

# A Novel SURE-Based Criterion for Parametric PSF Estimation

Feng Xue and Thierry Blu, *Fellow, IEEE*

**Abstract**—We propose an unbiased estimate of a filtered version of the mean squared error—the blur-SURE (Stein’s unbiased risk estimate)—as a novel criterion for estimating an unknown point spread function (PSF) from the degraded image only. The PSF is obtained by minimizing this new objective functional over a family of Wiener processings. Based on this estimated blur kernel, we then perform nonblind deconvolution using our recently developed algorithm. The SURE-based framework is exemplified with a number of parametric PSF, involving a scaling factor that controls the blur size. A typical example of such parametrization is the Gaussian kernel. The experimental results demonstrate that minimizing the blur-SURE yields highly accurate estimates of the PSF parameters, which also result in a restoration quality that is very similar to the one obtained with the exact PSF, when plugged into our recent multi-Wiener SURE-LET deconvolution algorithm. The highly competitive results obtained outline the great potential of developing more powerful blind deconvolution algorithms based on SURE-like estimates.

**Index Terms**—Parametric PSF estimation, SURE, blur-SURE, Wiener filtering.

## I. INTRODUCTION

AS A STANDARD linear inverse problem, image deconvolution has been an important image processing topic for several decades [1]–[3]. In many real applications, e.g., medical imaging [4], fluorescence microscopy [5], astronomical imaging [6], remote sensing [7] and photography [8], the point spread function (PSF) may not be easily and accurately obtained. Blind deconvolution amounts to estimating both the original image and the PSF, given the observed data only [9].

In order to address the ill-posedness of the problem, a standard approach consists in enforcing certain regularity conditions (possibly derived from Bayesian priors) on the original image and the PSF, and in formulating the problem as the optimization of a proper objective functional [10], [11]. Recent studies of the standard techniques can be found in [8] and [12]–[15]. The interested reader may refer

to [9] and [16]–[18] for a comprehensive review. Other optimization methods are also able to estimate simultaneously both the original image and the PSF (see [12], [19] for example).

It is, however, also possible to perform the estimation of the image and the PSF separately; i.e., first estimate the PSF, and then perform non-blind deconvolution to obtain the restored image. Compared to the joint estimation, the separate estimation seems more advantageous in practice, because it allows to apply any developed high-quality non-blind deconvolution algorithm, once the PSF is estimated [8], [20]. In this work, it is this approach that we have chosen and since we have already developed a very efficient non-blind deconvolution algorithm (see [21]), we are going to focus here on PSF estimation. Note that the accuracy of PSF estimation is very significant for the deconvolution performance: an inaccurately estimated kernel often results in substantial quality loss of the restored image.

### A. Non-Parametric PSF Estimation

If the PSF is represented by the discrete pixel values without any parametric form, the estimation is usually performed by regularization or Bayesian approach, depending on the assumption on the unknown PSF. The common choices of the regularization on the blur kernel include:

- Tikhonov  $\ell^2$ -Norm [11], [12]: Effective estimation of smooth PSF like Gaussian blur;
- TV (Total Variation)-Norm [10], [14]: Suitable for motion blur and out-of-focus blur;
- Sparsity Prior [8], [22]: Effective for camera shake and motion blur.

### B. Parametric PSF Estimation

In specific applications, the parametric forms of the PSF can be either theoretically available or practically assumed [23], [24]. In this context, the PSF is completely characterized by a small number of parameters, which dramatically reduces the degrees of freedom of PSF estimation [19], [25]. Typical examples of the parametric model can be found in the applications of linear motion blur [26]–[28], fluorescence microscopy [5], [19], [29], atmospheric turbulence [30], [31] and astronomical imaging [6].

For linear motion blur, the blur orientation and length (i.e. the PSF parameters) can be estimated by cepstrum analysis [32], steerable filter method [26], or Radon transform [28]. In the case of a Gaussian kernel, the Gaussian variance representing blur size is the only parameter to be estimated. APEX [24] estimated the blur size by fitting the Gaussian blur to the blurred image, utilizing the property of

Manuscript received May 10, 2014; revised October 4, 2014; accepted November 26, 2014. Date of publication December 12, 2014; date of current version January 8, 2015. This work was supported in part by the Research Grant Council, Hong Kong, under Grant CUHK410012 and Grant AoE/M-05/12 and in part by the National Natural Science Foundation of China under Grant 61401013. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Brendt Wohlberg.

F. Xue is with the National Key Laboratory of Science and Technology on Test Physics and Numerical Mathematics, Beijing 100076, China (e-mail: fxue2012@gmail.com).

T. Blu is with the Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong (e-mail: thierry.blu@m4x.org).

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.2014.2380174

the fast Fourier decrease of the Gaussian kernel. There are also several local estimation methods based on edge-based parametric representation [33], see [34]–[36] estimated the Gaussian blur size by applying a steerable second derivative of Gaussian operator and measuring the response of the edge. However, the local-based methods are very sensitive to the noise corruption: the edge localization and measurement become less accurate under higher noise level. Note that all the methods mentioned above were specifically designed for certain blur types, which precludes a wider range of degradation scenarios.

The regularization or Bayesian approach provides a general framework for parametric PSF estimation, by substituting the parametric form of PSF into the objective functional and performing the optimization. In microscopy deconvolution, [19] derived a maximum-likelihood-based method to estimate the PSF parameters, based on a diffraction model proposed in [37]. [29] performed the parametric estimation using Richardson-Lucy algorithm with TV regularization, assuming microscopy PSF as Gaussian function. In addition, there are some empirical approaches to the problem. [31] estimated the PSF parameters by kurtosis minimization of the restored image. In the work of [38], the estimated PSF parameter is selected to be at the maximum point of the differential coefficients of restored image Laplacian  $\ell^1$ -norm curve. However, the empirical methods need to manually adjust regularization parameter for restoration, and have no guarantee on the accuracy of PSF estimation. Note the work of [39], where the authors proposed GCV (generalized cross validation) as a criterion of PSF estimation and theoretically proved its validity. Unfortunately, the introduced auto-regressive and moving average models complicated the GCV minimization and brought about high computational cost. Moreover, the GCV criterion has many local minima in general. Hence, the minimization algorithm should be carefully performed to guarantee the global minimum.

### C. SURE-Based Approach

A statistical non-Bayesian approach based on the minimization of Stein’s unbiased risk estimate (SURE) [40] has been proposed for solving linear inverse problem under additive Gaussian noise assumption. SURE, as an unbiased estimate of the mean squared error (MSE), has been intensively used as a criterion for signal restoration, e.g. image denoising [41]–[43] and non-blind image deconvolution [21], [44]–[47]. The key advantage of the SURE-based approach is that it does not require any prior knowledge of the original image [42].

In the present paper, we further extend the SURE-based approach to blind PSF parameter estimation. More specifically, we propose a filtered version of the SURE—“blur-SURE”—as a new criterion for PSF estimation. To exemplify this framework, we consider the parametric forms of PSF, which involve a scaling factor that controls the blur size (refer to [14] and [23] for examples), and estimate the factor from the degraded image only. A typical example is the Gaussian function. Once the PSF is estimated, we carry out our developed non-blind deconvolution algorithm [21].

### D. Paper Organization

Section II proposes a novel criterion for PSF estimation: the blur-MSE—a filtered version of the MSE—incorporating a simple Wiener filtering. In Section III, we present the blur-SURE—an unbiased estimate of the blur-MSE defined in Section II, and formulate the PSF estimation as a blur-SURE minimization. Section IV exemplifies the blur-SURE framework with several types of PSF, and proposes an efficient minimization algorithm. Section V reports and discusses the experimental results.

Throughout this paper, we use boldface lowercase letters, e.g.  $\mathbf{x} \in \mathbb{R}^N$ , to denote  $N$ -dimensional real vectors, where  $N$  is typically the number of pixels in an image. The  $n$ -th element of  $\mathbf{x}$  is written as  $x_n$ . The linear (matrices) and non-linear transformations  $\mathbb{R}^N \rightarrow \mathbb{R}^M$  are denoted by boldface uppercase letters, e.g.  $\mathbf{H} \in \mathbb{R}^{M \times N}$ .  $\mathbf{H}^T \in \mathbb{R}^{N \times M}$  denotes the transpose of matrix  $\mathbf{H}$ . Also note that we use the subscript  $(\cdot)_0$  to denote the true (“ground truth”) quantity of  $(\cdot)$ ; for example, matrix  $\mathbf{H}_0$  is the true quantity of  $\mathbf{H}$ .

## II. THE BLUR-MSE AND ITS OPTIMIZATION

### A. Problem Statement

Consider the linear model

$$\mathbf{y} = \mathbf{H}_0 \mathbf{x} + \mathbf{b}, \quad (1)$$

where  $\mathbf{y} \in \mathbb{R}^N$  is the observed data of the original (unknown)  $\mathbf{x} \in \mathbb{R}^N$ , the ground truth (unknown) matrix  $\mathbf{H}_0$  denotes a linear distortion, the vector  $\mathbf{b} \in \mathbb{R}^N$  is a zero-mean additive Gaussian noise corruption with covariance matrix  $\mathbf{C} > 0$ . Our purpose is to estimate the matrix  $\mathbf{H}_0$ , such that the estimated  $\mathbf{H}$  is as close to the true  $\mathbf{H}_0$  as possible, from the observed data  $\mathbf{y}$  only.

The original signal  $\mathbf{x}$  will be considered *deterministic* in most of the paper, with the sole exception of Section II-C (and the related proof in Appendix), where it will be nevertheless independent of the noise  $\mathbf{b}$ . This will be outlined in the notation of the mathematical expectation:  $\mathcal{E}_{\mathbf{b}}\{\cdot\}$ ,  $\mathcal{E}_{\mathbf{x}}\{\cdot\}$  or  $\mathcal{E}_{\mathbf{x}, \mathbf{b}}\{\cdot\}$  indicate expectation over noise realizations, signal realizations, or both.

### B. Blur-MSE: An Oracle Criterion for the Estimation of $\mathbf{H}$

Denoting a function (or processing) by  $\mathbf{f} : \mathbb{R}^N \rightarrow \mathbb{R}^N$ , applied to the observed data  $\mathbf{y}$ , the MSE that results from this processing (assumed to provide an estimate of  $\mathbf{x}$ ) is defined as [41], [42], [44], [46]:

$$\text{MSE} = \frac{1}{N} \mathcal{E}_{\mathbf{b}} \left\{ \|\mathbf{f}(\mathbf{y}) - \mathbf{x}\|^2 \right\}. \quad (2)$$

Instead of the standard MSE, we consider the following “blurred” (filtered) version:

$$\text{blur-MSE} = \frac{1}{N} \mathcal{E}_{\mathbf{b}} \left\{ \|\mathbf{H}\mathbf{f}(\mathbf{y}) - \mathbf{H}_0 \mathbf{x}\|^2 \right\} \quad (3)$$

as an *oracle* criterion for estimating  $\mathbf{H}$ . We call this objective functional *blur-MSE*, since it measures the difference between two distorted (blurred) data:  $\mathbf{H}\mathbf{f}(\mathbf{y})$  and  $\mathbf{H}_0 \mathbf{x}$ . Assuming  $\mathbf{H}_0 \mathbf{x}$

is known, we formulate the estimation of  $\mathbf{H}$  as the following minimization problem:

$$\min_{\mathbf{H}} \frac{1}{N} \|\mathbf{H}\mathbf{f}_{\mathbf{H}}(\mathbf{y}) - \mathbf{H}_0\mathbf{x}\|^2$$

where the subscript  $\mathbf{H}$  in  $\mathbf{f}_{\mathbf{H}}$  emphasizes the dependence of the processing  $\mathbf{f}$  on the matrix  $\mathbf{H}$ . Particularly, in the context of deconvolution, the processing  $\mathbf{f}$  could be any non-blind deconvolution algorithms (e.g. multi-Wiener SURE-LET [21] or BM3D [48] algorithm), depending on the matrix  $\mathbf{H}$ . Experimentally, we found that given the processing  $\mathbf{f}_{\mathbf{H}}$ , the blur-MSE minimization yields a highly accurate estimate of  $\mathbf{H}_0$ .

However, since the ‘arbitrary’ processing  $\mathbf{f}_{\mathbf{H}}$  can be very complicated, our approach, to be described below, is to restrict  $\mathbf{f}$  to a subclass of processings—typically, Wiener filterings—that also depend on  $\mathbf{H}$ , and then minimize this blur-MSE (actually, an estimate of this blur-MSE) over  $\mathbf{H}$ . In this simplified setting, we will be able to demonstrate the link between the result of this minimization and the ground truth PSF.

### C. The Blur-MSE Minimization With Exact Wiener Filtering

In this Section only,  $\mathbf{x}$  will be assumed to be random, with covariance matrix  $\mathbf{A} = \mathcal{E}_{\mathbf{x}}\{\mathbf{x}\mathbf{x}^T\}$ . Let us consider the linear processing denoted by  $\mathbf{f}(\mathbf{y}) = \mathbf{W}_{\mathbf{H}}\mathbf{y}$ , where the notation  $\mathbf{W}_{\mathbf{H}}$  emphasizes that the matrix  $\mathbf{W}$  is related to  $\mathbf{H}$ . It is well-known that for the linear model (1) with the known matrix  $\mathbf{H}$ , the ideal linear processing  $\mathbf{W}_{\mathbf{H}}$  that minimizes the MSE averaged over realizations of the signal  $\mathcal{E}_{\mathbf{x}}\left\{\frac{1}{N}\mathcal{E}_{\mathbf{b}}\{\|\mathbf{W}_{\mathbf{H}}\mathbf{y} - \mathbf{x}\|^2\}\right\}$ , is Wiener filtering, expressed as [2]:

$$\mathbf{W}_{\mathbf{H}} = \mathbf{A}\mathbf{H}^T(\mathbf{H}\mathbf{A}\mathbf{H}^T + \mathbf{C})^{-1} \quad (4)$$

in matrix notation, where the covariance matrix of the noise is  $\mathbf{C} = \mathcal{E}_{\mathbf{b}}\{\mathbf{b}\mathbf{b}^T\}$ . Now, if we base our processing on Wiener filtering (4), then the following theorem shows that the solution  $\mathbf{H}$  that minimizes the blur-MSE (3) is related in a simple way to the true matrix  $\mathbf{H}_0$ .

*Theorem 2.1:* Consider only linear processings of the form  $\mathbf{f}(\mathbf{y}) = \mathbf{W}_{\mathbf{H}}\mathbf{y}$ , where  $\mathbf{W}_{\mathbf{H}}$  is defined as (4). Minimizing with respect to  $\mathbf{H}$  the averaged blur-MSE:

$$\min_{\mathbf{H}} \mathcal{E}_{\mathbf{x}} \left\{ \frac{1}{N} \mathcal{E}_{\mathbf{b}} \left\{ \|\mathbf{H}\mathbf{W}_{\mathbf{H}}\mathbf{y} - \mathbf{H}_0\mathbf{x}\|^2 \right\} \right\}, \quad (5)$$

yields  $\mathbf{H}\mathbf{A}\mathbf{H}^T = \mathbf{H}_0\mathbf{A}\mathbf{H}_0^T$ .

See Appendix A for a proof. Note that *Theorem 2.1* is valid for any linear distortion  $\mathbf{H}$ , not limited to convolution operation. In the case of convolution, we obtain the following corollary, as a Fourier description of *Theorem 2.1*.

*Corollary 2.1:* When  $\mathbf{H}$  is a convolution matrix, and when both the signal  $\mathbf{x}$  and the noise  $\mathbf{b}$  are stationary, the Wiener filtering  $\mathbf{W}_{\mathbf{H}}$  defined as (4) can be rewritten as:

$$W_{\mathbf{H}}(\omega) = \frac{H^*(\omega)}{|H(\omega)|^2 + C(\omega)/A(\omega)}$$

using the Fourier variable  $\omega = (\omega_1, \omega_2)$ . Here,  $H(\omega)$  is the Fourier representation of  $\mathbf{H}$ ,  $A(\omega)$  and  $C(\omega)$  are the

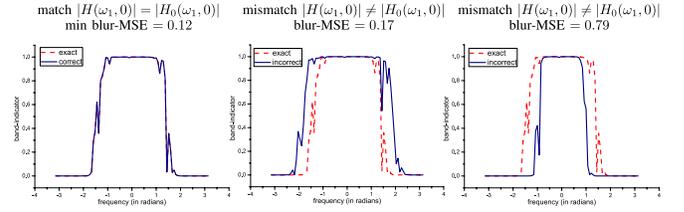


Fig. 1. Behaviour of the line  $\omega = (\omega_1, 0)$  of the band-indicator  $U(\omega)$  in (6): correct  $|H(\omega_1, 0)| = |H_0(\omega_1, 0)|$  yields the minimum blur-MSE in (5) (example: *Cameraman* blurred by Gaussian kernel).

power spectral densities of signal  $\mathbf{x}$  and noise  $\mathbf{b}$ , respectively. Thus, the blur-MSE minimization (5) yields that, for all  $\omega \in [-\pi, \pi]^2$

$$\begin{aligned} & \underbrace{H(\omega) \cdot \frac{H^*(\omega)}{|H(\omega)|^2 + C(\omega)/A(\omega)}}_{U(\omega)} \\ &= H_0(\omega) \cdot \underbrace{\frac{H_0^*(\omega)}{|H_0(\omega)|^2 + C(\omega)/A(\omega)}}_{U_0(\omega)} \end{aligned} \quad (6)$$

in the frequency domain, which is equivalent to  $|H(\omega)| = |H_0(\omega)|$  for all  $\omega \in [-\pi, \pi]^2$ .

This corollary is easily proved by the fact that under the assumptions of convolution operation and stationary process, all the matrices involved in the blur-MSE can be diagonalized by the discrete 2D Fourier transformation.

Let us call  $U(\omega) = H(\omega)W_{\mathbf{H}}(\omega)$  in (6) a *frequency-band indicator* or *band indicator* for short,<sup>1</sup> as it marks a certain frequency band as 0 or 1 with a narrow transition between the two values (see Fig. 1 for example). *Corollary 2.1* states that: (1) the blur-MSE minimization is essentially equivalent to matching the frequency band indicator  $U(\omega)$  to the ground truth  $U_0(\omega)$ ; (2) the blur-MSE minimization results in the magnitudes of the frequency responses alone to be matched:  $|H(\omega)| = |H_0(\omega)|$ . Obviously, the blur-MSE criterion cannot be used to identify frequency responses whose phase variations are not linked to amplitude variations (the same is true for GCV criterion [39]). Hence, we consider only zero-phase blur models in this paper. Since many real-life blurs—linear motion, out-of-focus, and atmospheric turbulence blurs—have zero phase, this assumption is rather unrestrictive [39].

To exemplify this corollary, Fig. 1 shows the band-indicator matching as in (6), when the image *Cameraman* (shown in Fig. 8) is blurred by Gaussian kernel. For the sake of clarity, we have chosen to show only one line  $\omega = (\omega_1, 0)$  of the band-indicator. We can see that the minimum blur-MSE is reached when the two band-indicators match perfectly; i.e.,  $U(\omega) = U_0(\omega)$ , and the derived equality  $|H(\omega)| = |H_0(\omega)|$  holds. If  $|H(\omega)| \neq |H_0(\omega)|$ , the corresponding band-indicator  $U(\omega)$  does not match the exact  $U_0(\omega)$  well, and the blur-MSE is not minimized.

<sup>1</sup>Regarding the terminology of  $U(\omega)$ , refer to [4], where the authors named  $U(\omega)$  as *combined response*.

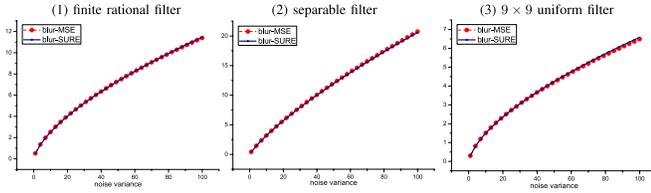


Fig. 2. Close match between the blur-SURE and the blur-MSE: *Cameraman* blurred by three filters, under various noise variances  $\sigma^2$ .

### III. BLUR-SURE: A NOVEL BLIND CRITERION FOR PSF ESTIMATION

#### A. Blur-SURE: An Unbiased Estimate of the Blur-MSE

Unfortunately, it is not possible to minimize directly the blur-MSE in practice, since  $\mathbf{H}_0\mathbf{x}$  is unknown. However, based on the linear model (1), the blur-MSE can be replaced by an unbiased statistical estimate—the “blur-SURE”—involving only the measurements  $\mathbf{y}$ , as summarized in the following theorem.

*Theorem 3.1:* Let  $\mathbf{f}(\mathbf{y}) = [f_n(\mathbf{y})]_{1 \leq n \leq N}$  be a (weakly) differentiable  $N$ -dimensional vector function which does not increase too fast,<sup>2</sup> and such that  $\mathcal{E}_{\mathbf{b}}\{|\partial_{y_k} f_n(\mathbf{y})|\}$  is finite for  $k, n = 1, 2, \dots, N$ .

Then, given the linear model (1), the following random variable:

$$\epsilon = \frac{1}{N} \left\{ \|\mathbf{H}\mathbf{f}(\mathbf{y}) - \mathbf{y}\|^2 + 2\text{div}_{\mathbf{y}}(\mathbf{C}\mathbf{H}\mathbf{f}(\mathbf{y})) - \text{Tr}(\mathbf{C}) \right\} \quad (7)$$

is an unbiased estimator of the blur-MSE (3), i.e.,  $\mathcal{E}_{\mathbf{b}}\{\epsilon\} = \frac{1}{N} \mathcal{E}_{\mathbf{b}}\{\|\mathbf{H}\mathbf{f}(\mathbf{y}) - \mathbf{H}_0\mathbf{x}\|^2\}$ , where  $\text{Tr}$  denotes matrix trace, and the divergence  $\text{div}_{\mathbf{y}}\mathbf{v} = \sum_{n=1}^N \partial v_n / \partial y_n$  for  $\forall \mathbf{v} \in \mathbb{R}^N$ .

See Appendix B for a sketch of the proof (for a rigorous proof of Stein’s Lemma with weak hypotheses, see [40], [43]). Note the difference with other flavours of SURE: instead of estimating (and minimizing) a distance between  $\mathbf{f}(\mathbf{y})$  and  $\mathbf{x}$ , we estimate a distance between  $\mathbf{H}\mathbf{f}(\mathbf{y})$  and  $\mathbf{H}_0\mathbf{x}$ —with different PSF. It is this key difference that allows to retrieve a reliable estimate of  $\mathbf{H}_0$ .

This theorem is valid for any linear distortion  $\mathbf{H}$  and any processing  $\mathbf{f}$ , under the assumption of additive Gaussian noise  $\mathbf{b}$ . Since the blur-SURE depends only on the observed data  $\mathbf{y}$ , it can be a practical substitute to the blur-MSE. In particular, if the processing  $\mathbf{f}$  is the Wiener filtering (4), and if the Gaussian noise  $\mathbf{b}$  is independent and identically distributed (i.i.d.) with variance  $\sigma^2$ , the blur-SURE becomes:

$$\epsilon = \frac{1}{N} \|\mathbf{H}\mathbf{W}\mathbf{H}\mathbf{y} - \mathbf{y}\|^2 + \frac{2\sigma^2}{N} \text{Tr}(\mathbf{H}\mathbf{W}\mathbf{H}) - \sigma^2 \quad (8)$$

which is the unbiased estimate of the blur-MSE given in (5). We obtain this equation from  $\text{div}_{\mathbf{y}}(\mathbf{H}\mathbf{W}\mathbf{H}\mathbf{y}) = \text{Tr}(\mathbf{H}\mathbf{W}\mathbf{H})$  and by specifying the covariance matrix  $\mathbf{C} = \sigma^2\mathbf{I}$  in (7). The blur-SURE is all the closer to the blur-MSE as the number of image pixels increases (law of large numbers). Fig. 2 shows the curves of the blur-MSE and the blur-SURE under different noise variance  $\sigma^2$ , taking three PSF as examples

<sup>2</sup>Typically,  $\|\mathbf{f}(\mathbf{y})\|^2 \leq \text{const} \times \exp(a\|\mathbf{y}\|^2)$  for some  $a < 1/(2\sigma^2)$ .

(the image size of *Cameraman* is  $256 \times 256$  pixels): (1) *finite rational filter*  $h_0(i, j) = (1 + i^2 + j^2)^{-1}$  for  $i, j = -7, \dots, 0, \dots, 7$ ; (2) *separable filter* with filter  $[1, 4, 6, 4, 1]/16$  along both horizontal and vertical directions; (3)  $9 \times 9$  *uniform filter*.

As can be seen in (8), the blur-SURE requires the estimation of the noise variance  $\sigma^2$ . In all that follows, we will always use the MAD (median absolute deviation) of the first high-high wavelet (Daubechies 6) subband of the blurred image [49] for that purpose, even when this variance is known.

#### B. Approximation of the Exact Wiener Filtering

In practice, the exact Wiener filtering  $\mathbf{W}_{\mathbf{H}}$  in the blur-SURE cannot be used, since the power spectral density  $A(\omega)$  is unknown. However a basic observation is that for natural images with strong low frequencies and weak high frequencies,  $C(\omega)/A(\omega)$  increases (roughly) quadratically with the frequency. For this reason, we choose to replace  $C(\omega)/A(\omega)$  by  $\lambda\|\omega\|^2$ , where  $\lambda$  is a parameter to be estimated. Here,  $\omega = (\omega_1, \omega_2)$  are the (zero-centered) 2D DFT frequency variables and  $\|\omega\|^2 = \omega_1^2 + \omega_2^2$ ; for instance, assuming that the size of the image is  $(K, L)$ , with  $K$  and  $L$  even, then  $\omega_1 = 2\pi k/K$ , for  $k = -K/2, -K/2 + 1, \dots, K/2 - 1$  and  $\omega_2 = 2\pi l/L$ , for  $l = -L/2, -L/2 + 1, \dots, L/2 - 1$ . Then, we obtain the approximated Wiener filtering  $W_{H,\lambda}(\omega)$  as:

$$W_{H,\lambda}(\omega) = \frac{H^*(\omega)}{|H(\omega)|^2 + \lambda\|\omega\|^2} \quad (9)$$

Due to the parameter  $\lambda$  introduced, we formulate the PSF estimation as the result of the minimization of the blur-SURE over both  $\mathbf{H}$  and  $\lambda$ , i.e.,

$$\min_{\mathbf{H}, \lambda} \underbrace{\frac{1}{N} \|\mathbf{H}\mathbf{W}_{H,\lambda}\mathbf{y} - \mathbf{y}\|^2 + \frac{2\sigma^2}{N} \text{Tr}(\mathbf{H}\mathbf{W}_{H,\lambda}) - \sigma^2}_{\text{blur-SURE: } \epsilon(\mathbf{H}, \lambda)} \quad (10)$$

Although we are not exactly in the conditions of *Corollary 2.1*, we still expect the solution  $H(\omega)$  of (10) to satisfy  $|H(\omega)| \approx |H_0(\omega)|$ , and also, with the obtained  $\lambda$ , to constitute the best approximation of the exact band indicator  $U_0(\omega)$ :

$$\begin{aligned} U_{H,\lambda}(\omega) &= \underbrace{\frac{|H(\omega)|^2}{|H(\omega)|^2 + \lambda\|\omega\|^2}}_{\text{approximated band-indicator}} \\ &\approx \underbrace{\frac{|H_0(\omega)|^2}{|H_0(\omega)|^2 + C(\omega)/A(\omega)}}_{\text{exact band-indicator}} = U_0(\omega) \end{aligned} \quad (11)$$

Fig. 3 shows the different approximations of the band-indicator, when *Cameraman* is blurred by a Gaussian kernel. We can see that it is only when both  $\mathbf{H}$  and  $\lambda$  are obtained by minimizing the blur-SURE that we obtain the best band indicator matching.

Finally, we summarize the blur-SURE framework in Fig. 4. Also note that the blur-MSE minimization serves as an *oracle* counterpart of the blur-SURE minimization (10).

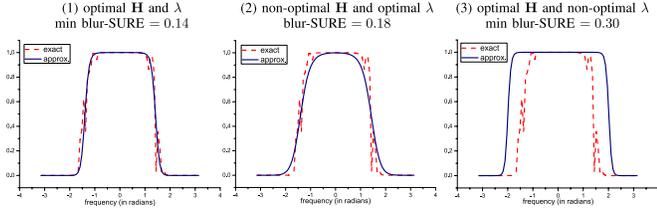


Fig. 3. Approximation of the band-indicator  $U_{H,\lambda}(\omega)$  (11): the blur-SURE minimization yields the best match between band-indicators (example: *Cameraman* blurred by Gaussian kernel). Only the line  $\omega = (\omega_1, 0)$  is shown here.

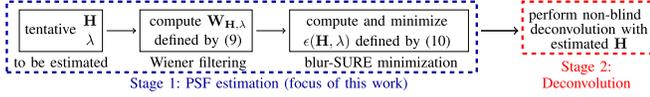


Fig. 4. PSF estimation flowchart: joint minimization of the blur-SURE over the parameters of  $\mathbf{H}$  and  $\lambda$ , as shown in (10).

#### IV. EXAMPLES OF PARAMETRIC PSF ESTIMATION

This section will exemplify the blur-SURE minimization framework with several typical parametric forms of PSF. Throughout all the tests in this section, the test image we used is *Cameraman* (see Fig. 8), and the variance of the Gaussian noise is  $\sigma^2 = 1.0$ .

##### A. Blur-SURE Parametric PSF Estimation

We assume that the PSF  $\mathbf{H} = \mathbf{H}_s$  can be completely represented by a small number of unknown parameters  $\mathbf{s} = [s_1, s_2, \dots, s_P]^T$  [19], [25]. We denote the ground truth parameter by  $\mathbf{s}_0$ . Therefore, following (10), we formulate the parameter estimation as:

$$\min_{\mathbf{s}, \lambda} \underbrace{\frac{1}{N} \|\mathbf{H}_s \mathbf{W}_{s, \lambda} \mathbf{y} - \mathbf{y}\|^2 + \frac{2\sigma^2}{N} \text{Tr}(\mathbf{H}_s \mathbf{W}_{s, \lambda}) - \sigma^2}_{\text{blur-SURE: } \epsilon(\mathbf{s}, \lambda)} \quad (12)$$

##### B. Examples of Parametric PSF Estimation

We consider the following typical PSF.

1) *Gaussian Kernel*: Characterized by:

$$\mathbf{h}_s(i, j; s) = M \cdot \exp\left(-\frac{i^2 + j^2}{2s^2}\right) \quad (13)$$

with variance  $s^2$ , where  $(i, j)$  denotes the 2D coordinates,  $M$  is a normalization coefficient such that  $\sum_{i,j} \mathbf{h}_s(i, j) = 1$ .  $s$  is the unknown parameter to be estimated.

To experimentally justify *Corollary 2.1*, we use the exact Wiener filtering (4) to perform the blur-MSE and blur-SURE minimizations. Fig. 5-(1) demonstrates the perfect estimation of  $s_0$ . If we apply the approximated Wiener filtering (9), Fig. 5-(2) shows that: (1) the blur-SURE is a reliable estimator of the blur-MSE, and the estimated  $s \approx s_0$ ; (2) the band-indicator is well approximated with the optimal  $\lambda$  and  $s$  obtained by minimizing the blur-SURE.

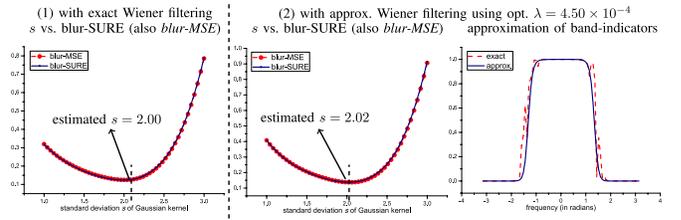


Fig. 5. Example of *Cameraman* blurred by Gaussian kernel with true  $s_0 = 2.0$  and noise variance  $\sigma^2 = 1.0$  (only one line of the band indicators are shown).

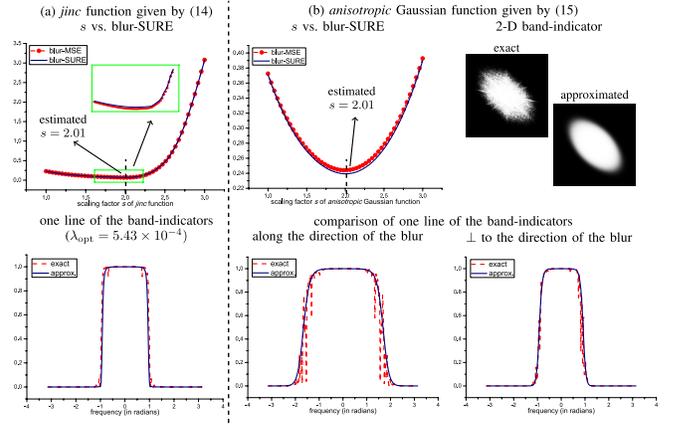


Fig. 6. Blur-SURE minimization with optimal  $\lambda$ : example of *Cameraman* blurred by two non-Gaussian kernels with true scaling factor  $s_0 = 2.0$  and noise variance  $\sigma^2 = 1.0$ .

2) *Non-Gaussian PSF With Scaling Factor s*: From (13), we can see that  $s$  plays the role of a dilation/scaling factor in a Gaussian function. Here, we consider the following two typical non-Gaussian functions with scaling factor  $s$ :

- *jinc* function,<sup>3</sup> which is frequently used for describing the optical diffraction [51], given as:

$$\mathbf{h}_s(r; s) = M \cdot \left[ \frac{2J_1(r/s)}{r/s} \right]^2 \quad (14)$$

where  $J_1(\cdot)$  is first-order Bessel function. It is an *isotropic* function of radius  $r = \sqrt{i^2 + j^2}$ .

- *anisotropic* Gaussian function given as:

$$\mathbf{h}_s(i, j; s) = M \cdot \exp\left(-\frac{(i \cos \theta - j \sin \theta)^2}{s \cdot \sigma_1^2} - \frac{(i \sin \theta + j \cos \theta)^2}{s \cdot \sigma_2^2}\right) \quad (15)$$

where  $\theta$  denotes the main direction w.r.t. the horizontal line,  $\sigma_1$  and  $\sigma_2$  stand for the blur sizes along two perpendicular directions. Here, we set  $\sigma_1 = 1.0$ ,  $\sigma_2 = 2.0$  and  $\theta = \pi/4$  in our example.

Fig. 6 shows the blur-SURE minimizations and band-indicator approximations under *jinc* function and *anisotropic* Gaussian kernel, respectively. We can see that the blur-SURE minimization yields a highly accurate estimate of the scaling factor and the best approximation of the band-indicator.

<sup>3</sup>The terminology *jinc* stems from optics, due to the structural similarity to sinc function, see [50].

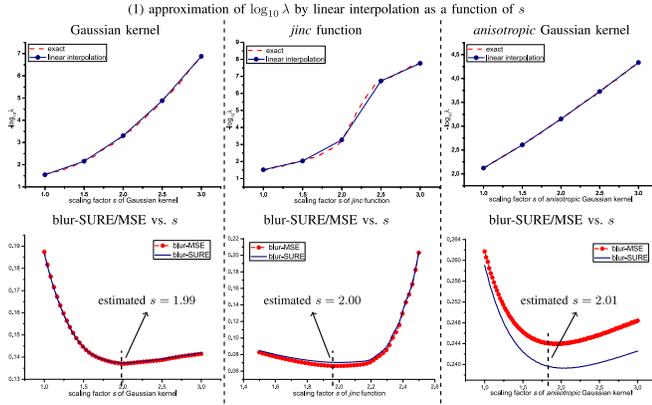


Fig. 7. Minimization of the blur-SURE  $\epsilon(s, \lambda)$  over  $s$  and  $\lambda$ . Top: given  $s$ , the optimal  $\lambda$  is a function of  $s$  whose  $\log_{10}$  is approximated by linear interpolation. Bottom: minimization of  $\epsilon(s, \lambda(s))$ . Experiment: *Cameraman* blurred by three kernels with true  $s_0 = 2.0$ : (left) Gaussian kernel, (middle) *jinc* function and (right) *anisotropic* Gaussian kernel.

Here, we stress that the blur-SURE approach also works for other types of PSF with two or more unknown parameters, for example, [52] estimated motion direction and blur length for linear motion deblurring, based on the blur-SURE criterion.

### C. An Efficient Minimization Algorithm

Considering the blur-SURE minimization over two scalar variables (PSF parameter  $s$  and regularization parameter  $\lambda$ ), the most straightforward way is to perform exhaustive search over all the possible values of  $s$  and  $\lambda$  in a certain range. If we take 50 discrete points for  $s$  and  $\lambda$  to process the image of size  $256 \times 256$ , the exhaustive search requires  $50 \times 50 = 2500$  computations of the blur-SURE.

The exhaustive search is essentially equivalent to what follows: for each fixed  $s$ , obtain the corresponding optimal  $\lambda$  by minimizing the blur-SURE, and then, insert the function  $\lambda = \lambda(s)$  into the blur-SURE. Thus, the minimization has to be performed over only one variable  $s$ . Experimentally, we found an approximate linear relation between the logarithm of  $\lambda$  and the associated optimal blur size  $s$ . Hence, we perform linear interpolation of  $\log_{10}(\lambda(s))$  over very few sampling points. We show the shape of this function (see the dashed curves) and its linear interpolation (see the navy blue curve) in Fig. 7 for the Gaussian, *jinc*, and *anisotropic* Gaussian kernels. It enables us to develop a more efficient algorithm than exhaustive search (see Algorithm 1). Fig. 7 (bottom) shows the line searches for minimizing the blur-SURE  $\epsilon(s, \lambda(s))$  w.r.t.  $s$ , where the corresponding optimal  $\lambda = \lambda(s)$  is given in Fig. 7 (top). Compared to the exhaustive search, Algorithm 1 requires only  $5 \times 50 + 50 = 300$  computations of the blur-SURE, i.e., roughly 8 times less. However, the efficiency of this algorithm reduces when the PSF has more than one parameter. In that case, quasi-Newton methods or alternating minimizations can be substituted.

Finally, we would like to note that the blur-SURE can be directly computed in the Fourier domain: there is no need to transform the Fourier coefficients back to the image domain. Thus, the computational complexity of the blur-SURE is further greatly reduced.

### Algorithm 1 Algorithm by Approximating the Function $\lambda = \lambda(s)$

**Input:**  $\epsilon(s, \lambda)$ : objective function given as (12);

**Output:** optimal  $\lambda$  and  $s$

- 1: take five sample-values  $s_k = 1.0, 1.5, 2.0, 2.5, 3.0$  for  $k = 1, 2, \dots, 5$ ;
- 2: for each  $s_k$ , find the optimal  $\lambda_k$  corresponding to the minimum  $\epsilon$ :  $\min_{\lambda_k} \epsilon(s_k, \lambda_k)$ ;
- 3: use linear interpolation to establish  $\lambda = \lambda(s)$ , based on the five pairs  $(s_k, \lambda_k)$ ;
- 4: incorporate  $\lambda = \lambda(s)$  into  $\epsilon(s, \lambda)$ , and minimize  $\epsilon(s, \lambda(s))$  over  $s$ .

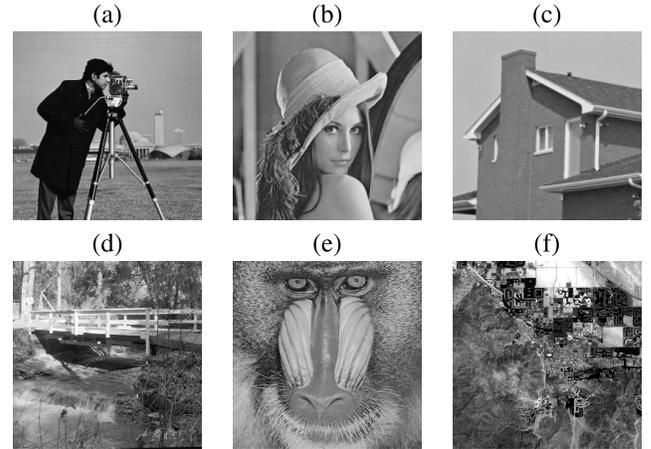


Fig. 8. Original images. (a) *Cameraman*  $256 \times 256$ ; (b) *Lena*  $256 \times 256$ ; (c) *House*  $256 \times 256$ ; (d) *Bridge*  $512 \times 512$ ; (e) *Mandrill*  $512 \times 512$ ; (f) *California*  $512 \times 512$ .

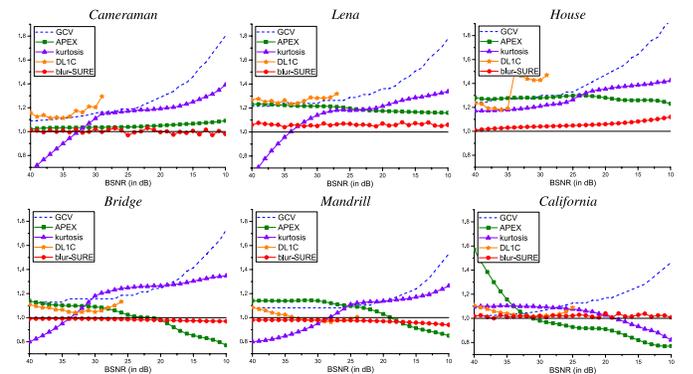


Fig. 9. Estimation of the standard deviation  $s$  (vertical axis) of a Gaussian PSF vs. BSNR (horizontal axis): comparison with state-of-the-art estimation methods (ground truth:  $s_0 = 1.0$ , horizontal black line).

## V. RESULTS AND DISCUSSIONS

We are going to evaluate the quality of our PSF parametric estimation. For this, we will consider the deconvolution efficiency, see Fig. 4. For computing it, we will use our multi-Wiener SURE-LET algorithm. We have demonstrated in [21] that this algorithm reaches the state of the art in deconvolution.

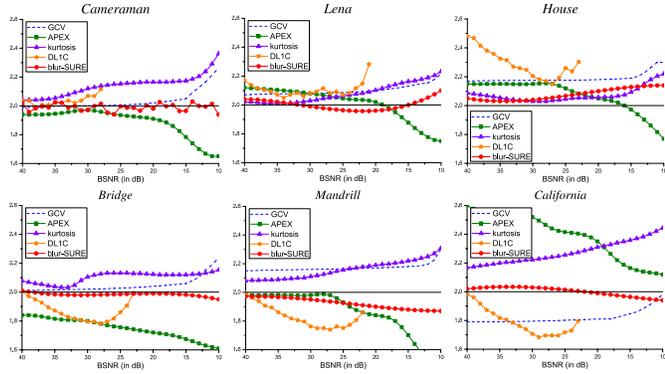


Fig. 10. Estimation of the standard deviation  $s$  (vertical axis) of a Gaussian PSF vs. BSNR (horizontal axis): comparison with state-of-the-art estimation methods (ground truth:  $s_0 = 2.0$ , horizontal black line).

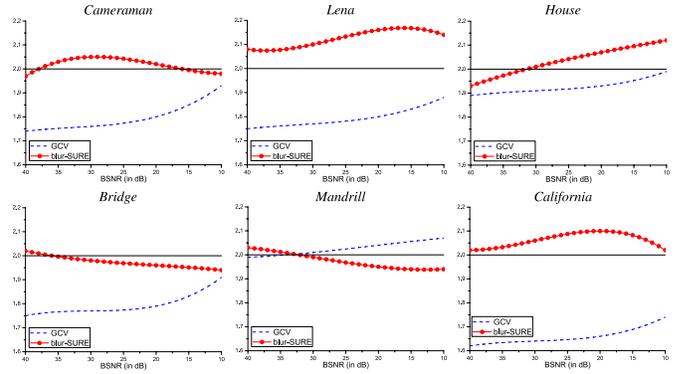


Fig. 12. Estimation of the scale factor  $s$  (vertical axis) of an *anisotropic* Gaussian PSF vs. BSNR (horizontal axis): comparison with GCV (ground truth:  $s_0 = 2.0$ , horizontal black line). Other methods are severely underperforming.

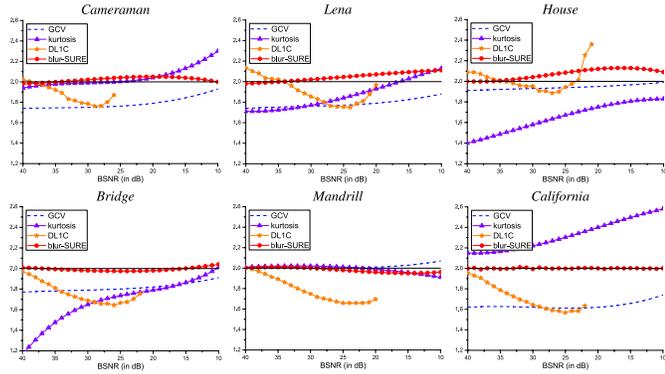


Fig. 11. Estimation of the scale factor  $s$  (vertical axis) of a *jinc* PSF vs. BSNR (horizontal axis): comparison with the GCV, kurtosis and DLIC methods (ground truth:  $s_0 = 2.0$ , horizontal black line). APEX is not applicable for *jinc* kernel.

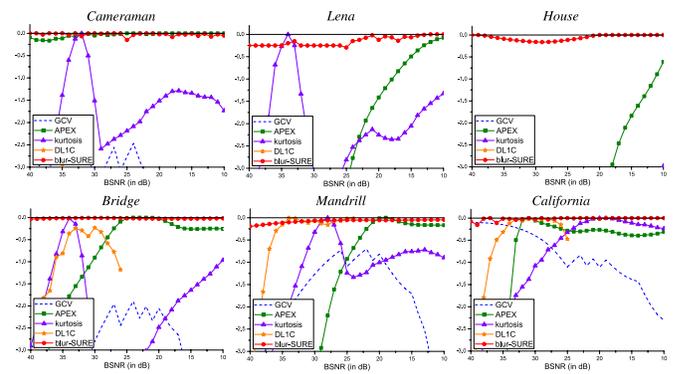


Fig. 13. Plugging the estimate of a Gaussian PSF (ground truth  $s_0 = 1.0$ ) into a state-of-the-art non-blind deconvolution algorithm [21]: comparison of PSNR loss (reference: deconvolution using  $s_0$ ) between PSF estimation methods. Note that some methods may be out of graph boundaries.

TABLE I

COMPARISON OF THE COMPUTATIONAL TIME OF VARIOUS PSF ESTIMATION TECHNIQUES (SECONDS)\*

Degradation scenario	GCV [39]	APEX [24]	kurtosis [31]	DLIC [38]	blur-SURE
<i>Cameraman</i> $256 \times 256$ Gaussian, $s_0 = 2.0$ BSNR=40dB	1.24	2.08	1.59	1.15	<b>0.73</b>
<i>California</i> $512 \times 512$ Gaussian, $s_0 = 1.0$ BSNR=20dB	3.74	5.72	5.89	3.17	<b>1.90</b>

\* Hardware environment used — Processor: 2.4GHz Intel Core i7; Memory: 8GB 1600MHz DDR3.

A. Experimental Setting

We consider the following convolution kernels described above, with scaling factor  $s$ :

- Gaussian function by (13);
- Non-Gaussian functions: *jinc* function by (14) and *anisotropic* Gaussian function by (15).

The blurred images are subsequently contaminated by i.i.d Gaussian noise with various variance  $\sigma^2$ , such that the blur signal-to-noise ratio (BSNR), defined as

$$BSNR = 10 \log_{10} \left( \frac{\|\mathbf{H}_0 \mathbf{x} - \text{mean}(\mathbf{H}_0 \mathbf{x})\|^2}{N \sigma^2} \right),$$

is 40, 30, 20 and 10 dB, respectively. The test dataset<sup>4</sup> contains six 8-bit images of size  $256 \times 256$  or  $512 \times 512$  displayed

<sup>4</sup>All  $512 \times 512$  images are available at <http://decsai.ugr.es/cvg/CG/base.htm>.

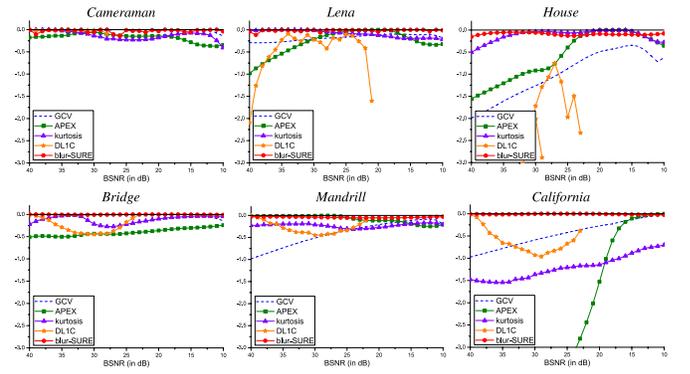


Fig. 14. Plugging the estimate of a Gaussian PSF (ground truth  $s_0 = 2.0$ ) into a state-of-the-art non-blind deconvolution algorithm [21]: comparison of PSNR loss (reference: deconvolution using  $s_0$ ) between PSF estimation methods. Note that some methods may be out of graph boundaries.

in Fig. 8, covering a wide range of natural images (people, animal, building, remote sensing, etc.). As emphasized above, we separate the PSF estimation and deconvolution. For the PSF estimation, we present our estimated PSF parameter, compared to the ground truth parameter  $s_0$ . The deconvolution performance is measured by the peak signal-to-noise ratio (PSNR), defined as

$$PSNR = 10 \log_{10} \left( \frac{255^2}{\|\hat{\mathbf{x}} - \mathbf{x}\|^2 / N} \right),$$

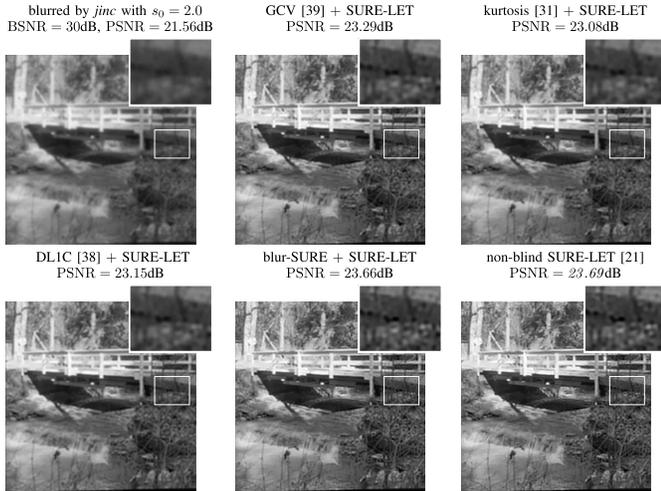


Fig. 15. Restoration of *Bridge*: a visual example for comparison between various PSF estimation methods, followed by the same SURE-LET deconvolution algorithm [21].

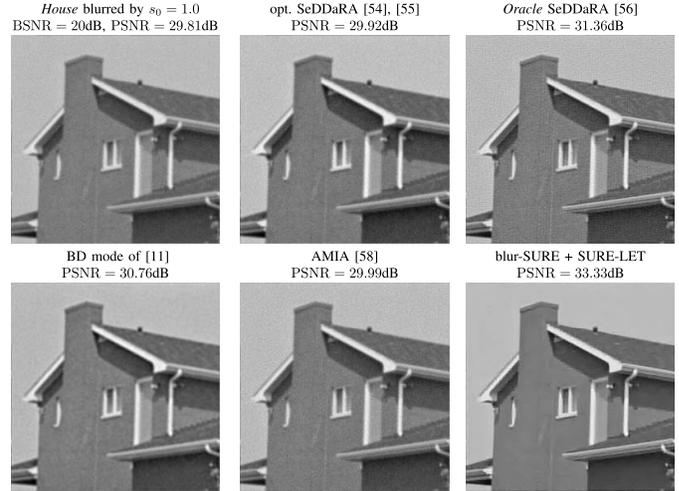


Fig. 17. Restoration of *House*: visual comparison between SeDDaRA [54], [55], *oracle* SeDDaRA [56], Molina *et al.* work [11], AMIA [58] and our proposed blur-SURE method.

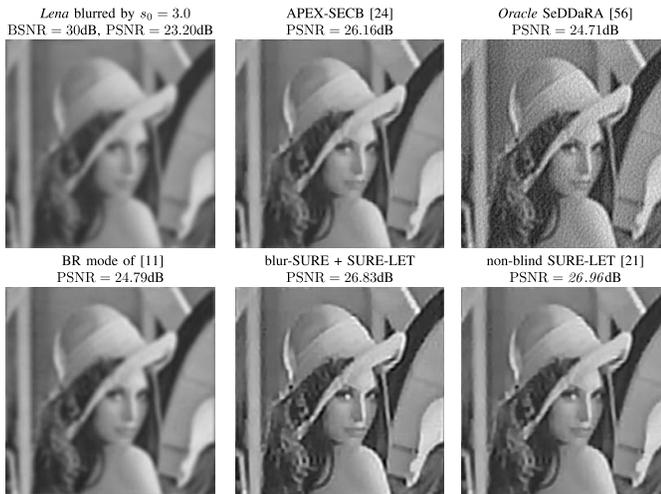


Fig. 16. Restoration of *Lena*: visual comparison between APEX-SECB [24], *oracle* SeDDaRA [56], Molina *et al.* work [11], our proposed blur-SURE method, and non-blind SURE-LET method [21].

where  $\hat{x}$  is the image deconvolved using the estimated PSF parameters. Note that all the results reported in this section have been averaged over 10 noise realizations.

### B. Gaussian Function: Estimation of Standard Deviation $s$

Figures 9–10 report the estimated Gaussian blur sizes under two true  $s_0 = 1.0$  and  $2.0$ . We also compare our estimation results with other state-of-the-art methods, including GCV-based [39], APEX [24], kurtosis-based [31] and DLIC [38]. Note that we optimized the configuration of GCV and APEX to get the best estimation results. For kurtosis and DLIC, we use the default settings as suggested by the authors. Unfortunately, DLIC fails to estimate the blur size at higher noise levels (e.g. BSNR  $< 20$ – $25$  dB). Hence, Figs. 9–11 are unable to show DLIC results over the whole BSNR range 10–40 dB. In addition, we could not compare our results with edge-based approaches [34]–[36], because



Fig. 18. *Text* restoration: the estimated Gaussian widths (modelization of the defocusing) are  $s = 1.62$ ,  $s = 2.33$  and  $s = 3.50$ , for images *a*, *b* and *c*, respectively.

these algorithms are too local, which results in substantial spatial variations: they do not provide a global blur size estimate.

From Figures 9–10, we can see that for all the test images: (1) our approach generally yields more accurate and consistent estimation of the blur size  $s$  than other methods, in terms of the error defined by  $e = |s - s_0|$ ; (2) without tuning any parameters, the blur-SURE approach automatically yields estimates that are within a few percents of the ground truth; (3) the blur-SURE approach is very robust to the noise corruption, even under high noise level.

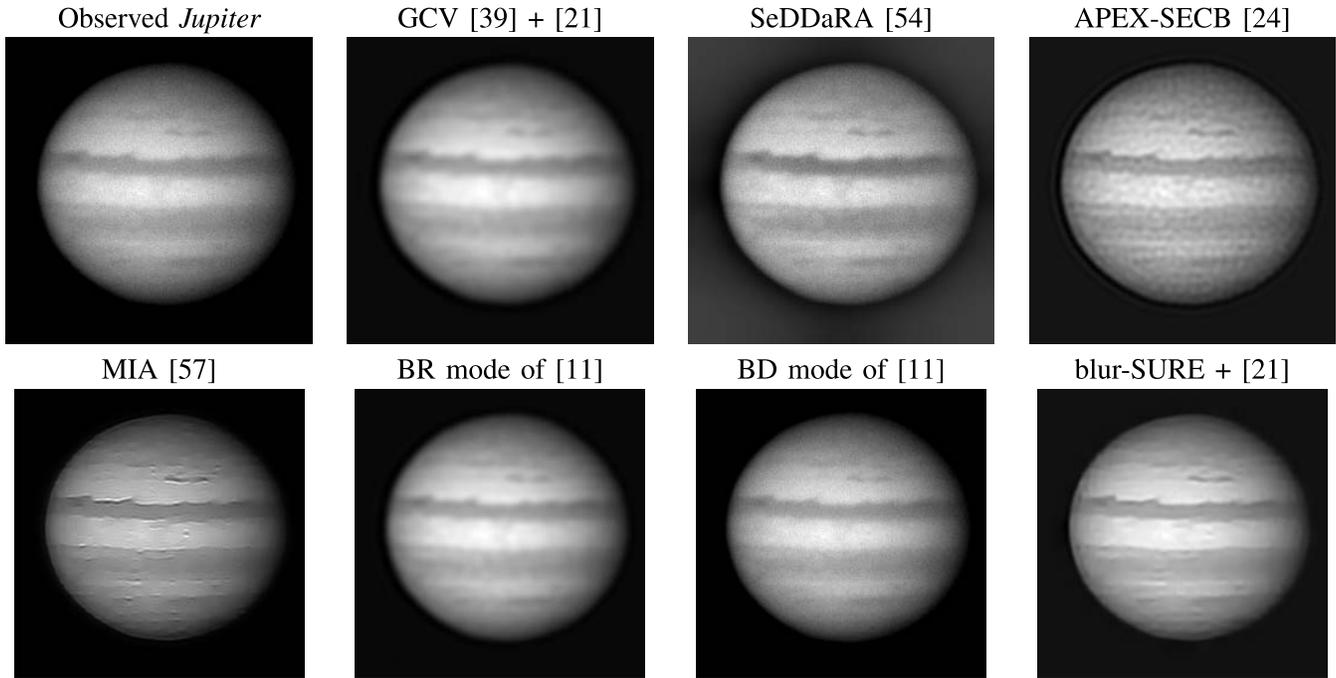


Fig. 19. Restoration of *Jupiter*: visual comparison between GCV [39], SeDDaRA [54], Molina *et al.* work [11], MIA [57], APEX-SECB [24] and our proposed blur-SURE method. The estimated Gaussian widths are  $s = 1.92$ ,  $s = 2.30$  and  $s = 2.41$ , by GCV, APEX and blur-SURE, respectively.

Table I shows the computation time comparisons. We can see that the blur-SURE approach is substantially faster than others.

### C. Non-Gaussian Functions: Estimation of the Scaling Factor $s$

Table II shows the estimation of the scaling factor of the (non-Gaussian) blur functions by the blur-SURE minimization, where the italic numbers stand for the *oracle* results (blur-MSE minimization). Similar to Figures 9–10, Figures 11–12 report comparisons between various methods, for  $s_0 = 2.0$  in the case of the *jinc* and *anisotropic* Gaussian PSF.

APEX [24] is not applicable for these kernels, while kurtosis-based [31] and DL1C [38] methods perform very poorly for estimating *anisotropic* Gaussian kernel. Hence, we do not report their results here. Again, the accuracy, consistency and robustness w.r.t. noise of our algorithm compares very favorably with other methods.

### D. PSF Estimation Accuracy and Blind Deconvolution Performance

We now evaluate the importance of PSF estimation on deconvolution quality. In order to be fair, we have applied the same non-blind deconvolution algorithm (multi-Wiener SURE-LET [21]) after performing PSF estimation using each method. Figures 13–14 show the corresponding deconvolution results in terms of PSNR loss with respect to the “oracle” multi-Wiener SURE-LET deconvolution (using the exact blur kernel). We can see that (1) more accurate PSF estimation always yields better deconvolution performance; (2) the blur-SURE approach, though not always superior, consistently

outperforms other methods on average: its PSNR loss is always kept within 0.1–0.2 dB. Other methods produce poor results from time to time. Extensive experiments (results not shown here) on other types of PSF (e.g. *jinc* and *anisotropic* Gaussian kernels) and other non-blind deconvolution algorithms (e.g. BM3D [48]) show the same trend as in Figures 13–14. Fig.15 shows a visual example of *Bridge*, blurred by *jinc* kernel with scaling factor  $s_0 = 2.0$ . We always perform the same SURE-LET deconvolution algorithm [21] after different PSF estimation methods.

### E. Comparison of Blind Deconvolution Algorithms Involving Parametric PSF

Our main motivation for PSF estimation is to enable us to perform blind deconvolution. For this reason, we would like to compare our results to the ones produced by state-of-the-art blind deconvolution algorithms. As mentioned earlier, after PSF estimation using the blur-SURE method, we use the multi-Wiener SURE-LET algorithm [21] to perform deconvolution. The other algorithms considered in our tests are:

- APEX-SECB [24] (*Parametric*): In the work of [24], the PSF estimation by APEX is followed by SECB deconvolution method [53].
- SeDDaRA (*Non-Parametric*):
  - 1) The work of [54] and [55]: An approach proposed in [55] is used to optimize the parameter involved in [54].
  - 2) The work of [56]: It uses the unknown original image  $\mathbf{x}$  to optimize the results, denoted by *oracle* SeDDaRA.

TABLE II  
ESTIMATION OF THE SCALING FACTOR  $s_0$  OF A NON-GAUSSIAN PSF

true $s_0$		$s_0 = 1.0$				$s_0 = 2.0$				$s_0 = 3.0$			
BSNR (in dB)		40	30	20	10	40	30	20	10	40	30	20	10
blur type		<i>jinc</i> function (14)											
<i>Cameraman</i>	blur-SURE	1.00	1.01	1.04	1.02	1.98	2.02	2.05	2.00	3.01	3.04	3.06	3.10
	<i>blur-MSE</i>	<i>1.00</i>	<i>1.00</i>	<i>1.04</i>	<i>1.10</i>	<i>2.00</i>	<i>2.00</i>	<i>2.05</i>	<i>2.06</i>	<i>2.98</i>	<i>3.02</i>	<i>3.04</i>	<i>3.12</i>
<i>Mandrill</i>	blur-SURE	1.00	1.00	0.99	0.94	2.00	2.00	1.96	1.96	3.00	3.01	2.99	3.01
	<i>blur-MSE</i>	<i>1.00</i>	<i>1.00</i>	<i>0.99</i>	<i>0.92</i>	<i>1.99</i>	<i>2.00</i>	<i>1.98</i>	<i>1.95</i>	<i>3.00</i>	<i>3.01</i>	<i>3.01</i>	<i>2.98</i>
<i>California</i>	blur-SURE	1.00	1.01	1.03	1.03	2.00	2.00	2.00	2.00	3.00	3.00	2.98	3.04
	<i>blur-MSE</i>	<i>1.00</i>	<i>1.01</i>	<i>1.03</i>	<i>1.07</i>	<i>2.00</i>	<i>2.01</i>	<i>2.00</i>	<i>2.00</i>	<i>3.00</i>	<i>3.00</i>	<i>3.01</i>	<i>3.02</i>
blur type		<i>anisotropic</i> Gaussian PSF (15)											
<i>Cameraman</i>	blur-SURE	1.04	1.06	1.06	1.07	1.97	2.08	2.02	1.98	3.13	2.91	2.95	2.95
	<i>blur-MSE</i>	<i>1.04</i>	<i>1.05</i>	<i>1.03</i>	<i>1.13</i>	<i>1.91</i>	<i>2.06</i>	<i>2.06</i>	<i>2.09</i>	<i>3.01</i>	<i>3.07</i>	<i>3.03</i>	<i>3.15</i>
<i>Mandrill</i>	blur-SURE	0.99	0.95	0.95	0.95	2.03	1.99	1.95	1.92	2.93	2.87	2.83	2.81
	<i>blur-MSE</i>	<i>0.99</i>	<i>0.95</i>	<i>0.95</i>	<i>0.95</i>	<i>2.02</i>	<i>1.98</i>	<i>1.95</i>	<i>1.93</i>	<i>2.93</i>	<i>2.87</i>	<i>2.93</i>	<i>2.84</i>
<i>California</i>	blur-SURE	1.01	1.07	1.07	1.07	2.02	2.06	2.10	2.02	3.01	3.03	3.07	2.99
	<i>blur-MSE</i>	<i>1.03</i>	<i>1.07</i>	<i>1.07</i>	<i>1.07</i>	<i>2.02</i>	<i>2.06</i>	<i>2.10</i>	<i>1.98</i>	<i>3.07</i>	<i>3.09</i>	<i>3.03</i>	<i>2.95</i>

TABLE III  
BLIND DECONVOLUTION: GAUSSIAN PSF WITH STANDARD DEVIATION  $s_0 = 2.0$

BSNR (in dB)		40	30	20	10	40	30	20	10	40	30	20	10
Image		<i>Cameraman</i> $256 \times 256$				<i>Lena</i> $256 \times 256$				<i>House</i> $256 \times 256$			
APEX-SECB [24]		25.60	24.72	23.78	22.41	29.54	28.65	27.35	25.56	30.70	29.97	28.84	26.72
SeD-DaRA	[54]+[55]	23.15	23.00	22.26	19.79	26.69	26.37	25.25	22.43	27.84	27.57	26.37	23.10
	<i>oracle</i> [56]	24.47	23.00	21.86	20.72	28.59	26.97	25.54	24.01	30.23	28.63	27.23	25.66
[11]	BR mode	22.54	23.74	22.96	21.71	25.52	27.15	26.29	24.77	26.78	29.30	27.99	25.84
	BD mode	23.86	23.39	22.38	17.85	27.04	26.76	25.77	21.87	26.78	28.60	27.04	21.23
MIA [57]		22.79	22.68	22.15	19.72	25.86	25.73	25.08	22.36	27.16	27.01	26.21	23.01
AMIA [58]		23.30	23.08	22.15	19.72	26.53	26.26	25.09	22.36	27.89	27.54	26.22	23.01
blur-SURE + [21]		<b>26.30</b>	<b>25.40</b>	<b>24.45</b>	<b>23.16</b>	<b>30.51</b>	<b>29.47</b>	<b>27.97</b>	<b>26.39</b>	<b>31.81</b>	<b>31.03</b>	<b>29.79</b>	<b>28.15</b>
Image		<i>Bridge</i> $512 \times 512$				<i>Mandrill</i> $512 \times 512$				<i>California</i> $512 \times 512$			
APEX-SECB [24]		25.59	24.84	23.95	22.61	23.27	22.64	22.00	21.17	20.37	21.65	21.84	21.03
SeD-DaRA	[54]+[55]	23.39	23.27	22.67	20.53	21.56	21.48	21.17	20.22	21.72	21.58	20.98	19.25
	<i>oracle</i> [56]	24.52	23.17	22.10	21.05	22.65	21.26	20.21	19.42	23.38	21.85	20.63	19.46
[11]	BR mode	22.86	23.67	23.02	21.74	21.14	21.59	21.23	20.56	20.99	21.92	21.26	19.95
	BD mode	22.86	23.37	22.64	18.08	21.14	21.46	21.07	19.26	20.99	21.60	20.89	17.22
MIA [57]		23.10	23.02	22.57	20.51	21.29	21.26	21.07	20.22	21.24	21.18	20.81	19.24
AMIA [58]		23.47	23.41	22.65	20.51	21.52	21.50	21.15	20.22	21.65	21.60	20.94	19.24
blur-SURE + [21]		<b>26.10</b>	<b>25.29</b>	<b>24.34</b>	<b>23.12</b>	<b>23.56</b>	<b>22.85</b>	<b>22.10</b>	<b>21.29</b>	<b>24.80</b>	<b>23.88</b>	<b>22.77</b>	<b>21.40</b>

- The work of [11] (*Non-Parametric*): It has two variants, depending on the assumptions of the distributions of original image and PSF: BR (both distributions are random) and BD (both distributions are degenerate). We use the parameters as recommended in [11].
- MIA [57] and AMIA [58] Algorithms (*Non-Parametric*): Two multiplicative iterative algorithms.

Note that in this paper, the blur-SURE approach is mainly exemplified with atmospheric/optical blurring, which are frequently encountered in microscopy, remote sensing and out-of-focus imaging. For this reason, we do not include *motion deblurring* algorithms (see [8], [22], [59]–[63]) in our comparisons, because their results would be poor when applied to these PSF.

From Table III, we observe that our proposed method outperforms the other methods substantially. Figs. 16–17 show two visual comparisons between the existing methods. We can

see that our novel PSF estimation method is the source a blind deconvolution algorithm that achieves similar visual quality as the non-blind algorithm [21].

Table IV shows the comparison of computational times. We can see that the proposed approach is substantially faster than all the other blind deconvolution methods.

#### F. Blind Deconvolution of Real Images

In our last set of experiments, the method is applied to two real observed images: *Text* captured by a digital camera and *Jupiter* captured by a telescope.

*Text Image Text*: Three images with different focal lengths (luminance shown in Fig. 18) were captured by a digital camera. We used the Gaussian assumption to modelize the out-of-focus blur (in agreement with [64]), and compared our results with the algorithms of [11]. The improved legibility of our result is obvious, even in the

TABLE IV  
COMPARISON OF COMPUTATIONAL TIMES FOR VARIOUS BLIND  
DECONVOLUTION TECHNIQUES (SECONDS)\*

Degradation scenario	APEX-SECB [24]	ScDDaRA [54]	[11] BR mode	MIA [57]	AMIA [58]	blur-SURE +[21]
<i>Camerman</i> 256 × 256 Gaussian, $s_0 = 2.0$ BSNR=40dB	2.18	4.01	5.63	3.15	6.98	<b>1.11</b>
<i>California</i> 512 × 512 Gaussian, $s_0 = 1.0$ BSNR=20dB	6.13	19.67	18.82	12.50	13.61	<b>3.71</b>

\* Hardware environment used — Processor: 2.4GHz Intel Core i7; Memory: 8GB 1600MHz DDR3.

case of substantial blurring, although artifacts start being important.

*Astronomical Image Jupiter*: There is no exact expression for the PSF of this image; however, as suggested in [6], [7], [11], and [14], the PSF can again be well approximated by Gaussian function. Fig. 19 shows the restored images by various state-of-the-art methods. We can easily see that our proposed approach yields a higher visual quality.

## VI. CONCLUSION

In this paper, we proposed a parametric PSF estimation method based on a new criterion—the “blur-SURE”. We have shown that in conjunction with Wiener filtering, the minimization of the blur-SURE yields highly accurate PSF estimations. The blur-SURE framework is exemplified by several parametric forms of PSF. For fast implementation, we also proposed an efficient algorithm for performing the blur-SURE minimization.

The results obtained show that, pipelining this novel PSF estimation method with our non-blind multi-Wiener SURE-LET deconvolution algorithm, results in a blind deconvolution algorithm which outperforms other approaches numerically, visually, and in terms of computational efficiency. The examples of blur kernels listed in this paper are only a small subset of possible models. It is worth emphasizing that the blur-SURE minimization itself does not specify any particular parametric form for the PSF.

In addition, it should be noted that the blur-SURE criterion is not limited to 2D signals: it can be extended painlessly to 3D (or any other dimension). It is not limited either to additive Gaussian noise: the blur-SURE expression can also be adapted to any noise type for which an unbiased risk estimate can be devised [65], [66]. Hence, despite its apparent phase limitation, we believe that this criterion has a huge potential in the development of specific algorithms for various applications; e.g. fluorescence microscopy [19], [29], [67].

## APPENDIX A PROOF OF THEOREM 2.1

*Proof*: First, consider the minimization of the signal-averaged blur-MSE over all possible linear processings  $\mathbf{U}$ :

$$\min_{\mathbf{U}} \frac{1}{N} \mathcal{E}_{\mathbf{x}, \mathbf{b}} \left\{ \|\mathbf{U}\mathbf{y} - \mathbf{H}_0\mathbf{x}\|^2 \right\} \quad (16)$$

Replacing  $\mathbf{y}$  by  $\mathbf{H}_0\mathbf{x} + \mathbf{b}$  as (1), the signal-averaged blur-MSE becomes:

$$\begin{aligned} \mathcal{E}_{\mathbf{x}}\{\text{blur-MSE}\} &= \frac{1}{N} \mathcal{E}_{\mathbf{x}, \mathbf{b}} \left\{ \|\mathbf{U}\mathbf{y} - \mathbf{H}_0\mathbf{x}\|^2 \right\} \\ &= \frac{1}{N} \mathcal{E}_{\mathbf{x}, \mathbf{b}} \left\{ \|\mathbf{U}(\mathbf{H}_0\mathbf{x} + \mathbf{b}) - \mathbf{H}_0\mathbf{x}\|^2 \right\} \\ &= \frac{1}{N} \mathcal{E}_{\mathbf{x}, \mathbf{b}} \left\{ \|(\mathbf{U} - \mathbf{I})\mathbf{H}_0\mathbf{x} + \mathbf{U}\mathbf{b}\|^2 \right\} \\ &= \frac{1}{N} \mathcal{E}_{\mathbf{x}} \left\{ \|(\mathbf{U} - \mathbf{I})\mathbf{H}_0\mathbf{x}\|^2 \right\} + \frac{1}{N} \mathcal{E}_{\mathbf{b}} \left\{ \|\mathbf{U}\mathbf{b}\|^2 \right\} \\ &= \frac{1}{N} \text{Tr} \left( (\mathbf{U} - \mathbf{I})\mathbf{H}_0\mathbf{A}\mathbf{H}_0^T (\mathbf{U} - \mathbf{I})^T \right) \\ &\quad + \frac{1}{N} \text{Tr}(\mathbf{U}\mathbf{C}\mathbf{U}^T) \end{aligned} \quad (17)$$

where  $\text{Tr}$  denotes matrix trace,  $\mathbf{I}$  is identity matrix, the covariance matrices are  $\mathbf{A} = \mathcal{E}_{\mathbf{x}}\{\mathbf{x}\mathbf{x}^T\}$  and  $\mathbf{C} = \mathcal{E}_{\mathbf{b}}\{\mathbf{b}\mathbf{b}^T\}$ . Thus, the minimization over  $\mathbf{U}$  yields that  $(\mathbf{U} - \mathbf{I})\mathbf{H}_0\mathbf{A}\mathbf{H}_0^T + \mathbf{U}\mathbf{C} = \mathbf{0}$ , which implies that:

$$\mathbf{U} = \mathbf{H}_0\mathbf{A}\mathbf{H}_0^T (\mathbf{H}_0\mathbf{A}\mathbf{H}_0^T + \mathbf{C})^{-1} \quad (18)$$

Considering (3), if we base our processing  $\mathbf{f}_{\mathbf{H}}$  on Wiener filtering  $\mathbf{W}_{\mathbf{H}}$  as (4), the blur-MSE minimization over  $\mathbf{H}$  becomes:

$$\min_{\mathbf{H}} \frac{1}{N} \mathcal{E}_{\mathbf{x}, \mathbf{b}} \left\{ \|\mathbf{H}\mathbf{W}_{\mathbf{H}}\mathbf{y} - \mathbf{H}_0\mathbf{x}\|^2 \right\}. \quad (19)$$

Compare the two minimization problems: (16) and (19). Obviously, we have:

$$\min_{\mathbf{H}} \frac{1}{N} \mathcal{E}_{\mathbf{x}, \mathbf{b}} \left\{ \|\mathbf{H}\mathbf{W}_{\mathbf{H}}\mathbf{y} - \mathbf{H}_0\mathbf{x}\|^2 \right\} \geq \min_{\mathbf{U}} \frac{1}{N} \mathcal{E}_{\mathbf{x}, \mathbf{b}} \left\{ \|\mathbf{U}\mathbf{y} - \mathbf{H}_0\mathbf{x}\|^2 \right\} \quad (20)$$

Considering the left-hand side, from (4), we obtains:  $\mathbf{H}\mathbf{W}_{\mathbf{H}} = \mathbf{H}\mathbf{A}\mathbf{H}^T (\mathbf{H}\mathbf{A}\mathbf{H}^T + \mathbf{C})^{-1}$ . Hence, if  $\mathbf{H} = \mathbf{H}_0$ , then, from (18), we obtain:

$$\frac{1}{N} \mathcal{E}_{\mathbf{x}, \mathbf{b}} \left\{ \|\mathbf{H}_0\mathbf{W}_{\mathbf{H}_0}\mathbf{y} - \mathbf{H}_0\mathbf{x}\|^2 \right\} = \min_{\mathbf{U}} \frac{1}{N} \mathcal{E}_{\mathbf{x}, \mathbf{b}} \left\{ \|\mathbf{U}\mathbf{y} - \mathbf{H}_0\mathbf{x}\|^2 \right\}$$

This implies that  $\mathbf{H}\mathbf{W}_{\mathbf{H}} = \mathbf{H}_0\mathbf{W}_{\mathbf{H}_0}$  is a minimizer to the problem (19), i.e.,

$$\min_{\mathbf{H}} \frac{1}{N} \mathcal{E}_{\mathbf{x}, \mathbf{b}} \left\{ \|\mathbf{H}\mathbf{W}_{\mathbf{H}}\mathbf{y} - \mathbf{H}_0\mathbf{x}\|^2 \right\} = \frac{1}{N} \mathcal{E}_{\mathbf{x}, \mathbf{b}} \left\{ \|\mathbf{H}_0\mathbf{W}_{\mathbf{H}_0}\mathbf{y} - \mathbf{H}_0\mathbf{x}\|^2 \right\}$$

which also indicates that the equality sign holds for the inequality (20).

Finally, from  $\mathbf{H}\mathbf{W}_{\mathbf{H}} = \mathbf{H}_0\mathbf{W}_{\mathbf{H}_0}$ , we conclude that any filter  $\mathbf{H}$  such that

$$\mathbf{H}\mathbf{A}\mathbf{H}^T (\mathbf{H}\mathbf{A}\mathbf{H}^T + \mathbf{C})^{-1} = \mathbf{H}_0\mathbf{A}\mathbf{H}_0^T (\mathbf{H}_0\mathbf{A}\mathbf{H}_0^T + \mathbf{C})^{-1}$$

minimizes the problem (19). It yields that  $\mathbf{H}\mathbf{A}\mathbf{H}^T = \mathbf{H}_0\mathbf{A}\mathbf{H}_0^T$ , which completes the proof. ■

## PROOF OF THEOREM 3.1

*Proof:* Substituting  $\mathbf{y} - \mathbf{b}$  for  $\mathbf{H}_0\mathbf{x}$  as (1), and expanding the blur-MSE (3), we obtain:

$$\begin{aligned}
\text{blur-MSE} &= \frac{1}{N} \mathcal{E}_{\mathbf{b}} \left\{ \left\| \mathbf{H}\mathbf{f}(\mathbf{y}) - \mathbf{H}_0\mathbf{x} \right\|^2 \right\} \\
&= \frac{1}{N} \mathcal{E}_{\mathbf{b}} \left\{ \left\| \mathbf{H}\mathbf{f}(\mathbf{y}) - \mathbf{y} + \mathbf{b} \right\|^2 \right\} \\
&= \frac{1}{N} \left( \mathcal{E}_{\mathbf{b}} \left\{ \left\| \mathbf{H}\mathbf{f}(\mathbf{y}) - \mathbf{y} \right\|^2 \right\} + 2\mathcal{E}_{\mathbf{b}} \left\{ \mathbf{b}^T \mathbf{H}\mathbf{f}(\mathbf{y}) \right\} \right) \\
&\quad - \frac{1}{N} \left( 2\mathcal{E}_{\mathbf{b}} \left\{ \mathbf{b}^T \mathbf{y} \right\} - \mathcal{E}_{\mathbf{b}} \left\{ \left\| \mathbf{b} \right\|^2 \right\} \right) \\
&= \frac{1}{N} \mathcal{E}_{\mathbf{b}} \left\{ \left\| \mathbf{H}\mathbf{f}(\mathbf{y}) - \mathbf{y} \right\|^2 \right\} + \frac{2}{N} \mathcal{E}_{\mathbf{b}} \left\{ \mathbf{b}^T \mathbf{H}\mathbf{f}(\mathbf{y}) \right\} \\
&\quad - \frac{1}{N} \mathcal{E}_{\mathbf{b}} \left\{ \left\| \mathbf{b} \right\|^2 \right\} \tag{21}
\end{aligned}$$

Consider the multivariate Gaussian probability density function  $q(\mathbf{b}) \propto \exp\left(-\frac{\mathbf{b}^T \mathbf{C}^{-1} \mathbf{b}}{2}\right)$ . It satisfies  $q(\mathbf{b})\mathbf{b} = -\mathbf{C}\nabla_{\mathbf{b}}q(\mathbf{b})$ , where  $\nabla_{\mathbf{b}}$  is the gradient operator w.r.t.  $\mathbf{b}$ . Hence,

$$\begin{aligned}
\mathcal{E}_{\mathbf{b}} \left\{ \mathbf{b}^T \mathbf{H}\mathbf{f}(\mathbf{y}) \right\} &= \int \mathbf{b}^T \mathbf{H}\mathbf{f}(\mathbf{y}) q(\mathbf{b}) d\mathbf{b} \\
&= - \int (\nabla_{\mathbf{b}} q(\mathbf{b}))^T \underbrace{\mathbf{C}\mathbf{H}\mathbf{f}(\mathbf{y})}_{\mathbf{v}(\mathbf{y})} d\mathbf{b} \\
&= - \sum_{n=1}^N \int \frac{\partial q(\mathbf{b})}{\partial b_n} v_n(\mathbf{y}) d\mathbf{b}
\end{aligned}$$

Noting that  $\int_{-\infty}^{\infty} \frac{\partial q(\mathbf{b})}{\partial b_n} v_n d\mathbf{b} = - \int_{-\infty}^{\infty} \frac{\partial v_n}{\partial b_n} q(\mathbf{b}) d\mathbf{b}$ , which follows from integration by parts, and the fact that  $|v_n q(\mathbf{b})| \rightarrow 0$  as  $|b_n| \rightarrow \infty$ , we have:

$$\begin{aligned}
\mathcal{E}_{\mathbf{b}} \left\{ \mathbf{b}^T \mathbf{H}\mathbf{f}(\mathbf{y}) \right\} &= \sum_{n=1}^N \int_{-\infty}^{\infty} \frac{\partial v_n}{\partial b_n} q(\mathbf{b}) d\mathbf{b} \\
&= \mathcal{E}_{\mathbf{b}} \left\{ \sum_{n=1}^N \frac{\partial v_n}{\partial b_n} \right\} = \mathcal{E}_{\mathbf{b}} \left\{ \sum_{n=1}^N \frac{\partial v_n}{\partial y_n} \right\} \\
&= \mathcal{E}_{\mathbf{b}} \left\{ \text{div}_{\mathbf{y}} \mathbf{v} \right\} \tag{22}
\end{aligned}$$

The last term in (21) is:

$$\begin{aligned}
\mathcal{E}_{\mathbf{b}} \left\{ \left\| \mathbf{b} \right\|^2 \right\} &= \mathcal{E}_{\mathbf{b}} \left\{ \mathbf{b}^T \mathbf{b} \right\} = \mathcal{E}_{\mathbf{b}} \left\{ \text{Tr}(\mathbf{b}\mathbf{b}^T) \right\} \\
&= \text{Tr} \left( \mathcal{E}_{\mathbf{b}} \left\{ \mathbf{b}\mathbf{b}^T \right\} \right) = \text{Tr}(\mathbf{C}) \tag{23}
\end{aligned}$$

Substituting (22) and (23) into (21) completes the proof. ■

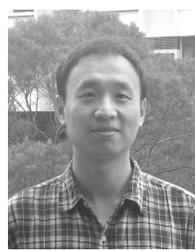
## ACKNOWLEDGMENT

The authors would like to thank Dr. Bae and Prof. Durand for providing their source code to implement [35], and Prof. Molina for sharing his code of the work [11]. Many thanks to Kim Miao LEE for providing the real images: *Jupiter* and *Text*. The authors also indebted to the anonymous reviewers for their helpful comments.

## REFERENCES

- [1] H. C. Andrews and B. R. Hunt, *Digital Image Restoration*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1977.
- [2] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. New York, NY, USA: Wiley, 1964.
- [3] A. Tikhonov and V. Arsenin, *Solutions of Ill-Posed Problems*. Washington, DC, USA: Winston, 1977.
- [4] O. Michailovich and A. Tannenbaum, "Blind deconvolution of medical ultrasound images: A parametric inverse filtering approach," *IEEE Trans. Image Process.*, vol. 16, no. 12, pp. 3005–3019, Dec. 2007.
- [5] P. Sarder and A. Nehorai, "Deconvolution methods for 3D fluorescence microscopy images," *IEEE Signal Process. Mag.*, vol. 23, no. 3, pp. 32–45, May 2006.
- [6] A. F. J. Moffat, "A theoretical investigation of focal stellar images in the photographic emulsion and application to photographic photometry," *Astron. Astrophys.*, vol. 3, pp. 455–461, Dec. 1969.
- [7] G. V. Poropat, "Effect of system point spread function, apparent size, and detector instantaneous field of view on the infrared image contrast of small objects," *Opt. Eng.*, vol. 32, no. 10, pp. 2598–2607, 1993.
- [8] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 787–794, 2006.
- [9] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Process. Mag.*, vol. 13, no. 3, pp. 43–64, May 1996.
- [10] T. F. Chan and C.-K. Wong, "Total variation blind deconvolution," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 370–375, Mar. 1998.
- [11] R. Molina, J. Mateos, and A. K. Katsaggelos, "Blind deconvolution using a variational approach to parameter, image, and blur estimation," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3715–3727, Dec. 2006.
- [12] H. Liao and M. K. Ng, "Blind deconvolution using generalized cross-validation approach to regularization parameter estimation," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 670–680, Mar. 2011.
- [13] M. S. C. Almeida and L. B. Almeida, "Blind and semi-blind deblurring of natural images," *IEEE Trans. Image Process.*, vol. 19, no. 1, pp. 36–52, Jan. 2010.
- [14] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Variational Bayesian blind deconvolution using a total variation prior," *IEEE Trans. Image Process.*, vol. 18, no. 1, pp. 12–26, Jan. 2009.
- [15] J. Bolte, P. L. Combettes, and J.-C. Pesquet, "Alternating proximal algorithm for blind image recovery," in *Proc. 17th IEEE Int. Conf. Image Process.*, Hong Kong, Sep. 2010, pp. 1673–1676.
- [16] A. S. Carasso, "Linear and nonlinear image deblurring: A documented study," *SIAM J. Numer. Anal.*, vol. 36, no. 6, pp. 1659–1689, 1999.
- [17] R. Molina, J. Núñez, F. J. Cortijo, and J. Mateos, "Image restoration in astronomy: A Bayesian perspective," *IEEE Signal Process. Mag.*, vol. 18, no. 2, pp. 11–29, Mar. 2001.
- [18] T. Bishop, S. Babacan, B. Amizic, A. Katsaggelos, T. F. Chan, and R. Molina, "Blind image deconvolution: Problem formulation and existing approaches," 2007.
- [19] J. Markham and J. A. Conchello, "Parametric blind deconvolution: A robust method for the simultaneous estimation of image and blur," *J. Opt. Soc. Amer. A*, vol. 16, no. 10, pp. 2377–2391, 1999.
- [20] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1964–1971.
- [21] F. Xue, F. Luisier, and T. Blu, "Multi-Wiener SURE-LET deconvolution," *IEEE Trans. Image Process.*, vol. 22, no. 5, pp. 1954–1968, May 2013.
- [22] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," *ACM Trans. Graph.*, vol. 27, no. 3, Aug. 2008, Art. ID 73.
- [23] L. Chen and K.-H. Yap, "A soft double regularization approach to parametric blind image deconvolution," *IEEE Trans. Image Process.*, vol. 14, no. 5, pp. 624–633, May 2005.
- [24] A. S. Carasso, "The APEX method in image sharpening and the use of low exponent Lévy stable laws," *SIAM J. Appl. Math.*, vol. 63, no. 2, pp. 593–618, 2002.
- [25] T. Kenig, Z. Kam, and A. Feuer, "Blind image deconvolution using machine learning for three-dimensional microscopy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2191–2204, Dec. 2010.
- [26] F. Kraher *et al.*, "Blind image deconvolution: Motion blur estimation," Univ. Minnesota, Minneapolis, MN, USA, Tech. Rep., 2006.
- [27] M. E. Moghaddam and M. Jamzad, "Motion blur identification in noisy images using mathematical models and statistical measures," *Pattern Recognit.*, vol. 40, no. 7, pp. 1946–1957, 2007.

- [28] J. P. Oliveira, M. A. T. Figueiredo, and J. M. Bioucas-Dias, "Parametric blur estimation for blind restoration of natural images: Linear motion and out-of-focus," *IEEE Trans. Image Process.*, vol. 23, no. 1, pp. 466–477, Jan. 2014.
- [29] P. Pankajakshan, B. Zhang, L. Blanc-Féraud, Z. Kam, J.-C. Olivo-Marín, and J. Zerubia, "Blind deconvolution for thin-layered confocal imaging," *Appl. Opt.*, vol. 48, no. 22, pp. 4437–4448, 2009.
- [30] B. J. Cooke, B. E. Laubscher, C. C. Borel, T. S. Lomheim, and C. F. Klein, "Methodology for rapid infrared multispectral electro-optical imaging system performance analysis and synthesis," *Proc. SPIE*, vol. 2743, pp. 52–86, Jun. 1996.
- [31] D. Li and S. Simske, "Atmospheric turbulence degraded-image restoration by kurtosis minimization," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 2, pp. 244–247, Apr. 2009.
- [32] J. Biemond, R. L. Lagendijk, and R. M. Mersereau, "Iterative methods for image deblurring," *Proc. IEEE*, vol. 78, no. 5, pp. 856–883, May 1990.
- [33] P. J. L. van Beek, "Edge-based image representation and coding," Ph.D. dissertation, Dept. Electr. Eng., Delft Univ. Technol., Delft, The Netherlands, 1995.
- [34] J. H. Elder and S. W. Zucker, "Local scale control for edge detection and blur estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 7, pp. 699–716, Jul. 1998.
- [35] S. Bae and F. Durand, "Defocus magnification," *Comput. Graph. Forum*, vol. 26, no. 3, pp. 571–579, 2007.
- [36] W. Zhang and W.-K. Cham, "Single-image refocusing and defocusing," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 873–882, Feb. 2012.
- [37] S. F. Gibson and F. Lanni, "Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy," *J. Opt. Soc. Amer. A*, vol. 8, no. 10, pp. 1601–1613, 1991.
- [38] F. Chen and J. Ma, "An empirical identification method of Gaussian blur parameter for image deblurring," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2467–2478, Jul. 2009.
- [39] S. J. Reeves and R. M. Mersereau, "Blur identification by the method of generalized cross-validation," *IEEE Trans. Image Process.*, vol. 1, no. 3, pp. 301–311, Jul. 1992.
- [40] C. M. Stein, "Estimation of the mean of a multivariate normal distribution," *Ann. Statist.*, vol. 9, no. 6, pp. 1135–1151, 1981.
- [41] F. Luisier, T. Blu, and M. Unser, "A new SURE approach to image denoising: Interscale orthonormal wavelet thresholding," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 593–606, Mar. 2007.
- [42] T. Blu and F. Luisier, "The SURE-LET approach to image denoising," *IEEE Trans. Image Process.*, vol. 16, no. 11, pp. 2778–2786, Nov. 2007.
- [43] C. Chaux, L. Duval, A. Benazza-Benyahia, and J. Pesquet, "A nonlinear Stein-based estimator for multichannel image denoising," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3855–3870, Aug. 2008.
- [44] C. Vonesch, S. Ramani, and M. Unser, "Recursive risk estimation for non-linear image deconvolution with a wavelet-domain sparsity constraint," in *Proc. 15th IEEE Int. Conf. Image Process.*, Oct. 2008, pp. 665–668.
- [45] J. Pesquet, A. Benazza-Benyahia, and C. Chaux, "A SURE approach for digital signal/image deconvolution problems," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4616–4632, Dec. 2009.
- [46] Y. C. Eldar, "Generalized SURE for exponential families: Applications to regularization," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 471–481, Feb. 2009.
- [47] R. Giryes, M. Elad, and Y. Eldar, "The projected GSURE for automatic parameter tuning in iterative shrinkage methods," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 3, pp. 407–422, 2010.
- [48] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image restoration by sparse 3D transform-domain collaborative filtering," *Proc. SPIE*, vol. 6812, pp. 681207-1–681207-12, Mar. 2008.
- [49] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.
- [50] Q. Cao, "Generalized Jinc functions and their application to focusing and diffraction of circular apertures," *J. Opt. Soc. Amer. A*, vol. 20, no. 4, pp. 661–667, 2003.
- [51] M. Born, E. Wolf, and A. Bhatia, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge, U.K.: Cambridge Univ. Press, 1999.
- [52] F. Xue and T. Blu, "Sure-based motion blur estimation," in *Proc. IEEE Int. Conf. Signal Process., Commun. Comput. (ICSPCC)*, Aug. 2012, pp. 373–377.
- [53] A. S. Carasso, "Direct blind deconvolution," *SIAM J. Appl. Math.*, vol. 61, no. 6, pp. 1980–2007, 2001.
- [54] J. N. Caron, N. M. Namazi, R. L. Lucke, C. J. Rollins, and P. R. Lynn, Jr., "Blind data restoration with an extracted filter function," *Opt. Lett.*, vol. 26, no. 15, pp. 1164–1166, 2001.
- [55] Y. Chuang, "Study on blind image separation and restoration," M.S. thesis, Nat. Cheng Kung Univ., Tainan, Taiwan, 2002.
- [56] J. N. Caron, N. M. Namazi, and C. J. Rollins, "Noniterative blind data restoration by use of an extracted filter function," *Appl. Opt.*, vol. 41, no. 32, pp. 6884–6889, 2002.
- [57] J. Zhang, Q. Zhang, and G. He, "Blind deconvolution: Multiplicative iterative algorithm," *Opt. Lett.*, vol. 33, no. 1, pp. 25–27, 2008.
- [58] J. Zhang, Q. Zhang, and G. He, "Blind image deconvolution by means of asymmetric multiplicative iterative algorithm," *J. Opt. Soc. Amer. A*, vol. 25, no. 3, pp. 710–717, 2008.
- [59] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in *Computer Vision*. Berlin, Germany: Springer-Verlag, 2010, pp. 157–170.
- [60] S. Cho and S. Lee, "Fast motion deblurring," *ACM Trans. Graph.*, vol. 28, no. 5, 2009, Art. ID 145.
- [61] M. Hirsch, C. J. Schuler, S. Harmeling, and B. Scholkopf, "Fast removal of non-uniform camera shake," in *Proc. IEEE Int. Conf. Comput. Vis.*, Nov. 2011, pp. 463–470.
- [62] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce, "Non-uniform deblurring for shaken images," *Int. J. Comput. Vis.*, vol. 98, no. 2, pp. 168–186, 2012.
- [63] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Efficient marginal likelihood optimization in blind deconvolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 2657–2664.
- [64] R. Fabian and D. Malah, "Robust identification of motion and out-of-focus blur parameters from blurred and noisy images," *CVGIP, Graph. Models Image Process.*, vol. 53, no. 5, pp. 403–412, 1991.
- [65] M. Raphan and E. P. Simoncelli, "Learning to be Bayesian without supervision," in *Proc. Adv. Neural Inf. Process. Syst.*, 2006, pp. 1145–1152.
- [66] F. Luisier, C. Vonesch, T. Blu, and M. Unser, "Fast interscale wavelet denoising of Poisson-corrupted images," *Signal Process.*, vol. 90, no. 2, pp. 415–427, Feb. 2010.
- [67] B. Zhang, J. Zerubia, and J.-C. Olivo-Marín, "Gaussian approximations of fluorescence microscope point-spread function models," *Appl. Opt.*, vol. 46, no. 10, pp. 1819–1829, 2007.



**Feng Xue** was born in Shandong, China, in 1983. He received the bachelor's degree in electronic science and technology and the master's degree in optical engineering from the Harbin Institute of Technology, Harbin, China, in 2005 and 2007, respectively, and the Ph.D. degree in electronics engineering from The Chinese University of Hong Kong, Hong Kong, in 2013, under the supervision of Prof. T. Blu.

He is currently with the National Key Laboratory of Science and Technology on Test Physics and Numerical Mathematics, Beijing, China. His research interests include image deconvolution, sparse reconstruction, and statistical learning.



**Thierry Blu** (M'96–SM'06–F'12) was born in Orléans, France, in 1964. He received the Dipl.-Ing. degree from the École Polytechnique, Paris, France, in 1986, and Télécom Paris (ENST), Paris, in 1988, and the Ph.D. degree in electrical engineering from ENST in 1996, with a focus on iterated rational filter banks, applied to wideband audio coding.

He was with the Biomedical Imaging Group, Swiss Federal Institute of Technology, Lausanne, Switzerland, from 1998 and 2007. He is currently a Professor with the Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong. His research interests include wavelets, approximation and sampling theory, sparse representations, image denoising, biomedical imaging, optics, and wave propagation.

Dr. Blu was a recipient of two best paper awards from the IEEE Signal Processing Society (2003 and 2006). He has co-authored a paper that received the Young Author Best Paper Award (2009) from the same society. He has been a member of the IEEE Signal Processing Theory and Methods Technical Committee (2008–2013), and an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING (2002–2006), the IEEE TRANSACTIONS ON SIGNAL PROCESSING (2006–2010), and *Elsevier Signal Processing* (2008–2011). He has been on the Board of the *EURASIP Journal on Image and Video Processing* since 2010.