SURE-BASED BLIND GAUSSIAN DECONVOLUTION

Feng Xue and Thierry Blu

Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong

ABSTRACT

We propose a novel blind deconvolution method that consisting of firstly estimating the variance of the Gaussian blur, then performing non-blind deconvolution with the estimated PSF. The main contribution of this paper is the first step — to estimate the variance of the Gaussian blur, by minimizing a novel objective functional: an unbiased estimate of a blur MSE (SURE). The optimal parameter and blur variance are obtained by minimizing this criterion over linear processings that have the form of simple Wiener filterings. We then perform non-blind deconvolution using our recent high-quality SURE-based deconvolution algorithm.

The very competitive results show the highly accurate estimation of the blur variance (compared to the ground-truth value) and the great potential of developing more powerful blind deconvolution algorithms based on the SURE-type principle.

Index Terms— Blind deconvolution, minimization of blur SURE, Wiener filtering, estimation of blur variance

1. INTRODUCTION

In many applications, the image formation can be mathematically represented as [1]:

$$\mathbf{y} = \mathbf{H}_0\mathbf{x} + \mathbf{b} \tag{1}$$

where $\mathbf{y} \in \mathbb{R}^N$ is the measurements of the original unknown data $\mathbf{x} \in \mathbb{R}^N$ (N is pixel number of the image), when $\mathbf{x}$ is distorted by unknown true point-spread function (PSF) $\mathbf{h}_0$ and corrupted by additive Gaussian noise $\mathbf{b} \sim \mathscr{N}(0, \sigma^2\mathbf{I})$, $\mathbf{H}_0 \in \mathbb{R}^{N\times N}$ is the corresponding convolution matrix of $\mathbf{h}_0$. The blind deconvolution task amounts to recovering both $\mathbf{x}$ and $\mathbf{h}_0$ from only the measured data $\mathbf{y}$ [1].

Since blind deconvolution is a highly ill-posed problem, a number of methods have been developed to introduce a certain constraints on $\mathbf{x}$ and $\mathbf{h}$ as regularization terms, namely $\mathscr{L}_1(\mathbf{x})$ and $\mathscr{L}_2(\mathbf{h})$, to represent the prior knowledge [1–3]. They formulate blind deconvolution as the following joint minimization problem [1, 2]:

$$\min_{\mathbf{x},\mathbf{h}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \lambda_1\mathscr{L}_1(\mathbf{x}) + \lambda_2\mathscr{L}_2(\mathbf{h}) \tag{2}$$

where $\|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2$ is data-fidelity term, λ_1 and λ_2 are regularization parameters.

In some particular applications, the PSF is of known specific parametric form and is completely characterized by a few number of parameters [4]. Of particular interest is the Gaussian PSF, which can be used for modeling many degradation scenarios in real applications [4–6]. In this context, the parametric blind deconvolution is to estimate the variance of Gaussian blur, by, for instance, minimizing (2) [4], or measuring the response of second derivative Gaussian filter to edges [7, 8].

Instead of jointly estimating both $\mathbf{h}$ and $\mathbf{x}$ as [1–4, 9], we focus on the estimation of the PSF only. This paper proposes a new method to estimate the variance of Gaussian blur, based on the minimization of an unbiased estimate of the blur MSE (blur SURE [1]), which has been recently used and shown to be a powerful tool for denoising and deconvolution [10]. Finally, we use our proposed SURE-LET approach to perform non-blind deconvolution with the estimated PSF [11].

2. GENERIC PSF ESTIMATION BASED ON THE MINIMIZATION OF BLUR SURE

2.1. New criterion — blur SURE

Denote the function (or processing) of the observed data $\mathbf{y}$ by $\mathbf{F}(\mathbf{y})$. Instead of the standard MSE given as $\frac{1}{N}\mathscr{E}\left\{\left\|\mathbf{F}(\mathbf{y}) - \mathbf{x}\right\|^2\right\}$, we consider the following blurred (filtered) version:

$$\text{blur MSE} = \frac{1}{N}\mathscr{E}\left\{\left\|\mathbf{H}\mathbf{F}(\mathbf{y}) - \mathbf{H}_0\mathbf{x}\right\|^2\right\} \tag{3}$$

as the objective functional to be minimized as shown in the next section, where $\mathbf{H}$ and $\mathbf{H}_0$ are the estimated (tentative) and the unknown true PSF, respectively, Thus, our purpose of is to find $\hat{\mathbf{H}}$ — the minimizer of the blur MSE, as our estimate of PSF.

Notice that we cannot directly minimize (3) to obtain the estimate $\mathbf{H}$, as $\mathbf{H}_0\mathbf{x}$ is unknown. However, based on the linear model (1), the quantity of (3) can be replaced by a statistical estimate — blur SURE, involving only the measurements $\mathbf{y}$, as summarized as the following theorem.

This work was supported by a General Research Fund CUHK 410110 from the HK Research Grant Council.

[1] SURE: acronym for Stein's Unbiased Risk Estimate.

Theorem 2.1 *Given the linear model (1), the following random variable:*

$$\tilde{\epsilon} = \frac{1}{N}\Big\{\|\mathbf{HF}(\mathbf{y}) - \mathbf{y}\|^2 + 2\sigma^2 \mathrm{div}_{\mathbf{y}}\big(\mathbf{HF}(\mathbf{y})\big)\Big\} - \sigma^2 \quad (4)$$

is an unbiased estimator of the blur MSE defined by (3), i.e. $\mathscr{E}\{\tilde{\epsilon}\} = \frac{1}{N}\mathscr{E}\Big\{\|\mathbf{HF}(\mathbf{y}) - \mathbf{H}_0\mathbf{x}\|^2\Big\}$, *where the divergence* $\mathrm{div}_{\mathbf{y}}\mathbf{u} = \sum_{n=1}^{N} \frac{\partial u_n}{\partial y_n}$ *for* $\forall \mathbf{u} \in \mathbb{R}^N$.

The proof can be completed by substituting $(\mathbf{y} - \mathbf{b})$ for $\mathbf{H}_0\mathbf{x}$ and using Stein's lemma [12]. Refer to [13] for the similar proof of the standard SURE. Now, $\tilde{\epsilon}$, the blur SURE, is our objective functional to be minimized, also a novel criterion for estimating $\mathbf{H}$.

2.2. Generic PSF estimation

Note that the minimization of (3) and (4) requires to specify the function $\mathbf{F}(\mathbf{y})$. We are now going to restrict ourselves to linear processing denoted by $\mathbf{F}(\mathbf{y}) = \mathbf{W}\mathbf{y}$. It is well known that for the linear model (1), the ideal linear processing $\mathbf{W}$ that minimizes $\frac{1}{N}\|\mathbf{W}\mathbf{y} - \mathbf{x}\|^2$ is Wiener filtering, given by:

$$W(\omega) = \frac{H_0^*(\omega)}{|H_0(\omega)|^2 + \sigma^2/S(\omega)} \quad (5)$$

in the Fourier domain, where $S(\omega)$ is the power spectrum density of $\mathbf{x}$ [14]. Then, based on Wiener filter as (5), we obtain the following theorem.

Theorem 2.2 *Consider the minimization of the blur MSE over tentative* $\mathbf{H}$*:*

$$\min_{\mathbf{H}} \frac{1}{N}\mathscr{E}\Big\{\|\mathbf{H}\mathbf{W}_{\mathbf{H}}\mathbf{y} - \mathbf{H}_0\mathbf{x}\|^2\Big\} \quad (6)$$

where $\mathbf{W}_{\mathbf{H}}$ *is defined as* $W_H(\omega) = \frac{H^*(\omega)}{|H(\omega)|^2 + \sigma^2/S(\omega)}$ *in Fourier domain. Then, the minimizer* $\hat{H}(\omega)$ *of (6) satisfies:*

$$U(\omega) = \frac{|H_0(\omega)|^2}{|H_0(\omega)|^2 + \sigma^2/S(\omega)} = \frac{|\hat{H}(\omega)|^2}{|\hat{H}(\omega)|^2 + \sigma^2/S(\omega)} \quad \text{for } \forall\omega \quad (7)$$

which implies that $\hat{H}(\omega) = H_0(\omega)$ *for* $\forall\omega$.

The theorem can be easily proved by Wiener theory for performing denoising/deconvolution [14]. Let us call $U(\omega)$ in (7) as "frequency-band indicator". This theorem shows the essential equivalence between minimization of the blur MSE (6) and equality of band indicator (7).

To conclude, it has been justified that the minimizer $\hat{H}(\omega)$ of the blur MSE is exactly the same as the latent true $H_0(\omega)$, which enables us to safely minimize $\tilde{\epsilon}$ — a good substitute for the blur MSE, to estimate $\mathbf{H}$.

2.3. Approximation of the Wiener filter

However, $S(\omega)$ in (7) is unknown in practice. If $\sigma^2/S(\omega)$ is replaced by a constant λ, such that $\frac{|H_0(\omega)|^2}{|H_0(\omega)|^2+\lambda}$ is a good approximation of the band indicator of (7) for $\forall\omega$, we obtain the approximated Wiener filter $\mathbf{W}_{\mathbf{H},\lambda}$ as:

$$\mathbf{W}_{\mathbf{H},\lambda} = (\mathbf{H}^{\mathrm{T}}\mathbf{H} + \lambda\mathbf{I})^{-1}\mathbf{H}^{\mathrm{T}} \qquad \text{i.e. } W_{H,\lambda}(\omega) = \frac{H^*(\omega)}{|H(\omega)|^2 + \lambda} \quad (8)$$

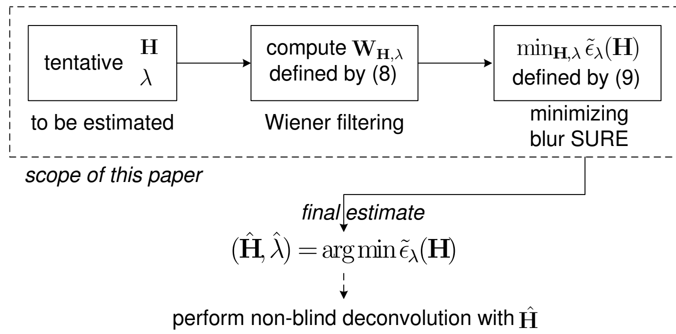

Fig. 1. The procedure of generic PSF estimation: joint minimization of blur SURE over $\mathbf{H}$ and λ, as shown in (9).

Substituting $\mathbf{F}(\mathbf{y}) = \mathbf{W}_{\mathbf{H},\lambda}\mathbf{y}$ into (4), we formulate PSF estimation as the following joint minimization problem:

$$\min_{\mathbf{H},\lambda} \tilde{\epsilon}_\lambda(\mathbf{H}) = \min_{\mathbf{H},\lambda} \frac{1}{N}\Big\{\|\mathbf{H}\mathbf{W}_{\mathbf{H},\lambda}\mathbf{y} - \mathbf{y}\|^2 + 2\sigma^2\mathrm{Tr}\big(\mathbf{H}\mathbf{W}_{\mathbf{H},\lambda}\big)\Big\} - \sigma^2 \quad (9)$$

Note that the corresponding approximated band indicator with $W_{H,\lambda}(\omega)$ is given by;

$$U_{\mathrm{appro}}(\omega) = H(\omega)W_{H,\lambda}(\omega) = \frac{|H(\omega)|^2}{|H(\omega)|^2 + \lambda} \quad (10)$$

The potential hypothesis of our approach is that the bandwidth of $H_0(\omega)$ is narrower than that of $S(\omega)$, so that $\sigma^2/S(\omega)$ can be approximated by the estimated $\hat{\lambda}$, which makes (10) to be a good approximation of the accurate band indicator (7). Thus, the procedure of PSF estimation can be summarized as Fig.1.

3. ESTIMATION OF VARIANCE OF GAUSSIAN BLUR BASED ON OPTIMAL WIENER FILTER

By taking Gaussian blur for example, this section is to exemplify the proposed approach for estimating PSF in Section 2.

3.1. Formulation of estimating variance of Gaussian blur

Denote Gaussian function with variance s^2 by $\mathbf{h}_s(i,j) = C \cdot \exp\Big(-\frac{i^2+j^2}{2s^2}\Big)$, where (i,j) denotes the 2-D coordinates, C is normalized coefficient such that $\sum_{i,j}\mathbf{h}_s(i,j) = 1$, also denote the corresponding convolution matrix of $\mathbf{h}_s$ by $\mathbf{H}_s$. Similar to (9), the estimation of blur variance s^2 is formulated as the following joint minimization problem:

$$\min_{s,\lambda} \tilde{\epsilon}_\lambda(s) = \min_{s,\lambda} \frac{1}{N}\|\mathbf{H}_s(\mathbf{H}_s^{\mathrm{T}}\mathbf{H}_s + \lambda\mathbf{I})^{-1}\mathbf{H}_s^{\mathrm{T}}\mathbf{y} - \mathbf{y}\|^2 + \frac{2\sigma^2}{N}\mathrm{Tr}\big(\mathbf{H}_s(\mathbf{H}_s^{\mathrm{T}}\mathbf{H}_s + \lambda\mathbf{I})^{-1}\mathbf{H}_s^{\mathrm{T}}\big) - \sigma^2 \quad (11)$$

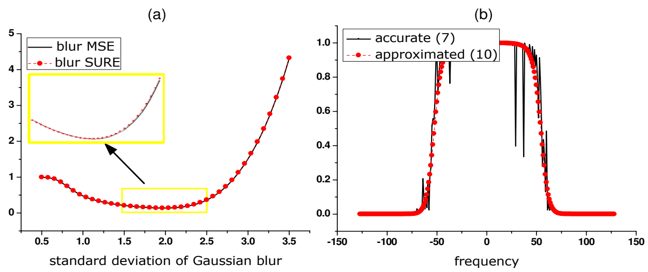

Fig. 2. Example of *Cameraman* blurred by true $\mathbf{H}_0$ with $s_0 = 2$ and noise level $\sigma^2 = 1$: the estimated $\hat{s} = 1.98$ and $\hat{\lambda} = 8 \times 10^{-4}$.

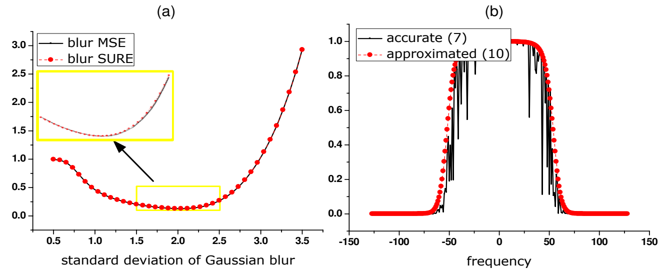

Fig. 3. Example of *Lena* blurred by true $\mathbf{H}_0$ with $s_0 = 2$ and and noise level $\sigma^2 = 1$: the estimated $\hat{s} = 2.04$ and $\hat{\lambda} = 1.2 \times 10^{-3}$.

3.2. Line search for alternating minimization

We can use the alternating minimization to solve (11), and the evolution of the algorithm can be described as:

$$\cdots \to \lambda^{(k)} = \arg\min_{\lambda} \tilde{\epsilon}_\lambda(s^{(k)}) \to s^{(k+1)} = \arg\min_{s} \tilde{\epsilon}_{\lambda^{(k)}}(s) \to \cdots \quad (12)$$

As the minimization problem involves only 2 variables: blur size s and Wiener parameter λ, the minimization in each direction can be efficiently performed by line search, until the algorithm reaches the convergence. The details of the alternating minimization are shown in Algorithm 1.

Algorithm 1 : Alternating Minimization Algorithm

Input: $\tilde{\epsilon}_\lambda(s)$: objective function given as (11);
Output: optimal $\hat{\lambda}$ and $\hat{s}$

1: initialize $s^{(0)}$;
2: **repeat** by $k := k+1$
3: given $s^{(k)}$, line search for $\lambda^{(k)} = \arg\min_\lambda \tilde{\epsilon}_\lambda\left(s^{(k)}\right)$;
4: given $\lambda^{(k)}$, line search for $s^{(k+1)} = \arg\min_s \tilde{\epsilon}_{\lambda^{(k)}}(s)$;
5: **until** $\left|s^{(k+1)} - s^{(k)}\right| \le \delta_1$ and $\left|\lambda^{(k+1)} - \lambda^{(k)}\right| \le \delta_2$ for some δ_1 and δ_2.

Figures (2-3) show that the estimated $\hat{\lambda}$ by minimizing (11) yields both the highly accurate approximation of the band indicator (7) and the accurate estimate $\hat{s} \approx s_0$. Sub-figures (a) show that both the blur MSE and SURE reach their minimums at the estimated $\hat{s} \approx s_0$; Sub-figures (b) show the good approximation as (10) of band indicator to the accurate one as (7) with the estimated $\hat{\lambda}$. The zoom-in of $s \in (1.5, 2.5)$ denoted by the yellow box is to show the really unique minimum of blur SURE. The figures also demonstrate the essential equivalence between approximation of band indicator and PSF estimation.

3.3. Additional remark

Note that the main purpose of this paper is to estimate the variance of Gaussian blur, rather than simultaneously estimating both PSF and original image, as emphasized in Fig.1. After estimating PSF, we can use existing non-blind deconvolution methods to obtain the deconvolved data with the estimated PSF. In this paper, we use the SURE-LET approach to perform the non-blind deconvolution [11]: the whole processing is linearly parametrized by using multiple Wiener filterings as elementary functions, followed by undecimated Haar-wavelet thresholding. The linear coefficients are obtained by minimizing standard SURE [13]. Refer to [11] for more details.

4. EXPERIMENTAL RESULTS

We perform experiments on 2 standard test images: *Lena* and *Cameraman* [2], blurred by Gaussian PSF with variances 9 and 5, respectively, and corrupted by Gaussian noise corresponding to BSNR = 40dB and 20dB [3] [1, 9].

4.1. The accuracy of our estimated blur variance

Table 1. The estimated $\hat{s}$ of the proposed method

true std	$s_0 = 3$		$s_0 = 2.24$	
BSNR	40dB	20dB	40dB	20dB
Lena	3.05	3.07	2.33	2.32
Cameraman	3.00	2.96	2.31	2.31

Table 1 reports the estimated $\hat{s}$. We can see that the estimated $\hat{s}$ is very close to the latent true s_0. One might wonder that how the non-blind deconvolution algorithm is sensitive to the accuracy of the PSF. We would like to note that for our proposed SURE-LET approach [11], the loss of using the estimated $\hat{\mathbf{H}}$ reported in this table, compared to that with accurate $\mathbf{H}_0$, is within 0.2dB in terms of PSNR of the deconvolution performance. It implies that with the estimated PSF, we can achieve comparable deconvolution performance to the non-blind version of SURE-LET approach.

4.2. Comparison with the state of the art

The results shown in Tables 2–3 are compared to the state-of-the-art blind deconvolution methods: TV-based algorithm [2], variational Bayesian algorithms: SAR1 and SAR2 [3], TV1 and TV2 [1], StStSt [9]. The performance

[2] The two images are shown in [1, 9]

[3] The blurred SNR is defined as: BSNR = $10 \log_{10} \frac{\mathrm{var}(\mathbf{Hx})}{\sigma^2}$ [1, 3]

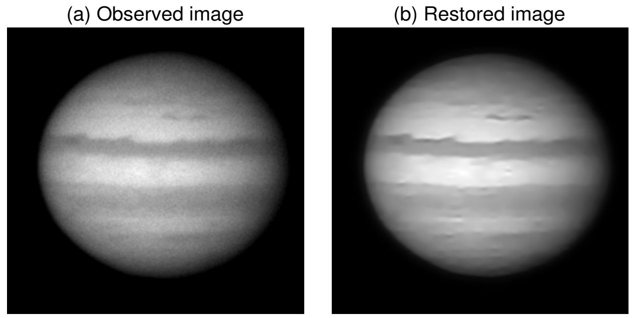

Fig. 4. Restoration of *Jupiter*: the estimated noise std is $\sigma = 4.68$ by using MAD (median absolute deviation) [15], the estimated Gaussian blur std is $\hat{s} = 2.41$.

is measured by Improved SNR [4]. We can observe that the proposed method outperforms the other methods.

Table 2. Improved SNR (in dB) under Gaussian blur with variance 9

Method	SAR1 [3]	SAR2 [3]	TV1 [1]	TV2 [1]	ours
BSNR	40dB				
Lena	1.35	1.43	2.53	2.59	**4.54**
Cameraman	1.03	1.01	1.82	1.73	**3.15**
BSNR	20dB				
Lena	1.62	-11.32	2.62	-32.50	**3.13**
Cameraman	1.16	-8.83	1.70	-40.89	**2.15**

Table 3. Improved SNR (in dB) on *Cameraman* under Gaussian blur with variance 5

BSNR	Method [2]	StStSt [9]	ours
40dB	1.32	2.82	**3.42**
20dB	1.17	1.57	**2.35**

4.3. Application to real astronomical images

In our last set of experiments, the method is applied to a real image of *Jupiter*, shown in Fig.4-(a). There is no exact expression of the PSF for this image; however, as suggested in [1, 3, 5, 6], the PSF can be well approximated by Gaussian function. Fig.4-(b) shows the restored image using Gaussian kernel with $\hat{s} = 2.41$, estimated by our algorithm.

5. CONCLUSIONS

In this paper, we proposed a new blind deconvolution method to estimate Gaussian blur variance based on a new criterion — blur SURE: a statistical estimate of blur MSE. The optimal Wiener filtering, as the linear solution of the minimization of blur MSE, can be well approximated using an optimal regularization parameter, which is jointly estimated by minimizing blur SURE, along with blur variance.

Results obtained show that the proposed method has significant improvement of quality both numerically and visually. Compared to the other methods of estimating blur variance, the main advantage of our approach is that 1) the accuracy of the estimated blur variance has been justified by the blur SURE: a novel objective functional; 2) it does not need to find particular edges and measure the zero-crossing of their particular responses to particular filter as [7, 8] did, which is easily affected by the noise corruption.

Although the discussion of this paper, limited to Gaussian blur, is but an exemplification of SURE-type approach to blind deconvolution problem, it is worth noting that SURE-type minimization itself does not specify any particular parametric form of PSF. There is huge potential to develop specific algorithms for various application, e.g. fluorescence microscopy [4], based on SURE-type principle.

[4]Improved SNR is defined as ISNR = $10\log_{10}\frac{\|\mathbf{y}-\mathbf{x}\|^2}{\|\hat{\mathbf{x}}-\mathbf{x}\|^2}$ [3, 9].

6. REFERENCES

[1] S.D. Babacan, R. Molina, and A.K. Katsaggelos, “Variational bayesian blind deconvolution using a total variation prior,” *Image Processing, IEEE Transactions on*, vol. 18, no. 1, pp. 12–26, 2009.

[2] T.F. Chan and C.K. Wong, “Total variation blind deconvolution,” *Image Processing, IEEE Transactions on*, vol. 7, no. 3, pp. 370–375, 1998.

[3] R. Molina, J. Mateos, and A.K. Katsaggelos, “Blind deconvolution using a variational approach to parameter, image, and blur estimation,” *Image Processing, IEEE Transactions on*, vol. 15, no. 12, pp. 3715–3727, 2006.

[4] P. Pankajakshan, L. Blanc-Féraud, B. Zhang, Z. Kam, J.C. Olivo-Marin, J. Zerubia, et al., “Parametric blind deconvolution for confocal laser scanning microscopy -proof of concept,” 2008.

[5] AFJ Moffat, “A theoretical investigation of focal stellar images in the photographic emulsion and application to photographic photometry,” *Astronomy and Astrophysics*, vol. 3, pp. 455–461, 1969.

[6] 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,” *Optical Engineering*, vol. 32, pp. 2598–2607, 1993.

[7] S. Bae and F. Durand, “Defocus magnification,” in *Computer Graphics Forum*. Wiley Online Library, 2007, vol. 26, pp. 571–579.

[8] J.H. Elder and S.W. Zucker, “Local scale control for edge detection and blur estimation,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 20, no. 7, pp. 699–716, 1998.

[9] D.G. Tzikas, A.C. Likas, and N.P. Galatsanos, “Variational bayesian sparse kernel-based blind image deconvolution with student’s-t priors,” *Image Processing, IEEE Transactions on*, vol. 18, no. 4, pp. 753–764, 2009.

[10] T. Blu and F. Luisier, “The SURE-LET approach to image denoising,” *Image Processing, IEEE Transactions on*, vol. 16, no. 11, pp. 2778–2786, 2007.

[11] F. Xue, F. Luisier, and T. Blu, “SURE-LET image deconvolution using multiple wiener filters,” *Image Processing, 2012 IEEE International Conference on,* accepted.

[12] C.M. Stein, “Estimation of the mean of a multivariate normal distribution,” *The annals of Statistics*, pp. 1135–1151, 1981.

[13] Y.C. Eldar, “Generalized SURE for exponential families: Applications to regularization,” *Signal Processing, IEEE Transactions on*, vol. 57, no. 2, pp. 471–481, 2009.

[14] N. Wiener, “Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications,” *Journal of the American Statistical Association*, vol. 47, no. 258, 1949.

[15] D.L. Donoho and J.M. Johnstone, “Ideal spatial adaptation by wavelet shrinkage,” *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.