Chapter 4
Application: Image Deblurring for Optical Imaging

The problem of reconstruction of digital images from their blurred and noisy measurements is unarguably one of the central problems in imaging sciences. Despite its ill-posed nature, this problem can often be solved in a unique and stable manner, provided appropriate assumptions on the nature of the images to be discovered. In this section, however, a more challenging setting is considered, in which accurate knowledge of the blurring operator is lacking, thereby transforming the reconstruction problem at hand into a problem of blind deconvolution [1, 2]. As a specific application, the current presentation focuses on reconstruction of short-exposure optical images measured through atmospheric turbulence. The latter is known to give rise to random aberrations in the optical wavefront, which are in turn translated into random variations of the point spread function (PSF) of the optical system in use. A standard way to track such variations involves using adaptive optics. For example, the Shack-Hartmann interferometer provides measurements of the optical wavefront through sensing its partial derivatives. In such a case, the accuracy of wavefront reconstruction is proportional to the number of lenslets used by the interferometer, and hence to its complexity. Accordingly, in this chapter, we show how to minimize the above complexity through reducing the number of the lenslets, while compensating for undersampling artifacts by means of derivative compressed sensing. Additionally, we provide empirical proof that the above simplification and its associated solution scheme result in image reconstructions, whose quality is comparable to the reconstructions obtained using conventional (dense) measurements of the optical wavefront.

4.1 Background

The necessity to recover digital images from their distorted and noisy observations is common for a variety of practical applications, with some specific examples including image denoising, super-resolution, image restoration, and watermarking, just to name

a few [3–6]. In such cases, it is conventional to assume that the observed image v is obtained as a result of convolution of its original counterpart u with a point spread function[1] (PSF) i. To account for measurement inaccuracies, it is also standard to contaminate the convolution output with an additive noise term ν, which is usually assumed to be white and Gaussian. Thus, formally,

$$v = i * u + \nu. \tag{4.1}$$

While u and v can be regarded as general members of the signal space $\mathbb{L}_2(\Omega)$ of real-valued functions on $\Omega \subseteq \mathbb{R}^2$, the PSF i is normally a much smoother function, with effectively band-limited spectrum. As a result, the convolution with i has a destructive effect on the informational content of u, in which case v typically has a substantially reduced set of features with respect to u. This makes the problem of reconstruction of u from v a problem of significant practical importance [8].

Reconstruction of the original image u from v can be carried out within the framework of image deconvolution, which is a specific instance of a more general class of inverse problems [9]. Most of such methods are Bayesian in nature, in which case the information lost in the process of convolution with i is recovered by requiring the optimal solution to reside within a predefined functional class [10–12]. Thus, for example, in the case when u is known to be an image of bounded variation, the above regularization leads to the famous Rudin-Osher-Fatemi reconstruction scheme, in which u is estimated as a solution to the following optimization problem [13, 14]

$$\hat{u} = \arg\min_u \left\{ \frac{1}{2} \|u * i - v\|_2^2 + \alpha \int |\nabla u| \, dx dy \right\}, \tag{4.2}$$

where $\alpha > 0$ is the regularization parameter. It should be noted that, if the PSF obeys $\int i \, dx dy \neq 0$, the problem (4.2) is strictly convex and therefore admits a unique minimizer, which can be computed using a spectrum of available algorithms [13–17].

In some applications, the knowledge of the PSF may be lacking, which results in the necessity to recover the original image from its blurred and noisy observations alone. Such a reconstruction problem is commonly referred to as the problem of blind deconvolution [9]. In the present study, however, we follow the philosophy of *hybrid* deconvolution [18], whose main idea is to leverage any *partial* information on the PSF to improve the accuracy of image restoration. In particular, in the algorithm described in this chapter, such partial information is derived from incomplete observations of the partial derivatives of the phase of the generalized pupil function (GPF) of the optical system in use, as detailed below.

Optical imaging is unarguably the field of applied sciences from which the notion of image deconvolution has originated [19–21]. In particular, in short-exposure turbulent imaging [2], acquired images are blurred with a PSF, which depends on a spatial distribution of the atmospheric refraction index along the optical path connecting

---

[1] Note that, in optical imaging, this function is also referred to as an impulse transfer function [7].

an object of interest and the observer. Due to the effect of turbulence, the above distribution is random and time-dependent, which implies that the PSF i cannot be known in advance.

A standard way to overcome the above limitation is through the use of adaptive optics (AO) [22]. As will be shown later, the PSF of a short-exposure optical system is determined by its corresponding *generalized pupil function* (GPF) P, which can be expressed in a polar form as $P = A\, e^{J\phi}$. While, in practice, the amplitude A can be either measured through calibration or computed as a function of the aperture geometry, the phase ϕ accounts for turbulence-induced aberrations of the optical wavefront, and hence is generally unknown at any given experimental time. Fortunately, the phase ϕ turns out to be a measurable quantity, and this is where the tools of AO come into play. One of such tools is the Shack-Hartmann interferometer (SHI) (aka Shack-Hartmann wavefront sensor) [23–25], which allows direct measurement of the gradient of ϕ over a predefined grid of spatial coordinates. Subsequently, these measurements are converted into a useful estimate of ϕ through numerically solving an associated Poisson equation.

Among some other factors, the accuracy of phase reconstruction by the SHI depends on the size of its sampling grid, which is in turn defined by the number of lenslets composing the wavefront sensor of the interferometer (see below). Unfortunately, the grid size and the complexity (and, hence, the cost) of the interferometer tend to increase *pro rata*, which creates an obvious practical limitation. Accordingly, to overcome this problem, we propose to modify the construction of the SHI through reducing the number of its lenslets. Although the advantages of such a simplification are immediate to see, its main shortcoming is obvious as well: the smaller the number of lenslets is, the stronger is the effect of undersampling and aliasing. These artifacts, however, can be compensated for by subjecting the output of the simplified SHI to the derivative compressed sensing (DCS) algorithm of [26], which is a special case of the problem, studied in Chap. 3. As will be shown below, DCS is particularly suitable for reconstruction of ϕ from incomplete measurements of its partial derivatives. The resulting estimates of ϕ can be subsequently combined with A to yield an estimate of the PSF i, which can in turn be used by a deconvolution algorithm. Thus, the proposed method for estimation of the PSF i and subsequent deconvolution of u can be regarded as a hybrid deconvolution technique, which comes to simplify the design and complexity of the SHI on one hand, and to make the process of reconstruction of optical images as automatic as possible, on the other hand.

4.2 Technical Preliminaries

In short exposure imaging, due to phase aberrations in the optical wavefront induced by atmospheric turbulence, the PSF of an imaging system in use is generally unknown [2]. To better understand the setup under consideration, we first note that, in optical imaging, the PSF i is obtained from an amplitude spread function (ASF) h as $i :=$ $|h|^2$. The ASF, in turn, is defined in terms of a generalized pupil function (GPF) $P(x, y)$ and is given by [27]

$$h(\xi, \eta) = \frac{1}{\lambda_w z_i} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(x, y) e^{-j \frac{2\pi}{\lambda z_i}(x\xi + y\eta)} \, dx dy, \qquad (4.3)$$

where z_i is the focal distance and λ_w is the optical wavelength. Being a complex-valued quantity, $P(x, y)$ can be represented in terms of its amplitude $A(x, y)$ and phase $\phi(x, y)$ as

$$P(x, y) = A(x, y) \, e^{J\phi(x,y)}. \qquad (4.4)$$

Here, the GPF amplitude $A(x, y)$ (which is sometimes simply referred to as the aperture function) is normally a function of the aperture geometry. Thus, for instance, in the case of a circular aperture, $A(x, y)$ can be defined as [28]

$$A(r) = \begin{cases} 1, & \text{if } r \leq \frac{D}{2} \\ 0, & \text{otherwise} \end{cases} \qquad (4.5)$$

where D denotes the pupil diameter. Thus, given $\phi(x, y)$, one could determine h and therefore i. Unfortunately, the phase $\phi(x, y)$ is influenced by the random effect of atmospheric turbulence, and as a result cannot be known ahead of time.

A standard way to overcome the uncertainty in $\phi(x, y)$ is to measure it using the tools of shearing interferometry, a particular example of which is the SHI [23, 29]. The latter is capable of sensing the partial derivatives of $\phi(x, y)$ over a predefined grid of spatial locations. In this case, an accurate reconstruction of $\phi(x, y)$ entails taking a fairly large number of the samples of $\nabla\phi(x, y)$, which is essential for minimizing the effect of aliasing on the estimation result [30]. Thus, in some applications, the number of sampling points (as defined by the number of SHI lenslets) reaches as many as a few thousands. It goes without saying that reducing the number of lenslets would have a positive impact on the SHI in terms of its cost and approachability. Alas, such a reduction is impossible without undersampling, which is likely to have a formidable effect on the overall quality of phase estimation.

To minimize the effect of phase undersampling, we exploit the DCS algorithm of [31]. The latter can be viewed as an extension of the conventional compressed sensing (CCS) scheme, in which the standard sparsity constraints are supplemented by additional constraints related to some intrinsic properties of partial derivatives. This "side information"—which are called the cross-derivative constraints—allows substantially improving the quality of reconstruction of $\phi(x, y)$, as compared to the case of CCS-based estimation.

4.2.1 Shack-Hartmann Interferometer

As it was mentioned the SHI can be used to measure the gradient $\nabla\phi(x, y)$ of the GPF phase $\phi(x, y)$, from which its values can be subsequently inferred. A standard approach to this reconstruction problem is to assume the unknown phase $\phi(x, y)$ to be expandable in terms of some basis functions $\{Z_k\}_{k=0}^{\infty}$, *viz.* [24]

$$\phi(x, y) = \sum_{k=0}^{\infty} a_k Z_k(x, y), \tag{4.6}$$

where the representation coefficients $\{a_k\}_{k=0}^{\infty}$ are supposed to be unique and stably computable. Note that, in this case, the datum of $\{a_k\}_{k=0}^{\infty}$ uniquely identifies $\phi(x, y)$, while the coefficients $\{a_k\}_{k=0}^{\infty}$ can be estimated due to the linearity of (4.6) which suggests

$$\nabla \phi(x, y) = \sum_{k=0}^{\infty} a_k \nabla Z_k(x, y). \tag{4.7}$$

In AO, it is conventional to define $\{Z_k\}_{k=0}^{\infty}$ to be Zernike polynomials (aka Zernike functions) [27]. These polynomials constitute an orthonormal basis in the space of square-integrable functions defined over the unit disk in $\mathbb{R}^2$. Zernike polynomials can be subdivided in two subsets of the even Z_n^m and odd Z_n^{-m} Zernike polynomials, which possess closed-form analytical definitions as given by

$$Z_n^m(\rho, \varphi) = R_n^m(\rho) \cos(m \varphi) \tag{4.8}$$

$$Z_n^{-m}(\rho, \varphi) = R_n^m(\rho) \sin(m \varphi), \tag{4.9}$$

where m and n are nonnegative integers with $n \geq m$, $0 \leq \varphi < 2\pi$ is the azimuthal angle, and $0 \leq \rho \leq 1$ is the radial distance. The radial polynomials R_n^m in (4.8) and (4.9) are defined as

$$R_n^m(\rho) = \sum_{k=0}^{(n-m)/2} \frac{(-1)^k (n-k)!}{k! ((n+m)/2 - k)! ((n-m)/2 - k)!} \rho^{n-2k}. \tag{4.10}$$

Note that, since the Zernike polynomials are defined using polar coordinates, it makes sense to re-express the phase ϕ and its gradient in the polar coordinate system as well. (Technically, this would amount to replacing x and y in (4.6), (4.7) by ρ and φ, respectively.) Moreover, due to the property of the Zernike polynomials to be an orthonormal basis, the representation coefficients $\{a_k\}_{k=0}^{\infty}$ in (4.6), (4.7) can be computed by orthogonal projection, namely

$$a_k = \int_0^{2\pi} \int_0^1 \phi(\rho, \varphi) Z_k(\rho, \varphi) \rho \, d\rho \, d\varphi. \tag{4.11}$$

In practice, however, $\phi(\rho, \varphi)$ is unknown and therefore the coefficients $\{a_k\}_{k=0}^{\infty}$ need to be estimated by other means. Thus, in the case of the SHI, the coefficients can be estimated from a finite set of discrete measurements of $\nabla \phi(\rho, \varphi)$.

The main function of the SHI is to acquire discrete measurements of $\nabla \phi$ by means of linearization. The linearization takes advantage of subdividing a (circular) aperture into rectangular blocks with their sides formed by a uniform rectangular
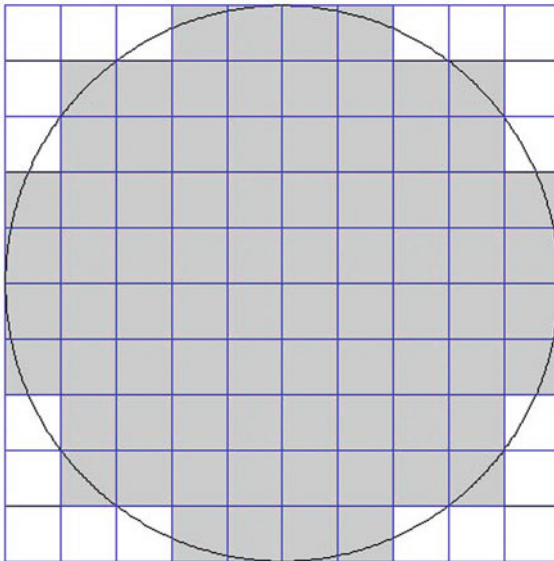
lattice. An example of such a subdivision is shown in Fig. 4.1 for the case of a 10×10
lattice grid. In general, the grid is assumed to be sufficiently fine to approximate ϕ by
a linear function over the extent of a single block. This results in a piecewise linear
approximation of ϕ, whose accuracy improves asymptotically when the lattice size
goes to infinity. Formally, let $\Omega := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq D^2\}$ be a circular
aperture of radius D and $\mathcal{S} = \{(x, y) \in \mathbb{R}^2 \mid \max\{|x|, |y|\} \leq D\}$ be a square subset
of $\mathbb{R}^2$ such that $\Omega \subset \mathcal{S}$. Then, for each polar coordinate $(\rho, \varphi) \in \Omega$ and an $N \times N$
grid of square blocks of size $2D/N \times 2D/N$, the phase ϕ can be expressed as

$$\phi(x, y) \approx ax + by + c, \tag{4.12}$$

for all (x, y) in a neighbourhood of $(\rho \cos \varphi, \rho \sin \varphi)$. The approximation in (4.12)
suggests that

$$\nabla \phi(x, y) \approx (a, b)^T \tag{4.13}$$

where $(\cdot)^T$ denotes matrix transposition. While c in (4.12) can be derived from
boundary conditions, coefficients a and b should be determined through direct mea-
surements. To this end, the SHI is endowed with an array of small focusing lenses
(i.e., lenslets), which are supported over each of the square blocks of the discrete grid,
thereby forming a wavefront sensor. In the absence of phase aberrations, the focal
points of the lenslets are spatially identified and registered using a high-resolution
CCD detector, whose imaging plane is aligned with the focal plane of the sensor.
Then, when the wavefront gets distorted by atmospheric turbulence, the focal points
are dislocated towards new spatial positions, which can also be pinpointed by the
same detector. The resulting displacements can be measured and subsequently related
to the values of $\nabla \phi$ at corresponding points of the sampling grid.

To explain how the above procedure can be performed, additional notations are in order. Let Ω_d denote a finite set of spatial coordinates defined according to

$$
\begin{aligned}
\Omega_d := \Big\{ (x_d, y_d) &\in \Omega \mid \\
x_d = -D + \frac{2D}{N} &\left(i + \frac{1}{2} \right), \quad i = 0, 1, \ldots, N-1 \\
y_d = -D + \frac{2D}{N} &\left(j + \frac{1}{2} \right), \quad j = 0, 1, \ldots, N-1 \\
&\text{and } x_d^2 + y_d^2 \leq D^2 \Big\}.
\end{aligned}
\tag{4.14}
$$

The set Ω_d can be thought of as a set of the spatial coordinates of the geometric centres of the SHI lenslets, restricted to the domain of its aperture Ω. Under the assumption of (4.12), one can then show [1] that the focal displacement $\Delta(x, y) = [\Delta_x(x, y), \Delta_y(x, y)]^T$ measured at some $(x, y) \in \Omega_d$ is related to the value of $\nabla \phi(x, y)$ according to

$$
\nabla \phi(x, y) \approx \frac{1}{F} \Delta \phi(x, y), \quad \forall (x, y) \in \Omega_d,
\tag{4.15}
$$

where F is the focal distance of the wavefront lenslets. An example of the above measurement setup is depicted in Fig. 4.2.

Now, provided a total of $M := \#\Omega_d$ ($\#\Omega_d$ denotes the cardinality of Ω_d) measurements of $\nabla \phi$ over Ω_d, one can approximate the coefficients $\{a_k\}_{k=1}^L$ of a truncated series expansion of ϕ as a solution to the least-square minimization problem given by
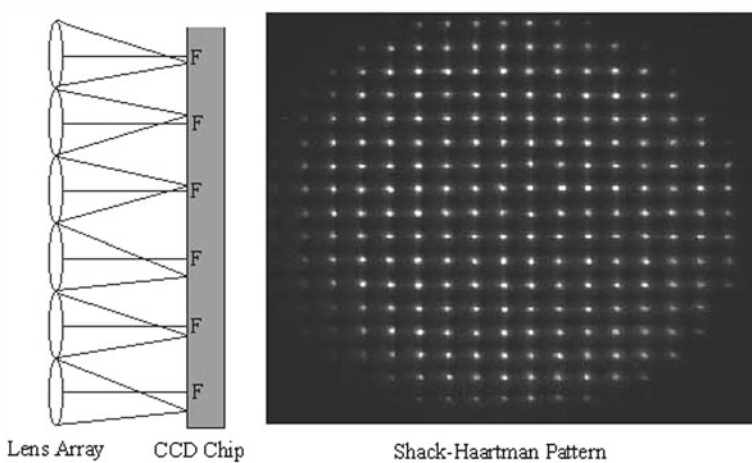

Lens Array   CCD Chip   Shack-Haartman Pattern

Fig. 4.2 Basic structure of the SHI and a resulting pattern of the focal points

$$\min_{\{a_k\}} \sum_{(x,y)\in\Omega_d} \Big\| \sum_{k=0}^{L} a_k \nabla Z_k(x,y) - F^{-1}\Delta(x,y) \Big\|_2^2, \qquad (4.16)$$

subject to appropriate boundary conditions. It is worthwhile noting that (4.16) can be rewritten in a vector-matrix form as

$$\min_{\mathbf{a}} \|\mathbf{Z}\mathbf{a} - \mathbf{d}\|_2^2, \quad \text{s.t. } \mathbf{a} \geq 0, \qquad (4.17)$$

where $\mathbf{Z}$ is a $2M \times L+1$ matrix of discrete values of the partial derivatives of the Zernike polynomials, $\mathbf{d}$ is a measurement (column) vector of length $2M$, and $\mathbf{a} = [a_0, a_1, \ldots, a_L]^T$ is a vector of the representation coefficients of ϕ. The constraint $\mathbf{a} \geq 0$ in (4.17) is optional and may be used to further regularize the solution by forcing $\mathbf{a}$ to belong to some convex set $\mathcal{K}_{\geq}$. Thus, for example, when the set coincides with the whole $\mathbb{R}^{L+1}$, the solution to (4.17) is given by

$$\mathbf{a} = \mathbf{Z}^{\#}\mathbf{d}, \qquad (4.18)$$

where $\mathbf{Z}^{\#}$ denotes the Moore-Penrose pseudo-inverse of $\mathbf{Z}$, whose definition is unique and stable as long as the row-rank of $\mathbf{Z}$ is greater or equal to $L + 1$ (hence suggesting that $2M \geq L + 1$). Having estimated $\mathbf{a}$, the phase ϕ can be approximated as

$$\phi(\rho, \varphi) \approx \sum_{k=0}^{L} a_k Z_k(\rho, \varphi). \qquad (4.19)$$

A higher accuracy of phase estimation requires using higher-order Zernike polynomials, which in turn necessitates a proportional increase in the number of wavefront lenses. Moreover, as required by the linearization procedure in the SHI, the lenses have to be of a relatively small sizes (sometimes, on the order of a few microns), which may lead to the use of a few thousand lenses per one interferometer. Accordingly, to simplify the construction and to reduce the cost of SHIs, we propose to reduce the number of wavefront lenslets, while compensating for the induced information loss through the use of DCS, which is detailed next.

4.3 Point Spread Function Estimation via Compressive Sampling

We now apply the proposed algorithm on this problem. First we show the side data, a source signal is a gradient field, can be transformed to (3.2) and then provide experiments that confirms that we can take advantage of the proposed scheme to improve the quality of image deblurring.

4.3.1 Derivative Compressed Sensing

Let the partial derivatives of ϕ evaluated at the points of set Ω_d be column-stacked into vectors $\mathbf{f}_x$ and $\mathbf{f}_y$ of length $M = \#\Omega_d$. In what follows, the partial derivatives $\mathbf{f}_x$ and $\mathbf{f}_y$ are assumed to be sparsely representable by an orthonormal basis in $\mathbb{R}^M$. Representing such a basis by an $M \times M$ unitary matrix W, the above assumption suggests the existence of two *sparse* vectors $\mathbf{c}_x$ and $\mathbf{c}_y$ such that $\mathbf{f}_x = W\mathbf{c}_x$ and $\mathbf{f}_y = W\mathbf{c}_y$. In the experimental studies of this section, the matrix W is constructed using the nearly symmetric orthogonal wavelets of I. Daubechies having five vanishing moments [32].

   The proposed simplification of the SHI amounts to reducing the number of wavefront lenslets. Formally, such a reduction can be described by two $n \times M$ sub-sampling matrices Ψ_x and Ψ_y, where $n < M$. Specifically, let $\mathbf{b}_x := \Psi_x \mathbf{f}_x$ and $\mathbf{b}_y := \Psi_y \mathbf{f}_y$ be incomplete (partial) observations of $\mathbf{f}_x$ and $\mathbf{f}_y$, respectively. Then, based on the theoretical guarantees of classical CS, the vectors $\mathbf{f}_x$ and $\mathbf{f}_y$ of the partial derivatives of ϕ can be approximated by $W\mathbf{c}_x^*$ and $W\mathbf{c}_y^*$, respectively, where $\mathbf{c}_x^*$ and $\mathbf{c}_y^*$ are obtained as

$$\mathbf{c}_x^* = \arg\min_{\mathbf{c}_x'} \left\{ \frac{1}{2} \|\Psi_x W \mathbf{c}_x' - \mathbf{b}_x\|_2^2 + \lambda_x \|\mathbf{c}_x'\|_1 \right\} \tag{4.20}$$

and

$$\mathbf{c}_y^* = \arg\min_{\mathbf{c}_y'} \left\{ \frac{1}{2} \|\Psi_y W \mathbf{c}_y' - \mathbf{b}_y\|_2^2 + \lambda_y \|\mathbf{c}_y'\|_1 \right\} \tag{4.21}$$

for some $\lambda_x, \lambda_y > 0$. Moreover, in the case when $\lambda_x = \lambda_y := \lambda$, computing the above estimates can be combined into a single optimization problem. Specifically, let $\mathbf{c} = [\mathbf{c}_x, \mathbf{c}_y]^T$, $\mathbf{b} = [\mathbf{b}_x, \mathbf{b}_y]^T$, and $A = \mathrm{diag}\{\Psi_x W, \Psi_y W\} \in \mathbb{R}^{2n \times 2M}$. Then,

$$\mathbf{c}^* = \arg\min_{\mathbf{c}'} \left\{ \frac{1}{2} \|A\mathbf{c}' - \mathbf{b}\|_2^2 + \lambda \|\mathbf{c}'\|_1 \right\}. \tag{4.22}$$

In this form, the problem (4.22) is identical to (1.4), in which case it can be solved by a variety of available tools of convex optimization [33, 34].

   The DCS algorithm augments classical CS by subjecting the minimization in (4.22) to an additional constraint which stems from the fact that [7]

$$\frac{\partial^2 \phi}{\partial x \, \partial y} = \frac{\partial^2 \phi}{\partial y \, \partial x}, \tag{4.23}$$

which is valid for all twice continuously differentiable functions ϕ. Thus, in the discrete setting, the above condition can be expressed using two partial differences matrices D_x and D_y, in which case it reads

$$D_x \mathbf{f}_y = D_y \mathbf{f}_x. \tag{4.24}$$

To further simplify the notations, let T_x and T_y be two coordinate-projection matrices, which map the composite vector $\mathbf{c}$ into $\mathbf{c}_x$ and $\mathbf{c}_y$ according to $T_x\mathbf{c} = \mathbf{c}_x$ and $T_y\mathbf{c} = \mathbf{c}_y$, respectively. Then (4.24) can be re-expressed in terms of $\mathbf{c}$ as

$$D_y W T_x \mathbf{c} = D_x W T_y \mathbf{c} \tag{4.25}$$

or, equivalently,

$$B\mathbf{c} = 0, \tag{4.26}$$

where $B := D_y W T_x - D_x W T_y$. Consequently, with the addition of the cross-derivative constraint (4.26), DCS solves the constrained minimization problem given by

$$\mathbf{c}^* = \arg\min_{\mathbf{c}'} \left\{ \frac{1}{2}\|A\mathbf{c}' - \mathbf{b}\|_2^2 + \lambda\|\mathbf{c}'\|_1 \right\}, \tag{4.27}$$
$$\text{s.t. } B\mathbf{c}' = 0.$$

The problem (4.27) is an instance of (3.8) and can be solved through the sequence of iterations produced by

$$\begin{cases} \mathbf{c}^{(t+1)} = \arg\min_{\mathbf{c}'} \left\{ \frac{1}{2}\|A\mathbf{c}' - \mathbf{b}\|_2^2 \right. \\ \qquad\qquad\qquad \left. + \lambda\|\mathbf{c}'\|_1 + \frac{\delta}{2}\|B\mathbf{c}' + p^{(t)}\|_2^2 \right\} \\ p^{(t+1)} = p^{(t)} + \delta B\mathbf{c}^{(t+1)}, \end{cases} \tag{4.28}$$

where $p^{(t)}$ is a vector of Bregman variables (or, equivalently, augmented Lagrange multipliers) and $\delta > 0$ is a user-defined parameter.[2]

Once an optimal $\mathbf{c}^*$ is recovered, it can be used to estimate the noise-free versions of $\mathbf{f}_x$ and $\mathbf{f}_y$ as $W T_x\mathbf{c}^*$ and $W T_y\mathbf{c}^*$, respectively. These estimates can be subsequently passed on to the fitting procedure to recover the values of ϕ, which, in combination with a known aperture function A, provide an estimate of the PSF i as an inverse discrete Fourier transform of the autocorrelation of $P = A e^{J\phi}$. Algorithm 1 below summarizes our method of estimation of the PSF.

The estimated PSF can be used to recover the original image u from v through the process of deconvolution as explained in the section that follows.

4.3.2 Deconvolution

The acquisition model 4.1 can be rewritten in an equivalent operator form as given by

---

[2] In this work, we use $\delta = 0.5$.

Algorithm 1: PSF estimation via DCS

---

1. *Data:* $\mathbf{b}_x$, $\mathbf{b}_y$, and $\lambda > 0$

2. *Initialization:* For a given transform matrix W and matrices/operators Ψ_x, Ψ_y, D_x, D_y, T_x and T_y, preset the procedures of multiplication by A, A^T, B and B^T.

3. *Phase recovery:* Starting with an arbitrary $\mathbf{c}^{(0)}$ and $p^{(0)} = 0$, iterate (4.28) until convergence to result in an optimal $\mathbf{c}^*$. Use the estimated (full) partial derivatives $WT_x\mathbf{c}^*$ and $WT_y\mathbf{c}^*$ to recover the values of ϕ over Ω.

4. *PSF estimation:* Using a known aperture function A, compute the inverse Fourier transform of $P = A\,e^{j\phi}$ to result in a corresponding ASF h. Estimate the PSF i as $i = |h|^2$.

$$v = \mathcal{H}\{u\} + \nu, \tag{4.29}$$

where $\mathcal{H}$ denote the operator of convolution with the estimated PSF i. Note that, in this case, the noise term ν accounts for both measurement noise as well as the inaccuracies related to estimation error in i.

The deconvolution problem of finding a useful approximation of u given its distorted measurement v can be addressed in many way, using a multitude of different techniques [35–39]. In this work, we use the ROF model and recover a regularized approximation of the original image u as

$$u^* = \arg\min_u \left\{ \frac{1}{2}\|\mathcal{H}\{u\} - v\|_2^2 + \gamma\,\|u\|_{TV} \right\}, \tag{4.30}$$

where $\|u\|_{TV} = \int\int |\nabla u|\,dx\,dy$ denotes the total variation (TV) semi-norm of u.

The minimization problem in (4.30) can be solved using a magnitude of possible approaches. One particularly efficient way to solve (4.30) is to substitute a direct minimization of the cost function in (4.30) by recursively minimizing a sequence of its local quadratic majorizers [38]. In this case, the optimal solution u^* can be obtained as the stationary point of a sequence of intermediate solutions produced by

$$\begin{cases} w^{(t)} = u^{(t)} + \mu\,\mathcal{H}^*\left\{v - \mathcal{H}\{u^{(t)}\}\right\} \\ u^{(t+1)} = \arg\min_u \left\{\frac{1}{2}\|u - w^{(t)}\|_2^2 + \gamma\,\|u\|_{TV}\right\}, \end{cases} \tag{4.31}$$

where $\mathcal{H}^*$ is the adjoint of $\mathcal{H}$ and μ is chosen to satisfy $\mu > \|\mathcal{H}^*\mathcal{H}\|$. In this work, the TV denoising at the second step of (4.31) has been performed using the fixed-point algorithm of Chambolle [14]. The convergence of (4.31) can be further improved by using the same FISTA algorithm of [38]. The resulting procedure is summarized below in Algorithm 2.

Algorithm 2: TV deconvolution using FISTA

---

1. *Initialize:* Select an initial value $u^{(0)}$; set $y^{(0)} = u^{(0)}$ and $\tau^{(0)} = 1$

2. *Repeat until convergence:*

   - $w^{(t)} = y^{(t)} + \mu\, \mathcal{H}^* \left\{ v - \mathcal{H}\{y^{(t)}\} \right\}$
   - $u^{(t+1)} = \arg\min_u \left\{ \frac{1}{2} \|u - w^{(t)}\|_2^2 + \gamma\, \|u\|_{TV} \right\}$
   - $\tau^{(t+1)} = 0.5 \left(1 + \sqrt{1 + 4\,(\tau^{(t)})^2} \right)$
   - $y^{(t+1)} = u^{(t+1)} + (\tau^{(t)}/\tau^{(t+1)})(u^{(t+1)} - u^{(t)})$

In summary, Algorithms 1 and 2 represent the essence of the proposed algorithm for hybrid deconvolution of short-exposure optical images. Next, experimental results are provided which further support the value and applicability of the proposed methodology.

4.4 Experiments

To demonstrate the viability of the proposed approach, its performance has been compared against two reference methods. The first reference method used a dense sampling of the phase (as it would have been the case with a conventional design of the SHI), thereby eliminating the need for a CS-based phase reconstruction. The resulting method is referred below to as the dense sampling (DS) approach. Second, to assess the importance of incorporation of the cross-derivative constraints, we have used both classical CS and DCS for phase recovery. In what follows, comparative results for phase estimation and subsequent deconvolution are provided for all the above methods.

4.4.1 Phase Recovery

To assess the performance of the proposed and reference methods under controllable conditions, simulation data was used. The random nature of atmospheric turbulence necessitates the use of statistical methods to model its effect on a wavefront propagation. Specifically, in this work, the effect of atmospheric turbulence was simulated based on the modified Von Karman model [40]. This model is derived based on Kolmogorov's theory of turbulence which models atmospheric turbulence using temperature fluctuations [41]. In particular, under some general assumptions on the velocity of turbulent medium and the distribution of its refraction index, the Von Karman power spectrum density is given by
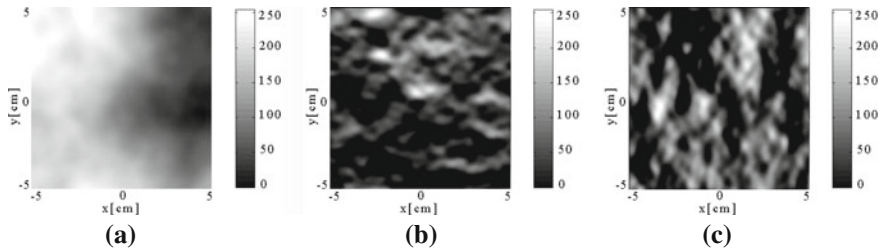
Fig. 4.3 An example of a simulated phase ϕ (**a**) along with its partial derivatives w.r.t. x (**b**) and y (**c**)

$$Q(t) = 0.033\, C_n^2 \, \frac{e^{(-t^2/t_m^2)}}{(t^2 + t_0^2)^{11/6}}, \tag{4.32}$$

where C_n^2 is the refractive-index and t_m, t_0 are chosen to match the high frequency and low frequency behaviour of turbulence. The model of (4.32) can be used to generate random realizations of the GPF phase, as described, e.g., in [2].

A typical example of the GPF phase ϕ is shown in subplot (a) of Fig. 4.3. In this case, the size of the phase screen was set to be equal to 10×10 cm, while the sampling was performed over a 128×128 uniform grid (which would have corresponded to the use of 16384 lenslets of a SHI). The corresponding values of the (discretized) partial derivatives $\partial\phi/\partial x$ and $\partial\phi/\partial y$ are shown in subplots (b) and (c) of Fig. 4.3, respectively.

The subsampling matrices Ψ_x and Ψ_y were obtained from an identity matrix I through a random subsampling of its rows by a factor resulting in a required compression ratio r. To sparsely represent the partial derivatives of ϕ, W was defined to correspond to a four-level orthogonal wavelet transform using the nearly symmetric wavelets of I. Daubechies with five vanishing moments [32] and periodic boundary condition.

To demonstrate the value of using the cross derivative constraint for phase reconstruction, the classical CS and DCS algorithms have been compared in terms of the mean squared errors (MSE) of their corresponding phase estimates. The results of this comparison are summarized in Fig. 4.4 for different compression ratios (or, equivalently, (sub)sampling densities) and SNR = 40 dB.

As expected, one can see that DCS results in lower values of MSE as compared to classical CS, which implies a higher accuracy of phase reconstruction. Moreover, the difference in the performances of classical CS and DCS appears to be more significant for lower sampling rates, while both algorithms tend to perform similarly when the sampling density approaches the DS case. Specifically, for the sampling density of $r = 0.3$, DCS results in a ten times smaller value of MSE as compared to the case of classical CS, whereas both algorithms have comparable performance for $r = 0.83$. This result suggests that, at higher compression rates, DCS is likely to result in more accurate reconstructions of the GPF phase as compared to the case of classical CS.
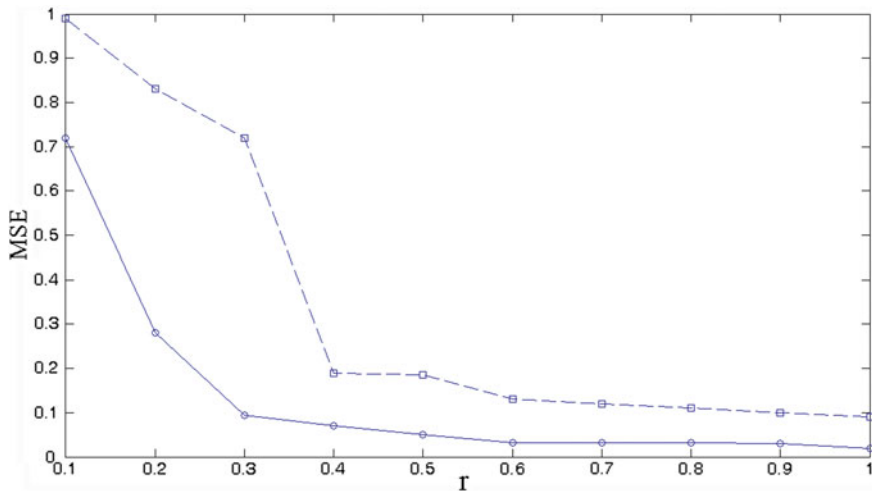
Fig. 4.4 MSE of phase reconstruction obtained with different methods as a function of r. Here, the *dashed* and *solid lines* correspond to classical CS and DCS, respectively, and SNR is equal to 40 dB

A number of typical reconstruction results are shown in Fig. 4.5, whose left and right subplots depict the phase estimates obtained using the classical CS and DCS algorithms, respectively, for the case of $r = 0.5$. The error maps of the two estimates are shown in subplot (c) and (d) of the same figure, which allows us to see the difference in the performance of these methods more clearly. Also, a close comparison with the original phase (as shown in subplot (a) of Fig. 4.3) reveals that DCS provides a more accurate recovery of the original ϕ, which further supports the value of using the cross-derivative constraints. In fact, exploiting these constraints effectively amounts to using additional "measurements", which are ignored in the case of classical CS.

As an additional comparison, Fig. 4.6 illustrates the convergence of the MSE as a function of the number of iterations, for both classical CS and DCS algorithms. One can see that DCS results in a substantially faster convergence as compared to classical CS. This behaviour could be explained by considering the cross-derivative constraints exploited by DCS to be effectively equivalent to noise-free measurements. To further investigate this argument, Fig. 4.7 compares the convergence of the cross-derivative fidelity term $\|D_y f_x - D_x f_y\|^2$ for both methods under comparison. One can see that, in the case of DCS, this term converges considerably faster than in the case of classical CS, which improves to the overall speed of convergence of DCS, making it superior to that of classical CS.

To investigate the robustness of the compared algorithms towards measurement noises, their performances have been compared for a range of SNR values. The results of this comparison are summarized in Fig. 4.8. Since the cross-derivative constraints exploited by DCS effectively restrict the feasibility region for an optimal solution, the algorithm exhibits an improved robustness to the effect of additive noise as compared
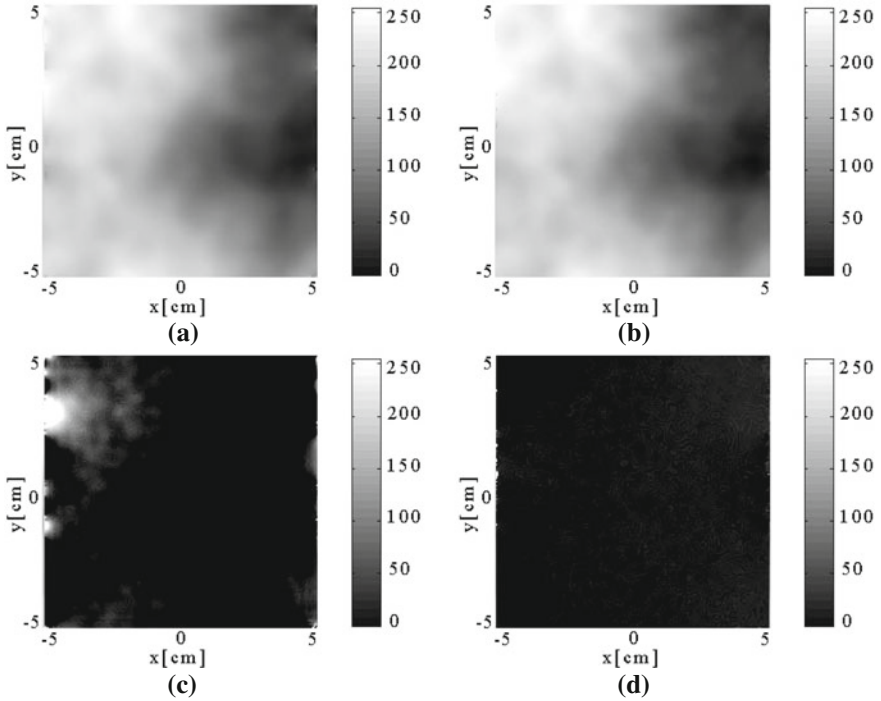
Fig. 4.5 **a** Phase reconstructed obtained by means of classical CS for SNR = 40 dB and $r = 0.5$; **b** Phase reconstructed obtained by means of DCS for the same values of SNR and r; **c** and **d** Corresponding error maps for classical CS and DCS

to the case of classical CS. This fact represents another advantage of incorporating the cross-derivative constraints in the process of phase recovery.

From the viewpoint of statistical estimation theory, the data fidelity terms in (4.27) suggests a Gaussian noise model, which may not be natural for all optical systems. In fact, this is the Poisson noise model, which is considered to be a more standard one in optical imagery. It turns out, however, that the use of the cross-derivative constraints by DCS makes it robust towards the inconsistency in noise modeling. This argument is supported by the results of Fig. 4.9, which summarizes the values of MSE obtained by classical CS and DCS reconstructions for different levels of Poisson noise. One can see that, in this case, the MSE values are comparable to the Gaussian case, while being substantially smaller in comparison to the CCS-based reconstruction.

It should be taken into account that, although the shape of ϕ does not change the energy of the PSF i, it plays a crucial role in determining its spatial behaviour. In the section that follows, it will be shown that even small inaccuracies in reconstruction of ϕ could be translated into dramatic difference in the quality of image deconvolution.
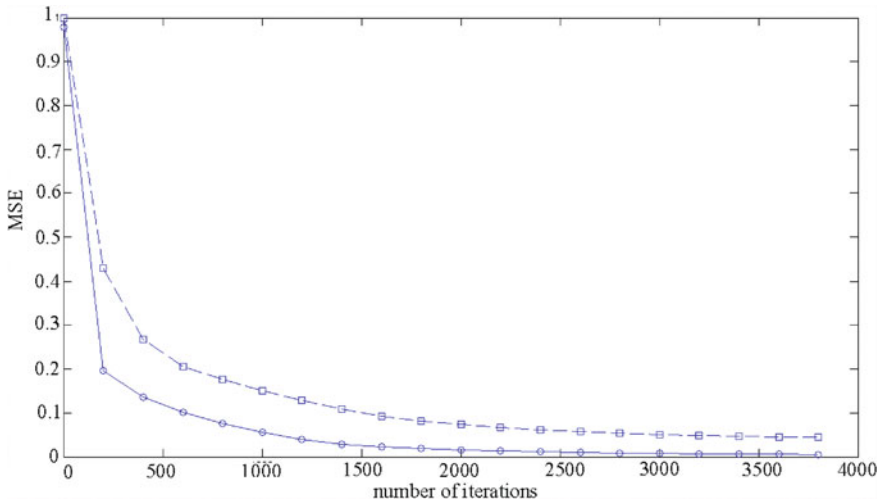
Fig. 4.6 Convergence analysis of phase reconstruction obtained with different methods as a function of iterations. Here, the *dashed* and *solid lines* correspond to classical CS and DCS, respectively, $SNR = 40$, and $r = 0.5$
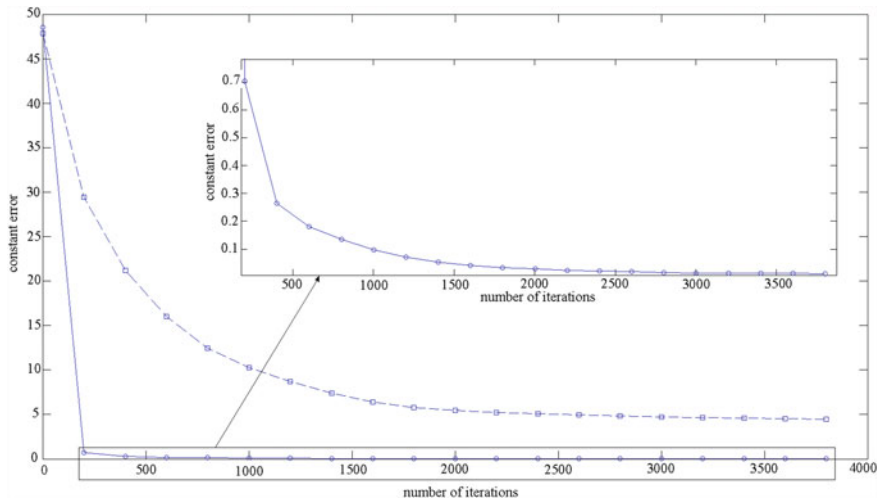

Fig. 4.7 Convergence analysis of derivative constraint obtained with different methods as a function of iterations. Here, the *dashed* and *solid lines* correspond to classical CS and DCS, respectively, $SNR = 40$, and $r = 0.5$

4.4.2 Deblurring

As a next step, the phase estimates obtained using the CCS- and DCS-based methods for $r = 0.5$ were combined with the aperture function A to result in their respective estimates of the PSF i. These estimates were subsequently used to deconvolve a
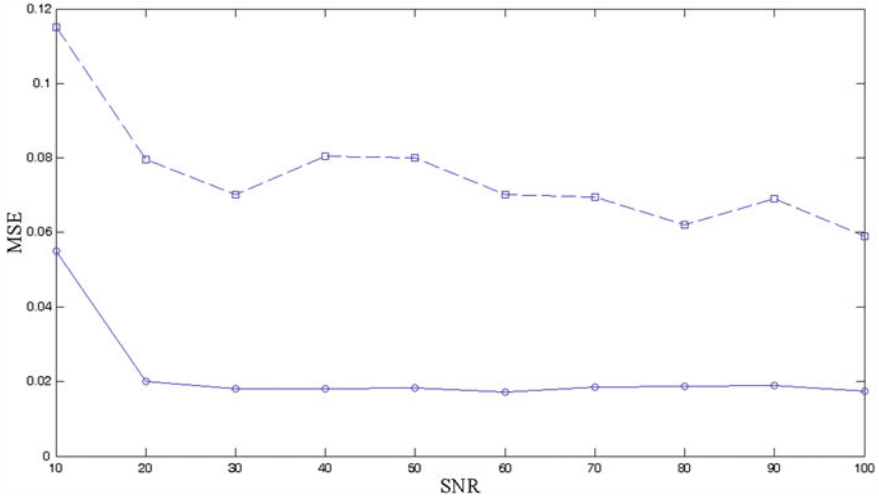
Fig. 4.8 MSE of phase reconstruction obtained with different methods as a function of SNR. Here, the *dashed* and *solid lines* correspond to classical CS and DCS, respectively, and $r = 0.5$
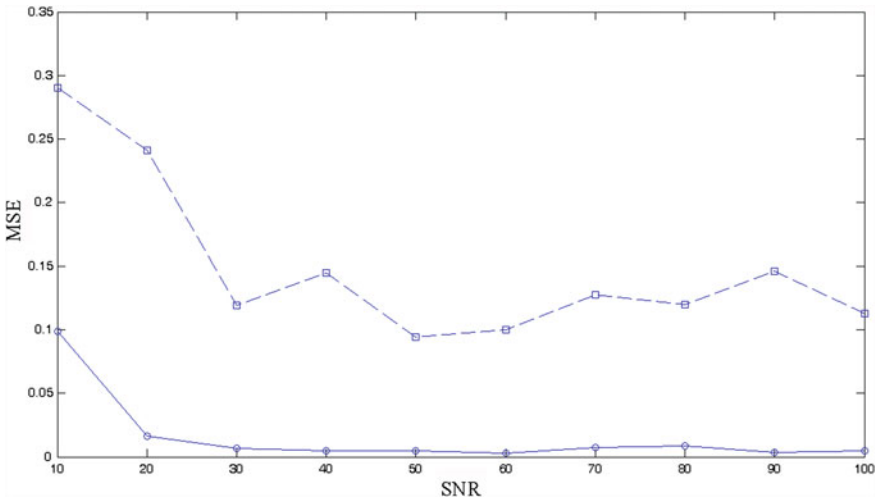

Fig. 4.9 MSE of phase reconstruction obtained with different methods as a function of SNR where the noise model is Poisson. Here, the *dashed* and *solid lines* correspond to classical CS and DCS, respectively, and $r = 0.5$

number of test images such as "Satellite", "Saturn", "Moon" and "Galaxy". All the test images were blurred with an original PSF, followed by their contamination with additive Gaussian noise of different levels which is controlled by the variance of noise distribution. As an example, Fig. 4.10 shows the "Satellite" image [subplot (a)] along with its blurred and noisy version [subplot (b)].
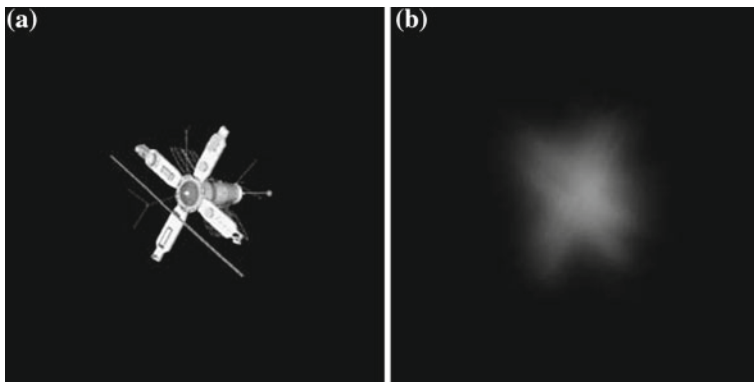
Fig. 4.10 Satellite image (**a**) and its blurred and noisy version (**b**)

Table 4.1 SSIM and PSNR comparisons of phase recovery results

Image	PSNR comparison (dB)				SSIM comparison			
	Noise std	Blurred	DS	CS	DCS	Blurred	DS	CS	DCS
	10^{-5}	14.06	27.97	17.06	27.42	0.200	0.730	0.349	0.674
Satellite	0.001	14.06	27.75	16.93	27.22	0.200	0.720	0.344	0.667
	0.003	14.06	25.97	16.54	25.56	0.199	0.554	0.306	0.519
	0.005	14.05	22.43	15.63	22.22	0.197	0.269	0.206	0.263
	10^{-5}	17.78	31.49	23.42	31.02	0.226	0.688	0.424	0.656
Saturn	0.001	17.78	31.08	23.38	30.65	0.226	0.66	0.416	0.641
	0.003	17.78	28.50	22.80	28.30	0.226	0.506	0.348	0.483
	0.005	17.78	23.89	20.55	23.72	0.175	0.228	0.212	0.223
	10^{-5}	19.98	25.06	22.36	25.00	0.512	0.645	0.539	0.643
Moon	0.001	19.97	25.04	22.38	24.99	0.512	0.642	0.538	0.64
	0.003	19.97	24.83	22.30	24.78	0.509	0.607	0.493	0.604
	0.005	19.97	21.76	19.73	21.73	0.504	0.552	0.488	0.549
	10^{-5}	18.79	23.58	21.16	23.52	0.257	0.493	0.348	0.490
Galaxy	0.001	18.79	23.60	21.12	23.54	0.257	0.495	0.347	0.491
	0.003	18.78	23.38	20.64	23.32	0.257	0.501	0.326	0.501
	0.005	18.78	20.93	18.46	20.86	0.254	0.397	0.224	0.393

   Using the PSF estimates, the deconvolution was carried out using the method
detailed in [14]. For the sake of comparison, the deconvolution was also performed
using the PSF recovered from dense sampling (DS) of ϕ. Note that this reconstruction
is expected to have the best accuracy, since it neither involves undersampling nor
requires a CS-based phase estimation. All the deconvolved images have been com-
pared with their original counterparts in terms of PSNR as well as of the structural
similarity index (SSIM) of [42], which is believed to be a better indicator of perceptual
image quality [43]. The resulting values of the comparison metrics are summarized
in Table 4.1, while Fig. 4.11 shows the deconvolution results produced by the CCS-
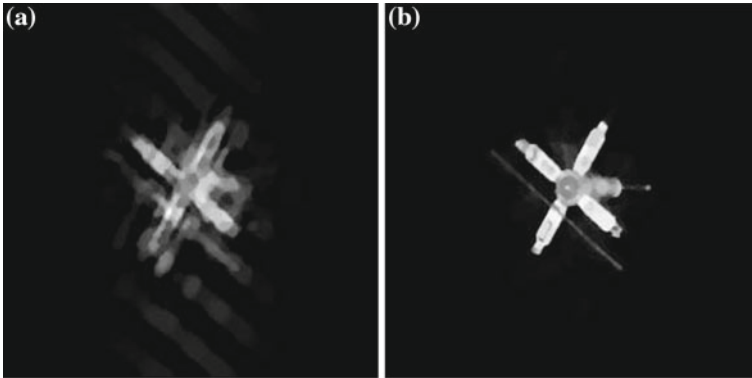and DCS-based methods.

Fig. 4.11 a Image estimate obtained with the CCS-based method for phase recovery (SSIM = 0.781). **b** Image estimate obtained with the DCS-based method for phase recovery (SSIM = 0.917)
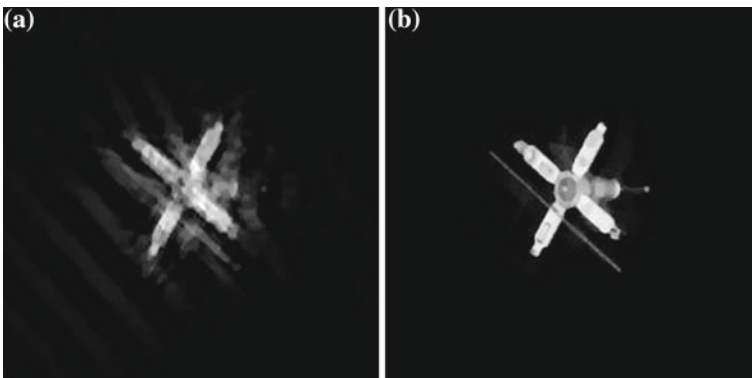

Fig. 4.12 a Image estimate obtained with the CCS-based method for phase recovery (SSIM = 0.732). **b** Image estimate obtained with the DCS-based method for phase recovery (SSIM = 0.888) where the noise model is assumed to be Poisson

The above results demonstrate the importance of accurate phase recovery, where even a relatively small phase error can have a dramatic effect on the quality of image deconvolution. Under such conditions, the proposed method produces image reconstructions of a superior quality as compared to the case of classical CS. Moreover, comparing the results of Table 4.1, one can see that DS only slightly outperforms DCS in terms of PSNR and SSIM, while in many practical cases, the difference between the performances of these methods are hard to detect visually.

Finally, Fig. 4.12 shows the results of CCS-based and DCS-based image reconstruction for the case of Poisson noise contamination. A close comparison of these results reveals a noticeable degradation in the performance of the CCS-based algorithm, while the DCS-based results are virtually indistinguishable from those obtained in the Gaussian case.

4.5 Summary

In this chapter, the applicability of the proposed scheme to the practical problem of image deblurring in optical imaging was studied. It was shown that, in the presence of atmospheric turbulence, the phase ϕ of the GPF $P = A\,e^{J\phi}$ is a random function, which needs to be measured using adaptive optics. To simplify the complexity of the latter, a CS-based approach was proposed. As opposed to classical CS, however, the proposed method performs phase reconstruction subject to an additional constraint, which stems from the property of $\nabla\phi$ to be a potential field. The DCS algorithm has been shown to yield phase estimates of substantially better quality as compared to the case of classical CS. our main focus has been on simplifying the structure of the SHI through reducing the number of its wavefront lenslets, while compensating for the effect of undersampling by means of DCS. The resulting phase estimates were used to recover their associated PSF, which was subsequently used for image deconvolution. It was shown that the DCS-based estimation of ϕ with $r = 0.3$ results in image reconstructions of the quality comparable to that of DS, while substantially outperforming the results obtained with classical CS.

References

1. J. Primot, G. Rousset, J.C. Fontanella, Deconvolution from wave-front sensing: A new technique for compensating turbulence-degraded images. J. Opt. Soc. Am. **7**, 1598–1608 (1990)
2. M.C. Roggemann, B.M. Welsh, *Imaging Through Turbulence* (CRC Press, Boca Raton, 1996)
3. J. Yang, J. Wright, T.S. Huang, Y. Ma, Image super-resolution via sparse representation. IEEE Trans. Image Processing **19**, 2861–2873 (2010)
4. M. Elad, M. Aharon, Image denoising via sparse and redundant representations over learned dictionaries. IEEE Trans. Image Processing **17**, 3736–3745 (2006)
5. I.J. Mairal, G. Sapiro, M. Elad, Learning multiscale sparse representations for image and video restoration. Multiscale Modeling and Simulation **7**, 214–241 (2008)
6. R.T. Paul, Review of robust video watermarking techniques. IJCA Special Issue on Computational Science **3**, 90–95 (2011)
7. G.D. Boreman, *Modulation Transfer Function in Optical and Electro-Optical Systems* (SPIE Optical Engineering Press, Bellingham, Washington, 2001)
8. R.T. Paul, Blind deconvolution via cumulant extrema. IEEE Signal Processing Magazine **3**, 24–42 (1996)
9. D. Kundur, D. Hatzinakos, Blind image deconvolution. IEEE Signal Processing Magazine **3**, 43–64 (1996)
10. J.K. Kauppinen, D.J. Moffatt, H.H. Mantsch, D.G. Cameron, Fourier self-deconvolution: A method for resolving intrinsically overlapped bands. Applied Spectroscopy **35**, 271–276 (1981)
11. S. Geman and D. Geman. Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. IEEE Trans. Pattern Analysis and Machine Intelligence, PAMI-6:721–741, 1984
12. T. Poggio, V. Torre, C. Koch, Computational vision and regularization theory. Nature **317**, 314–319 (1985)
13. L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms. Phys. D **60**, 259–268 (November 1992)
14. A. Chambolle, An algorithm for total variation minimization and applications. Journal of Mathematical Imaging and Vision **20**, 89–97 (2004)

15. T. Goldstein, S. Osher, The split Bregman method for l_1-regularized problems. SIAM J. Img. Sci. **2**, 323–343 (2009)
16. A. Marquina, Nonlinear inverse scale space methods for total variation blind deconvolution. SIAM J. Img. Sci. **2**, 64–83 (2009)
17. L. He, A. Marquina, S. Osher, Blind deconvolution using TV regularization and Bregman iteration. International Journal of Imaging Systems and Technology **15**, 74–83 (2005)
18. O. Michailovich, A. Tannenbaum, Blind deconvolution of medical ultrasound images: Parametric inverse filtering approach. IEEE Trans. Image Processing **16**(12), 3005–3019 (December 2007)
19. W.H. Richardson, Bayesian-based iterative method of image restoration. J. Opt. Soc. Am. A **62**(1), 55–59 (1972)
20. L.B. Lucy, An iterative technique for the rectification of observed distributions. Astron. J. **79**(6), 745–754 (1974)
21. P.A. Jansson, Deconvolution of Images and Spectra. Opt. Eng. **36**, 3224 (1997)
22. M. J. Cullum. Adaptive Optics. European Southern, Observatory, 1996
23. D. Dayton, B. Pierson, B. Spielbusch, J. Gonglewski, Atmospheric structure function measurements with a Shack-Hartmann wave-front sensor. Journal of Mathematical Imaging and Vision **20**, 89–97 (2004)
24. R.G. Lane, M. Tallon, Wave-front reconstruction using a Shack Hartmann sensor. Applied Optics **31**, 6902–6908 (1992)
25. R. Irwan, R.G. Lane, Analysis of optimal centroid estimation applied to Shack-Hartmann sensing. Applied Optics **38**(32), 6737–6743 (1999)
26. Y. Eldar, P. Kuppinger, H. Bölcskei, Block-sparse signals: Uncertainty relations and efficient recovery. IEEE Trans. Signal Process **58**(6), 3042–3054 (2010)
27. D.L. Fried, Statistics of a geometric representation of wavefront distortion. J. Opt. Soc. Am. **55**, 1427–1431 (1965)
28. V. Stanković, L. Stanković, and S. Cheng. Compressive image sampling with side information. In Proceedings of the 16th IEEE International Conference on Image Processing, ICIP'09, pages 3001–3004, 2009
29. T.O. Salmon, L.N. Thibos, A. Bradley, Comparison of the eyes wave-front aberration measured psychophysically and with the ShackHartmann wave-front sensor. Journal of the Optical Society of America A **15**, 2457–2465 (2007)
30. O. Michailovich, A. Tannenbaum, A fast approximation of smooth functions from samples of partial derivatives with application to phase unwrapping. Signal Processing **88**, 358–374 (2008)
31. M. Hosseini, O. Michailovich, *Derivative compressive sampling with application to phase unwrapping* (In Proceedings of EUSIPCO, Glasgow, UK, August, 2009)
32. I. Daubechies, *Ten Lectures on Wavelets* (SIAM, CBMS-NSF Reg. Conf. Series in Applied Math, 1992)
33. D. L. Donoho and Y. Tsaig. Fast solution of l_1-norm minimization problems when the solution may be sparse. Technical Report 2006–18, Stanford, 2006
34. S. Osher, M. Burger, D. Goldfarb, J. Xu, W. Yin, An iterative regularization method for total variation-based image restoration. Simul **4**, 460–489 (2005)
35. A. Savitzky, M.J.E. Golay, Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. **36**, 1627–1639 (1964)
36. A.N. Tikhonov, V.Y. Arsenin, *Solutions of Ill-Posed Problem*, vol. H (Winston, Washington, D.C., 1977)
37. Å. Björck, *Numerical methods for least squares problems* (SIAM, Philadelphia, 1996)
38. A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM Journal on Imaging Sciences **2**, 183–202 (2009)
39. O. Michailovich, An iterative shrinkage approach to total-variation image restoration. IEEE Trans. Image Process **20**(5), 1281–1299 (2011)
40. J.D. Schmidt, *Numerical Simulation of Optical Wave Propagation with Examples in MATLAB* (SPIE, Washington, 2010)

41. I. Daubchies, M. Defrise, C.D. Mol, An iterative thresholding algorithm for linear inverse
    problems with sparsity constraint. Comm. Pure Appl. Math. **75**, 1412–1457 (2009)
42. Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: From error
    visibility to structural similarity. IEEE Trans. Image Process **13**(4), 600–612 (2004)
43. Z. Wang, A.C. Bovik, Mean squared error: Love it or leave it? - A new look at signal fidelity
    measures. IEEE Signal Processing Magazine **26**(1), 98–117 (2009)