Bivariate Shrinkage With Local Variance Estimation

Levent Şendur and Ivan W. Selesnick, *Member, IEEE*

***Abstract*—The performance of image-denoising algorithms using wavelet transforms can be improved significantly by taking into account the statistical dependencies among wavelet coefficients as demonstrated by several algorithms presented in the literature. In two earlier papers by the authors, a simple bivariate shrinkage rule is described using a coefficient and its parent. The performance can also be improved using simple models by estimating model parameters in a local neighborhood. This letter presents a locally adaptive denoising algorithm using the bivariate shrinkage function. The algorithm is illustrated using both the orthogonal and dual tree complex wavelet transforms. Some comparisons with the best available results will be given in order to illustrate the effectiveness of the proposed algorithm.**

***Index Terms*—Bivariate shrinkage, image denoising, statistical modeling, wavelet transforms.**

I. Introduction

SOME RECENT research has addressed the development of statistical models of wavelet coefficients of natural images and application of these models to image denoising [5], [11], [14], [15]. Recently, highly effective yet simple schemes mostly based on soft thresholding have been developed [1], [2], [10]. In [10], the wavelet coefficients are modeled with a Gaussian *a priori* density, and locally adaptive estimation is done for coefficient variances. Also, prior knowledge is taken into account to estimate coefficient variances more accurately. In [1], the interscale dependencies are used to improve the performance. In [2], the simple soft-thresholding idea is used for each of the wavelet subbands, and the threshold value is estimated to minimize the mean-square error.

The models that exploit the dependency between coefficients give better results compared to the ones using an independence assumption [5], [11], [14], [15]. However, some of these models are complicated and result in high computational cost. In [12], a bivariate probability density function (pdf) is proposed to model the statistical dependence between a coefficient and its parent, and the corresponding bivariate shrinkage function is obtained. This new rule maintains the simplicity, efficiency, and intuition of soft thresholding. An explicit multivariate shrinkage function for wavelet denoising is also presented in [16].

In this letter, the local adaptive estimation of necessary parameters for the bivariate shrinkage function will be described. Also, the performance of this system will be demonstrated on both the orthogonal wavelet transform and the dual-tree complex wavelet transform (CWT) [7], [8], and some comparisons with the best available wavelet-based image-denoising results will be given in order to illustrate the effectiveness of the system.

Manuscript received May 3, 2002; revised August 6, 2002. This research was supported by the National Science Foundation under CAREER grant CCR-9875452. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Yu-Hen Hu.

The authors are with the Electrical and Computer Engineering, Polytechnic University, Brooklyn, NY 11201 USA (e-mail: levent@taco.poly.edu; selesi@taco.poly.edu).

Digital Object Identifier 10.1109/LSP.2002.806054

II. Local Adaptive Algorithm

In this letter, the denoising of an image corrupted by additive independent white Gaussian noise with variance σ_n^2 will be considered. Let w_{2k} represent the parent of w_{1k}. (w_{2k} is the wavelet coefficient at the same position as the kth wavelet coefficient w_{1k}, but at the next coarser scale.) We formulate the problem in wavelet domain as $y_{1k} = w_{1k} + n_{1k}$ and $y_{2k} = w_{2k} + n_{2k}$ to take into account the statistical dependencies between a coefficient and its parent. y_{1k} and y_{2k} are noisy observations of w_{1k} and w_{2k}; and n_{1k}, and n_{2k} are noise samples. We can write

$$\mathbf{y}_k = \mathbf{w}_k + \mathbf{n}_k, \qquad k = 1..\text{ no. of wavelet coefs} \tag{1}$$

where $\mathbf{w_k} = (w_{1k}, w_{2k})$, $\mathbf{y_k} = (y_{1k}, y_{2k})$, and $\mathbf{n_k} = (n_{1k}, n_{2k})$. From this point, the coefficient index k will be omitted from the equations in order to improve the readability of the equations.

The standard MAP estimator for $\mathbf{w}$ given the corrupted observation $\mathbf{y}$ is

$$\hat{\mathbf{w}}(\mathbf{y}) = \arg\max_{\mathbf{w}} p_{\mathbf{w}|\mathbf{y}}(\mathbf{w}|\mathbf{y}). \tag{2}$$

After some manipulations, this equation can be written as

$$\hat{\mathbf{w}}(\mathbf{y}) = \arg\max_{\mathbf{w}} \left[p_{\mathbf{n}}(\mathbf{y}-\mathbf{w}) \cdot p_{\mathbf{w}}(\mathbf{w})\right]. \tag{3}$$

In [12], we proposed a non-Gaussian bivariate pdf for the coefficient and its parent as

$$p_{\mathbf{w}}(\mathbf{w}) = \frac{3}{2\pi\sigma^2} \cdot \exp\left(-\frac{\sqrt{3}}{\sigma}\sqrt{w_1^2 + w_2^2}\right). \tag{4}$$

The marginal variance σ^2 is also dependent on the coefficient index k. Using (4) with (3), the MAP estimator of w_1 is derived to be

$$\hat{w}_1 = \frac{\left(\sqrt{y_1^2 + y_2^2} - \frac{\sqrt{3}\sigma_n^2}{\sigma}\right)_+}{\sqrt{y_1^2 + y_2^2}} \cdot y_1 \tag{5}$$

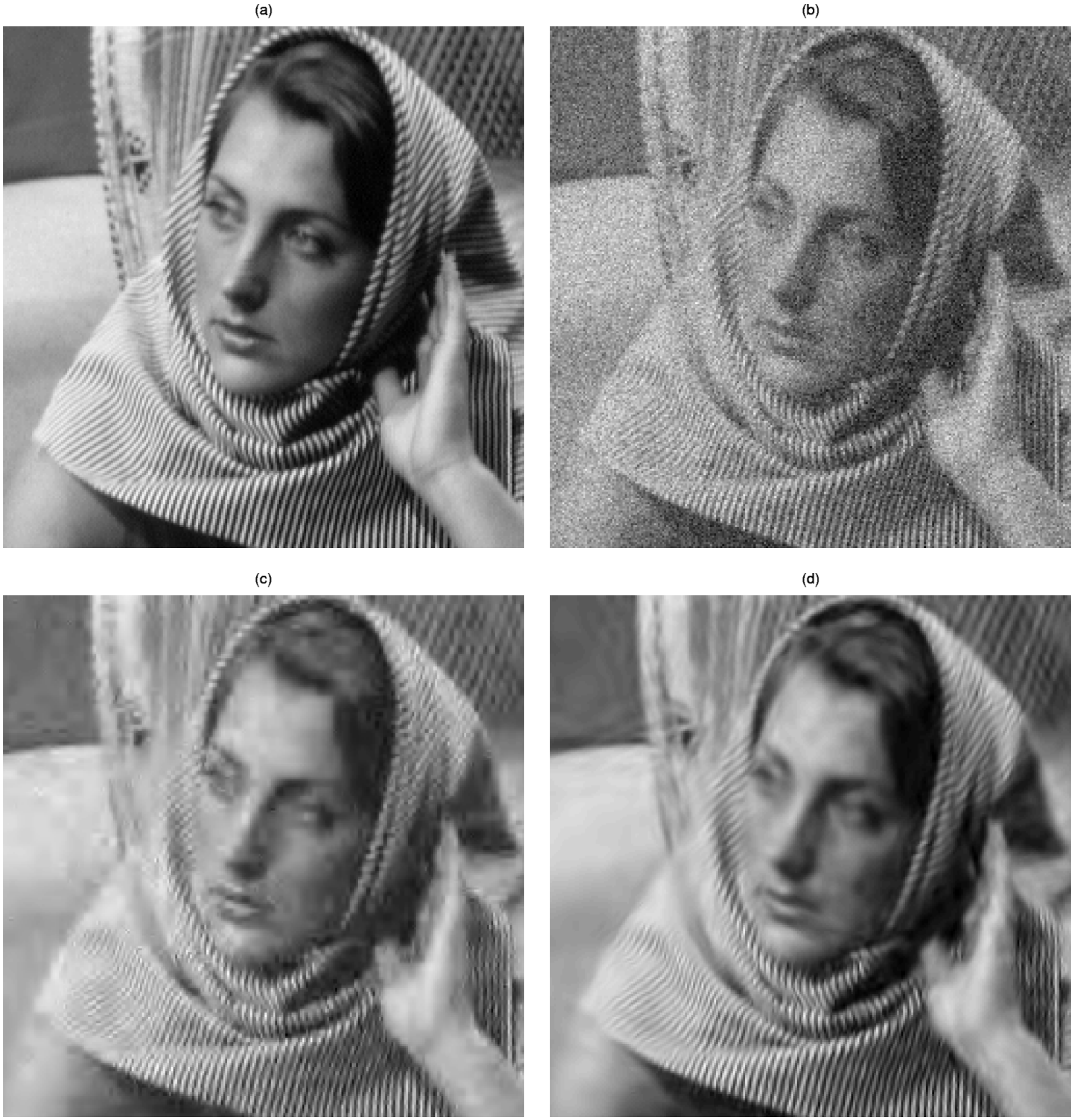

Fig. 1. (a) Original image. (b) Noisy image with PSNR $=$ 20.18 dB ($\sigma_n = 25$). (c) Denoised image using critically sampled transform PSNR $=$ 27.16 dB. (d) Denoised image using dual-tree CWT PSNR $=$ 28.61 dB.

which can be interpreted as a bivariate shrinkage function. Here $(g)_+$ is defined as

$$(g)_+ = \begin{cases} 0, & \text{if } g < 0 \\ g, & \text{otherwise.} \end{cases} \tag{6}$$

This estimator requires the prior knowledge of the noise variance σ_n^2 and the marginal variance σ^2 for each wavelet coefficient. In our algorithm, the marginal variance for the kth coefficient will be estimated using neighboring coefficients in the region $N(k)$. Here $N(k)$ is defined as all coefficients within a square-shaped window that is centered at the kth coefficient as illustrated in Fig. 2.

To estimate the noise variance σ_n^2 from the noisy wavelet coefficients, a robust median estimator is used from the finest scale wavelet coefficients [6]

$$\hat{\sigma}_n^2 = \frac{\text{median}\,(|y_i|)}{0.6745}, \quad y_i \in \text{subband} \quad HH. \tag{7}$$

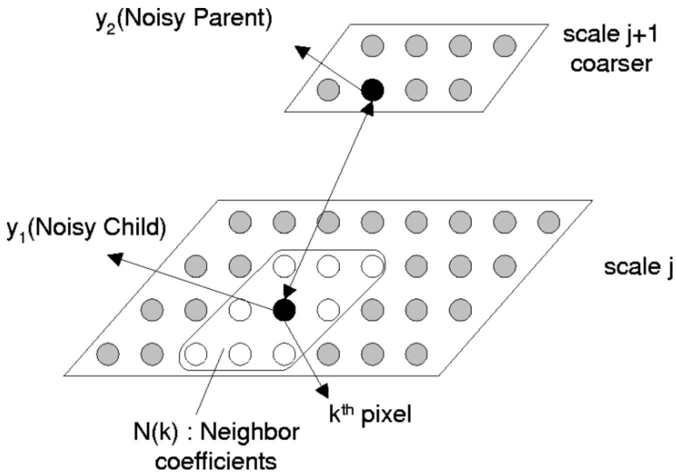

Fig. 2. Illustration of neighborhood $N(k)$.

Let us assume we are trying to estimate the marginal variance σ^2 for the kth wavelet coefficient. From our observation model, one gets $\sigma_y^2 = \sigma^2 + \sigma_n^2$ where σ_y^2 is the marginal variance of noisy observations y_1 and y_2. Since y_1 and y_2 are modeled as zero mean, σ_y^2 can be found empirically by

$$\hat{\sigma}_y^2 = \frac{1}{M} \sum_{y_i \in N(k)} y_i^2 \tag{8}$$

where M is the size of the neighborhood $N(k)$. Then, σ can be estimated as

$$\hat{\sigma} = \sqrt{\left(\hat{\sigma}_y^2 - \hat{\sigma}_n^2\right)_+}. \tag{9}$$

The algorithm is summarized as follows.

1) Calculate the noise variance $\hat{\sigma}_n^2$ using (7).
2) For each each wavelet coefficient ($k = 1..$ number of wavelet coefficients):
   a) Calculate $\hat{\sigma}_y^2$ using (8).
   b) Calculate $\hat{\sigma}$ using (9).
   c) Estimate each coefficient using $\hat{\sigma}$ and $\hat{\sigma}_n^2$ in (5).

III. Results

We compared the proposed algorithm using the orthogonal wavelet transform to other effective systems in the literature, namely BayesShrink [2], AdaptShrink [3], locally adaptive window-based denoising using MAP (LAWMAP) estimator [10], and the hidden Markov tree (HMT) model [5]. We also made a comparison with our subband adaptive algorithm described in [13]. (As seen from the results in Table I, the locally adaptive estimation may improve the performance over the subband adaptive estimation by 1 dB for some images. i.e., Barbara image.) The Daubechies length-eight filter and a 7×7 window size [$N(k)$] are used. (We have also investigated different window sizes. A 5×5 window size can also be a good choice. However, using a 3×3 window size resulted in a slight performance loss. In this letter, we have not considered different shapes for $N(k)$.) The peak signal-to-noise ratio (PSNR) values of these systems are tabulated in Table I. The performance of this system is tested using the PSNR measure. Let s, d denote the original and the denoised image. The rms error is given by

$$\epsilon = \sqrt{\frac{1}{N} \sum_k (s_k - d_k)^2} \tag{10}$$

where N is the number of pixels. The PSNR in decibels is given by

$$\text{PSNR} = 20 \log_{10} \left(\frac{256}{\epsilon} \right). \tag{11}$$

The PSNR values for AdaptShrink are taken from the paper [3]. The Lena, Boat, and Barbara images are used for this purpose with different noise levels σ_n^2.

In our examples, in addition to the orthogonal wavelet transform, the dual-tree CWT will be used in order to take advantage of this transform (near shift-invariance and directional selectivity). The new bivariate shrinkage function will be applied to the magnitude of the complex coefficients since the real and imaginary parts are not shift invariant individually but the magnitudes are. (We have also applied the same algorithm to the real and imaginary parts separately but we have observed significant performance loss.) We assume the magnitudes of the coefficients are corrupted by additive Gaussian noise although this is an approximation. The performance of our algorithm suggests it is an acceptable assumption. The PSNR values, when the CWT is used, are listed in the last column of Table I.

We also compared the proposed algorithm with other published results [3], [4], [9], [11] that use various redundant wavelet transforms, i.e., the undecimated wavelet transform in [3] and [9], the steerable pyramid in [11], and the dual-tree CWT in [4]. The PSNR values are listed in Table II (the values are taken from the corresponding papers). In [11], it is noted that the Matlab implementation for that algorithm takes 12.8 min for a 512×512 image on 900-MHz Pentium III, although the Matlab program for the proposed algorithm takes 25 s for a 512×512 image on 450-MHz Pentium II.

One example using a 512×512 Barbara image is given in Fig. 1. The original and the noisy images are illustrated in Fig. 1(a) and (b). The denoised images obtained using the data-driven denoising algorithm described above with the orthogonal wavelet transform and the dual-tree CWT are illustrated in Fig. 1(c) and (d), respectively, and have PSNR values of 27.23 and 28.61 dB, respectively.

In [17] and [18], very high quality image denoising algorithms are presented using newly developed multiscale representation systems, namely the ridgelet and curvelet transform. In their experiments, the Lena image is used with the Gaussian noise standard deviation 20. They reported the denoised PSNR value is 31.95 in [17] and and 32.72 in [18]. The result here lies between these two values.

TABLE I
PSNR VALUES OF DENOISED IMAGES FOR DIFFERENT TEST IMAGES AND NOISE LEVELS (σ_N) OF NOISY, BAYESSHRINK [2], ADAPTSHRINK [3], HMT SYSTEM [5], LAWMAP [10], PROPOSED ALGORITHM IN [13], PROPOSED SYSTEM, AND PROPOSED SYSTEM FOR DUAL-TREE COMPLEX DISCRETE WAVELET TRANSFORM

	Noisy	BayesShrink in [2]	AdaptShr in [3]	HMT in [5]	LAWMAP in [10]	Proposed in [13]	Proposed	Proposed (complex)
Lena								
$\sigma = 10$	28.18	33.32	-	33.84	34.10	33.94	34.36	35.34
$\sigma = 15$	24.65	31.41	32.39	31.76	32.23	32.06	32.51	33.67
$\sigma = 20$	22.14	30.17	31.07	30.39	30.89	30.73	31.19	32.40
$\sigma = 25$	20.17	29.22	30.70	29.24	29.89	29.81	30.15	31.40
$\sigma = 30$	18.62	28.48	-	28.35	29.05	28.94	29.41	30.54
Boat								
$\sigma = 10$	28.16	31.80	-	32.28	32.25	32.25	32.42	33.10
$\sigma = 15$	24.65	29.87	-	30.31	30.40	30.25	30.55	31.36
$\sigma = 20$	22.15	28.48	-	28.84	29.00	28.93	29.18	30.08
$\sigma = 25$	20.15	27.40	-	27.68	27.91	27.91	28.14	29.06
$\sigma = 30$	18.62	26.60	-	26.83	27.06	27.11	27.29	28.31
Barbara								
$\sigma = 10$	28.16	30.86	-	31.36	31.99	31.13	32.25	33.35
$\sigma = 15$	24.63	28.51	29.96	29.23	29.60	28.71	29.97	31.31
$\sigma = 20$	22.14	27.13	28.36	27.80	27.94	27.25	28.36	29.80
$\sigma = 25$	20.18	26.01	27.23	25.99	26.75	25.97	27.16	28.61
$\sigma = 30$	18.62	25.16	-	25.11	25.80	25.21	26.28	27.65

TABLE II
PSNR VALUES OF DENOISED IMAGES FOR DIFFERENT TEST IMAGES AND NOISE LEVELS (σ_n) OF NOISY, THE SYSTEM IN [9], SI-ADAPTSHR [3], CHMT [4], THE SYSTEM IN [11], AND PROPOSED SYSTEM FOR DUAL-TREE COMPLEX DISCRETE WAVELET TRANSFORM

	Noisy Noisy	The system in [9]	SI-AdaptShr in [3]	CHMT in [4]	The system in [11]	Proposed (complex)
Lena						
$\sigma = 10$	28.18	34.96	-	34.9	35.31	35.34
$\sigma = 15$	24.65	33.05	33.41	-	33.55	33.67
$\sigma = 20$	22.14	31.72	32.12	-	32.31	32.40
$\sigma = 25$	20.17	30.64	31.11	29.9	31.33	31.40
Barbara						
$\sigma = 10$	28.16	33.35	-	-	33.45	33.35
$\sigma = 15$	24.63	31.10	31.14	-	31.22	31.31
$\sigma = 20$	22.14	29.44	29.52	-	29.71	29.80
$\sigma = 25$	20.18	28.23	28.33	-	28.57	28.61

IV. CONCLUSION

This letter presents an effective and low-complexity image-denoising algorithm using the joint statistics of the wavelet coefficients of natural images. We presented our result for both orthogonal and dual-tree CWTs and compared it with the other published results in order to illustrate the effectiveness of the proposed algorithm. The comparison suggests the new denoising results are competitive with the best wavelet-based results reported in the literature.

REFERENCES

[1] Z. Cai, T. H. Cheng, C. Lu, and K. R. Subramanian, "Efficient wavelet-based image denoising algorithm," *Electron. Lett.*, vol. 37, no. 11, pp. 683–685, May 2001.

[2] S. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. Image Processing*, vol. 9, pp. 1532–1546, Sept. 2000.

[3] S. G. Chang, B. Yu, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Trans. Image Processing*, vol. 9, pp. 1522–1531, Sept. 2000.

[4] H. Choi, J. K. Romberg, R. G. Baraniuk, and N. G. Kingsbury, "Hidden Markov tree modeling of complex wavelet transforms," in *Proc. ICASSP*, vol. 1, Istanbul, Turkey, June 2000, pp. 133–136.

[5] M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based signal processing using hidden Markov models," *IEEE Trans. Signal Processing*, vol. 46, pp. 886–902, Apr. 1998.

[6] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

[7] N. G. Kingsbury, "Image processing with complex wavelets," *Phil. Trans. R. Soc. London A*, Sept. 1999.

[8] ——, "Complex wavelets for shift invariant analysis and filtering of signals," *Appl. Compt. Harmon. Anal.*, pp. 234–253, May 2001.

[9] X. Li and M. T. Orchard, "Spatially adaptive image denosing under overcomplete expansion," in *Proc. ICIP*, Sept. 2000.

[10] M. K. Mihcak, I. Kozintsev, K. Ramchandran, and P. Moulin, "Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Processing Lett.*, vol. 6, pp. 300–303, Dec. 1999.

[11] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli, "Adaptive Wiener denoising using a Gaussian scale mixture model," in *Proc. ICIP*, 2001.

[12] L. Sendur and I. W. Selesnick, "A bivariate shrinkage function for wavelet based denoising," in *IEEE ICASSP*, 2002.

[13] ——, "Bivariate shrinkage functions for wavelet-based denoising," *IEEE Trans. Signal Processing*, vol. 50, pp. 2744–2756, Nov. 2002.

[14] M. J. Wainwright and E. P. Simoncelli, "Scale mixtures of Gaussians and the statistics of natural images," *Adv. Neural Inform. Process. Syst.*, vol. 12, May 2000.

[15] G. Fan and X. G. Xia, "Image denoising using a local contextual hidden Markov model in the wavelet domain," *IEEE Signal Processing Lett.*, vol. 8, pp. 125–128, May 2001.

[16] T. Cai and B. W. Silverman, "Incorporating information on neighboring coefficients into wavelet estimation," *Sankhya*, vol. 63, pp. 127–148, 2001.

[17] J. L. Starck, E. J. Candes, and D. L. Donoho, "The curvelet transform for image denoising," *IEEE Trans. Image Processing*, vol. 11, pp. 670–684, June 2002.

[18] J. L. Starck, D. L. Donoho, and E. J. Candes, "Very high quality image restoration by combining wavelets and curvelets," in *Proc. SPIE Conf. Signal and Image Processing: Wavelet Applications in Signal and Image Processing IX*, vol. 4478, M. A. Unser and A. Aldroubi, Eds., 2001.