

Image Deblurring Using Derivative Compressed Sensing for Optical Imaging Application

Mohammad Rostami, *Student Member, IEEE*, Oleg Michailovich, *Member, IEEE*, and Zhou Wang, *Member, IEEE*

Abstract—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 recovered. In this paper, 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. 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 of the optical system in use. A standard way to track such variations involves using adaptive optics. Thus, 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 paper, 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.

Index Terms—Deconvolution, derivative compressive sampling, inverse problem, Shack–Hartmann interferometer (SHI).

I. INTRODUCTION

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 [1]–[4]. 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 function¹ (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

$$v = i * u + \nu. \quad (1)$$

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 [6].

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 [7]. 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 [8], [9]. Thus, for example, in the case when u is known to be an image of bounded variation, the given regularization leads to the famous Rudin–Osher–Fatemi reconstruction scheme, in which u is estimated as a solution to the following optimization problem [10], [11]:

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

where $\alpha > 0$ is the regularization parameter. It should be noted that, if the PSF obeys $\int I \, dx dy \neq 0$, problem (2) is strictly convex and therefore admits a unique minimizer, which can be computed using a spectrum of available algorithms [10]–[14].

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 [7]. In this paper, however, we follow the philosophy of *hybrid* deconvolution [15], 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 paper, 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 in the following.

¹Note that, in optical imaging, this function is also referred to as an impulse transfer function [5].

Manuscript received July 28, 2011; revised January 17, 2012; accepted February 29, 2012. Date of publication March 12, 2012; date of current version June 13, 2012. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada and in part by the Ontario Early Researcher Award Program. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Pascal Frossard.

The authors are with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: m2rostam@uwaterloo.ca).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2012.2190610

Optical imaging is unarguably the field of applied sciences from which the notion of image deconvolution has originated [16]–[18]. In particular, in short-exposure turbulent imaging [19], acquired images are blurred with a PSF, which depends on a spatial distribution of the atmospheric refraction index along the optical path connecting 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) [20]. As will be shown later in this paper, the PSF of a short-exposure optical system is determined by its corresponding 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, phase ϕ accounts for turbulence-induced aberrations of the optical wavefront and, hence, is generally unknown at any given experimental time. Fortunately, 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) (also known as the Shack–Hartmann wavefront sensor) [21], [22], 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 the following). 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 [23]. As will be shown in the following, 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 PSF i , which can in turn be used by a deconvolution algorithm. Thus, the proposed method for estimation of 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.

The rest of this paper is organized as follows. Some necessary technical preliminaries are summarized in Section II. In Section III, the main principles of SHI-based phase reconstruction are presented. The DCS algorithm, along with its efficient numerical implementation, is introduced in Section IV. In Section V, the resulting estimates of the PSF are used for

deconvolution of short-exposure turbulent images. Experimental results are presented in Section VI, whereas Section VII finalizes this paper with a discussion and conclusions.

II. 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 [19]. To better understand the setup under consideration, we first note that, in optical imaging, 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 GPF $P(x, y)$ that is given by [24]

$$h(\xi, \eta) = \frac{1}{\lambda_w z_i} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(x, y) e^{-j \frac{2\pi}{\lambda_w z_i} (x\xi + y\eta)} dx dy \quad (3)$$

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

$$P(x, y) = A(x, y) e^{j\phi(x, y)}. \quad (4)$$

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 [19]

$$A(r) = \begin{cases} 1, & \text{if } r \leq \frac{D}{2} \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

where D denotes the pupil diameter. Thus, given $\phi(x, y)$, one could determine h and, therefore, i . Unfortunately, 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 [21]. 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 [25]. 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.

In this paper, to minimize the effect of phase undersampling, we exploit the DCS algorithm of [23]. 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. Using this “side information,” which are called the cross-derivative constraints, allows substantially improving the quality of reconstruction of $\phi(x, y)$, as compared with the case of CCS-based estimation.

III. SHI

As it was mentioned earlier in this paper, 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 $\{Z_k\}_{k=0}^{\infty}$, as shown in the following [22]:

$$\phi(x, y) = \sum_{k=0}^{\infty} a_k Z_k(x, y) \quad (6)$$

where the representation coefficients $\{a_k\}_{k=0}^{\infty}$ are supposed to be unique and stably computable. Note that, in this case, the data of $\{a_k\}_{k=0}^{\infty}$ uniquely identify $\phi(x, y)$, whereas the coefficients $\{a_k\}_{k=0}^{\infty}$ can be estimated due to the linearity of (6) that suggests

$$\nabla\phi(x, y) = \sum_{k=0}^{\infty} a_k \nabla Z_k(x, y). \quad (7)$$

In AO, it is conventional to define $\{Z_k\}_{k=0}^{\infty}$ to be Zernike polynomials (also known as Zernike functions) [24]. 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

$$Z_n^m(\rho, \varphi) = R_n^m(\rho) \cos(m\varphi) \quad (8)$$

$$Z_n^{-m}(\rho, \varphi) = R_n^m(\rho) \sin(m\varphi) \quad (9)$$

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 (8) and (9) are defined as

$$R_n^m(\rho) = \sum_{k=0}^{(n-m)/2} \frac{(-1)^k (n-k)}{k!((n+m)/2-k)!((n-m)/2-k)!} \rho^{n-2k}. \quad (10)$$

Note that, since the Zernike polynomials are defined using polar coordinates, it makes sense to reexpress the phase ϕ and its gradient in the polar coordinate system as well (technically, this would amount to replacing x and y in (6) and (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 (6) and (7) can be computed by orthogonal projection, namely

$$a_k = \int_0^{2\pi} \int_0^1 \phi(\rho, \varphi) Z_k(\rho, \varphi) \rho d\rho d\varphi. \quad (11)$$

In practice, however, $\phi(\rho, \varphi)$ is unknown; 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 $\nabla\phi(\rho, \varphi)$.

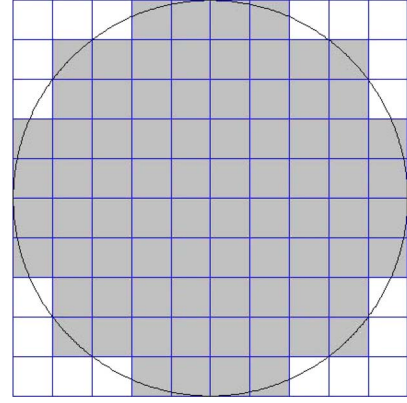


Fig. 1. Example of a 10×10 SHI array on a circular aperture. The shading indicates those blocks (i.e., lenslets) that are rendered active.

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. 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 asymptotically improves 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

$$\phi(x, y) \approx ax + by + c \quad (12)$$

for all (x, y) in a neighborhood of $(\rho \cos \varphi, \rho \sin \varphi)$. The approximation in (12) suggests that

$$\nabla\phi(x, y) \approx (a, b)^T \quad (13)$$

where $(\cdot)^T$ denotes matrix transposition. While c in (12) can be derived from boundary conditions, coefficients a and b should be determined through direct measurements. 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 toward 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.

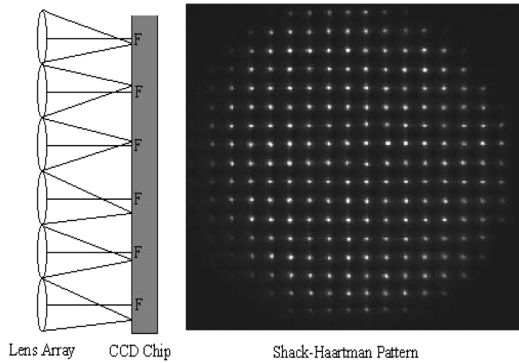


Fig. 2. Basic structure of the SHI and a resulting pattern of the focal points.

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

$$\begin{aligned} \Omega_d := & \left\{ (x_d, y_d) \in \Omega \right. \\ & x_d = -D + \frac{2D}{N} \left(i + \frac{1}{2} \right), \quad i = 0, 1, \dots, N-1 \\ & y_d = -D + \frac{2D}{N} \left(j + \frac{1}{2} \right), \quad j = 0, 1, \dots, N-1 \\ & \left. \text{and } x_d^2 + y_d^2 \leq D^2 \right\}. \end{aligned} \quad (14)$$

The set Ω_d can be thought of as a set of the spatial coordinates of the geometric centers of the SHI lenslets, which are restricted to the domain of its aperture Ω . Under the assumption of (12), one can then show [26] 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

$$\nabla\phi(x, y) \approx \frac{1}{F} \Delta\phi(x, y) \quad \forall (x, y) \in \Omega_d \quad (15)$$

where F is the focal distance of the wavefront lenslets. An example of the given measurement setup is depicted in Fig. 2.

Now, provided a total of $M := \#\Omega_d$ 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 squares minimization problem, as given by

$$\min_{\{a_k\}} \sum_{(x,y) \in \Omega_d} \left\| \sum_{k=0}^L a_k \nabla Z_k(x, y) - F^{-1} \Delta(x, y) \right\|_2^2 \quad (16)$$

subject to appropriate boundary conditions. It is worthwhile noting that (16) can be rewritten in a vector–matrix form as

$$\min_{\mathbf{a}} \|\mathbf{Z}\mathbf{a} - \mathbf{d}\|_2^2, \quad \text{s.t. } \mathbf{a} \succeq 0 \quad (17)$$

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 $\mathbf{a} = [a_0, a_1, \dots, a_L]^T$ is a vector of the representation coefficients of ϕ . The constraint $\mathbf{a} \succeq 0$ in (17) is optional and may be used to further regularize the solution by forcing \mathbf{a} to belong to some convex set

\mathcal{K}_{\succeq} . Thus, for example, when the set coincides with the whole \mathbb{R}^{L+1} , the solution to (17) is given by

$$\mathbf{a} = \mathbf{Z}^\# \mathbf{d} \quad (18)$$

where $\mathbf{Z}^\#$ denotes the Moore–Penrose pseudoinverse 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} , phase ϕ can be approximated as

$$\phi(\rho, \varphi) \approx \sum_{k=0}^L a_k Z_k(\rho, \varphi). \quad (19)$$

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 micrometers), 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 in the following.

IV. DCS

A. Classical Compressed Sensing

Central to signal processing is the Shannon–Nyquist theorem [27], which specifies conditions on which a band-limited signal can be stably and uniquely recovered from its discrete measurements. However, in around 2005, a different sampling theorem was formulated that, in some cases, abrogates the fundamentals of its predecessor. This new theory, nowadays known as compressed sensing (also known as compressive sampling), asserts that signals, which admit a sparse representation in a predefined basis/frame, can be recovered from their discrete measurements, whose number is proportional to the ℓ_0 -norm of the coefficients of the sparse representation. In such a case, the sparser the representation of the signal is, the smaller the number of measurements required for signal reconstruction can be. As a result, cases are numerous in which the sampling efficiency of compressed sensing far supersedes that of the classical Shannon–Nyquist sampling [28], [29].

The CCS setting assumes a linear measurement model, which the observations $\mathbf{y} \in \mathbb{R}^m$ of some unknown quantity of interest $\mathbf{x} \in \mathbb{R}^n$ are acquired according to

$$\mathbf{y} = \Psi \mathbf{x} \quad (20)$$

where $\Psi \in \mathbb{R}^{m \times n}$ stands for an observation (sampling) matrix with $n > m$. Because of the property of the linear system in (20) of being underdetermined, the recovery of \mathbf{x} from \mathbf{y} is impossible unless it is known that \mathbf{x} is sparse and hence has a relatively low value of $\|\mathbf{x}\|_0$. In such a case, if the sampling matrix Ψ obeys the restricted isometry property [28], [29] with respect to a predefined class of sparse signals (to which \mathbf{x} is believed

to belong), CCS allows \mathbf{x} to be recovered as a solution to [30], [31]

$$\mathbf{x} = \arg \min_{\mathbf{x}'} \{ \|\mathbf{x}'\|_1 | \Psi \mathbf{x}' = \mathbf{y} \} \quad (21)$$

which is a convex minimization problem, which is straightforward to reformulate in terms of linear programming. Moreover, in the case when the measurements \mathbf{y} are error prone, a more robust version of CCS is to recover \mathbf{x} according to

$$\mathbf{x} = \arg \min_{\mathbf{x}'} \{ \|\mathbf{x}'\|_1 | \|\Psi \mathbf{x}' - \mathbf{y}\|_2^2 \leq \epsilon \} \quad (22)$$

where $\epsilon > 0$ is a parameter controlling the size of measurement noise. Moreover, it was shown in [28] and [29] that the estimation error in the signal reconstructed according to (22) can be bounded by a linear function of ϵ . This implies robustness of the CCS reconstruction toward the presence of measurement noise.

It should be finally noted that the optimization problem (22) can be reformulated in its equivalent Lagrangian form, in which case one has

$$\mathbf{x} = \arg \min_{\mathbf{x}'} \left\{ \frac{1}{2} \|\Psi \mathbf{x}' - \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}'\|_1 \right\} \quad (23)$$

with $\lambda > 0$ being an optimal Lagrange multiplier [30]. In what follows, it is assumed that an optimal value of λ is known. (For more details on this subject, the reader is referred to [30] as well as to the later sections of this paper).

B. DCS

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 given 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 reported in this paper, matrix W is constructed using the nearly symmetric orthogonal wavelets of 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 $n \times M$ subsampling 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 CCS, 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

$$\mathbf{c}_x^* = \arg \min_{\mathbf{c}_x'} \left\{ \frac{1}{2} \|\Psi_x W \mathbf{c}_x' - \mathbf{b}_x\|_2^2 + \lambda_x \|\mathbf{c}_x'\|_1 \right\} \quad (24)$$

$$\mathbf{c}_y^* = \arg \min_{\mathbf{c}_y'} \left\{ \frac{1}{2} \|\Psi_y W \mathbf{c}_y' - \mathbf{b}_y\|_2^2 + \lambda_y \|\mathbf{c}_y'\|_1 \right\} \quad (25)$$

for some $\lambda_x, \lambda_y > 0$. Moreover, in the case when $\lambda_x = \lambda_y := \lambda$, computing the given 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 $A = \text{diag}\{\Psi_x W, \Psi_y W\} \in \mathbb{R}^{2n \times 2M}$. Then

$$\mathbf{c}^* = \arg \min_{\mathbf{c}'} \left\{ \frac{1}{2} \|A\mathbf{c}' - \mathbf{b}\|_2^2 + \lambda \|\mathbf{c}'\|_1 \right\}. \quad (26)$$

In this form, problem (26) is identical to (23), in which case it can be solved by a variety of available tools of convex optimization [30], [31].

The DCS algorithm augments CCS by subjecting the minimization in (26) to an additional constraint that stems from the fact that [23]

$$\frac{\partial^2 \phi}{\partial x \partial y} = \frac{\partial^2 \phi}{\partial y \partial x} \quad (27)$$

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

$$D_x \mathbf{f}_y = D_y \mathbf{f}_x. \quad (28)$$

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, (31) can be reexpressed in terms of \mathbf{c} as

$$D_y W T_x \mathbf{c} = D_x W T_y \mathbf{c} \quad (29)$$

or, equivalently

$$B\mathbf{c} = 0 \quad (30)$$

where $B := D_y W T_x - D_x W T_y$. Consequently, with the addition of the cross-derivative constraint (30), DCS solves the constrained minimization problem given by

$$\mathbf{c}^* = \arg \min_{\mathbf{c}'} \left\{ \frac{1}{2} \|A\mathbf{c}' - \mathbf{b}\|_2^2 + \lambda \|\mathbf{c}'\|_1 \right\}, \quad \text{s.t. } B\mathbf{c}' = 0. \quad (31)$$

A solution to (31) can be found, for instance, by means of the Bregman algorithm [33], in which case \mathbf{c}^* is obtained as a stationary point of the sequence of iterations produced by

$$\begin{cases} \mathbf{c}^{(t+1)} = \arg \min_{\mathbf{c}'} \\ \quad \times \left\{ \frac{1}{2} \|A\mathbf{c}' - \mathbf{b}\|_2^2 + \lambda \|\mathbf{c}'\|_1 + \frac{\delta}{2} \|B\mathbf{c}' + p^{(t)}\|_2^2 \right\} \\ p^{(t+1)} = p^{(t)} + \delta B\mathbf{c}^{(t+1)} \end{cases} \quad (32)$$

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

Note that the \mathbf{c} -update step in (32) has the format of a standard basis pursuit denoising problem [34], which can be solved by a variety of optimization methods [35]. In this paper, we used the FISTA algorithm in [36] due to the simplicity of its implementation as well as for its remarkable convergence properties. It should be noted that the algorithm does not require explicitly

²In this paper, we use $\delta = 0.5$.

defining the matrices A and B . Only the *operations* of multiplication by these matrices and their transposes need to be known, which can be implemented in an implicit and computationally efficient manner.

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 $WT_x\mathbf{c}^*$ and $WT_y\mathbf{c}^*$, respectively. These estimates can be subsequently passed on to the fitting procedure in Section III 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 $P = Ae^{j\phi}$. Algorithm 1 below summarizes our method of estimation of the PSF.

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 (32) 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 $P = Ae^{j\phi}$ to result in a corresponding ASF h . Estimate the PSF i as $i = |h|^2$.
-

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.

V. DECONVOLUTION

The acquisition model (1) can be rewritten in an equivalent operator form as given by

$$v = \mathcal{H}\{u\} + \nu \quad (33)$$

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 [36]–[38]. In this paper, we use the ROF model and recover a regularized approximation of the original image u as

$$u^* = \arg \min_u \left\{ \frac{1}{2} \|\mathcal{H}\{u\} - v\|_2^2 + \gamma \|u\|_{TV} \right\} \quad (34)$$

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

The minimization problem in (34) can be solved using a magnitude of possible approaches. One particularly efficient way to solve (34) is to substitute a direct minimization of the cost function in (34) by recursively minimizing a sequence of its local quadratic majorizers [36]. In this case, the optimal solution u^*

can be obtained as the stationary point of a sequence of intermediate solutions produced by

$$\begin{cases} w^{(t)} = u^{(t)} + \mu \mathcal{H}^* \{v - \mathcal{H}\{u^{(t)}\}\} \\ u^{(t+1)} = \arg \min_u \left\{ \frac{1}{2} \|u - w^{(t)}\|_2^2 + \gamma \|u\|_{TV} \right\} \end{cases} \quad (35)$$

where \mathcal{H}^* is the adjoint of \mathcal{H} and μ is chosen to satisfy $\mu > \|\mathcal{H}^*\mathcal{H}\|$. In this paper, the TV denoising at the second step of (35) has been performed using the fixed-point algorithm in [11]. The convergence of (35) can be further improved by using the same FISTA algorithm in [36]. The resulting procedure is summarized 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:*
 - $w^{(t)} = y^{(t)} + \mu \mathcal{H}^* \{v - \mathcal{H}\{y^{(t)}\}\}$
 - $u^{(t+1)} = \arg \min_u \left\{ \frac{1}{2} \|u - w^{(t)}\|_2^2 + \gamma \|u\|_{TV} \right\}$
 - $\tau^{(t+1)} = 0.5(1 + \sqrt{1 + 4(\tau^{(t)})^2})$
 - $y^{(t+1)} = u^{(t+1)} + (\tau^{(t)}/\tau^{(t+1)})(u^{(t+1)} - u^{(t)})$
-

In summary, Algorithms 1 and 2 represent the essence of the proposed algorithm for hybrid deconvolution of short-exposure optical images. The next section provides experimental results that further support the value and applicability of the proposed methodology.

VI. RESULTS

To demonstrate the viability of the proposed approach, its performance has been compared against reference methods. The first reference method used a dense sampling (DS) 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 DS approach. Second, to assess the importance of incorporation of the cross-derivative constraints, we have used both CCS and DCS for phase recovery. In what follows, comparative results for phase estimation and subsequent deconvolution are provided for all the given methods.

A. Phase Recovery

To assess the performance of the proposed and reference methods under controllable conditions, simulation data were used. The random nature of atmospheric turbulence necessitated the use of statistical methods to model its effect on a wavefront propagation. Specifically, in this paper, the effect of atmospheric turbulence was simulated based on the modified Von Karman model [39]. This model is derived based on Kolmogorov's theory of turbulence, which models atmospheric turbulence using temperature fluctuations [39]. 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

$$Q(t) = 0.033 C_n^2 \frac{e^{(-t^2/t_m^2)}}{(t^2 + t_0^2)^{11/6}} \quad (36)$$

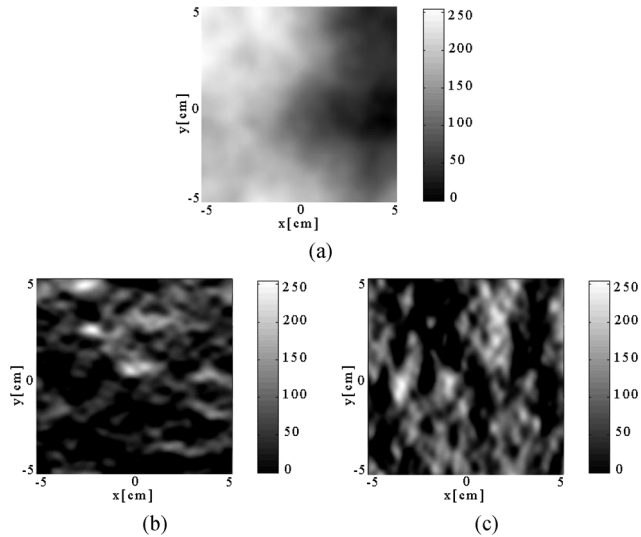


Fig. 3. Example of a simulated phase ϕ (a) along with its partial derivatives with respect to (b) x and (c) y .

where C_n^2 is the refractive index and t_m and t_0 are chosen to match the high-frequency and low-frequency behaviors of turbulence, respectively. The model of (36) can be used to generate random realizations of the GPF phase, as described, e.g., in [19].

A typical example of the GPF phase ϕ is shown in Fig. 3(a). In this case, the size of the phase screen was set to be equal to 10×10 cm, whereas the sampling was performed over a 128×128 uniform grid (which would have corresponded to the use of 16 384 lenslets of a SHI). The corresponding values of the (discretized) partial derivatives $\partial\phi/\partial x$ and $\partial\phi/\partial y$ are shown in Fig. 3(b) and (c), respectively.

In this paper, 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 , \bar{W} was defined to correspond to a four-level orthogonal wavelet transform using the nearly symmetric wavelets of Daubechies with five vanishing moments [40] and periodic boundary condition.

To demonstrate the value of using the cross-derivative constraint for phase reconstruction, the CCS and DCS algorithms have been compared in terms of the MSEs of their corresponding phase estimates. The results of this comparison are summarized in Fig. 4 for different compression ratios (or, equivalently, (sub)sampling densities) and $\text{SNR} = 40$ dB.

As expected, one can see that DCS results in lower values of MSE, as compared with CCS, which implies higher accuracy of phase reconstruction. Moreover, the difference in the performances of CCS and DCS appears to be more significant for lower sampling rates, whereas 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 with the case of CCS, 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 with the case of CCS.

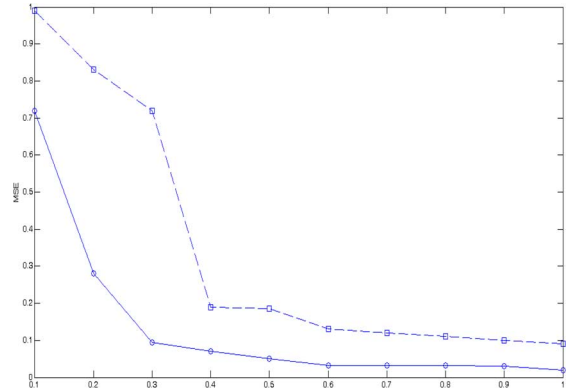


Fig. 4. MSE of phase reconstruction obtained with different methods as a function of r . Here, the dashed and solid lines correspond to CCS and DCS, respectively, and SNR is equal to 40 dB.

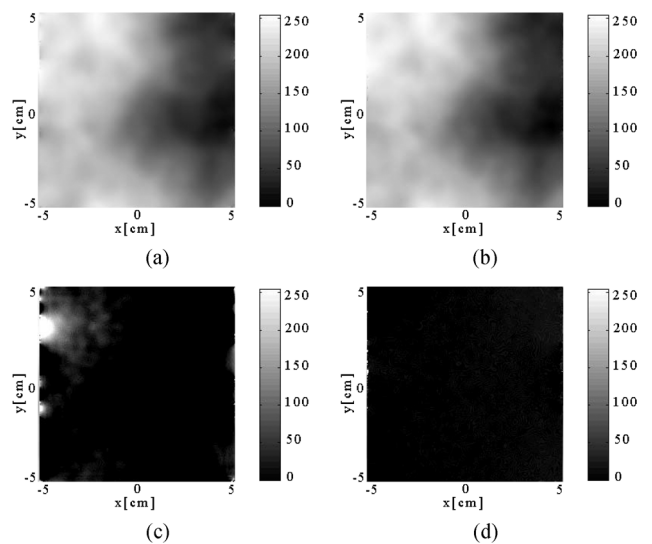


Fig. 5. (a) Phase reconstructed obtained by means of CCS for $\text{SNR} = 40$ dB and $r = 0.5$. (b) Phase reconstructed obtained by means of DCS for the same values of SNR and r . (c) and (d) Corresponding error maps for CCS and DCS.

A number of typical reconstruction results are shown in Fig. 5, whose left and right subplots depict the phase estimates obtained using the CCS 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. In addition, a close comparison with the original phase [as shown in Fig. 3(a)] 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 CCS.

As an additional comparison, The convergence of the MSE as a function of the number of iterations, for both CCS and DCS algorithms, is illustrated in Fig. 6. One can see that DCS results in a substantially faster convergence, as compared with CCS. This behavior 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, the convergence of the cross-derivative fidelity term $\|D_y f_x -$

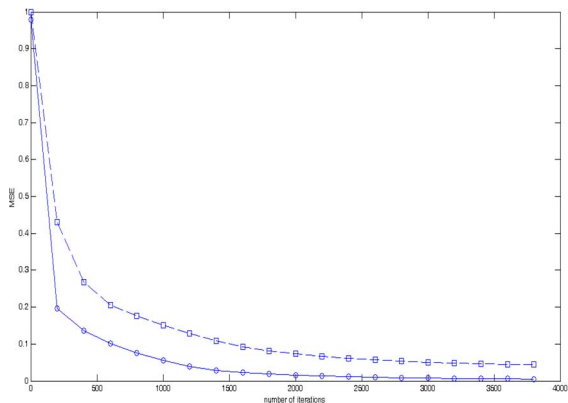


Fig. 6. Convergence analysis of phase reconstruction obtained with different methods as a function of iterations. Here, the dashed and solid lines correspond to CCS and DCS, respectively, SNR = 40, and $r = 0.5$.

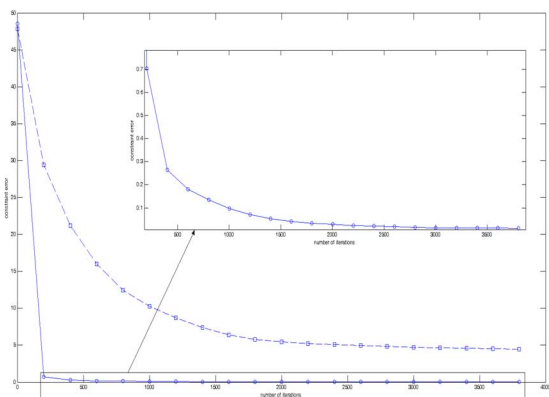


Fig. 7. Convergence analysis of derivative constraint obtained with different methods as a function of iterations. Here, the dashed and solid lines correspond to CCS and DCS, respectively, SNR = 40, and $r = 0.5$.

$\|D_x f_y\|^2$ for both methods under comparison is compared in Fig. 7. One can see that, in the case of DCS, this term converges considerably faster than in the case of CCS, which improves to the overall speed of convergence of DCS, making it superior to that of CCS.

To investigate the robustness of the compared algorithms toward measurement noise, their performances have been compared for a range of SNR values. The results of this comparison are summarized in Fig. 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 with the case of CCS. 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 (23)–(25) suggest 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 toward the inconsistency in noise modeling. This argument is supported by the results in Fig. 9, which summarizes the values of MSE obtained by CCS and DCS reconstructions for different levels of Poisson noise. One can see that, in this case, the MSE values

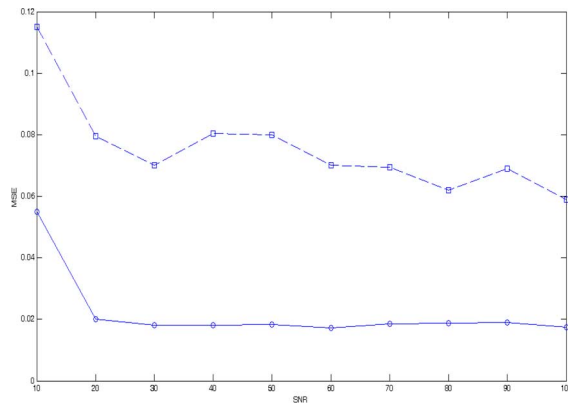


Fig. 8. MSE of phase reconstruction obtained with different methods as a function of SNR. Here, the dashed and solid lines correspond to CCS and DCS, respectively, and $r = 0.5$.

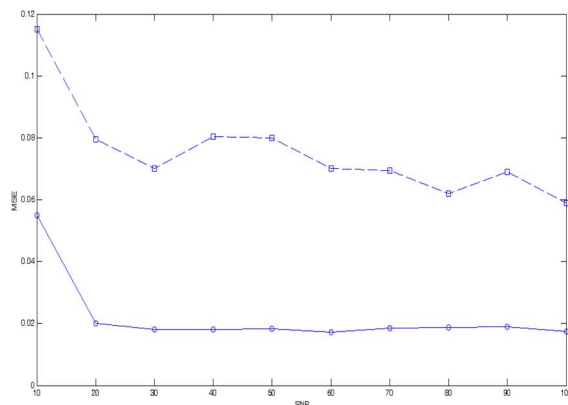


Fig. 9. MSE of phase reconstruction obtained with different methods as a function of SNR where the noise model is Poisson. Here, the dashed and solid lines correspond to CCS and DCS, respectively, and $r = 0.5$.

are comparable with 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 behavior. 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.

B. Image Deconvolution

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 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, the “Satellite” image along with its blurred and noisy version are shown in Fig. 10(a) and (b), respectively.

Using the PSF estimates, the deconvolution was carried out using the method detailed in [11]. For the sake of comparison, the deconvolution was also performed using the PSF recovered

TABLE I
SSIM AND PSNR COMPARISONS OF PHASE RECOVERY RESULTS

Image	Satellite				Saturn				Moon				Galaxy			
	10^{-5}	0.001	0.003	0.005	10^{-5}	0.001	0.003	0.005	10^{-5}	0.001	0.003	0.005	10^{-5}	0.001	0.003	0.005
Noise std	PSNR comparison (in dB)															
Blurred	14.06	14.06	14.06	14.05	17.78	17.78	17.78	17.78	19.98	19.97	19.97	19.97	18.79	18.79	18.78	18.78
DS	27.97	27.75	25.97	22.43	31.49	31.08	28.50	23.89	25.06	25.04	24.83	21.76	23.58	23.60	23.38	20.93
CS	17.06	16.93	16.54	15.63	23.42	23.38	22.80	20.55	22.36	22.38	22.30	19.73	21.16	21.12	20.64	18.46
DCS	27.42	27.22	25.56	22.22	31.02	30.65	28.30	23.72	25.00	24.99	24.78	21.73	23.52	23.54	23.32	20.86
	SSIM comparison															
Blurred	0.200	0.200	0.199	0.197	0.226	0.226	0.226	0.175	0.512	0.512	0.509	0.504	0.257	0.257	0.257	0.254
DS	0.730	0.720	0.554	0.269	0.688	0.660	0.506	0.228	0.645	0.642	0.607	0.552	0.493	0.495	0.501	0.397
CS	0.349	0.344	0.306	0.206	0.424	0.416	0.348	0.212	0.539	0.538	0.493	0.488	0.348	0.347	0.326	0.224
DCS	0.674	0.667	0.519	0.263	0.656	0.641	0.483	0.223	0.643	0.640	0.604	0.549	0.490	0.491	0.501	0.393

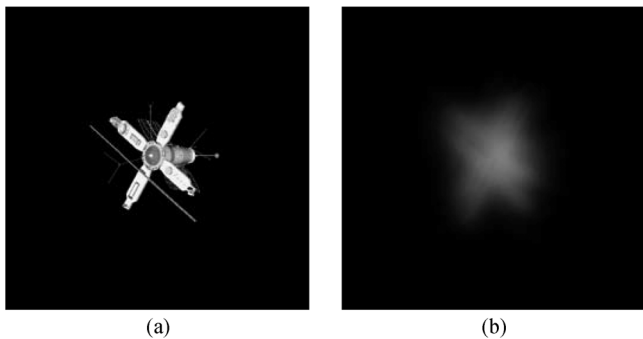


Fig. 10. (a) Satellite image and (b) its blurred and noisy version.

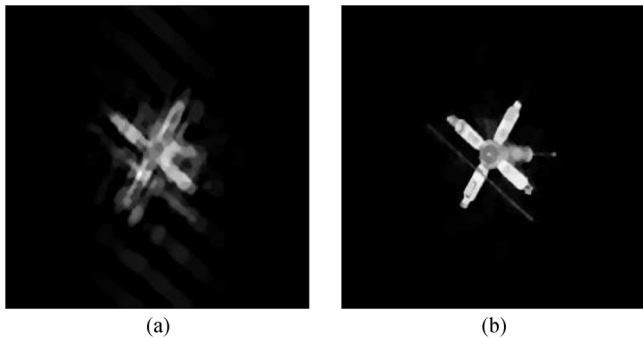


Fig. 11. (a) Image estimate obtained with the CCS-based method for phase recovery (SSIM = 0.781). (b) Image estimate obtained with the DCS-based method for phase recovery (SSIM = 0.917).

from 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 compared with their original counterparts in terms of PSNR as well as of the structural similarity index (SSIM) of [41], which is believed to be a better indicator of perceptual image quality [42]. The resulting values of the comparison metrics are summarized in Table I, whereas the deconvolution results produced by the CCS- and DCS-based methods are shown in Fig. 11.

The given 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 with the case of CCS. Moreover, comparing the results in Table I, one can see that DS only slightly outperforms DCS in terms of PSNR and

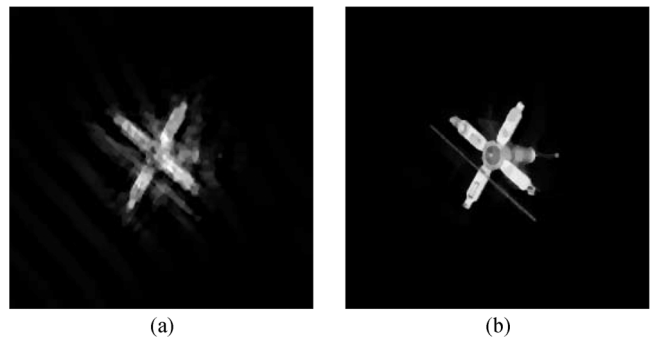


Fig. 12. (a) Image estimate obtained with the CCS-based method for phase recovery (SSIM = 0.732). (b) Image estimate obtained with the DCS-based method for phase recovery (SSIM = 0.888) where the noise model is assumed to be Poisson.

SSIM, whereas in many practical cases, the difference between the performances of these methods are hard to detect visually.

Finally, the results of CCS-based and DCS-based image reconstructions for the case of Poisson noise contamination are shown in Fig. 12. A close comparison of these results reveals a noticeable degradation in the performance of the CCS-based algorithm, whereas the DCS-based results are virtually indistinguishable from those obtained in the Gaussian case.

VII. DISCUSSION AND CONCLUSION

In this paper, the applicability of DCS to the problem of reconstruction of optical images has been demonstrated. It was shown that, in the presence of atmospheric turbulence, the phase ϕ of GPF $P = A e^{j\phi}$ is a random function, which needs to be measured using AO. To simplify the complexity of the latter, a CS-based approach has been proposed. As opposed to CCS, 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 with the case of CCS.

In this paper, 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 solution was computed using the Bregman algorithm, which provides a computationally efficient framework to carry out the constrained phase recovery. Moreover, 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 CCS.

While the proposed method offers a practical solution to the problem of phase estimation in AO, some interesting questions about the theoretical aspects of DCS still lay open. In particular, the question of theoretical performance of CS in the presence of side information on the source signal needs to be addressed through future research. For practical purpose one can also take benefit of this algorithm to modify the SHI. Instead of working with the measurements of the phase gradient, their linear combination can be used, e.g., Bernoulli weights. The resulting sensing basis might have smaller coherence with respect to the basis of wavelets, thereby offering the possibility of more accurate and stable reconstruction.

ACKNOWLEDGMENT

The authors would like to thank S. Dolui for his helpful comments as well as for providing deconvolution codes.

REFERENCES

- [1] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.
- [2] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.
- [3] I. J. Mairal, G. Sapiro, and M. Elad, "Learning multiscale sparse representations for image and video restoration," *Multiscale Model. Simul.*, vol. 7, pp. 214–241, 2008.
- [4] R. T. Paul, "Review of robust video watermarking techniques," *IJCA Spec. Issue Comput. Sci.*, no. 3, pp. 90–95, 2011.
- [5] G. D. Boreman, *Modulation Transfer Function in Optical and Electro-Optical Systems*. Bellingham, WA: SPIE, 2001.
- [6] J. A. Cadzow, "Blind deconvolution via cumulant extrema," *IEEE Signal Process. Mag.*, vol. 13, no. 3, pp. 24–42, May 1996.
- [7] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Process. Mag.*, vol. 13, no. 3, pp. 43–64, May 1996.
- [8] S. Geman and D. Geman, "Stochastic relaxation, gibbs distribution and the bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.
- [9] V. Torre, T. Poggio, and C. Koch, "Computational vision and regularization theory," *Nature*, vol. 317, pp. 314–319, Sep. 1985.
- [10] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenom.*, vol. 60, no. 1–4, pp. 259–268, Nov. 1992.
- [11] A. Chambolle, "An algorithm for total variation minimization and applications," *J. Math. Imag. Vis.*, vol. 20, no. 1, pp. 89–97, Jan. 2004.
- [12] T. Goldstein and S. Osher, "The split bregman method for l1-regularized problems," *SIAM J. Imag. Sci.*, vol. 2, no. 2, pp. 323–343, 2009.
- [13] A. Marquina, "Nonlinear inverse scale space methods for total variation blind deconvolution," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 64–83, 2009.
- [14] L. He, A. Marquina, and S. Osher, "Blind deconvolution using tv regularization and bregman iteration," *Int. J. Imag. Syst. Technol.*, vol. 15, no. 1, pp. 74–83, 2005.
- [15] O. Michailovich and A. Tannenbaum, "Blind deconvolution of medical ultrasound images: Parametric inverse filtering approach," *IEEE Trans. Image Process.*, vol. 16, no. 12, pp. 3005–3019, Dec. 2007.
- [16] W. H. Richardson, "Bayesian-based iterative method of image restoration," *J. Opt. Soc. Amer. A*, vol. 62, no. 1, pp. 55–59, 1972.
- [17] L. B. Lucy, "An iterative technique for the rectification of observed distributions," *Astron. J.*, vol. 79, no. 6, pp. 745–754, Jun. 1974.
- [18] P. A. Jansson, "Deconvolution of images and spectra," *Opt. Eng.*, vol. 36, p. 3224, 1997.
- [19] M. C. Roggemann and B. M. Welsh, *Imaging Through Turbulence*. Boca Raton, FL: CRC Press, 1996.
- [20] M. J. Cullum, *Adaptive Optics*. Garching, Germany: Eur. Southern Observatory, 1996.
- [21] D. Dayton, B. Pierson, B. Spielbusch, and J. Gonglewski, "Atmospheric structure function measurements with a Shack–Hartmann wave-front sensor," *J. Math. Imag. Vis.*, vol. 20, pp. 89–97, 2004.
- [22] R. G. Lane and M. Tallon, "Wave-front reconstruction using a Shack–Hartmann sensor," *Appl. Opt.*, vol. 31, pp. 6902–6908, 1992.
- [23] M. Hosseini and O. Michailovich, "Derivative compressive sampling with application to phase unwrapping," in *Proc. EUSIPCO*, Glasgow, U.K., Aug. 2009.
- [24] D. L. Fried, "Statistics of a geometric representation of wavefront distortion," *J. Opt. Soc. Amer.*, vol. 55, no. 11, pp. 1427–1431, 1965.
- [25] O. Michailovich and A. Tannenbaum, "A fast approximation of smooth functions from samples of partial derivatives with application to phase unwrapping," *Signal Process.*, vol. 88, pp. 358–374, Aug. 2008.
- [26] J. Primot, G. Rousset, and J. C. Fontanella, "Deconvolution from wave-front sensing: A new technique for compensating turbulence-degraded images," *J. Opt. Soc. Amer. A*, vol. 7, no. 9, pp. 1598–1608, 1990.
- [27] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, no. 1, pp. 10–21, Jan. 1949.
- [28] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [29] Y. Tsaig and D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [30] D. L. Donoho and Y. Tsaig, "Fast solution of l_1 -norm minimization problems when the solution may be sparse," 2006.
- [31] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, "Bregman iterative algorithms for l_1 -minimization with applications to compressed sensing," *SIAM J. Imag. Sci.*, vol. 1, no. 1, pp. 143–168, 2008.
- [32] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1995.
- [33] J. Cai, S. Osher, and Z. Shen, "Split bregman methods and frame based image restoration," *Multiscale Model. Simul.*, vol. 8, no. 2, pp. 337–369, 2009.
- [34] S. S. Chen and D. L. Donoho, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, 1998.
- [35] I. Daubechies, M. Defrise, and C. D. Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.*, vol. 57, pp. 1413–1457, 2004.
- [36] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, 2009.
- [37] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Anal. Chem.*, vol. 36, pp. 1627–1639, 1964.
- [38] A. N. Tikhonov and V. Y. Arsenin, "Solutions of ill-posed problems," 1977.
- [39] J. D. Schmidt, *Numerical Simulation of Optical Wave Propagation With Examples in MATLAB*. Washington, DC: SPIE, 2010.
- [40] I. Daubechies, "Ten lectures on wavelets," in *Proc. CBMS-NSF Reg. Conf. Ser. Appl. Math.*, 1992.
- [41] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [42] Z. Wang and A. C. Bovik, "Mean squared error: love it or leave it?—A new look at signal fidelity measures," *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98–117, Jan. 2009.



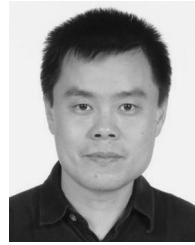
Mohammad Rostami (S'11) received the B.S. degree in electrical engineering and pure mathematics from Sharif University of Technology, Tehran, Iran, in 2010. He is currently working toward the M.A.Sc. degree from the University of Waterloo, Waterloo, ON, Canada.

His research interests include image and video processing/quality assessment, compressed sensing, and machine learning.



Oleg Michailovich (S'02–M'03) was born in Saratov, Russia, in 1972. He received the M.Sc. degree in electrical engineering from Saratov State University, Saratov, Russia, in 1994, and the M.Sc. and Ph.D. degrees in biomedical engineering from the Technion Israel Institute of Technology, Haifa, Israel, in 2003.

He is currently with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. His research interests include the application of image processing to various problems of image reconstruction, segmentation, inverse problems, nonparametric estimations, approximation theory, and multiresolution analysis.



Zhou Wang (S'97–A'01–M'02) received the Ph.D. degree in electrical and computer engineering from The University of Texas at Austin, Austin, in 2001.

He is currently an Associate Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. His research interests include image processing, coding, and quality assessment; computational vision and pattern analysis; multimedia communications; and biomedical signal processing. He has more than 100 publications in these fields with over 9000 citations.

He is the author or coauthor of *Modern Image Quality Assessment* (Morgan & Claypool, 2006).

Dr. Wang has served as an Associate Editor of IEEE TRANSACTIONS ON IMAGE PROCESSING (2009–present), IEEE SIGNAL PROCESSING LETTERS (2006–2010), and *Pattern Recognition* (2006–present) and a Guest Editor of IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING (2007–2009), *EURASIP Journal of Image and Video Processing* (2009–2010), and *Signal, Image and Video Processing* (2011–Present). He was a recipient of the 2009 IEEE Signal Processing Best Paper Award, ICIP 2008 IBM Best Student Paper Award (as senior author), and 2009 Ontario Early Researcher Award.