SURE-LET IMAGE DECONVOLUTION USING MULTIPLE WIENER FILTERS

Feng Xue[1], Florian Luisier[2] and Thierry Blu[1]

[1]Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong
[2]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA

ABSTRACT

We propose a novel deconvolution algorithm based on the minimization of Stein's unbiased risk estimate (SURE). We linearly parametrize the deconvolution process by using multiple Wiener filterings as elementary functions, followed by undecimated Haar-wavelet thresholding. The key contributions of our approach are: 1) the linear combination of several Wiener filters with different (but fixed) regularization parameters, which avoids the manual adjustment of a single nonlinear parameter; 2) the use of linear parameterization, which makes the SURE minimization finally boil down to solving a linear system of equations, leading to a very fast and exact optimization of the whole deconvolution process.

The results obtained on standard test images show that our algorithm favorably compares with the other state-of-the-art deconvolution methods in both speed and quality.

Index Terms— Deconvolution, SURE minimization, linear parametrization, Wiener filtering, undecimated Haar wavelet thresholding

1. INTRODUCTION

Image deconvolution is a standard linear inverse problem with various applications [1]. Of most practical interest is that singular or ill-conditioned convolution operator yields highly noise-sensitive solution, which makes deconvolution particularly challenging [2].

To cope with the ill-posedness of the problem, a large number of techniques have been developed, by incorporating some regularization term to express *a priori* knowledge of the original image [3]: e.g. Tikhonov regularization [4], piecewise smoothness [1] and sparsity in transform domain [5]. Particularly, the sparse representation can be further enhanced by, for example, block matching technique [6], which currently achieves the state-of-the-art deconvolution performance.

Recently, we have developed an alternative approach for denoising application [7], based on Stein's unbiased risk estimate (SURE) — an unbiased estimator of the Mean Squared Error (MSE) [8]. In the context of image deconvolution, SURE can be used 1) to optimize non-linear parameter in the estimator [9], 2) to monitor the SNR improvement in iterative

algorithms [10], 3) as a minimization criterion for designing the estimators [2]. This paper follows the last point.

One of the main contributions of our work is to overcome the difficulty of adjusting a nonlinear regularization parameter, by linearly parametrizing the deconvolution processing. This is achieved by considering the linear combination of a number of Wiener filterings with different (but fixed) regularization parameters, optimized by SURE minimization. Due to the quadratic nature of SURE and the linear parameterization of the processing, the deconvolution problem finally boils down to solving a linear system of equations. Its solution, i.e. the linear coefficients of these Wiener filterings, automatically constitutes the combination with minimum MSE as our final estimate.

This paper is organized as follows: first, we introduce the theoretical background of SURE-LET principle to deconvolution problem; then, we exemplify the principle by describing a typical structure of the elementary functions, i.e. multiple Wiener filterings with different regularization parameters followed by undecimated wavelet thresholding; finally, we present our experimental results and draw the conclusion.

2. SURE-LET PRINCIPLE TO DECONVOLUTION PROBLEM

2.1. Problem statement

Consider the linear observation model:

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

where $\mathbf{y} \in \mathbb{R}^N$ is the collection of the N pixels of the measured image, $\mathbf{H} \in \mathbb{R}^{N \times N}$ is a distortion matrix (typically a convolution), $\mathbf{x} \in \mathbb{R}^N$ is the original unknown image, $\mathbf{b} \in \mathbb{R}^N$ is additive white Gaussian noise with variance σ^2.

Our approach to estimate $\mathbf{x}$ consists in finding a function (or processing) of the measured data $\mathbf{F}(\mathbf{y}) = [f_1(\mathbf{y}), f_2(\mathbf{y}), \ldots, f_N(\mathbf{y})]^T \in \mathbb{R}^N$ such that the MSE (or actually, an estimate of this MSE, typically the SURE):

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

is minimal, where $\mathscr{E}\{\cdot\}$ denotes expected value. We would like to insist here that the estimate $\hat{\mathbf{x}} = \mathbf{F}(\mathbf{y})$ is only the outcome of the processing. The main spirit of our approach is to estimate

the function $\mathbf{F}$ that transforms $\mathbf{y}$ into $\hat{\mathbf{x}}$, indeed, rather than the solution $\hat{\mathbf{x}}$ itself.

2.2. Convolution SURE

Notice that we cannot directly minimize 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 Gaussian noise assumption, the value of the MSE can be replaced by a statistical estimate — Stein's Unbiased Risk Estimate (SURE), involving only the measurements $\mathbf{y}$, which is summarized in the following theorem.

Theorem 2.1 *Given the linear model (1), assuming $f_n(\mathbf{y})$ is at least weakly differentiable , the following random variable:*

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

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

This theorem is a natural extension of Stein's lemma [8–10] to the distortion model (1). Considering the possible ill-posedness of the matrix $\mathbf{H}$, we approximate $\mathbf{H}^{-1}$ by a Tikhonov regularized inverse [4]:

$$\mathbf{H}_\beta^{-1} = \left(\mathbf{H}^T\mathbf{H} + \beta\mathbf{S}^T\mathbf{S}\right)^{-1}\mathbf{H}^T \tag{3}$$

for some parameter β and matrix $\mathbf{S}$, to stabilize ϵ. Thus, we replace the SURE by a stabilized version [10]:

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

where the last term $\frac{1}{N}\|\mathbf{x}\|^2$ is a constant w.r.t. $\mathbf{F}(\mathbf{y})$, and is thus irrelevant for minimizing ϵ_β. For detailed derivation of (4), refer to [10], which has also taken the similar stabilization strategy. Empirically, we choose $\mathbf{S}$ as the discrete Laplacian operator and $\beta = 1 \times 10^{-5}\sigma^2$ to achieve a good balance between the approximation accuracy and the stability of ϵ_β.

2.3. Linear parametrization of $\mathbf{F}(\mathbf{y})$ — LET

Now, we can minimize the stabilized SURE denoted by ϵ_β (4) in practice, instead of minimizing the MSE (2). The next question naturally arises: how to choose the function $\mathbf{F}$, such that $\mathbf{F}(\mathbf{y})$ is sufficiently close to $\mathbf{x}$? Here, we adopt the LET (Linear Expansion of Thresholds) strategy [7] to deconvolution problems, by parametrizing $\mathbf{F}(\mathbf{y})$ as a linear combination of a small number of known basic processings $\mathbf{F}_k(\mathbf{y}) \in \mathbb{R}^N$ for $k = 1, 2, ..., K$, i.e.

$$\mathbf{F}(\mathbf{y}) = \sum_{k=1}^{K} a_k\mathbf{F}_k(\mathbf{y}) = \mathbf{Fa} \tag{5}$$

where K is the number of linear coefficients a_k to be determined, and generally $K \ll N$. $\mathbf{Fa}$ is the matrix notation, where $\mathbf{F} = [\mathbf{F}_1(\mathbf{y}), \mathbf{F}_2(\mathbf{y}), ..., \mathbf{F}_K(\mathbf{y})] \in \mathbb{R}^{N \times K}$, $\mathbf{a} = [a_1, a_2, ..., a_K]^T \in \mathbb{R}^K$. This parametrization dramatically reduces the problem size from pixel number N to the number of basis functions K. The

underlying hypothesis here is that the processing $\mathbf{F}$ does not change "too much" when the input $\mathbf{y}$ changes "slowly", such that $\mathbf{F}(\mathbf{y})$ can be represented by only a few number of basis functions as (5).

2.4. SURE-LET optimization

Now, the deconvolution problem essentially amounts to finding the linear coefficients a_k in (5) such that ϵ_β is minimized. Substituting (5) into (4), and performing the derivative of ϵ_β w.r.t. a_k: $\frac{1}{2}\frac{\partial\epsilon_\beta}{\partial a_k} = 0$, we obtain the following linear system of equations:

$$\sum_{k'=1}^{K}\left(\underbrace{\mathbf{F}_k^T(\mathbf{y})\mathbf{F}_{k'}(\mathbf{y})}_{\mathbf{M}_{k,k'}}\right)a_{k'} - \left(\underbrace{\mathbf{y}^T\mathbf{H}_\beta^{-T}\mathbf{F}_k(\mathbf{y}) - \sigma^2\mathrm{div}(\mathbf{H}_\beta^{-T}\mathbf{F}_k(\mathbf{y}))}_{c_k}\right) = 0 \tag{6}$$

Eq.(6) can be summarized in matrix form as $\mathbf{Ma} = \mathbf{c}$, where $\mathbf{M} = \mathbf{F}^T\mathbf{F} \in \mathbb{R}^{K \times K}$ and $\mathbf{c} = [c_1, c_2, ..., c_K]^T \in \mathbb{R}^K$. Due to the possible singularity of $\mathbf{M}$, we use the pseudo-inverse of $\mathbf{M}$ to solve the linear system of equations.

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 be verified by comparing it with *MSE-LET*.

3. CONSTRUCTION OF ELEMENTARY FUNCTION

3.1. Multi-Wiener deconvolutions

To exemplify the SURE-LET approach to deconvolution problems, we begin with simple Wiener deconvolutions with parameters λ_k, for $k = 1, 2, ..., K$:

$$\mathbf{F}_k(\mathbf{y}) = \left(\mathbf{H}^T\mathbf{H} + \lambda_k\mathbf{S}^T\mathbf{S}\right)^{-1}\mathbf{H}^T\mathbf{y} \stackrel{\mathrm{def}}{=\!=} \mathbf{H}_{\lambda_k}^{-1}\mathbf{y} \tag{7}$$

as the elementary functions, where $\mathbf{S}$ is the discrete Laplacian operator. Thus, the LET representation (5) is a linear combination of Wiener filters with different λ_k, whose weights a_k are obtained by solving (6). Fig.1 shows an example of taking three Wiener filters: we can see that different values of λ_k capture different details and features of the image. The SURE-optimized linear coefficients a_k will produce the combined estimate with a similar balance between noise reduction and edge preservation, as a single optimized Wiener filter. The advantage of the LET approach in this example is that there is no non-linear regularization parameter tuning.

3.2. Multi-Wiener wavelet-thresholding deconvolution

Further developing the structure of 3.1, we consider multiple Wiener filterings followed by thresholding in the undecimated Haar wavelet domain as the elementary function, shown in Fig.2. Denoting the decomposition and reconstruction of undecimated wavelet transform by $\mathbf{D} =$
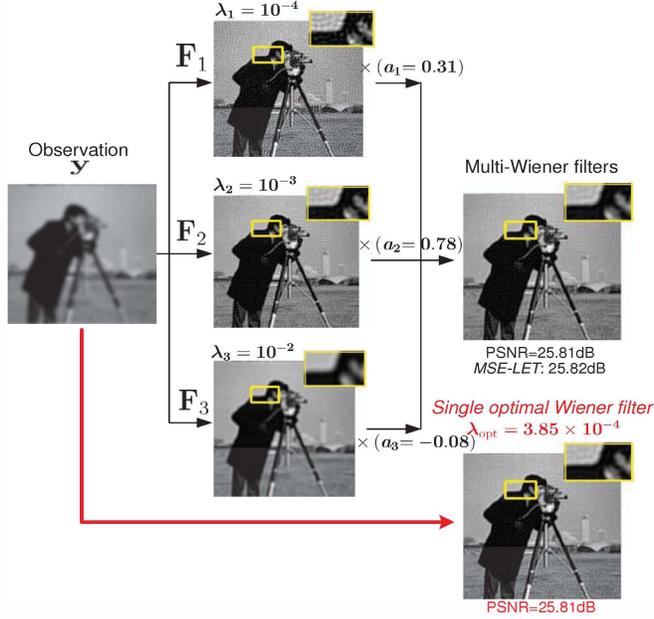
Fig. 1. Example of linearly combining three Wiener filters with regularization parameters $\lambda_1 = 1 \times 10^{-4}$, $\lambda_2 = 1 \times 10^{-3}$, $\lambda_3 = 1 \times 10^{-2}$, balanced by their weights $a_1 = 0.31$, $a_2 = 0.78$, $a_3 = -0.08$ — the solution to (6). The linearly-combined estimate is nearly equivalent to single Wiener filtering with optimal $\lambda = 3.85 \times 10^{-4}$ in terms of PSNR.

$\left[\mathbf{D}_1^T, \mathbf{D}_2^T, ..., \mathbf{D}_J^T, \mathbf{D}_{J+1}^T\right]^T$ and $\mathbf{R} = [\mathbf{R}_1, \mathbf{R}_2, ..., \mathbf{R}_J, \mathbf{R}_{J+1}]$, respectively ($J+1$ refers to the low-pass wavelet subband, which is not thresholded), the LET representation (5) is expressed as:

$$\mathbf{F}(\mathbf{y}) = \underbrace{\sum_{m=1}^{M} \sum_{l=1}^{L} \sum_{j=1}^{J} a_{m,l,j} \mathbf{R}_j \theta_l\left(\mathbf{w}_{m,j}\right)}_{\text{highpass subbands}} + \underbrace{\sum_{m=1}^{M} a_{m,J+1} \mathbf{R}_{J+1} \mathbf{w}_{m,J+1}}_{\text{lowpass subbands}} \quad (8)$$

where the wavelet coefficients $\mathbf{w}_{m,j} = \mathbf{D}_j \mathbf{H}_{\lambda_m}^{-1} \mathbf{y}$ in j-th subband of m-th Wiener deconvolution. M, L and J refer to the number of Wiener regularizations (typically $M = 3$), elementary non-linear point-wise thresholding functions $\theta_l(\cdot)$ (typically $L = 2$), and high-pass wavelet subbands (typically $J = 9$ for three decomposition levels), respectively. In this paper, we choose the following thresholding functions:

$$\begin{cases} \theta_1(\mathbf{w}_{m,j}) = \mathbf{w}_{m,j}\left\{1 - \exp\left(-\left(\frac{\mathbf{w}_{m,j}}{4\sigma_{m,j}}\right)^4\right)\right\} \\ \theta_2(\mathbf{w}_{m,j}) = \mathbf{w}_{m,j}\left\{1 - \exp\left(-\left(\frac{\mathbf{w}_{m,j}}{9\sigma_{m,j}}\right)^4\right)\right\} \end{cases} \quad (9)$$

where $\sigma_{m,j}^2$ stands for the variance of the colored noise $\mathbf{b}_{m,j} = \mathbf{D}_j \mathbf{H}_{\lambda_m}^{-1} \mathbf{b}$: $\sigma_{m,j}^2 = \text{Tr}(\mathbf{H}_{\lambda_m}^{-T} \mathbf{D}_j^T \mathbf{D}_j \mathbf{H}_{\lambda_m}^{-1})$.

As (8) indicated, we have $K = MJL + M$ elementary functions in all and so, as many weights $a_{m,l,j}$ to determine. Again, the SURE minimization finally boils down to solving a linear system of equations of order K.
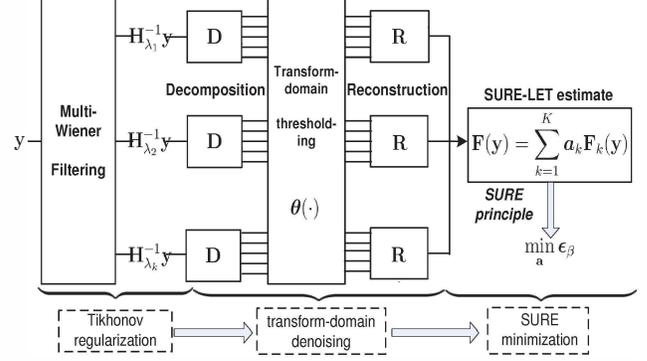

Fig. 2. Flowchart of SURE-LET approach: multi-Wiener filtering followed by transform-domain thresholding.

4. EXPERIMENTAL RESULTS

4.1. Parameter setting of the proposed method

In this section, we use the structure of Eq.(8) and Fig.2 as the elementary function $\mathbf{F}_k(\mathbf{y})$. We apply $M = 3$ Wiener filters with $\lambda_1 = 1 \times 10^{-4}\sigma^2$, $\lambda_2 = 1 \times 10^{-3}\sigma^2$, $\lambda_3 = 3 \times 10^{-3}\sigma^2$ and undecimated Haar wavelet transform. The thresholding functions are specified as (9). Typically, if we use $J = 9$ (i.e. 3 decomposition level), we will have $K = 3 \times 2 \times 9 + 3 = 57$ coefficients to be found via solving (6).

4.2. Comparison with the state of the art

We perform experiments on two test images (other images also show the same trends): *House* (256×256) and *Mixture* (512×512) [1], under three noise levels ($\sigma = 1, 10, 50$) and five benchmark convolution kernels [1,6]: 1) Gaussian filter with std 3; 2) Rational filter $h(i, j) = (1 + i^2 + j^2)^{-1}$ for $i, j = -7, ..., 7$; 3) Separable filter 5×5 filter with weights $[1, 4, 6, 4, 1]/16$ along both horizontal and vertical directions; 4) 5×5 uniform blur; 5) 9×9 uniform blur. The experimental performance is measured by PSNR [2]. The results shown in Tables 1–2 are compared with the state-of-the-art deconvolution methods: BM3D [6], C-SALSA [5] and TVMM [1]. Fig.3 shows a visual result. Note that all the PSNR results (in dB) reported below have been averaged over 10 noise realizations, and the best scores within a 0.1dB margin are highlighted in Tables 1–2.

It can be seen from the reported results that the proposed SURE-LET algorithm uniformly outperforms all of the other techniques. We also wish to note that our algorithm is very robust to a wide range of noise levels, and particularly, significant improvements are frequently observed for large noise variance over other algorithms. The tables also show in italic the result of minimizing the MSE in lieu of the SURE, demon-

---

[1] All of the images are available at: http://decsai.ugr.es/cvg/CG/base.htm
[2] For 8-bit images, we define PSNR = $10\log_{10} \frac{255^2}{\|\hat{\mathbf{x}} - \mathbf{x}\|^2/N}$ [6,7].

strating that the SURE minimization is a good substitute to the MSE minimization. Table 3 reports the computational time of the deconvolution methods, which shows that our approach is substantially faster than other approaches. We would also like to stress that our implementation uses only standard Matlab code without any compiled routine.

Table 1. Comparison of some state-of-the-art deconvolution methods (under Gaussian blur with std 3)

σ	1	10	50	1	10	50
Image	*House* 256×256			*Mixture* 512×512		
Input	24.22	22.73	13.73	14.84	14.64	11.46
BM3D	29.19	**27.05**	23.95	**16.70**	15.59	14.60
TVMM	**29.45**	26.72	23.08	16.37	15.20	14.46
C-SALSA	29.25	26.55	22.97	16.52	**15.65**	14.70
SURE-LET	**29.40**	**27.10**	**24.20**	**16.73**	**15.74**	**14.85**
MSE−LET	*29.74*	*27.44*	*24.53*	*16.80*	*15.88*	*14.98*

* Best scores within a 0.1dB margin are highlighted.

Table 2. Image *Mixture* deconvolution for various blurs and noise levels

σ	1	10	50	1	10	50
Blur	*Rational filter*			*Separable filter*		
Input	15.96	15.70	11.95	18.38	17.94	12.76
BM3D	28.53	17.25	14.85	26.54	20.04	16.15
TVMM	28.28	17.30	13.78	27.17	20.64	15.25
C-SALSA	27.26	18.04	15.02	26.58	20.16	16.19
SURE-LET	**30.00**	**18.53**	**15.50**	**28.08**	**21.18**	**16.94**
MSE−LET	*30.38*	*18.70*	*15.58*	*28.43*	*21.53*	*17.06*
Blur	5×5 *uniform blur*			9×9 *uniform blur*		
Input	15.92	15.67	11.94	14.58	14.40	11.35
BM3D	25.53	17.90	15.56	20.66	16.01	14.60
TVMM	25.70	18.30	14.52	20.70	15.64	13.66
C-SALSA	24.26	18.15	15.37	20.04	16.30	14.29
SURE-LET	**26.41**	**18.94**	**16.04**	**21.70**	**16.65**	**15.01**
MSE−LET	*26.71*	*19.06*	*16.13*	*21.96*	*16.76*	*15.07*

* Best scores within a 0.1dB margin are highlighted.

Table 3. Comparisons of computational time (units: sec.) *

Degradation scenario	BM3D	TV-MM	C-SALSA	SURE-LET
House 256×256 Gaussian blur, $\sigma = 1$	2.96	146.13	48.04	**1.36**
Mixture 512×512 Rational blur, $\sigma = 10$	13.65	289.14	55.72	**5.34**

* Hardware environment used: Intel(R) Core(TM)2 Duo CPU E7400 @2.80GHz, memory size: 3GB

5. CONCLUSION

In this paper, we have presented a new deconvolution method based on the SURE-LET principle, which linearly parametrizes the processing and minimizes SURE — an estimate of the resulting MSE, to find the best linear combination of the basis functions. The originality of our approach is to use multiple Wiener filterings with different but fixed regularization parameters, to avoid empirical adjustment.

Note that the proposed algorithm is but a simple exemplification of the SURE-LET principle, where we did not try to take advantage of all the degrees of freedom in the design of the elementary functions $\mathbf{F}_k(\mathbf{y})$. And yet, the obtained results are already quite outstanding. The great flexibility and the limited computational cost of the proposed approach allow to develop more sophisticated forms of the basic processings, for example, by taking into account multiscale or spatial dependencies, or electing more directional deconvolutions. It
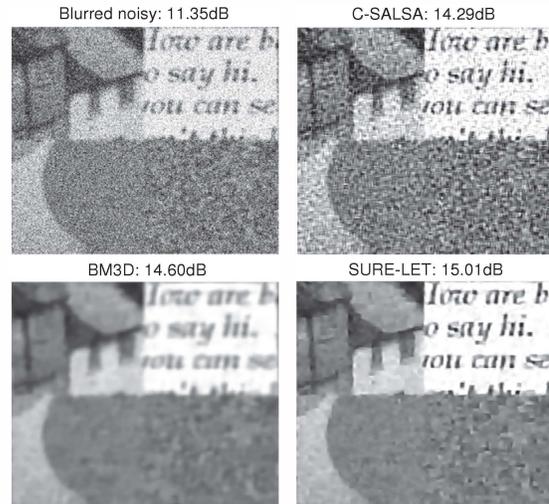

Fig. 3. Restoration of *Mixture* degraded by 9×9 uniform blur with noise std $\sigma = 50$

is likely that there is a substantial margin of improvement of SURE-LET type algorithms.

6. REFERENCES

[1] J.M. Bioucas-Dias, M.A.T. Figueiredo, and J.P. Oliveira, "Total variation-based image deconvolution: a majorization-minimization approach," in *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*. IEEE, 2006, vol. 2, pp. II–II.

[2] J.C. Pesquet, A. Benazza-Benyahia, and C. Chaux, "A SURE approach for digital signal/image deconvolution problems," *Signal Processing, IEEE Transactions on*, vol. 57, no. 12, pp. 4616–4632, 2009.

[3] L. Mancera, S.D. Babacan, R. Molina, and AK Katsaggelos, "Image restoration by mixture modelling of an overcomplete linear representation," in *16th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2009, pp. 3949–3952.

[4] A.N. Tikhonov, V.I.A. Arsenin, and F. John, *Solutions of ill-posed problems*, Winston Washington, DC:, 1977.

[5] M.V. Afonso, J.M. Bioucas-Dias, and M.A.T. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *Image Processing, IEEE Transactions on*, vol. 20, no. 3, pp. 681–695, 2011.

[6] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image restoration by sparse 3D transform-domain collaborative filtering," in *SPIE Electronic Imaging*. Citeseer, 2008, vol. 6812.

[7] 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.

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

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

[10] C. Vonesch, S. Ramani, and M. Unser, "Recursive risk estimation for non-linear image deconvolution with a wavelet-domain sparsity constraint," in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*. IEEE, 2008, pp. 665–668.