Topics cover some different aspects of the theory of image restoration, but this book is also an occasion to highlight some new topics of research related to the emergence of some original imaging devices. From this arise some real challenging problems related to image reconstruction/restoration that open the way to some new fundamental scientific questions closely related with the world we interact with.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
