Multi-Wiener SURE-LET Deconvolution

Feng Xue, *Student Member, IEEE*, Florian Luisier, *Member, IEEE*, and Thierry Blu, *Fellow, IEEE*

Abstract—In this paper, we propose a novel deconvolution algorithm based on the minimization of a regularized Stein's unbiased risk estimate (SURE), which is a good estimate of the mean squared error. We linearly parametrize the deconvolution process by using multiple Wiener filters as elementary functions, followed by undecimated Haar-wavelet thresholding. Due to the quadratic nature of SURE and the linear parametrization, the deconvolution problem finally boils down to solving a linear system of equations, which is very fast and exact. The linear coefficients, i.e., the solution of the linear system of equations, constitute the best approximation of the optimal processing on the Wiener–Haar-threshold basis that we consider. In addition, the proposed multi-Wiener SURE-LET approach is applicable for both periodic and symmetric boundary conditions, and can thus be used in various practical scenarios. The very competitive (both in computation time and quality) results show that the proposed algorithm, which can be interpreted as a kind of nonlinear Wiener processing, can be used as a basic tool for building more sophisticated deconvolution algorithms.

Index Terms—Deconvolution, multi-Wiener filtering, Stein's unbiased risk estimate (SURE) minimization, undecimated Haar-wavelet thresholding.

I. INTRODUCTION

IMAGE DECONVOLUTION is a standard linear inverse problem with applications in medical imaging [2], seismology [3], astronomy [4], remote sensing [5], optical imaging [6], etc. The observed image is mathematically described as a blurring operator convolved with the original (unknown) image, further degraded by a zero-mean additive white Gaussian noise (AWGN) [7]. In particular, singular or ill-conditioned convolution operators yield highly noise-sensitive solutions, which makes the deconvolution problem particularly challenging [8]. We briefly review hereafter the most popular deconvolution approaches.

Regularization is a standard technique to cope with the ill-posed nature of the deconvolution problem. It formulates the problem as a convex constrained minimization problem [9], by imposing certain regularity conditions on the original image. The commonly used regularizers are:

1) Wiener filtering/Tikhonov regularization [10], [11]: This regularization leads to a closed-form linear solution due to the ℓ_2-norm nature of the regularizer. While straightforward to implement, this type of regularization often produces over-smoothed edges.

2) Transform-domain sparsity constraints [12]–[16]: Here, the coefficients of a frame-based representation of the original image are estimated under a sparsity-inducing regularizer such as the ℓ_1-norm [17]. Algorithmic implementations are frequently based on iterative shrinkage/thresholding (IST) [15], [16], [18], and [19].

3) Total-Variation (TV) regularization [20], [21]: TV regularization assumes that the ℓ_1-norm of the gradient of the original image is small. It is well-suited for piecewise smooth images, and remarkably effective at preserving edges [22]–[25].

A. Regularized Inversion Followed by Denoising

Except for a few works, e.g. [26], that merge both deblurring and denoising steps, deconvolution is usually decomposed as a two-step procedure that consists of regularized inverse followed by additive Gaussian noise reduction. Some representatives of this popular deconvolution approach are:

1) ForWaRD [27]: This two-stage shrinkage procedure successively operates in Fourier and wavelet domains with an optimal (in terms of an approximate MSE metric) balance between the amount of Fourier and wavelet regularization. To ensure shift-invariance, the results obtained after all possible shifts of an orthogonal discrete wavelet transform (DWT) are averaged. ForWaRD also considers different Fourier shrinkage parameters at different DWT scales.

2) More sophisticated denoising techniques: In the SV-GSM [28], SA-DCT [29] and BM3D [30] image restoration algorithms, the regularized inverse is followed by a more sophisticated denoising technique (BLS-GSM [31], Shape-Adaptive DCT [32] and BM3D [33], respectively). These deconvolution strategies achieve state-of-the-art results, usually at the expense of a high computational complexity.

B. Bayesian Approach

Another class of deconvolution algorithms [34]–[44] are derived in a Bayesian framework, where some *a priori* statistical knowledge of the original image are assumed. Some regularization approaches can be re-interpreted in this framework as maximum a posteriori (MAP) or penalized likelihood.

We refer the interested reader to [37], [45] for a more detailed description of the Bayesian approach.

Either regularization techniques or Bayesian methods need to find a reasonable value of the regularization parameter, to keep a good balance between data-fidelity and regularization terms [46]–[49].

C. SURE-Based Approaches

A statistical, yet non-Bayesian, approach to image denoising based on Stein's unbiased risk estimate (SURE) [50] has been recently revitalized [51], [52]. SURE is an unbiased estimate of the MSE under additive Gaussian noise assumption. Although the first applications of SURE were restricted to denoising only, it is possible to extend it to any linear restoration problem [53]–[56]. Since SURE depends on the observed data only, it can be practically used for the following purposes:

1) To optimize the parameters of the processing, e.g., the Tikhonov regularization parameter [55] or the optimal parameters involved in the non-local means (NLM) denoising technique [57];
2) To monitor the PSNR improvement during the iterations of IST algorithms, without referring to the original unknown data, e.g., automatic determination of the number of iterations [53], optimal update of the parameters at each iteration [56];
3) As a minimization criterion for designing denoising algorithms expressed as a linear expansion of thresholds (LET), the so-called SURE-LET approach [52], [58]. It requires the basic structure or parametric form of the estimator to be determined in advance [52], [54], and [55].

The present paper follows the last point. Similar to the work of [54], we further extend the SURE-LET approach to regularized inversion followed by denoising. Note that [54] empirically chooses the regularization parameter of the regularized inverse filter. The main contribution of our work, also the major difference from [54], is that besides applying a linear combination of non-linear thresholding functions for the denoising step as in [54], we also linearly parametrize the regularized inversion into a number of basic Wiener filters with different (but fixed) regularization parameters. Instead of optimizing or empirically choosing non-linear regularization parameters [47], [48], [54], and [55], the proposed multi-Wiener SURE-LET approach finds the best linear combination of multiple Wiener deconvolutions. The linear parametrization reduces the deconvolution problem to solving a linear system of equations. Its solution, i.e., the weights in the linear combination, automatically constitutes the minimum MSE/SURE combination as final estimate. As a result, the proposed deconvolution algorithm is fast and of low computational complexity[1].

The paper is organized as follows. Section II introduces the theoretical background of the SURE-LET approach in the general linear model. In Section III, we exemplify the SURE-LET approach by describing the deconvolution process as the linear combination of elementary functions consisting of Wiener filters followed by transform-domain thresholding, with the linear coefficients automatically obtained by minimizing SURE. Section IV discusses the computational aspects of the SURE-LET approach for periodic and symmetric boundary conditions. Section V reports and discusses the experimental results. Some concluding remarks are finally given in Section VI.

II. THEORETICAL BACKGROUND

This section presents the SURE-LET approach in the general linear framework—i.e., not limited to convolution operators. 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 $\mathbf{H}$. The identity matrix is denoted as $\mathbf{I}$.

A. Problem Statement

Consider the linear model

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{b}, \tag{1}$$

where $\mathbf{y} \in \mathbb{R}^N$ is a distorted observation of the original (unknown) data $\mathbf{x} \in \mathbb{R}^N$. The matrix $\mathbf{H} \in \mathbb{R}^{N \times N}$ implements a linear distortion[2], while the vector $\mathbf{b} \in \mathbb{R}^N$ is a zero-mean additive Gaussian noise corruption with covariance matrix $\mathbf{C} \succ 0$. Our objective is to find a function (or operation) $\mathbf{f} : \mathbb{R}^N \rightarrow \mathbb{R}^N$ of the measured data $\mathbf{y}$ such that the MSE

$$\text{MSE} = \frac{1}{N} \mathscr{E}\{\|\mathbf{f}(\mathbf{y}) - \mathbf{x}\|^2\} \tag{2}$$

is minimized. Here, $\mathscr{E}\{\cdot\}$ denotes the mathematical expectation operator. We would like to insist that the estimate $\hat{\mathbf{x}} = \mathbf{f}(\mathbf{y})$ is only the outcome of the processing. A key feature of our approach is to estimate the function $\mathbf{f}$ that transforms $\mathbf{y}$ into $\hat{\mathbf{x}}$, rather than the estimate $\hat{\mathbf{x}}$ itself.

B. Stein's Unbiased Risk Estimate (SURE)

Notice that we cannot directly minimize the MSE given by (2) to obtain the estimate $\hat{\mathbf{x}}$, as we have no access to the original data $\mathbf{x}$. However, based on the linear model (1) and the additive Gaussian noise assumption, the MSE can be replaced by a statistical estimate, SURE, involving the measurements $\mathbf{y}$ only. This is summarized in the following theorem [55], [59] (see the proof in Appendix A).

Theorem 2.1: Given the linear model (1) with any invertible matrix $\mathbf{H} \in \mathbb{R}^{N \times N}$, the following random variable

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

---

[1]A demo software is available at www.ee.cuhk.edu.hk/~tblu/demos.

[2]The theory that follows can be applied to rectangular matrices, but for the sake of simplicity, we restrict ourselves to square matrices.
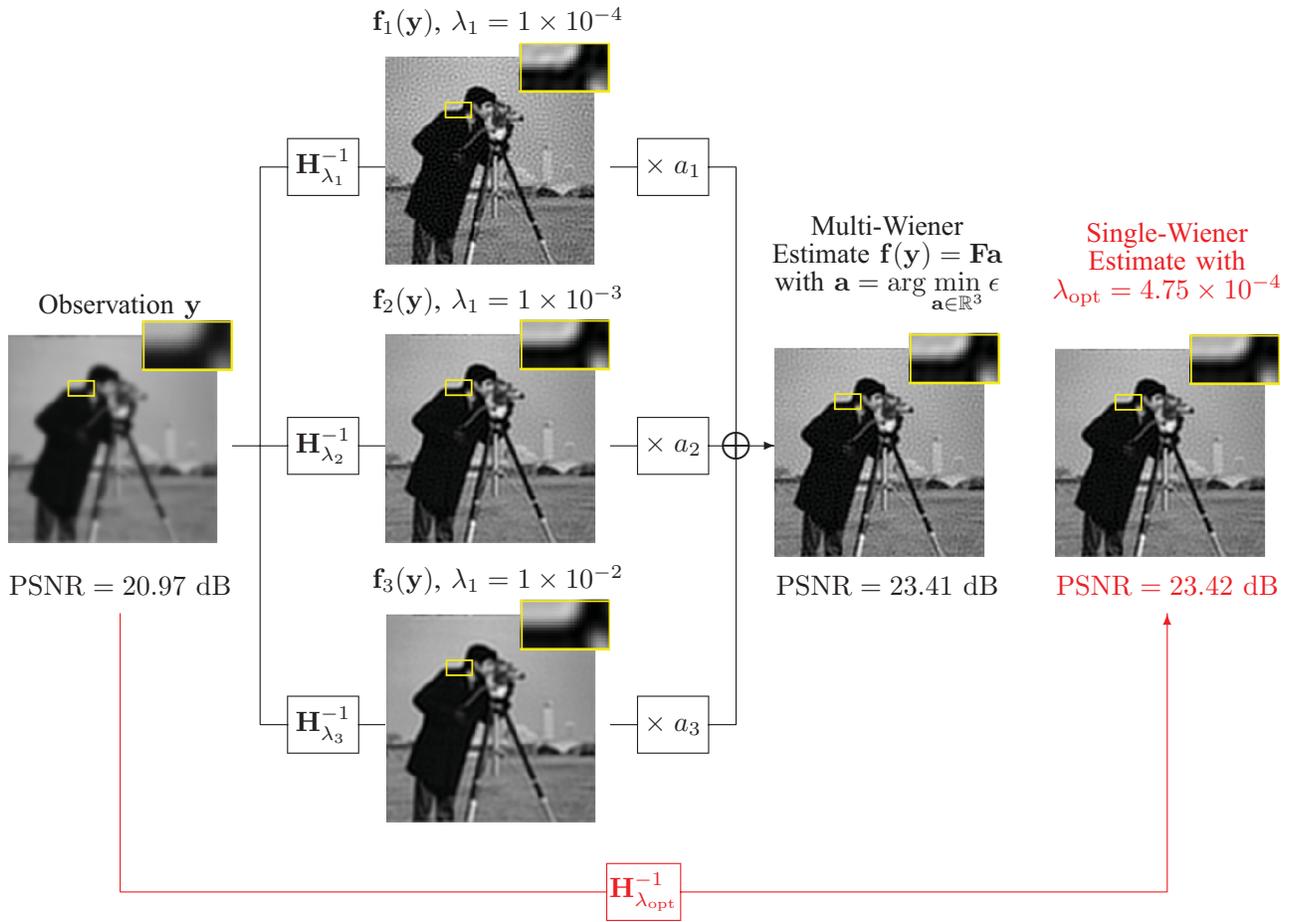
Fig. 1.   Example of linearly combining three Wiener filters with regularization parameters λ_1, λ_2, and λ_3, balanced by their weights $a_1 = 0.43$, $a_2 = 0.61$, and $a_3 = -0.04$ given by (8). The linearly-combined estimate is equivalent to a single optimal Wiener filtering in terms of PSNR.

is an unbiased estimator of the MSE defined in (2); i.e.,

$$\mathscr{E}\{\epsilon\} = \frac{1}{N}\mathscr{E}\left\{\|\mathbf{f}(\mathbf{y}) - \mathbf{x}\|^2\right\}.$$

In particular, if the Gaussian noise $\mathbf{b}$ is independent and identically distributed (i.i.d.) with variance σ^2, SURE becomes

$$\epsilon = \frac{1}{N}\left(\|\mathbf{f}(\mathbf{y})\|^2 - 2\mathbf{y}^{\mathrm{T}}\mathbf{H}^{-\mathrm{T}}\mathbf{f}(\mathbf{y}) + 2\sigma^2\mathrm{div}_{\mathbf{y}}\left\{\mathbf{H}^{-\mathrm{T}}\mathbf{f}(\mathbf{y})\right\}\right)$$
$$+ \frac{1}{N}\|\mathbf{x}\|^2, \tag{3}$$

which results from specifying the covariance matrix $\mathbf{C} = \sigma^2\mathbf{I}$.

Corollary 2.1: In the particular case of an i.i.d. Gaussian noise with variance σ^2, if the processing $\mathbf{f}(\cdot)$ can be expressed as $\mathbf{f}(\mathbf{y}) = \mathbf{H}^{\mathrm{T}}\tilde{\mathbf{f}}(\mathbf{y})$, then SURE becomes

$$\epsilon = \frac{1}{N}\left(\left\|\mathbf{H}^{\mathrm{T}}\tilde{\mathbf{f}}(\mathbf{y})\right\|^2 - 2\mathbf{y}^{\mathrm{T}}\tilde{\mathbf{f}}(\mathbf{y}) + 2\sigma^2\mathrm{div}_{\mathbf{y}}\left\{\tilde{\mathbf{f}}(\mathbf{y})\right\}\right) + \frac{1}{N}\|\mathbf{x}\|^2. \tag{4}$$

C. Regularized Approximation of SURE

If the matrix $\mathbf{H}$ is ill-conditioned or singular, ϵ in (3) fails to be a reliable estimate of the MSE. To keep the stability of ϵ, we may intuitively replace the unstable inverse $\mathbf{H}^{-1}$ by a

Tikhonov regularized inverse [8]

$$\mathbf{H}_\beta^{-1} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{H} + \beta\mathbf{S}^{\mathrm{T}}\mathbf{S}\right)^{-1}\mathbf{H}^{\mathrm{T}}, \tag{5}$$

for some parameter $\beta > 0$ and matrix $\mathbf{S} \in \mathbb{R}^N \times \mathbb{R}^N$. We can then approximate SURE as

$$\epsilon_\beta = \frac{1}{N}\left(\|\mathbf{f}(\mathbf{y})\|^2 - 2\mathbf{y}^{\mathrm{T}}\mathbf{H}_\beta^{-\mathrm{T}}\mathbf{f}(\mathbf{y}) + 2\sigma^2\mathrm{div}\left\{\mathbf{H}_\beta^{-\mathrm{T}}\mathbf{f}(\mathbf{y})\right\}\right)$$
$$+ \frac{1}{N}\left\|\mathbf{H}_\beta^{-1}\mathbf{H}\mathbf{x}\right\|^2. \tag{6}$$

This approximated estimator of the MSE turns out to be an unbiased estimate of $\frac{1}{N}\left\|\mathbf{f}(\mathbf{y}) - \mathbf{H}_\beta^{-1}\mathbf{H}\mathbf{x}\right\|^2$ (see the proof in Appendix B). Hence, if $\mathbf{H}_\beta^{-1}\mathbf{H}\mathbf{x}$ is close enough to $\mathbf{x}$ (hypothesis on $\mathbf{x}$), then, ϵ_β, the approximated SURE, is a good estimate of the MSE.

In imaging application, a good choice for $\mathbf{S}$ is a high-pass operator. A reasonable value for β should achieve a good balance between the approximation accuracy and the stability of ϵ_β; i.e., β should be neither too large, to avoid significant loss of high-frequency features, nor too small, to keep the stability of ϵ_β. In this paper, we choose $\mathbf{S}$ as a Laplacian operator and $\beta = 10^{-5}\sigma^2$, as extensive tests on various natural images have shown any value of $\beta \in [5 \times 10^{-6}\sigma^2, 5 \times 10^{-5}\sigma^2]$ to be consistently close to the minimum MSE one. We also define

a regularized inverse for $\mathbf{S}$ as $\mathbf{S}_\beta^{-1} = \beta \left(\mathbf{H}^T\mathbf{H} + \beta\mathbf{S}^T\mathbf{S}\right)^{-1} \mathbf{S}^T$, which leads to the following decomposition of the identity matrix: $\mathbf{H}_\beta^{-1}\mathbf{H} + \mathbf{S}_\beta^{-1}\mathbf{S} = \mathbf{I}$.

D. SURE-LET Approach

In practice, we will minimize the approximated SURE (6) instead of the actual MSE (2). The next question that naturally arises is: how to choose the function $\mathbf{f}$, such that $\mathbf{f}(\mathbf{y})$ is sufficiently close to $\mathbf{x}$? Here, we adopt the LET method [51], [52], and [58], which parametrizes $\mathbf{f}(\mathbf{y})$ as a linear combination of a small number of pre-defined basic functions (or processing) $\mathbf{f}_k$ for $k = 1, 2, \ldots, K$; i.e.,

$$\mathbf{f}(\mathbf{y}) = \sum_{k=1}^{K} a_k \mathbf{f}_k(\mathbf{y}) = \underbrace{[\mathbf{f}_1(\mathbf{y})\, \mathbf{f}_2(\mathbf{y})\, \ldots\, \mathbf{f}_K(\mathbf{y})]}_{\mathbf{F}(\mathbf{y})} \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_K \end{bmatrix}}_{\mathbf{a}} = \mathbf{Fa},$$

(7)

where $K \ll N$ is the number of linear coefficients a_k. Here, $\mathbf{f}(\mathbf{y}) = \mathbf{Fa}$ is a shorthand matrix notation to outline the linearity of the representation. The deconvolution problem then amounts to find the linear coefficients a_k that minimize ϵ_β. This parametrization dramatically reduces the size of the estimation problem from N number of pixels to K number of basis functions. Note that the LET decomposition (7) does not imply any hypothesis on $\mathbf{x}$ itself. Yet, a "bad" basis of processing $\mathbf{f}_k$ will obviously lead to poor deconvolution performances compared to a "good" basis (See however the note in Section V-C on the robustness w.r.t. thresholding functions).

Substituting (7) into (6), and performing the differentiation of ϵ_β with respect to a_k for all $k = 1, 2, \ldots, K$, we obtain the linear system of equations

$$\sum_{k'=1}^{K} \underbrace{\frac{1}{N}\mathbf{f}_k^T(\mathbf{y})\mathbf{f}_{k'}(\mathbf{y})}_{\mathbf{M}_{k,k'}} a_{k'} = \underbrace{\frac{1}{N}\left(\mathbf{y}^T\mathbf{H}_\beta^{-T}\mathbf{f}_k(\mathbf{y}) - \sigma^2\mathrm{div}\left\{\mathbf{H}_\beta^{-T}\mathbf{f}_k(\mathbf{y})\right\}\right)}_{c_k}$$

(8)

for $k = 1, 2, \ldots, K$. These equations can be summarized in matrix form as $\mathbf{Ma} = \mathbf{c}$, where $\mathbf{M} = [\mathbf{M}_{k,k'}]_{k,k'=1,2,\ldots,K}$ and $\mathbf{c} = [c_1\ c_2\ \cdots\ c_K]^T$. Due to the possible singularity of $\mathbf{M}$, we solve the following ℓ_2-regularized linear system of equations

$$(\mathbf{M} + \mu\mathbf{I})\mathbf{a} = \mathbf{c} \quad \Leftrightarrow \quad \mathbf{a} = (\mathbf{M} + \mu\mathbf{I})^{-1}\mathbf{c} \tag{9}$$

instead of (8), where $\mu > 0$ is a regularization parameter. We empirically found that any value of $\mu \in [0.01, 0.1]$ achieves very similar deconvolution performance. In this paper, we empirically set $\mu = 5 \times 10^{-2}$.

It is also worth noting that the corresponding MSE minimization leads to solving $\mathbf{Ma}_{\mathrm{MSE}} = \mathbf{F}^T\mathbf{x}$, with the solution, namely *MSE-LET*, serving as a counterpart to SURE-LET. The accuracy of using the SURE-LET estimate can then be verified by comparing it with *MSE-LET*.
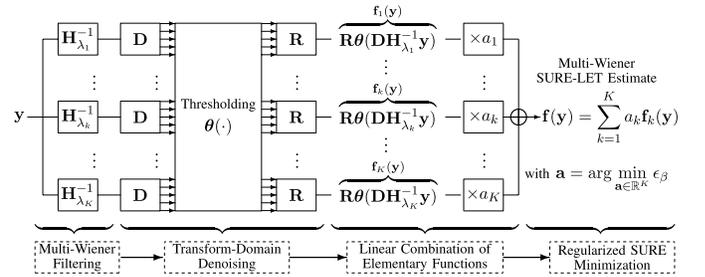

Fig. 2. Flowchart of the proposed multi-Wiener SURE-LET approach.

III. CONSTRUCTION OF ELEMENTARY FUNCTION $\mathbf{f}_k$

This section shows how to choose the elementary processing $\mathbf{f}_k$ in the context of image deconvolution.

A. Linear Deconvolution: Multi-Wiener Filtering

We first choose each $\mathbf{f}_k(\mathbf{y})$ to be a Wiener filter with a given regularization parameter λ_k:

$$\mathbf{f}_k(\mathbf{y}) = \underbrace{\left(\mathbf{H}^T\mathbf{H} + \lambda_k\mathbf{S}^T\mathbf{S}\right)^{-1}\mathbf{H}^T}_{\mathbf{H}_{\lambda_k}^{-1}}\mathbf{y}, \text{ for } k = 1, 2, \ldots, K. \tag{10}$$

In the standard case of a shift-invariant convolution operator, the matrix products with $\mathbf{H}$ or $\mathbf{H}^T$ are commutative. Each elementary processing $\mathbf{f}_k(\mathbf{y})$ can then be written as $\mathbf{f}_k(\mathbf{y}) = \mathbf{H}^T\tilde{\mathbf{f}}(\mathbf{y})$ where $\tilde{\mathbf{f}}(\mathbf{y}) = \left(\mathbf{H}^T\mathbf{H} + \lambda_k\mathbf{S}^T\mathbf{S}\right)^{-1}\mathbf{y}$ and Corollary 2.1 applies. Consequently, we do not need to use the regularized SURE ϵ_β defined in (6). The SURE-LET method (7) consists in finding the minimum SURE/MSE combination of these Wiener filters, which is achieved by solving (8). The flowchart of the SURE-optimized multi-Wiener deconvolution is shown in Fig. 1. We observe that different values of λ_k capture different details and features of the image, and the optimal linear coefficients a_k produce the combined estimate with the best balance between noise reduction and edge preservation. The key advantage of the multi-Wiener SURE optimization is that it avoids the empirical adjustment of a unique non-linear regularization parameter λ, contrary to [47], [48], and [55].

B. Non-Linear Deconvolution: Multi-Wiener Filtering Followed by Transform-Domain Thresholding

We now present the construction of elementary functions $\mathbf{f}_k(\mathbf{y})$ for the non-linear deconvolution approach which consists of multi-Wiener filtering followed by transform-domain thresholding[3]. Both steps are linearly parametrized using the LET strategy (7).

An illustrative description of the proposed non-linear deconvolution approach is shown in Fig. 2 and 3. In Fig. 2, the matrices $\mathbf{D}$ and $\mathbf{R}$ represent a pair of linear decomposition/reconstruction that satisfies the perfect reconstruction condition $\mathbf{RD} = \mathbf{I}$. Typically, $\mathbf{D}$ and $\mathbf{R}$ implement a $(J + 1)$-band filterbank of decimated or undecimated filters. They are structured as $\mathbf{D} = \left[\mathbf{D}_1^T\, \mathbf{D}_2^T\, \cdots\, \mathbf{D}_J^T\, \mathbf{D}_{J+1}^T\right]^T \in \mathbb{R}^{Q \times N}$ and

---

[3]By thresholding, we denote any non-linear operation (smooth or non-smooth).
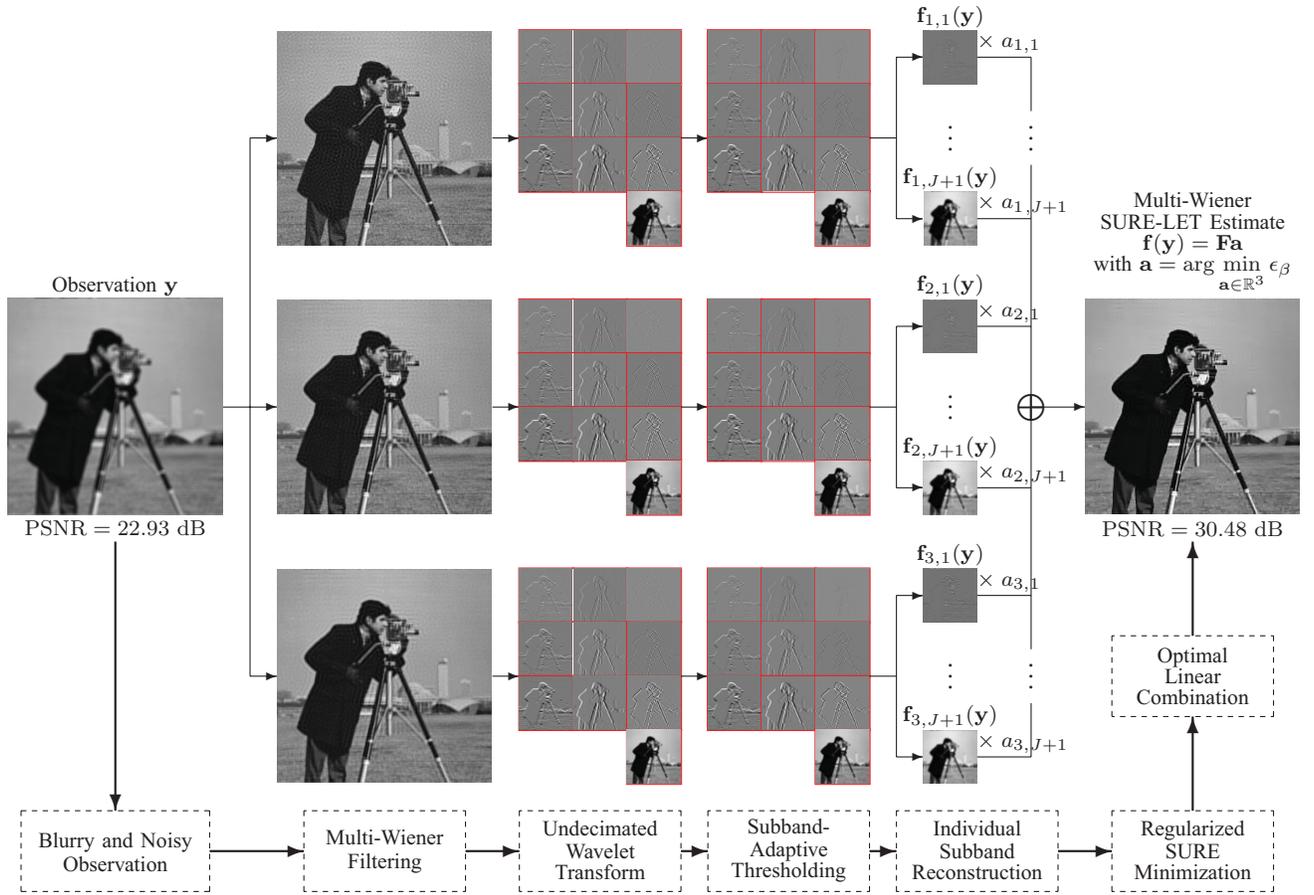
Fig. 3. Illustration of the proposed multi-Wiener SURE-LET approach. Note that: (1) The undecimated wavelet subbands and their processed reconstructions are displayed in reduced size for convenience. (2) The reconstruction $\mathbf{R}_j$ is performed to the specific j-th subband only, by setting all the other subbands to zero [52]. (3) The thresholding function $\boldsymbol{\theta}(\cdot)$ can also be linearly parametrized as $\boldsymbol{\theta}(\cdot) = \sum_{l=1}^{L} a_l \boldsymbol{\theta}_l(\cdot)$.

$\mathbf{R} = \begin{bmatrix} \mathbf{R}_1 \ \mathbf{R}_2 \ \cdots \ \mathbf{R}_J \ \mathbf{R}_{J+1} \end{bmatrix} \in \mathbb{R}^{N \times Q}$, where $\mathbf{D}_j \in \mathbb{R}^{N_j \times N}$ and $\mathbf{R}_j \in \mathbb{R}^{N \times N_j}$ for $j = 1, 2, \ldots, J, J+1$. We consider $(J+1)$ bands for the convenience of the future discussion. The dimensions Q and N_j depend on the transform being performed. Mathematically, the flowchart of Fig. 2 can be described by the following function

$$\mathbf{f}(\mathbf{y}) = \mathbf{R}\boldsymbol{\theta}\left(\mathbf{D}\mathbf{H}_\lambda^{-1}\mathbf{y}\right). \tag{11}$$

Corollary 3.1: Given the processing $\mathbf{f}(\cdot)$ defined by (11) and considering a pointwise function $\boldsymbol{\theta}(\cdot)$, the regularized approximation of SURE ϵ_β introduced in (6) can be further derived as

$$\epsilon_\beta = \frac{1}{N}\left(\|\mathbf{f}(\mathbf{y})\|^2 - 2\mathbf{y}^{\mathrm{T}}\mathbf{H}_\beta^{-\mathrm{T}}\mathbf{f}(\mathbf{y}) + 2\sigma^2\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\theta}'(\mathbf{w})\right)$$
$$+ \frac{1}{N}\|\mathbf{H}_\beta^{-1}\mathbf{H}\mathbf{x}\|^2. \tag{12}$$

In (12), the vector $\mathbf{w} = \mathbf{D}\mathbf{H}_\lambda^{-1}\mathbf{y} = \begin{bmatrix} \mathbf{w}_1^{\mathrm{T}} \ \mathbf{w}_2^{\mathrm{T}} \ \cdots \ \mathbf{w}_J^{\mathrm{T}} \ \mathbf{w}_{J+1}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^Q$, with $\mathbf{w}_j = \mathbf{D}_j\mathbf{H}_\lambda^{-1}\mathbf{y} \in \mathbb{R}^{N_j}$ for $j = 1, 2, \ldots, J+1$, denotes the transform coefficients. The vector $\boldsymbol{\theta}'(\mathbf{w}) = \begin{bmatrix} \theta_n'(w_n) \end{bmatrix}_{n \in [1; Q]}$ represents the first derivative of the pointwise function $\boldsymbol{\theta}$. The vector $\boldsymbol{\alpha} \in \mathbb{R}^Q$ is defined by

$$\boldsymbol{\alpha} = \mathrm{diag}\left\{\mathbf{D}\mathbf{H}_\lambda^{-1}\mathbf{H}_\beta^{-\mathrm{T}}\mathbf{R}\right\} = \begin{bmatrix} \mathbf{P}_{n,n} \end{bmatrix}_{n \in [1; Q]}$$

where $\mathbf{P} = \mathbf{D}\mathbf{H}_\lambda^{-1}\mathbf{H}_\beta^{-\mathrm{T}}\mathbf{R} \in \mathbb{R}^{Q \times Q}$. Consequently, the vector $\boldsymbol{\alpha}$ is structured as $\boldsymbol{\alpha} = \begin{bmatrix} \boldsymbol{\alpha}_1^{\mathrm{T}} \ \boldsymbol{\alpha}_2^{\mathrm{T}} \ \cdots \ \boldsymbol{\alpha}_J^{\mathrm{T}} \ \boldsymbol{\alpha}_{J+1}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$, where

$$\boldsymbol{\alpha}_j = \mathrm{diag}\left\{\mathbf{D}_j\mathbf{H}_\lambda^{-1}\mathbf{H}_\beta^{-\mathrm{T}}\mathbf{R}_j\right\} \in \mathbb{R}^{N_j} \tag{13}$$

for $j = 1, 2, \ldots, J+1$.

The proof of Corollary 3.1 is very similar to that of Corollary 1 in [52]. We thus omit it here.

When $\mathbf{D}$ and $\mathbf{R}$ implement an orthogonal transform (e.g., an orthonormal wavelet transform), the transform-domain minimization of ϵ_β is equivalent to its image-domain minimization, thanks to the MSE preservation property of orthogonal transforms. If the applied transform is not orthogonal (e.g., an undecimated wavelet transform), the minimization of ϵ_β has to be performed in the image domain to ensure a global minimum MSE optimality. In the remainder of this paper, we confine ourselves to undecimated wavelet transforms, due to their superior denoising abilities. Consequently, the minimization of the regularized approximation of SURE ϵ_β is performed in the image domain.

Suppose that the decomposition $\mathbf{D}$ produces J high-pass subbands and one lowpass subband (indexed as the $(J+1)$-th subband) that is not thresholded, then $\mathbf{f}(\mathbf{y})$ of (11)

can be linearly parametrized as

$$\mathbf{f}(\mathbf{y}) = \underbrace{\sum_{m=1}^{M} \sum_{l=1}^{L} \sum_{j=1}^{J} a_{m,l,j} \mathbf{R}_j \boldsymbol{\theta}_l \left(\mathbf{w}_{m,j}\right)}_{\text{highpass subbands}}$$

$$+ \underbrace{\sum_{m=1}^{M} a_{m,J+1} \mathbf{R}_{J+1} \mathbf{w}_{m,J+1}}_{\text{lowpass subband}}, \quad (14)$$

where M denotes the number of Wiener filters (typically $M = 3$, which is validated by Table I in Section V-B), L the number of elementary pointwise thresholding functions (typically $L = 2$), and J the number of highpass wavelet subbands (typically $J = 9$ for three decomposition levels).

As (14) shows, $K = M \cdot J \cdot L + M$ (typically, $K = 3 \cdot 9 \cdot 2 + 3 = 57$) weights $a_{m,l,j}$ and L thresholding function need to be determined. The weights $a_{m,l,j}$ are obtained by minimizing the regularized approximation of SURE ϵ_β, which boils down to solving a linear system of K equations (9). To be efficient, the pointwise thresholding functions $\theta_l(\cdot)$ have to satisfy some desirable properties that are discussed in [51], [52]. We empirically found that a linear combination of the following two functions (see Fig. 4)

$$\begin{cases} \theta_1(w) = w\left(1 - \exp\left(-\left(\frac{w}{T_1}\right)^4\right)\right) \\ \theta_2(w) = w\left(1 - \exp\left(-\left(\frac{w}{T_2}\right)^4\right)\right) \end{cases} \quad (15)$$

yields satisfactory results. Other thresholding functions, such as the ones reported in [52], [54], do not result in significant PSNR loss (maximum 0.2 dB loss). The thresholds T_1 and T_2 are subband-adaptive. We empirically found that the proposed algorithm is able to keep similar deconvolution quality for a wide range of value T_1 and T_2. In this paper, we choose $T_1 = 4\sigma_j$ and $T_2 = 9\sigma_j$ for the j-th wavelet subband, where σ_j^2 is the variance of the colored noise $\mathbf{b}_j = \mathbf{D}_j \mathbf{H}_\lambda^{-1} \mathbf{b} \in \mathbb{R}^{N_j}$; i.e.,

$$\sigma_j^2 = \frac{1}{N_j} \mathscr{E}\left\{\mathbf{b}_j^{\mathrm{T}} \mathbf{b}_j\right\}$$

$$= \frac{1}{N_j} \mathscr{E}\left\{\mathbf{b}^{\mathrm{T}} \mathbf{H}_\lambda^{-\mathrm{T}} \mathbf{D}_j^{\mathrm{T}} \mathbf{D}_j \mathbf{H}_\lambda^{-1} \mathbf{b}\right\}$$

$$= \frac{\sigma^2}{N_j} \mathrm{Tr}\left(\mathbf{H}_\lambda^{-\mathrm{T}} \mathbf{D}_j^{\mathrm{T}} \mathbf{D}_j \mathbf{H}_\lambda^{-1}\right). \quad (16)$$

The variance σ_j^2 in (16) and the vector $\boldsymbol{\alpha}_j$ in (13) can be analytically computed, as shown in the next section.

IV. COMPUTATIONAL ASPECTS

The proposed method requires to compute several Wiener filters (10), the subband-dependent noise variances σ_j^2 (16) and the subband-dependent vectors $\boldsymbol{\alpha}_j$ (13), listed here below:

1) Wiener filter: $\mathbf{H}_\lambda^{-1} = \left(\mathbf{H}^{\mathrm{T}}\mathbf{H} + \lambda \mathbf{S}^{\mathrm{T}}\mathbf{S}\right)^{-1} \mathbf{H}^{\mathrm{T}}$;
2) Variance σ_j^2 of the colored noise $\mathbf{b}_j = \mathbf{D}_j \mathbf{H}_\lambda^{-1}\mathbf{b}$;
3) Vector $\boldsymbol{\alpha}_j = \mathrm{diag}\left\{\mathbf{D}_j \mathbf{H}_\lambda^{-1}\mathbf{H}_\beta^{-\mathrm{T}}\mathbf{R}_j\right\}$.
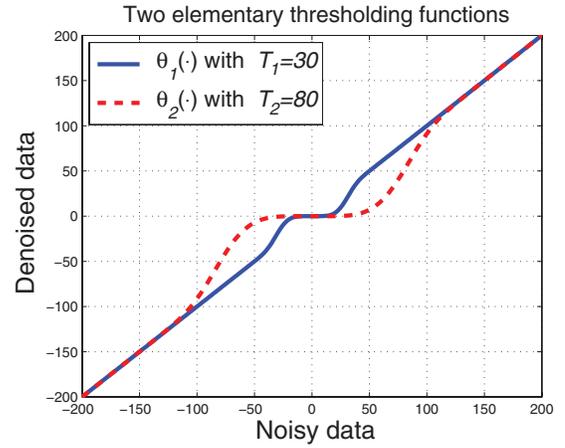

Fig. 4. Shape of our two basic thresholding functions $\theta_1(\cdot)$ and $\theta_2(\cdot)$, given by (15).

In this section, we show how to perform these computation analytically in the Fourier domain. For the sake of brevity of the discussion and formulation, we stick to the one-dimensional case. The 2-D computations can then be straightforwardly deduced. We define the coefficients of a 1-D filter as $g(n)$, the coefficients of the input of the 1-D filter as $x(n)$ and the coefficients of the output of the 1-D filter by $y(n)$, for $n = 0, 1, \ldots, N - 1$. All these coefficients are assumed to be zero outside the domain $[0, N - 1]$. Note that the implementation of the proposed approach needs the specification of boundary conditions. The most commonly used are periodic extensions. Yet, periodic extensions may generate unwanted artifacts (discontinuity at the boundaries), so symmetric boundary extensions are often preferred in practice [25], [60], and [61]. In the next two sections, we discuss the computations of the above mentioned terms for these two particular choices of boundary conditions.

A. Periodic Boundary Extensions

1) Circular Convolution: Under periodic boundary conditions, the convolution is circular and reads as

$$y(n) = \sum_{k \in \mathbb{Z}} g(n - k)x(k) = \sum_{k=0}^{N-1} g_N(n - k)x(k)$$

where

$$g_N(n) = \sum_{n_0 \in \mathbb{Z}} g(n + n_0 N). \quad (17)$$

Consequently, the matrices $\mathbf{H}_\lambda^{-1}$, $\mathbf{D}_j$, $\mathbf{R}_j$, and $\mathbf{H}_\beta^{-\mathrm{T}}$ are all circulant. They can thus be efficiently computed by discrete Fourier transform (DFT) with period N [62], where the Fourier coefficient of $g_N(n)$ is given as

$$G\left(e^{j\frac{2\pi k}{N}}\right) = \sum_{n=0}^{N-1} g_N(n)e^{-j\frac{2\pi k}{N}n}. \quad (18)$$

2) Analytic Computation of σ_j^2 Given in (16):
Theorem 4.1: Under periodic boundary conditions, given a zero-mean white Gaussian noise $b(n)$ with variance σ^2 and

a filter $g(n)$, the variance of $p(n) = \sum_{k \in \mathbb{Z}} g(n-k)b(k)$ is given by

$$\mathscr{E}\left\{p(n)^2\right\} = \sigma^2 \sum_{n=0}^{N-1} g_N(n)^2 = \frac{\sigma^2}{N} \sum_{k=0}^{N-1} \left|G(e^{j\frac{2\pi k}{N}})\right|^2$$

for $n = 0, 1, \ldots, N-1$, where $g_N(n)$ and $G(e^{j\frac{2\pi k}{N}})$ are given by (17) and (18).

Proof: The second equality comes from Parseval's theorem [62]. The i.i.d. condition of the zero-mean white Gaussian noise $b(n)$ implies that

$$\mathscr{E}\left\{b(n)b(n')\right\} = \sigma^2 \delta(n - n') = \begin{cases} \sigma^2, & \text{for } n = n' \\ 0, & \text{otherwise.} \end{cases} \quad (19)$$

Thus, we obtain that the variance of the filtered noise $p(n)$ is:

$$\begin{aligned} \mathscr{E}\left\{p(n)^2\right\} &= \sum_{k \in \mathbb{Z}} \sum_{k' \in \mathbb{Z}} g_N(n-k)g_N(n-k')\mathscr{E}\left\{b(k)b(k')\right\} \\ &\overset{(19)}{=} \sigma^2 \sum_{k \in \mathbb{Z}} g(n-k)^2 \\ &= \sigma^2 \sum_{n=0}^{N-1} g_N(n)^2. \end{aligned}$$

∎

Note that Theorem 4.1 is a straightforward application of the result obtained for filtered wide-sense stationary process (see for instance [63, Theorem 3.14]). It is restated and proved here for later comparison with that obtained under symmetric boundary conditions.

3) Analytical Computation of $\boldsymbol{\alpha}_j$ Given by (13):

Theorem 4.2: Under periodic boundary conditions, the vector formed by the diagonal elements of the convolution matrix $\mathbf{G} \in \mathbb{R}^{N \times N}$ is given by

$$\text{diag}\{\mathbf{G}\} = g_N(0) \cdot \mathbf{1} = \left(\frac{1}{N}\sum_{k=0}^{N-1} G(e^{j\frac{2\pi k}{N}})\right) \cdot \mathbf{1}$$

where $g_N(n)$ given by (17) is the filter implemented by $\mathbf{G}$, $G(e^{j\frac{2\pi k}{N}})$ is given by (18).

Proof: The second equality is essentially the inverse DFT. In matrix form, each element of the diagonal of $\mathbf{G}$ can be obtained as $\mathbf{e}_{k'}^{\text{T}}\mathbf{G}\mathbf{e}_{k'}$, where the indicator vector $\mathbf{e}_{k'}$ is defined as $\mathbf{e}_{k'} = [0 \cdots 0\, 1\, 0 \cdots 0]^{\text{T}} \in \mathbb{R}^N$ with k'-th element 1. In convolution form, it reads as

$$\mathbf{e}_{k'}^{\text{T}}\mathbf{G}\mathbf{e}_{k'} = \sum_{k=1}^{N} \delta(k-k')\left(\sum_{n=1}^{N} g_N(k-n)\delta(n-k')\right) = g_N(0)$$

where $\delta(k-k')$ is defined as in (19).

We conclude the proof by using the fact that, under periodic boundary conditions, the convolution matrix $\mathbf{G}$ is circulant and thus its diagonal elements are all equal. ∎

B. Symmetric Boundary Extensions

1) Symmetric Convolution: The term "symmetric convolution" was first introduced by S. Martucci in [64], to describe the convolution under symmetric boundary conditions. In [64], he discussed 64 possible types of symmetric convolutions,

depending on half-point or whole-point symmetry. The symmetric convolution can be performed by DCT or Types I–VIII DST, according to specific types of symmetry. Among all these, half-point symmetric boundary extensions corresponding to Type II DCT deserves a particular consideration [25]. The following computations are based on this typical case.

Lemma 4.1: Given a signal $x(n)$ and filter $g(n)$, under half-point symmetric boundary conditions, convolution is defined as

$$y(n) = \sum_{k=0}^{N-1} \left(g_{2N}(n-k) + g_{2N}(n+k+1)\right)x(k)$$

for $n = 0, 1, \ldots, N-1$, where

$$g_{2N}(n) = \sum_{n_0 \in \mathbb{Z}} g(n + 2n_0 N). \quad (20)$$

Proof: Under half-point symmetric extension, the signal $x(n)$ is periodic with period $2N$. Thus, the convolution become circular with period $2N$:

$$y(n) = \sum_{k \in \mathbb{Z}} g(n-k)x(k) = \sum_{k=0}^{2N-1} g_{2N}(n-k)x(k) \quad (21)$$

where $g_{2N}(n)$ is given as (20). Using the property of half-point symmetry that $x(n) = x(2N - 1 - n)$ and changing variables in (21) completes the proof. ∎

Since the half-point symmetric extension can be considered as the periodic condition with period $2N$, besides from DCT, the filtering $\mathbf{H}_\lambda^{-1}$, $\mathbf{D}_j$, $\mathbf{R}_j$, and $\mathbf{H}_\beta^{-\text{T}}$ can also be efficiently computed by DFT with period $2N$, by means of the Fourier coefficient of $g_{2N}(n)$ given as:

$$G(e^{j\frac{\pi k}{N}}) = \sum_{n=0}^{2N-1} g_{2N}(n)e^{-j\frac{\pi k}{N}n}. \quad (22)$$

Note that $G(e^{j\frac{\pi k}{N}})$ will be used for computing σ_j^2 and α_j.

2) Analytic Computation of σ_j^2 Given in (16):

Theorem 4.3: Under half-point symmetric boundary conditions, given a zero-mean white Gaussian noise $b(n)$ with variance σ^2 and a filter $g(n)$, the variance of $p(n) = \sum_{k \in \mathbb{Z}} g(n-k)b(k)$ is given by

$$\mathscr{E}\left\{p(n)^2\right\} = \sigma^2 \sum_{k=0}^{N-1} \left[g_{2N}(n-k) + g_{2N}(n+k+1)\right]^2$$

in spatial domain, where $g_{2N}(n)$ is given as (20), and is also expressed as:

$$\begin{aligned} \mathscr{E}\left\{p(n)^2\right\} = \frac{\sigma^2}{2N} \sum_{k=0}^{2N-1} \Big\{ &\left|G(e^{j\frac{\pi k}{N}})\right|^2 \\ &+ Re\big\{(G(e^{j\frac{\pi k}{N}}))^2 e^{j\frac{\pi k}{N}(2n+1)}\big\}\Big\} \end{aligned}$$

in Fourier domain, where $G(e^{j\frac{\pi k}{N}})$ is given as (22). See the proof in Appendix C.

TABLE I
PSNR RESULTS OF OUR MULTI-WIENER SURE-LET FOR VARIOUS NUMBER OF WIENER FILTERS

| Case		*House*, Separable Filter, $\sigma = 1$ (Input: 30.92 dB; BM3D: 35.80 dB)			*Cameraman*, Gaussian With std 3, $\sigma = 10$ (Input: 20.22 dB; BM3D: 22.60 dB)			*Coco*, Gaussian With std 3, $\sigma = 1$ (Input: 26.45 dB; BM3D: 31.27 dB)		
Number of Deconvolutions	Regularization Parameter	λ/σ^2	SURE-LET	*MSE-LET*	λ/σ^2	SURE-LET	*MSE-LET*	λ/σ^2	SURE-LET	*MSE-LET*
One	λ/σ^2	1×10^{-4}	34.33	*34.59*	1×10^{-4}	22.30	*22.47*	1×10^{-4}	31.02	*31.20*
Two	λ_1/σ^2 λ_2/σ^2	1×10^{-4} 1×10^{-2}	35.78	*36.04*	1×10^{-4} 1×10^{-2}	22.37	*22.65*	1×10^{-4} 1×10^{-2}	31.40	*31.68*
Three	λ_1/σ^2 λ_2/σ^2 λ_3/σ^2	1×10^{-4} 1×10^{-3} 1×10^{-2}	**36.31**	*36.47*	1×10^{-4} 1×10^{-3} 1×10^{-2}	**22.52**	*22.70*	1×10^{-4} 1×10^{-3} 1×10^{-2}	**31.70**	*32.00*
Four	λ_1/σ^2 λ_2/σ^2 λ_3/σ^2 λ_4/σ^2	1×10^{-5} 1×10^{-4} 1×10^{-3} 1×10^{-2}	36.27	*36.51*	1×10^{-5} 1×10^{-4} 1×10^{-3} 1×10^{-2}	22.54	*22.74*	1×10^{-5} 1×10^{-4} 1×10^{-3} 1×10^{-2}	**31.70**	*32.03*
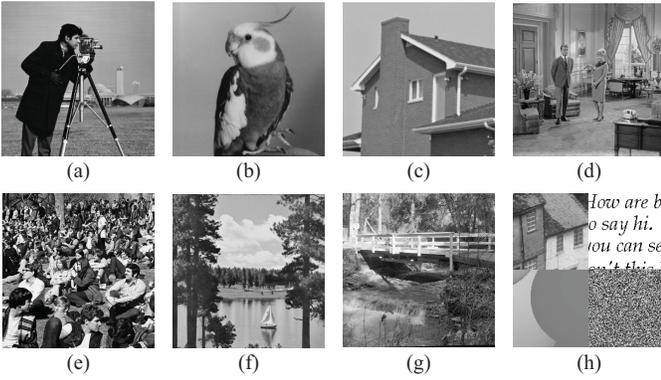
Fig. 5. Original images. (a) *Cameraman* 256×256. (b) *Coco* 256×256. (c) *House* 256×256. (d) *Couple* 512×512. (e) *Crowd* 512×512. (f) *Lake* 512×512. (g) *Bridge* 512×512. (h) *Mixture* 512×512.

3) Analytical Computation of $\boldsymbol{\alpha}_j$ Given in (13)*:*

Theorem 4.4: Under half-point symmetric boundary conditions, the vector formed by the diagonal elements of the convolution matrix $\mathbf{G} \in \mathbb{R}^{N \times N}$ is given by

$$\big[\text{diag}\,\{\mathbf{G}\}\big]_n = g_{2N}(0) + g_{2N}(2n + 1)$$

for $n = 0, 1, \ldots, N - 1$, where $g_{2N}(n)$ is given by (20). It can also be expressed in the Fourier domain as

$$\big[\text{diag}\,\{\mathbf{G}\}\big]_n = \frac{1}{2N} \sum_{k=0}^{2N-1} G\big(e^{j\frac{\pi k}{N}}\big)\big(1 + e^{j\frac{\pi k}{N}(2n+1)}\big)$$

where $G(e^{j\frac{\pi k}{N}})$ is given as (22).

Since the proof is similar to that of Theorem 4.2, it is omitted here. Contrary to the periodic boundary conditions, the diagonal elements of $\mathbf{G}$ are not equal, as $\mathbf{G}$ is not a circulant matrix under symmetric boundary conditions. However, it is not in contradiction to the fact that if we consider the problem as periodic extension with period $2N$, the matrix $\mathbf{G}^{2N \times 2N}$ is still circulant.

V. RESULTS AND DISCUSSION

A. Experimental Setting

We consider the following benchmark convolution kernels commonly used in [12], [25], [30], and [65]:

1) Rational filter: $h(i, j) = (1 + i^2 + j^2)^{-1}$ for $i, j = -7, \ldots, 7$;
2) Separable filter: 5×5 filter with weights $[1, 4, 6, 4, 1]/16$ along both horizontal and vertical directions;
3) 5×5 uniform blur;
4) 9×9 uniform blur;
5) Gaussian blur with standard deviation (std) 3.

The blurred images are subsequently contaminated by i.i.d Gaussian noise with various variance σ^2. The test dataset contains eight 8-bit images of size 256×256 or 512×512 displayed in Fig. 5[4], covering a wide range of natural images. The experimental performance is measured by the peak signal-to-noise ratio (PSNR), defined as $\text{PSNR} = 10 \log_{10}(255^2/(\|\hat{\mathbf{x}} - \mathbf{x}\|^2/N))$. Note that all the PSNR results (in dB) reported in this section have been averaged over 10 noise realizations.

B. Influence of the Number of Wiener Filters

Our multi-Wiener SURE-LET approach involves several Wiener filters as elementary processing in the linear combination. In this section, we evaluate the influence of the retained number of Wiener filters on the deconvolution performance. Based on the observation that a reasonable value of the regularization parameter λ should be proportional to the noise variance σ^2 [24], [30], we focus on the ratio λ/σ^2. Table I shows the results obtained when considering one to four Wiener filters for various images, blur kernels and noise levels. We observe that one or two Wiener filters may not always produce the best performance, whereas three or four are able to cope with various scenarios (different images and different blur kernels). Moreover, using four Wiener filterings does not bring any significant performance improvement, compared to using three. Besides, we also found that for one Wiener filter, the optimal value of λ/σ^2 typically varies in the range 2×10^{-4} to 2×10^{-3}. For these reasons, in the following sections, we use three Wiener filters (i.e. $M = 3$) with $\lambda_1 = 1 \times 10^{-4}\sigma^2$, $\lambda_2 = 1 \times 10^{-3}\sigma^2$ and $\lambda_3 = 1 \times 10^{-2}\sigma^2$. Extensive tests show that the deconvolution performance is largely insensitive to the actual choice of λ's.

[4]All 512×512 images are available at http://decsai.ugr.es/cvg/CG/base.htm

TABLE II
PARAMETERS SETTING OF THE PROPOSED MULTI-WIENER
SURE-LET ALGORITHM

$\mathbf{S}$ in (5) and (10)	Discrete Laplacian Operator
β in (5)	$\beta = 1 \times 10^{-5}\sigma^2$
λ_k in (10)	$\lambda_1 = 10^{-4}\sigma^2;\ \lambda_2 = 10^{-3}\sigma^2;\ \lambda_3 = 10^{-2}\sigma^2$
T_l in (15)	$T_1 = 4\sigma_j;\ T_2 = 9\sigma_j$
Transforms $\mathbf{D}$ and $\mathbf{R}$	Undecimated Haar wavelet transform
μ in (9)	$\mu = 5 \times 10^{-2}$

Fig. 6. Restoration of *House* degraded by Gaussian blur of variance 9 with noise std $\sigma = 1$.
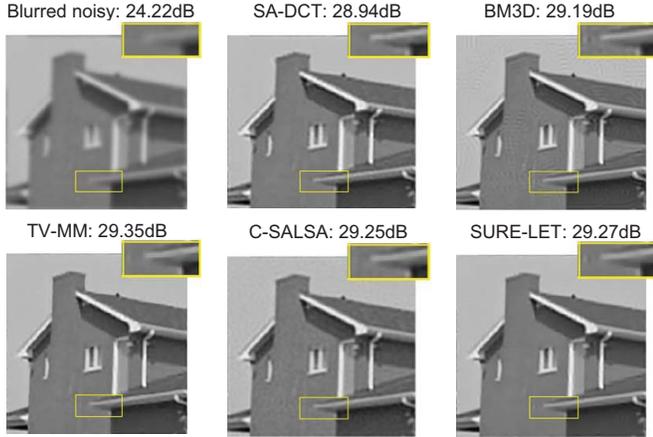

Fig. 7. Restoration of *Crowd* degraded by Gaussian blur of variance 9 with noise std $\sigma = 100$.

In this paper, we apply an undecimated Haar wavelet to perform wavelet-domain thresholding, since the redundant Haar substantially outperforms other types of redundant wavelets for image denoising [52]. This observation is also true for image deconvolution.

C. Influence of the Parameters on the Deconvolution Performance

All the parameters involved in the proposed multi-Wiener SURE-LET algorithm are specified in Table II. According to this table, we have $K = MJL + M = 6J + 3$ elementary functions, which yields K weights to be optimized using (9). Typically, $J = 9$ or 12, for three or four levels of an undecimated Haar wavelet decomposition. Table II shows that only 7 parameters involved in our algorithm need to be manually tuned, whereas we have typically up to $K = 75$ (for 4 wavelet iterations) parameters that are automatically adjusted to the image. By comparison, BM3D [30], [33] has more than 60 non-linear parameters to be manually determined for an efficient block matching and collaborative filtering.

Apart from the case of λ_1, λ_2 and λ_3 already discussed above (Section V-B), we report the following experimental observations:

1) Choosing β in $[5 \times 10^{-6}\sigma^2, 5 \times 10^{-5}\sigma^2]$ and μ in $[0.01, 0.1]$ yields very similar deconvolution results with PSNR loss within 0.05 dB, compared to the maximum PSNR with optimal values.
2) Any combinations of T_1 in $[3\sigma, 6\sigma]$ and T_2 in $[7\sigma, 10\sigma]$ achieve very similar PSNR performance (the PSNR variations are generally within 0.1 dB).

3) Applying the other two choices of thresholding functions θ_1 and θ_2 reported in [52], [54] does not result in significant PSNR loss (maximum 0.2 dB), compared to using (15).

Based on these observations, we now draw the conclusion that the proposed algorithm is very robust and largely insensitive to the parameters. No optimization is needed here. The main reason for the robustness is that any changes on the parameters or functions (β, μ, θ_1, θ_2, T_1 and T_2) are eventually counterbalanced by an (optimal) adaptation of the (at least 57) LET coefficients a_k (final PSNR variations of 0.1~0.2 dB). For this reason, the proposed algorithm becomes highly adaptive to variousscenarios, even if the parameters in Table II are not well-tuned.

D. Influence of the Boundary Conditions: Periodic versus Symmetric

Table III displays the PSNR results obtained for different boundary conditions (periodic, half-point symmetric or whole-point symmetric). As observed, the type of boundary conditions has an impact on the deconvolution performance. From the results, we can conclude that symmetric boundary conditions perform uniformly better than periodic. This is mainly because symmetric boundary conditions ensure slowly-varying changes at the boundaries, unlike periodic conditions which may introduce discontinuities. The deconvolution gain obtained by symmetric boundary conditions over periodic ones depends on the image and blur scenario. Table III reports the PSNR results obtained on *House* and *Cameraman* as two extreme cases: for *Cameraman*, the boundary conditions have almost no influence on the deconvolution performance, whereas significant differences are noticed for *House*.

E. Comparison With the State-of-the-Art

In Tables IV–VI, we present several comparisons between the proposed multi-Wiener SURE-LET method and some state-of-the-art deconvolution techniques whose softwares are

TABLE III

PSNR RESULTS OF OUR MULTI-WIENER SURE-LET FOR DIFFERENT BOUNDARY CONDITIONS ($\sigma^2 = 1$)

Blur Kernel	Rational Filter			Separable Filter			9 × 9 uniform			Gaussian Blur With std 3		
Boundary Conditions	Periodic	Symmetric		Periodic	Symmetric		Periodic	Symmetric		Periodic	Symmetric	
		Half-point	Whole-point		Half-point	Whole-point		Half-point	Whole-point		Half-point	Whole-point
Image	*House* 256 × 256											
Input	25.64	26.30	26.29	30.92	31.83	31.82	24.09	24.75	24.75	24.22	24.86	24.86
SURE-LET	35.33	**35.50**	35.46	36.40	**36.51**	36.47	32.97	**32.98**	32.97	29.37	**30.03**	30.01
MSE-LET	*35.48*	*35.61*	*35.58*	*36.58*	*36.82*	*36.77*	*33.19*	*33.26*	*33.20*	*29.74*	*30.32*	*30.32*
Image	*Cameraman* 256 × 256											
Input	22.24	22.37	22.36	25.67	25.79	25.79	20.76	20.89	20.89	20.97	21.10	21.10
SURE-LET	30.91	**30.97**	30.96	30.83	**30.93**	30.90	27.40	**27.47**	27.46	23.97	**24.14**	24.14
MSE-LET	*31.12*	*31.17*	*31.16*	*31.10*	*31.17*	*31.16*	*27.75*	*27.77*	*27.76*	*24.25*	*24.41*	*24.40*

TABLE IV

COMPARISON OF SOME STATE-OF-THE-ART DECONVOLUTION METHODS UNDER GAUSSIAN BLUR WITH VARIANCE 9*

σ	1	5	10	30	50	100	1	5	10	30	50	100
Image	*Cameraman* 256 × 256						*Coco* 256 × 256					
Input	20.97	20.78	20.22	16.61	13.33	7.91	26.45	25.78	24.19	17.92	13.89	8.06
ForWaRD	23.76	22.88	22.40	21.18	20.35	18.79	31.18	29.48	28.51	25.91	24.22	20.80
SA-DCT	23.73	**23.10**	22.50	20.80	19.39	16.86	31.27	**29.82**	28.58	24.76	22.16	18.09
BM3D	**24.05**	**23.10**	**22.61**	**21.46**	20.78	**19.74**	31.27	29.70	**28.72**	26.61	25.48	23.80
TV-MM	23.93	22.80	22.06	21.10	20.24	18.65	**31.65**	29.59	28.56	26.30	24.75	21.81
C-SALSA	23.75	22.81	22.26	21.27	20.47	18.95	31.35	29.07	28.47	26.13	24.31	21.50
[54]	23.32	22.44	21.83	20.78	20.25	18.95	30.39	28.58	27.70	25.81	24.82	16.88
SURE-LET	**23.97**	**23.01**	**22.52**	**21.50**	**20.91**	**19.80**	**31.57**	**29.80**	**28.80**	**26.95**	**25.78**	**24.02**
MSE-LET	*24.25*	*23.30*	*22.70*	*21.72*	*21.11*	*20.08*	*31.88*	*30.10*	*29.12*	*27.28*	*26.10*	*24.46*
Image	*House* 256 × 256						*Couple* 512 × 512					
Input	24.22	23.81	22.73	17.52	13.73	8.02	23.55	23.20	22.26	17.39	13.68	8.01
ForWaRD	28.87	27.43	26.63	24.27	22.87	20.23	26.40	25.25	24.62	23.12	22.25	19.64
SA-DCT	28.94	27.86	26.75	23.51	21.32	17.69	26.43	**25.35**	24.62	22.75	21.20	18.04
BM3D	29.19	**27.90**	**27.10**	**25.17**	23.98	**22.32**	**26.60**	25.39	**24.76**	23.42	22.72	21.37
TV-MM	**29.35**	27.50	26.72	24.60	23.08	20.63	26.12	24.92	24.20	23.02	21.79	19.53
C-SALSA	29.25	27.46	26.55	24.51	22.97	20.50	26.36	25.20	24.59	23.30	22.04	19.91
[54]	28.19	26.64	25.90	23.93	22.72	15.57	26.29	25.06	24.49	23.20	22.44	21.48
SURE-LET	**29.27**	**27.81**	**27.00**	**25.20**	**24.10**	**22.35**	26.56	**25.40**	**24.80**	**23.59**	**22.90**	**21.88**
MSE-LET	*29.62*	*28.01*	*27.24*	*25.54*	*24.33*	*22.62*	*26.64*	*25.47*	*24.88*	*23.68*	*22.99*	*21.96*
Image	*Crowd* 512 × 512						*Lake* 512 × 512					
Input	16.56	16.49	16.27	14.45	12.19	7.55	22.98	22.67	21.83	17.24	13.61	7.99
ForWaRD	**19.97**	18.83	18.22	17.21	16.57	15.25	26.80	25.39	24.65	22.94	21.72	18.84
SA-DCT	19.87	**18.97**	**18.33**	17.02	16.30	14.97	**26.82**	**25.49**	24.57	22.44	20.97	18.05
BM3D	**20.01**	18.80	18.15	16.83	16.15	14.93	**26.90**	25.48	24.68	22.97	22.10	20.42
TV-MM	19.43	18.77	18.17	17.10	16.18	14.87	26.58	25.25	24.38	22.83	21.46	19.12
C-SALSA	19.92	**18.88**	**18.31**	**17.33**	**16.73**	15.65	26.74	**25.40**	24.67	23.05	21.73	19.50
[54]	19.80	18.61	18.04	16.91	16.36	15.55	26.67	25.20	24.44	22.89	22.04	20.74
SURE-LET	**20.05**	**18.90**	**18.30**	17.25	16.67	**15.79**	**26.90**	25.48	**24.80**	**23.33**	**22.48**	**21.18**
MSE-LET	*20.21*	*19.06*	*18.48*	*17.40*	*16.81*	*15.89*	*26.97*	*25.62*	*24.92*	*23.49*	*22.64*	*21.32*
Image	*Bridge* 512 × 512						*Mixture* 512 × 512					
Input	21.39	21.18	20.57	16.77	13.41	7.94	14.84	14.79	14.64	13.31	11.46	7.29
ForWaRD	23.73	22.79	22.28	21.32	20.53	18.41	16.63	15.94	**15.72**	**15.15**	**14.77**	14.08
SA-DCT	23.75	22.84	22.32	21.27	20.50	18.63	16.64	**16.13**	**15.80**	**15.19**	**14.75**	13.64
BM3D	**23.85**	**22.88**	**22.38**	21.38	20.78	19.61	**16.74**	15.95	15.59	14.91	14.60	14.02
TV-MM	23.56	22.69	21.99	20.51	19.80	18.12	16.37	15.69	15.20	14.72	14.46	13.52
C-SALSA	23.78	**22.87**	22.09	21.38	20.52	18.89	16.52	15.96	15.65	15.07	14.70	13.92
[54]	23.70	22.73	22.24	21.27	20.71	19.87	16.56	15.88	15.57	14.98	14.67	14.27
SURE-LET	**23.90**	**22.95**	**22.44**	**21.54**	**20.95**	**20.01**	**16.70**	**16.07**	**15.74**	**15.15**	**14.80**	**14.43**
MSE-LET	*23.97*	*23.01*	*22.53*	*21.61*	*21.05*	*20.14*	*16.78*	*16.12*	*15.81*	*15.28*	*14.93*	*14.53*

*Best PSNR results within a 0.1 dB margin are highlighted. The results have been averaged over 10 noise realizations.

available online. Various degradation scenarios have been considered. In order to compare with other methods in exactly the same experimental situations, we have considered only periodic boundary condition. The state-of-the-art methods include

TABLE V

DECONVOLUTION OF THE *Bridge* IMAGE FOR VARIOUS BLURS AND NOISE LEVELS[*]

σ	1	5	10	30	50	100	1	5	10	30	50	100
Blur	Rational Filter						Separable Filter					
Input	22.60	22.33	21.55	17.15	13.58	7.98	25.78	25.22	23.82	17.85	13.88	8.06
ForWaRD	28.32	24.65	23.37	21.48	20.20	18.60	29.09	26.47	25.04	22.77	21.74	19.75
SA-DCT	**28.94**	25.00	**23.74**	21.99	20.99	19.02	29.23	**26.97**	25.68	**23.43**	22.40	20.77
BM3D	**28.95**	**25.12**	23.70	21.87	21.05	19.67	**29.40**	**27.00**	**25.72**	23.37	22.36	20.43
TV-MM	28.64	**25.10**	23.42	20.16	18.71	17.32	28.95	**26.93**	**25.70**	22.98	20.87	19.28
C-SALSA	28.63	24.90	23.49	20.45	18.82	18.31	28.96	26.57	25.34	22.44	20.46	16.46
[54]	28.44	24.87	23.53	21.71	20.91	19.92	29.03	26.70	25.48	23.25	22.16	20.61
SURE-LET	**28.87**	**25.10**	**23.81**	**22.12**	**21.32**	**20.19**	29.37	**27.00**	**25.70**	**23.50**	**22.52**	**21.04**
MSE-LET	*28.90*	*25.23*	*23.95*	*22.22*	*21.44*	*20.32*	*29.48*	*27.07*	*25.80*	*23.58*	*22.59*	*21.16*
Blur	5×5 Uniform Blur						9×9 Uniform Blur					
Input	23.24	22.92	22.04	17.33	13.66	8.01	21.16	20.96	20.38	16.69	13.37	7.93
ForWaRD	28.31	25.15	23.89	22.06	20.11	18.82	25.77	23.48	22.65	21.21	20.08	18.63
SA-DCT	27.96	24.76	23.82	22.32	21.43	19.86	25.83	23.31	22.48	20.82	19.71	18.23
BM3D	**28.57**	**25.50**	**24.46**	22.65	21.88	20.21	**26.12**	**23.75**	22.82	21.50	20.90	19.70
TV-MM	28.31	**25.52**	**24.41**	21.74	19.97	18.24	25.85	23.57	22.43	19.95	18.39	16.44
C-SALSA	28.36	25.16	24.15	21.78	18.59	17.48	25.74	23.42	22.58	20.67	19.52	18.00
[54]	28.25	25.20	24.17	22.52	21.74	20.41	25.82	23.39	22.47	21.32	20.83	19.96
SURE-LET	**28.65**	**25.55**	**24.43**	**22.78**	**22.00**	**20.78**	**26.10**	**23.79**	**22.95**	**21.71**	**21.10**	**20.13**
MSE-LET	*28.69*	*25.64*	*24.56*	*22.90*	*22.11*	*20.88*	*26.17*	*23.88*	*23.05*	*21.81*	*21.19*	*20.24*

[*]Best PSNR results within a 0.1 dB margin are highlighted. The results have been averaged over 10 noise realizations.

TABLE VI

DECONVOLUTION OF THE *Mixture* IMAGE FOR VARIOUS BLURS AND NOISE LEVELS[*]

σ	1	5	10	30	50	100	1	5	10	30	50	100
Blur	Rational Filter						Separable Filter					
Input	15.96	15.89	15.70	14.07	11.95	7.47	18.38	18.27	17.94	15.47	12.76	7.74
ForWaRD	27.67	19.95	17.77	15.76	15.01	13.99	26.01	22.03	20.42	17.50	16.30	14.96
SA-DCT	28.54	20.64	18.20	15.99	15.32	14.40	24.35	22.25	20.95	17.93	16.70	15.46
BM3D	28.53	19.24	17.25	15.45	14.85	14.10	26.54	22.26	20.04	17.14	16.15	14.86
TVMM	28.28	20.40	17.30	14.41	13.78	13.18	27.17	22.80	20.64	17.15	15.25	13.80
C-SALSA	27.26	20.14	18.04	15.80	15.02	13.99	26.58	21.75	20.16	17.51	16.19	13.80
[54]	25.92	19.52	17.74	15.77	15.12	14.42	24.67	21.40	20.01	17.67	16.67	15.41
SURE-LET	**29.90**	**21.10**	**18.53**	**16.12**	**15.50**	**14.71**	**27.80**	**23.13**	**21.08**	**18.05**	**16.94**	**15.70**
MSE-LET	*30.28*	*21.34*	*18.70*	*16.27*	*15.58*	*14.78*	*28.11*	*23.40*	*21.43*	*18.21*	*17.06*	*15.79*
Blur	5×5 Uniform Blur						9×9 Uniform Blur					
Input	15.92	15.86	15.67	14.04	11.94	7.46	14.58	14.54	14.40	13.13	11.35	7.24
ForWaRD	24.76	19.80	18.26	16.29	15.58	14.42	20.42	17.21	16.16	15.20	14.68	13.79
SA-DCT	25.35	19.40	17.57	16.03	15.48	14.74	21.29	17.10	16.13	15.06	14.60	13.93
BM3D	25.53	19.78	17.90	16.18	15.56	14.58	20.66	17.07	16.01	15.00	14.68	14.09
TVMM	25.70	20.39	18.31	15.66	14.52	13.64	20.70	17.02	15.65	14.23	13.66	12.87
C-SALSA	24.29	19.67	18.15	15.77	15.37	14.00	20.24	17.23	16.30	14.95	14.29	13.30
[54]	23.19	19.14	17.83	16.37	15.78	15.00	19.39	16.67	15.92	14.99	14.70	14.36
SURE-LET	**26.34**	**20.75**	**18.90**	**16.74**	**16.04**	**15.25**	**21.60**	**17.75**	**16.65**	**15.32**	**15.01**	**14.52**
MSE-LET	*26.60*	*21.00*	*19.06*	*16.85*	*16.13*	*15.35*	*21.90*	*17.89*	*16.76*	*15.42*	*15.07*	*14.60*

[*]Best PSNR results within a 0.1 dB margin are highlighted. The results have been averaged over 10 noise realizations.

ForWaRD[5] [27], SA-DCT[6] [29], BM3D[7] [30], TV-MM[8] [24], C-SALSA[9] [16] and Pesquet's *et al.* work[10] [54]. We have run all the source codes by default throughout all the experiments performed. For TV-MM [24] and C-SALSA [16], we used the

[5]The source code of ForWaRD is available at http://dsp.rice.edu/software/forward.

[6]The source code of SA-DCT is available at http://www.cs.tut.fi/~foi/SA-DCT/ref_software.

[7]The source code of BM3D is available at http://www.cs.tut.fi/~foi/GCF-BM3D/.

[8]The source code of TV-MM is available at http://www.lx.it.pt/~bioucas/code.htm.

[9]The source code of C-SALSA is available at http://cascais.lx.it.pt/~mafonso/salsa.html.

[10]The source code of [54] is available at http://www-syscom.univ-mlv.fr/~chaux/toolbox/TOOLBOX_Surelet_deconvolution_v1.0.zip.

TABLE VII

COMPARISON OF THE COMPUTATIONAL TIME OF VARIOUS DECONVOLUTION TECHNIQUES (UNITS: SEC.)*

Degradation Scenario	ForWaRD	SA-DCT	BM3D	TV-MM	C-SALSA	Pesquet [54]	SURE-LET
Cameraman 256×256 Gaussian blur, $\sigma = 1$	3.62	2.81	1.90	103.91	33.52	114.12	**0.69**
Mixture 512×512 Rational filter, $\sigma = 10$	31.11	12.28	8.00	93.50	34.32	375.74	**2.69**
Crowd 512×512 9×9 uniform, $\sigma = 50$	45.53	12.33	7.80	59.08	25.18	353.52	**2.69**

* Hardware environment used: Intel(R) Core(TM) i3-2100 CPU @3.10GHz, memory size: 2GB.
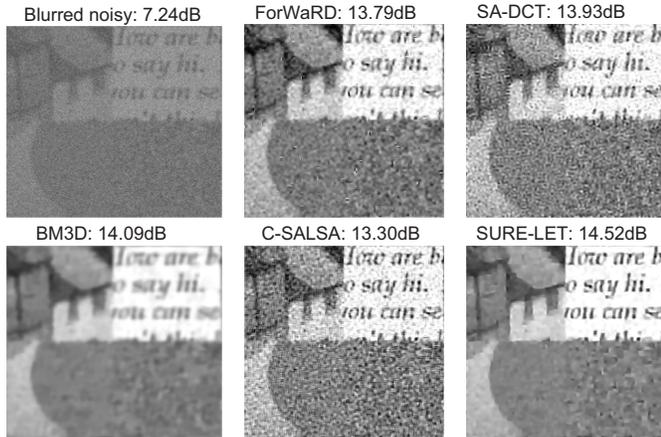

Fig. 8. Restoration of *Mixture* degraded by 9×9 uniform blur with noise std $\sigma = 100$.

default stopping criteria suggested by their respective authors. For [54], we used symlet-8 translation invariant wavelet, as it gives slightly better performance than using Haar undecimated wavelet (generally by 0.2∼0.3 dB).

The "MSE-LET" rows of Tables IV–VI are the results of minimizing the actual MSE in lieu of the SURE, demonstrating that the latter is a good substitute to the MSE minimization.

It can be seen from the reported PSNR scores that the proposed SURE-LET algorithm uniformly achieves the highest performance. Although TV-MM is well-known for its outstanding performance on regularly-structured images such as *House*, *Coco*, and the left-bottom of *Mixture*, it is substantially outperformed by the proposed algorithm. For more complicated images like *Crowd* with lots of irregular edges and disordered features, the proposed method also shows better PSNR results than the other techniques. ForWaRD is more effective for these less structured images than for cartoon-like images. BM3D, which achieves the best performance among the other techniques on average, is not as efficient on these irregular structures. Note that the proposed multi-Wiener SURE-LET algorithm is very robust to a wide range of noise levels from $\sigma^2 = 1$ to $\sigma^2 = 1 \times 10^4$. In particular, significant improvements over other deconvolution algorithms are frequently observed for large noise variance.

Regarding the subjective visual quality (see Figs. 6–8), the proposed deconvolution algorithm preserves various image details, while introducing very few artifacts.

For *House*, it achieves good preservation of uniform areas and regularly-sharp edges, while for *Crowd*, it preserves the finer details of the irregularities. In Fig. 8, the differences between the various algorithms are clearly visible: the proposed algorithm introduces fewer artifacts than the other techniques. In particular, the text part of *Mixture* is easier to read in the image restored by the proposed multi-Wiener SURE-LET.

F. Analysis of Computational Complexity

The most computationally-intensive part of the proposed algorithm is the construction of the elementary functions $\mathbf{f}_k(\mathbf{y})$, due to the computation of the thresholding $\boldsymbol{\theta}(\cdot)$ and of its first-order derivative, as well as the independent reconstructions of all the processed subbands. Compared to these steps, the construction and resolution of the linear system of equations (9) of low order K has a negligible computational cost.

Based on these observations, the computational complexity of the proposed multi-Wiener SURE-LET is roughly evaluated as KN, which is linearly proportional to both the number of basis functions K and pixel number N. Since the processing time of the proposed deconvolution algorithm is independent of the content of the input data, it grows linearly with the data size. Table VII reports the execution time of the various algorithms, measured on an Intel(R) Core(TM) i3-2100 CPU @3.10 GHz, with 2 GB of memory. As expected, the processing time of a 512×512 image is roughly four times that of a 256×256 image for the proposed algorithm.

Finally, we would like to emphasize the low complexity of the proposed algorithm. Since our deconvolution approach merely boils down to solving a linear system of equations, it is substantially faster than other state-of-the-art techniques, as confirmed in Table VII. We would also like to stress that our implementation uses standard MATLAB code only, without any compiled routines, and can be easily parallelized for even faster processing.

VI. CONCLUSION

In this paper, we have presented a new image deconvolution method based on the SURE-LET approach initially developed for image denoising applications. The main originality of the proposed deconvolution approach is the use of multiple Wiener filters with different but fixed regularization parameters, to avoid the empirical and fastidious adjustment of the non-linear regularization parameter.

Although the proposed multi-Wiener SURE-LET deconvolution algorithm has low complexity, its performances are already quite competitive with the state-of-the-art deconvolution techniques, both quantitatively and visually. The great flexibility and the limited computational cost of the proposed approach suggest that it is possible to develop more sophisticated forms of basic processings; e.g., performing a multivariate thresholding by taking into account inter-scale and/or intra-scale dependencies between the wavelet coefficients as in [51]. Electing more directional deconvolutions instead of using classic Wiener filters might also lead to improved performances. There is thus a substantial margin of improvement for SURE-LET type deconvolution algorithms.

APPENDIX A
PROOF OF THEOREM 2.1

Proof: Expanding the MSE (2) and using $\mathbf{x} = \mathbf{H}^{-1}(\mathbf{y} - \mathbf{b})$, we obtain (ignoring the factor $\frac{1}{N}$)

$$
\begin{aligned}
\mathscr{E}\left\{\|\mathbf{f}(\mathbf{y}) - \mathbf{x}\|^2\right\} &= \mathscr{E}\left\{\|\mathbf{f}(\mathbf{y})\|^2\right\} - 2\mathscr{E}\left\{\mathbf{x}^\mathrm{T}\mathbf{f}(\mathbf{y})\right\} + \mathscr{E}\left\{\|\mathbf{x}\|^2\right\} \\
&= \mathscr{E}\left\{\|\mathbf{f}(\mathbf{y})\|^2\right\} - 2\mathscr{E}\left\{\mathbf{y}^\mathrm{T}\mathbf{H}^{-\mathrm{T}}\mathbf{f}(\mathbf{y})\right\} \\
&\quad + 2\mathscr{E}\left\{\mathbf{b}^\mathrm{T}\mathbf{H}^{-\mathrm{T}}\mathbf{f}(\mathbf{y})\right\} + \mathscr{E}\left\{\|\mathbf{x}\|^2\right\}. \quad (23)
\end{aligned}
$$

Consider the multivariate Gaussian probability density function $q(\mathbf{b}) \propto \exp(-\frac{\mathbf{b}^\mathrm{T}\mathbf{C}^{-1}\mathbf{b}}{2})$. 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,

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

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

$$
\begin{aligned}
\mathscr{E}\left\{\mathbf{b}^\mathrm{T}\mathbf{H}^{-\mathrm{T}}\mathbf{f}(\mathbf{y})\right\} &= \sum_{n=1}^{N}\int_{-\infty}^{\infty}\frac{\partial u_n}{\partial b_n}q(\mathbf{b})\partial\mathbf{b} \\
&= \mathscr{E}\left\{\sum_{n=1}^{N}\frac{\partial u_n}{\partial b_n}\right\} \\
&= \mathscr{E}\left\{\sum_{n=1}^{N}\frac{\partial u_n}{\partial y_n}\right\} \\
&= \mathscr{E}\left\{\mathrm{div}_\mathbf{y}\mathbf{u}\right\}. \quad (24)
\end{aligned}
$$

Substituting (24) into (23) completes the proof. ∎

APPENDIX B
PROOF OF (6) IN SECTION II-C

Proof: Similar to Appendix A, using $\mathbf{Hx} = \mathbf{y} - \mathbf{b}$, we expand the modified MSE as (ignoring the factor $\frac{1}{N}$)

$$
\left\|\mathbf{f}(\mathbf{y}) - \mathbf{H}_\beta^{-1}\mathbf{Hx}\right\|^2 = \|\mathbf{f}(\mathbf{y})\|^2 - 2\mathbf{x}^\mathrm{T}\mathbf{H}^\mathrm{T}\mathbf{H}_\beta^{-\mathrm{T}}\mathbf{f}(\mathbf{y}) + \left\|\mathbf{H}_\beta^{-1}\mathbf{Hx}\right\|^2 \quad (25)
$$

where the second term is:

$$
\mathbf{x}^\mathrm{T}\mathbf{H}^\mathrm{T}\mathbf{H}_\beta^{-\mathrm{T}}\mathbf{f}(\mathbf{y}) = \mathbf{y}^\mathrm{T}\mathbf{H}^\mathrm{T}\mathbf{H}_\beta^{-\mathrm{T}}\mathbf{f}(\mathbf{y}) - \mathbf{b}^\mathrm{T}\mathbf{H}^\mathrm{T}\mathbf{H}_\beta^{-\mathrm{T}}\mathbf{f}(\mathbf{y}). \quad (26)
$$

According to Appendix A, we have

$$
\mathscr{E}\left\{\mathbf{b}^\mathrm{T}\mathbf{H}_\beta^{-\mathrm{T}}\mathbf{f}(\mathbf{y})\right\} = \mathscr{E}\left\{\mathrm{div}_\mathbf{y}\left\{\mathbf{C}\mathbf{H}_\beta^{-\mathrm{T}}\mathbf{f}(\mathbf{y})\right\}\right\}. \quad (27)
$$

Substituting (26) and (27) into (25) completes the proof. ∎

APPENDIX C
PROOF OF THEOREM 4.3

Proof: From Lemma 4.1, $p(n)$ can be expressed as:

$$
p(n) = \sum_{k=0}^{N-1}\big[\underbrace{g_{2N}(n-k) + g_{2N}(n+k+1)}_{u(k)}\big]b(k)
$$

then, using the i.i.d. condition of $b(n)$ as in (19), the variance of $p(n)$ is

$$
\mathscr{E}\left\{p(n)^2\right\} = \mathscr{E}\left\{\left(\sum_{k=0}^{N-1}u(k)b(k)\right)^2\right\} = \sigma^2\sum_{k=0}^{N-1}u(k)^2. \quad (28)
$$

Due to the fact that $u(-k) = u(k-1)$ and $u(k+2N) = u(k)$, by change of variable and Parseval's theorem, we have

$$
\sum_{k=0}^{N-1}u(k)^2 = \frac{1}{2}\sum_{k=0}^{2N-1}u(k)^2 = \frac{1}{4N}\sum_{k=0}^{2N-1}|U(k)|^2 \quad (29)
$$

where $U(k)$ is the DFT of $u(k')$:

$$
\begin{aligned}
U(k) &= \sum_{k'=0}^{2N-1}u(k')e^{-j\frac{2\pi kk'}{2N}} \\
&= e^{-j\frac{\pi kn}{N}}\sum_{k'=0}^{2N-1}g_{2N}(k')e^{j\frac{\pi kk'}{N}} \\
&\quad + e^{j\frac{\pi k}{N}(n+1)}\sum_{k'=0}^{2N-1}g_{2N}(k')e^{-j\frac{\pi kk'}{N}}.
\end{aligned}
$$

Recalling Eq.(22): $G\left(e^{j\frac{\pi k}{N}}\right) = \sum_{n=0}^{2N-1}g_{2N}(n)e^{-j\frac{\pi kn}{N}}$, and combining with $U(k)$, $U(k)$ is related to $G\left(e^{j\frac{\pi k}{N}}\right)$ through:

$$
U(k) = e^{-j\frac{\pi kn}{N}}G\left(e^{-j\frac{\pi k}{N}}\right) + e^{j\frac{\pi k}{N}(n+1)}G\left(e^{j\frac{\pi k}{N}}\right). \quad (30)
$$

Substituting (30) into (29), and combining with (28) complete the proof. ∎

REFERENCES

[1] F. Xue, F. Luisier, and T. Blu, "SURE-LET image deconvolution using multiple Wiener filters," in *Proc. 19th IEEE Int. Conf. Image Process.*, Orlando, FL, USA, Oct. 2012, pp. 3037–3040.

[2] S. Wan, B. Raju, and M. Srinivasan, "Robust deconvolution of high-frequency ultrasound images using higher-order spectral analysis and wavelets," *IEEE Trans. Ultrason., Ferroelectr. Frequency Control*, vol. 50, no. 10, pp. 1286–1295, Oct. 2003.

[3] T. Sjoeberg, L. Gelius, and I. Lecomte, "2-D deconvolution of seismic image blur," in *Proc. Soc. Explorat. Geophys. Annu. Meeting*, Oct. 2003, pp. 1–4.

[4] J. Starck, E. Pantin, and F. Murtagh, "Deconvolution in astronomy: A review," *Publicat. Astronom. Soc. Pacific*, vol. 114, no. 800, pp. 1051–1069, 2002.

[5] G. 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.

[6] P. Sarder and A. Nehorai, "Deconvolution methods for 3-D fluorescence microscopy images," *IEEE Signal Process. Mag.*, vol. 23, no. 3, pp. 32–45, May 2006.

[7] H. Andrews and B. Hunt, *Digital Image Restoration*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1977.

[8] A. Tikhonov and V. Arsenin, *Solutions of Ill-Posed Problems*. Washington, DC, USA: Winston, 1977.

[9] P. Combettes and V. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Model. Simulat.*, vol. 4, no. 4, pp. 1168–1200, 2006.

[10] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*, New York, USA: Wiley, 1964.

[11] A. N. Tikhonov, "Solution of incorrectly formulated problems and the regularization method," *Soviet Math. Doklady*, vol. 4, no. 4, pp. 1035–1038, 1963.

[12] M. Figueiredo and R. Nowak, "An EM algorithm for wavelet-based image restoration," *IEEE Trans. Image Process.*, vol. 12, no. 8, pp. 906–916, Aug. 2003.

[13] M. Figueiredo, J. Bioucas-Dias, and R. Nowak, "Majorization—minimization algorithms for wavelet-based image restoration," *IEEE Trans. Image Process.*, vol. 16, no. 12, pp. 2980–2991, Dec. 2007.

[14] E. Esser, "Applications of Lagrangian-based alternating direction methods and connections to split Bregman," Univ. California, Los Angeles, CA, USA, Tech Rep. 09-31, 2009, vol. 9.

[15] C. Vonesch and M. Unser, "A fast thresholded Landweber algorithm for wavelet-regularized multidimensional deconvolution," *IEEE Trans. Image Process.*, vol. 17, no. 4, pp. 539–549, Apr. 2008.

[16] M. Afonso, J. Bioucas-Dias, and M. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, Mar. 2011.

[17] M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," *Inverse Problems*, vol. 23, no. 3, pp. 947–968, 2007.

[18] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.*, vol. 57, no. 11, pp. 1413–1457, 2004.

[19] H. Pan and T. Blu, "Sparse image restoration using iterated linear expansion of thresholds," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2011, pp. 1905–1908.

[20] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, Nov. 1992.

[21] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regularization method for total variation-based image restoration," *Multiscale Model. Simulat.*, vol. 4, no. 2, pp. 460–489, 2006.

[22] A. Chambolle, "An algorithm for total variation minimization and Applications," *J. Math. Imag. Vis.*, vol. 20, no. 1, pp. 89–97, 2004.

[23] P. Combettes and J. Pesquet, "Image restoration subject to a total variation constraint," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1213–1222, Sep. 2004.

[24] J. Bioucas-Dias, M. Figueiredo, and J. Oliveira, "Total variation-based image deconvolution: A majorization-minimization approach," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, vol. 2. May 2006, pp. 861–864.

[25] O. Michailovich, "An iterative shrinkage approach to total-variation image restoration," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1281–1299, May 2011.

[26] H. Takeda, S. Farsiu, and P. Milanfar, "Deblurring using regularized locally adaptive kernel regression," *IEEE Trans. Image Process.*, vol. 17, no. 4, pp. 550–563, Apr. 2008.

[27] R. Neelamani, H. Choi, and R. Baraniuk, "ForWaRD: Fourier-wavelet regularized deconvolution for Ill-conditioned systems," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 418–433, Apr. 2004.

[28] J. Guerrero-Colón, L. Mancera, and J. Portilla, "Image restoration using space-variant Gaussian scale mixtures in overcomplete pyramids," *IEEE Trans. Image Process.*, vol. 17, no. 1, pp. 27–41, Jan. 2008.

[29] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Inverse halftoning by pointwise shape-adaptive DCT regularized deconvolution," in *Proc. Int. Tampere Int. Center Signal Process. Workshop Spectral Meth. Multirate Signal Process.*, Sep. 2006, pp. 1–6.

[30] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image restoration by sparse 3-D transform-domain collaborative filtering," in *Proc. Int. Soc. Opt. Eng. Electron. Imag.*, vol. 6812. Jan. 2008.

[31] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1338–1351, Nov. 2003.

[32] A. Foi, V. Katkovnik, and K. Egiazarian, "Pointwise shape-adaptive dct for high-quality denoising and deblocking of grayscale and color images," *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1395–1411, May 2007.

[33] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[34] A. Jain, *Fundamentals of digital Image Processing* (Information and System Sciences). Englewood Cliffs, NJ, USA: Prentice-Hall, 1989.

[35] A. Jalobeanu, L. Blanc-Feraud, and J. Zerubia, "An adaptive Gaussian model for satellite image deblurring," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 613–621, Apr. 2004.

[36] F. Jeng and J. Woods, "Compound Gauss-Markov random fields for image estimation," *IEEE Trans. Signal Process.*, vol. 39, no. 3, pp. 683–697, Mar. 1991.

[37] R. Molina, J. Núñez, F. Cortijo, and J. Mateos, "Image restoration in astronomy: A Bayesian perspective," *IEEE Signal Process. Mag.*, vol. 18, no. 2, pp. 11–29, Mar. 2001.

[38] P. de Rivaz and N. Kingsbury, "Bayesian image deconvolution and denoising using complex wavelets," in *Proc. Int. Conf. Image Process.*, vol. 2. Oct. 2001, pp. 273–276.

[39] G. Chantas, N. Galatsanos, A. Likas, and M. Saunders, "Variational Bayesian image restoration based on a product of *t*-distributions image prior," *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1795–1805, Oct. 2008.

[40] M. Figueiredo and R. Nowak, "Wavelet-based image estimation: An empirical Bayes approach using Jeffrey's noninformative prior," *IEEE Trans. Image Process.*, vol. 10, no. 9, pp. 1322–1331, Sep. 2001.

[41] J. Bioucas-Dias, "Bayesian wavelet-based image deconvolution: A GEM algorithm exploiting a class of heavy-tailed priors," *IEEE Trans. Image Process.*, vol. 15, no. 4, pp. 937–951, Apr. 2006.

[42] M. Mignotte, "A non-local regularization strategy for image deconvolution," *Pattern Recognit. Lett.*, vol. 29, no. 16, pp. 2206–2212, Dec. 2008.

[43] L. Mancera, S. Babacan, R. Molina, and A. Katsaggelos, "Image restoration by mixture modelling of an overcomplete linear representation," in *Proc. 16th IEEE Int. Conf. Image Process.*, Nov. 2009, pp. 3949–3952.

[44] S. Babacan, R. Molina, and A. Katsaggelos, "Sparse bayesian image restoration," in *Proc. 17th IEEE Int. Conf. Image Process.*, Sep. 2010, pp. 26–29.

[45] A. Mohammad-Djafari, "A full Bayesian approach for inverse problems," in *Proc. Max. Entr. Bayes.Meth.*, K. Hanson and R. Silver. Eds., Kluwer Acad. Publ., Norwell, MA, USA, 1996, pp. 135–143.

[46] A. Neumaier, "Solving ill-conditioned and singular linear systems: A tutorial on regularization," *Soc. Ind. Appl. Math. Rev.*, vol. 40, no. 3, pp. 636–666, 1998.

[47] N. Galatsanos and A. Katsaggelos, "Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation," *IEEE Trans. Image Process.*, vol. 1, no. 3, pp. 322–336, Jul. 1992.

[48] A. Thompson, J. Brown, J. Kay, and D. Titterington, "A study of methods of choosing the smoothing parameter in image restoration by regularization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 4, pp. 326–339, Apr. 1991.

[49] M. Lukas, "Asymptotic optimality of generalized cross-validation for choosing the regularization parameter," *Numer. Math.*, vol. 66, no. 1, pp. 41–66, 1993.

[50] C. Stein, "Estimation of the mean of a multivariate normal distribution," *Ann. Stat.*, vol. 9, no. 6, pp. 1135–1151, 1981.

[51] 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.

[52] 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.

[53] 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.*, San Diego, CA, USA, Oct. 2008, pp. 665–668.

[54] 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.

[55] Y. Eldar, "Generalized SURE for exponential families: Applications to regularization," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 471–481, Feb. 2009.

[56] 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.

[57] D. Van De Ville and M. Kocher, "SURE-based non-local means," *IEEE Signal Process. Lett.*, vol. 16, no. 11, pp. 973–976, Nov. 2009.

[58] M. Raphan and E. Simoncelli, "Optimal denoising in redundant representations," *IEEE Trans. Image Process.*, vol. 17, no. 8, pp. 1342–1352, Aug. 2008.

[59] T. Blu, "The SURE-LET methodology—a prior-free approach to signal and image denoising," in *Proc. Plenary Present. 8th Int. Workshop Sampl. Theory Appl.*, May 2009.

[60] G. Strang, "The discrete cosine transform," *Soc. Ind. Appl. Math. Rev.*, vol. 41, no. 1, pp. 135–147, 1999.

[61] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Cambridge, MA, USA: Wellesley Cambridge, 1996.

[62] A. Oppenheim, A. Willsky, and S. Hamid, *Signals and Systems*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1997.

[63] T. Ha, *Theory and Design of Digital Communication Systems*. Cambridge, MA, USA: Cambridge Univ. Press, 2010.

[64] S. Martucci, "Symmetric convolution and the discrete sine and cosine transforms," *IEEE Trans. Signal Process.*, vol. 42, no. 5, pp. 1038–1051, May 1994.

[65] S. Babacan, R. Molina, and A. Katsaggelos, "Variational Bayesian blind deconvolution using a total variation prior," *IEEE Trans. Image Process.*, vol. 18, no. 1, pp. 12–26, Jan. 2009.

Feng Xue (S'12) was born in Shandong, China, in 1983. He received the B.Eng. degree in electronic science and technology and the M.Eng. degree in optical engineering from Harbin Institute of Technology, Harbin, China, in 2005 and 2007, respectively. He is currently pursuing the Ph.D. degree in electronic engineering under the guidance of Prof. Thierry Blu with The Chinese University of Hong Kong, Hong Kong.

His current research interests include inverse problems and image deconvolution.

Florian Luisier (S'05–M'10) was born in Bagnes, Switzerland, in 1981. He received the Master's degree in microengineering and the Ph.D. degree in computer, communication, and information sciences from the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, in 2005 and 2010, respectively.

He is currently a Post-Doctoral Fellow with the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA. From 2005 to 2010, he was with the Biomedical Imaging Group, EPFL. His current research interests include statistical signal, image, and video processing, image restoration, biomedical imaging, multiresolution representations, estimation and detection theory, and inverse problems.

Dr. Luisier was a recipient of the 2009 Young Author Best Paper Award from the IEEE Signal Processing Society for a paper co-authored with Prof. Thierry Blu and Prof. Michael Unser.

Thierry Blu (M'96–SM'06–F'12) was born in Orléans, France, in 1964. He received the Diplôme d'Ingénieur degree from the École Polytechnique, Palaiseau, France, in 1986 and from the Télécom Paris (ENST), Paris, France, where he received the Ph.D. degree in electrical engineering with research on iterated rational filterbanks, applied to wideband audio coding in 1988.

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

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