LOCALLY ADAPTIVE WAVELET-BASED IMAGE DENOISING USING THE GRAM-CHARLIER PRIOR FUNCTION

S. M. Mahbubur Rahman, M. Omair Ahmad, Fellow, IEEE*, M. N. S. Swamy,* Fellow, IEEE

Department of Electrical and Computer Engineering
Concordia University, 1455 De Maisonneuve Blvd. West
Montreal, Quebec, Canada, H3G 1M8
e-mail: {mahb_rah, omair, swamy}@ece.concordia.ca

ABSTRACT

Statistical estimation techniques for the wavelet-based image denoising use suitable probability density functions (PDFs) as prior functions for the image coefficients. Due to the intrascale dependency of the local neighboring image wavelet coefficients, the prior functions are assumed to be stationary. In this paper, it is shown that the stationary Gram-Charlier (GC) PDF models the image coefficients better than the traditional ones, such as the stationary Gaussian and stationary generalized Gaussian PDFs. A Bayesian wavelet-based maximum *a posteriori* estimator is then developed by using the proposed GC prior function. Experimental results on standard images show that the proposed estimator provides a denoising performance, which is better than that of several existing denoising methods in terms of signal-to-noise ratio and visual quality.

Index Terms— Denoising, image wavelet coefficients, Gram-Charlier, MAP estimator

1. INTRODUCTION

Due to the compaction and decorrelation properties of the multiscale discrete wavelet transform (DWT) coefficients [1], a great deal of success has been achieved in the past few years for estimating a signal corrupted by additive noise using the technique of thresholding or shrinkage of the wavelet coefficients. Improved estimation can be achieved by taking advantage of the shift-invariance and directional properties of similar types of multiscale redundant representations, such as the shift-invariant (SI) form of the DWT [1], dual-tree complex wavelet transform (DT-CWT) [2], curvelet, contourlet, wedgelet, bandlet, steerable pyramid (SP) [3], matching pursuit and basis pursuit. Almost all the statistical estimators concentrate on choosing an appropriate probability density function (PDF) for modeling the multiscale wavelet coefficients, and deriving a shrinkage function using the minimum mean squared error or maximum *a posteriori* (MAP) criterion. It is well-known that the PDF of the image wavelet coefficients is non-Gaussian and symmetric [4]. The standard univariate PDFs for modeling the subband image wavelet coefficients include the generalized Gaussian (GG) [5], Bessel K-form [4], α-stable [6], Gaussian scale mixture (GSM) [3], normal inverse Gaussian [7] and Gauss-Hermite [8]. The image wavelet coefficients have strong intrascale and weak interscale dependencies [9]. Assuming that the image wavelet coefficients are locally stationary, the parameters of some of the above mentioned PDFs are refined with respect to the local spatial contexts. A number of univariate PDFs have also been extended to the bivariate [6], [10] and multivariate cases or refined in the hidden Markov tree framework to take into consideration the interscale dependency.

In this paper, a novel statistical model is proposed for the intrascale local neighboring image wavelet coefficients by using the stationary Gram-Charlier (GC) PDF. It is shown that the GC PDF matches with the empirical one better than other standard PDFs such as the Gaussian and the GG. An expression for the MAP estimator is then derived assuming that local neighboring image wavelet coefficients follow the GC PDF. It is shown that introduction of the proposed PDF in the MAP estimation shows an improved denoising performance as compared to that of several existing ones. The paper is organized as follows. In Section 2, the modeling of the local neighboring image wavelet coefficients by the GC PDF is presented. The statistical MAP estimator that exploits the proposed PDF is described in Section 3. In Section 4, the performance of the proposed estimator is compared with that of the existing ones. Finally, Section 5 provides conclusions.

2. GC PDF FOR LOCAL NEIGHBORING COEFFICIENTS

Let f_{ij} represent an image wavelet coefficient in the decomposition level ℓ at spatial location (i, j) and $\boldsymbol{f}$ represent the vector containing all the local neighboring coefficients at that reference location. Let $p_{\boldsymbol{f}}(\boldsymbol{f})$ denote the PDF of the random variable $\boldsymbol{f}$ that has a zero mean and variance $\sigma_{\boldsymbol{f}}^2$. This PDF may be expressed in terms of the GC density function in the form [11]

$$p_{\boldsymbol{f}}(\boldsymbol{f}) = \mathcal{G}(\sigma_{\boldsymbol{f}}) \left[1 + \frac{S}{3!} H_3\left(\frac{\boldsymbol{f}}{\sigma_{\boldsymbol{f}}}\right) + \frac{K-3}{4!} H_4\left(\frac{\boldsymbol{f}}{\sigma_{\boldsymbol{f}}}\right)\right] \quad (1)$$

where $S = M_{3\boldsymbol{f}}/\sigma_{\boldsymbol{f}}^3$ and $K = M_{4\boldsymbol{f}}/\sigma_{\boldsymbol{f}}^4$, respectively, are the skewness and kurtosis, $M_{k\boldsymbol{f}}$ being the k-th order moment, $\mathcal{G}(\sigma_{\boldsymbol{f}}) = \left(\sigma_{\boldsymbol{f}}\sqrt{2\pi}\ \right)^{-1} e^{-\boldsymbol{f}^2/2\sigma_{\boldsymbol{f}}^2}$ and $H_r(\boldsymbol{f})$ is the Hermite polynomial of order r satisfying the recurrance relation

$$\begin{aligned} H_0(\boldsymbol{f}) &= 1 \\ H_1(\boldsymbol{f}) &= \boldsymbol{f} \\ H_{r+1}(\boldsymbol{f}) &= \boldsymbol{f} H_r(\boldsymbol{f}) - r H_{r-1}(\boldsymbol{f}) \quad r \geq 1 \end{aligned} \quad (2)$$

It is observed that the skewness of the image wavelet coefficients are very small ($S \ll 1$). Hence, the PDF of the local neighboring coefficients can be treated as symmetric by considering S to be zero. Table 1 shows the average values of K for the local neighboring DWT coefficients of the first and second level decompositions of a few test images using the window sizes of 5×5 and 11×11. Note that similar results have been obtained for K using other test images given in the database[1], but are not included here for lack of space. It can

[1] http://decsai.ugr.es/cvg/dbimagenes/index.php

Table 1. The average values of kurtosis for the test images obtained from the local neighboring DWT coefficients.

Test image	Kurtosis, K			
	Local (5×5)		Local (11×11)	
	$\ell = 1$	$\ell = 2$	$\ell = 1$	$\ell = 2$
Lena	3.51	4.51	4.25	6.19
Barbara	3.41	4.60	4.53	6.36
Boat	3.57	4.80	4.41	6.30
Bridge	4.08	4.10	4.71	4.44
Baboon	3.29	3.54	3.76	4.17

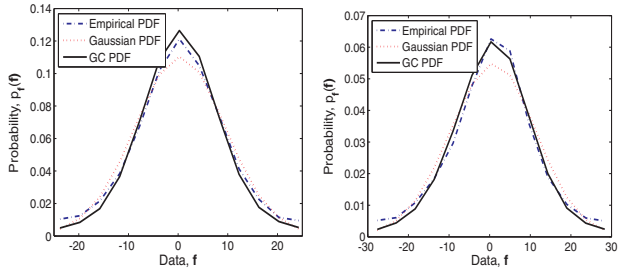

Fig. 1. Gaussian and GC PDFs for modeling the local neighboring DWT coefficients using a 7×7 window in the $LH1$ subband. Test images are (a) *Lena* and (b) *Barbara*.

be observed from Table 1 that the kurtosis of the local neighboring coefficients is usually greater than 3 and increases with the increasing window size. Therefore, the Gaussian distribution is not a good probabilistic model for the local neighboring coefficients. It may be mentioned that the dependency of the local neighboring coefficients is significant only when the window size is 5×5 or 7×7, and that a further increase in the window size generally does not yield any appreciable improvement in the denoising performance. The symmetric GC PDF is ensured to be positive, if $3 \leq K \leq 7$ [12]. Extensive simulations reveal that the kurtosis of the local neighboring coefficients for the window size of interest lies in this range. Hence, the symmetric GC is a valid PDF for modeling the local neighboring coefficients. Fig. 1 shows the empirical, Gaussian and GC PDFs to model the local neighboring coefficients of the $LH1$ subband for the test images *Lena* and *Barbara* using a window size of 7×7. Since the images are of size 256×256 and the $LH1$ subband is of size 128×128, each of the PDFs is obtained by averaging the $128 \times 128 = 16384$ PDFs corresponding to each of the coefficients in the $LH1$ subband. It is evident from Fig. 1 that the proposed PDF matches better with the empirical one than the Gaussian PDF does. It is to be noted that similar results are observed for other window sizes using various subbands of different test images. Table 2 shows the modeling performance of the GG and GC PDFs in terms of the metrics Kolmogorov-Smirnov distance (KSD) and Kullback-Leibler divergence (KLD) [11] of a few test images. It is seen from Table 2 that both the KSD and KLD of the GC PDF are lower than that of the GG PDF, showing that the former provides a better model than the latter for the local neighboring coefficients. Similar results have been obtained for the other test images. In order to demonstrate the effectiveness of the proposed GC PDF, we consider the case of Bayesian denoising as an application.

3. MAP ESTIMATOR FOR DENOISING

In this section, we assume that the image pixels are corrupted by additive white Gaussian noise (AWGN) with known variance $\sigma_{\boldsymbol{\varepsilon}}^2$. If $\sigma_{\boldsymbol{\varepsilon}}^2$

Table 2. Results concerning the metrics KSD and KLD (standard deviations in parentheses) for prior function modeling of the local neighboring DWT coefficients using a 7×7 window.

Prior model	KSD		KLD	
	$\ell = 1$	$\ell = 2$	$\ell = 1$	$\ell = 2$
	Lena			
GG PDF	0.082	0.086	0.203 (0.824)	0.352 (1.234)
GC PDF	**0.072**	**0.080**	**0.179 (0.615)**	**0.294 (0.921)**
	Barbara			
GG PDF	0.087	0.099	0.210 (0.960)	0.430 (1.511)
GC PDF	**0.076**	**0.089**	**0.184 (0.734)**	**0.349 (1.132)**
	Boat			
GG PDF	0.076	0.095	0.210 (0.815)	0.378 (1.226)
GC PDF	**0.072**	**0.084**	**0.179 (0.596)**	**0.308 (0.925)**
	Bridge			
GG PDF	0.077	0.080	0.263 (1.028)	0.256 (0.958)
GC PDF	**0.075**	**0.075**	**0.218 (0.778)**	**0.209 (0.755)**
	Baboon			
GG PDF	0.078	0.079	0.222 (1.506)	0.289 (1.574)
GC PDF	**0.065**	**0.074**	**0.194 (1.295)**	**0.241 (1.300)**

is unknown, it may be estimated by applying the median-absolute-deviation method [13] in the $HH1$ subband of the noisy wavelet coefficients. Since the noise is uncorrelated with the true image signal, the wavelet coefficient of the noisy image at the spatial location (i, j) is given by

$$g_{ij} = f_{ij} + \varepsilon_{ij} \tag{3}$$

where ε_{ij} is the wavelet coefficient of the additive noise at the reference location. We propose the MAP-based Bayesian denoising, which is locally adaptive in nature. In a Bayesian framework, g_{ij}, f_{ij}, and ε_{ij} in (3) are considered as samples of the independent random variables $\boldsymbol{g}$, $\boldsymbol{f}$, and $\boldsymbol{\varepsilon}$, respectively. The random variable $\boldsymbol{f}$ is modelled by symmetric GC PDF, and $\boldsymbol{\varepsilon}$ by Gaussian PDF. The Bayes-risk estimator for the denoised wavelet coefficient $\hat{f}_{ij}$ using the MAP criterion is given by [10]

$$\hat{f}_{ij}(\boldsymbol{g}) = \arg\max_{\boldsymbol{f}} \left[-\frac{(\boldsymbol{g} - \boldsymbol{f})^2}{2\sigma_{\boldsymbol{\varepsilon}}^2} + \ln p_{\boldsymbol{f}}(\boldsymbol{f}) \right] \tag{4}$$

Using the approach proposed by Hyvarinen [14], an approximate solution of (4) can be obtained as

$$\hat{f}_{ij}(\boldsymbol{g}) = \text{sign}(g_{ij}) \max\left(|g_{ij}| - \sigma_{\boldsymbol{\varepsilon}}^2 |\Phi(\boldsymbol{g})|, 0 \right) \tag{5}$$

where $\Phi(\boldsymbol{f}) = -\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{f}} \left[\ln p_{\boldsymbol{f}}(\boldsymbol{f}) \right]$ is called the score function of $\boldsymbol{f}$. Using the identity $\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{f}} \left[H_r(\boldsymbol{f}) \right] = r H_{r-1}(\boldsymbol{f})$, the score function of the symmetric GC PDF can be obtained as

$$\Phi(\boldsymbol{f}) = \frac{\boldsymbol{f}}{\sigma_{\boldsymbol{f}}^2} - \frac{\frac{K-3}{3!} H_3\left(\frac{\boldsymbol{f}}{\sigma_{\boldsymbol{f}}}\right)}{1 + \frac{K-3}{4!} H_4\left(\frac{\boldsymbol{f}}{\sigma_{\boldsymbol{f}}}\right)} \cdot \frac{1}{\sigma_{\boldsymbol{f}}} \tag{6}$$

The proposed MAP estimator requires that the variance and kurtosis of the noise-free wavelet coefficients be calculated from the noisy condition. In other words, it is necessary to estimate the second and fourth order moments of the noise-free image coefficients from the noisy coefficients. Since the image is corrupted by AWGN, the second and fourth order moments can be estimated as

$$\begin{aligned} \hat{M}_{2\boldsymbol{f}} &= \max(\hat{M}_{2\boldsymbol{g}} - \sigma_{\boldsymbol{\varepsilon}}^2, 0) \\ \hat{M}_{4\boldsymbol{f}} &= \max(\hat{M}_{4\boldsymbol{g}} - 6\hat{M}_{2\boldsymbol{f}}\sigma_{\boldsymbol{\varepsilon}}^2 - 3\sigma_{\boldsymbol{\varepsilon}}^4, 0) \end{aligned} \tag{7}$$

where $\hat{M}_{2\boldsymbol{g}}$ and $\hat{M}_{2\boldsymbol{g}}$, respectively, are the second and fourth order sample moments [11] of the noisy wavelet coefficients.

Table 3. Output PSNR, $20\log_{10}(255/\sigma_e)$ in dB, where σ_e is the error standard deviation, for different locally adaptive denoising methods that use the decimated DWT

Denoising algorithms	Noise standard deviation $\sigma_{\boldsymbol{\varepsilon}}$				
	10	15	20	25	30
	Lena				
BivariateShrink [10]	34.47	32.63	31.30	30.30	29.49
LAWMAP [15]	34.35	32.40	31.06	30.02	29.22
NeighShrink [16]	34.46	32.52	31.04	29.88	28.88
ProbShrink [17]	34.30	32.41	31.05	30.02	29.25
GSM [3]	34.57	32.72	31.41	30.36	29.57
Proposed method	**34.67**	**32.84**	**31.55**	**30.52**	**29.75**
	Barbara				
BivariateShrink [10]	32.69	30.35	28.75	27.58	26.63
LAWMAP [15]	32.58	30.27	28.73	27.60	26.72
NeighShrink [16]	32.86	30.44	28.78	27.53	26.54
ProbShrink [17]	32.51	30.10	28.50	27.31	26.38
GSM [3]	32.84	30.44	28.81	27.68	26.72
Proposed method	**33.00**	**30.62**	**29.02**	**27.90**	**26.96**
	Boat				
BivariateShrink [10]	32.48	30.61	29.28	28.24	27.40
LAWMAP [15]	32.36	30.48	29.09	28.07	27.16
NeighShrink [16]	32.72	30.64	29.15	28.05	27.08
ProbShrink [17]	32.53	30.57	29.17	28.10	27.27
GSM [3]	**32.78**	**30.83**	**29.44**	28.38	27.52
Proposed method	32.76	30.82	**29.44**	**28.40**	**27.56**
	Bridge				
BivariateShrink [10]	30.41	27.93	26.40	25.33	24.58
LAWMAP [15]	30.42	28.06	26.58	25.55	24.77
NeighShrink [16]	30.35	27.91	26.36	25.31	24.50
ProbShrink [17]	30.14	27.75	26.27	25.25	24.50
GSM [3]	30.31	27.83	26.37	25.34	24.58
Proposed method	**30.47**	**28.01**	**26.57**	**25.55**	**24.81**

4. EXPERIMENTAL RESULTS

Extensive experimentations have been carried out for comparing the performance of the proposed denoising method with that of the others. Due to the limitation of space, however, we give in this section results concerning four 512×512 grayscale images, namely, *Lena*, *Barbara*, *Boat*, and *Bridge* and two types of wavelet representations, namely, the decimated DWT and redundant DT-CWT. The test images are obtained from the same sources as mentioned in [3]. The DWT-based experiments use the orthogonal Daubechies-8 wavelet filter and the DT-CWT-based ones use the $(11, 17)$-tap biorthogonal filters at level $\ell = 1$, and 6-tap Q-shift orthogonal filters for $\ell \geq 2$, which have been proposed in [2]. The proposed method is applied on the coefficients of the 4-level DWT and 3-level DT-CWT, since any further decomposition level does not produce a significant increase in the denoising performance.

We evaluate the performance of the proposed method by comparing it with that of five other locally adaptive wavelet-based denoising methods, namely, the BivariateShrink [10], locally adaptive window-based MAP (LAWMAP) [15], NeighShrink [16], ProbShrink [17] and GSM [3]. The first one is an example, wherein both the inter- and intrascale dependencies are taken into account. The second, third, and fourth methods inherently use only the intrascale dependency. In the case of the fifth, we provide the results obtained from the intrascale dependency only in order to make a fair comparison with our proposed method. The results for the BivariateShrink method are obtained using the codes provided by I. Selesnick[2], the ProbShrink method using the codes provided by A. Pizurica[3] and the GSM method using the codes provided by J. Portilla[4]. We report the results provided by our proposed denoising method using a 7×7 window, since this window size provides the highest output peak signal-to-noise ratio (PSNR) in most of the cases. The window sizes and parameters for the other five methods mentioned above are chosen to produce the maximum output PSNR.

Table 3 shows the output PSNR values obtained from the various denoising methods that use the decimated DWT for the four test images considered here. From the tabular results, it can be observed that except for two instances of the GSM method [3] in the *Boat* image, the proposed method always provides the highest output PSNR values. Further, for these two instances, the proposed method provides the PSNR values of about the same level as that provided by the GSM method. Fig. 2 shows the original cropped image of *Barbara* along with its noisy version with $\sigma_{\boldsymbol{\varepsilon}} = 20$, and the decimated DWT-based denoised versions provided by the BivariateShrink method [10], ProbShrink method [17], GSM method [3] and proposed method. From this figure, it can be observed that the proposed method reduces the noise significantly with the least amount of artifacts (e.g., in the eyes and scarf of *Barbara*) as compared to the other methods. It is to be noted that the proposed method is computationally very fast. For example, in a 3 GHz and 512 MB machine, the processing time of the proposed method is 1.6 seconds, while that of the GSM method [3] and ProbShrink method [17] are 12.7 and 14.2 seconds, respectively.

Table 4 shows the output PSNR values obtained from various denoising methods using the different types of redundant representations of the DWT. The ProbShrink method [17] uses the SI form of the DWT, the BivariateShrink method [10] and the proposed method use the DT-CWT, and the GSM method [3] uses the SP. It can be observed from the table that the proposed method always provides higher PSNR values than that provided by the BivariateShrink method [10] and the ProbShrink method [17]. Except for the *Bridge* image, although the GSM method [3] provides a small amount of higher PSNR values than that provided by the proposed one, the computational burden of the GSM method [3] that uses the SP is not insignificant. Specifically, the redundancy of the three types of wavelet transform that have been considered here, namely, the SI form of the DWT [5], the DT-CWT [2], and the SP [3] are, respectively, $4^{\mathcal{L}} : 1$, $4 : 1$, and $(7\eta_o/3) : 1$, where $\mathcal{L}$ is the total number of wavelet decompositions and η_o is the total number orientations (the GSM method in [3] uses $\eta_o = 8$) in the SP. Hence, the proposed method enjoys a reduced amount of computational complexity by a factor of $7\eta_o/12$, while providing a PSNR performance comparable to that of the GSM method.

5. CONCLUSION

In this paper, we have proposed the Gram-Charlier PDF for the probabilistic modeling of the local neighboring image wavelet coefficients. It has been shown that the proposed PDF provides a better match to the empirical one than that provided by some of the traditional PDFs, such as the stationary Gaussian and the stationary GG. The proposed PDF is then used to design a Bayesian MAP estimator for the wavelet-based image denoising. The experimental results on standard images have shown a superior performance of the proposed estimator over several existing denoising methods.

6. REFERENCES

[1] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, CA, second edition, 1999.

[2] http://taco.poly.edu/selesi/index.html

[3] http://telin.ugent.be/∼sanja/

[4] http://decsai.ugr.es/∼javier/index.html

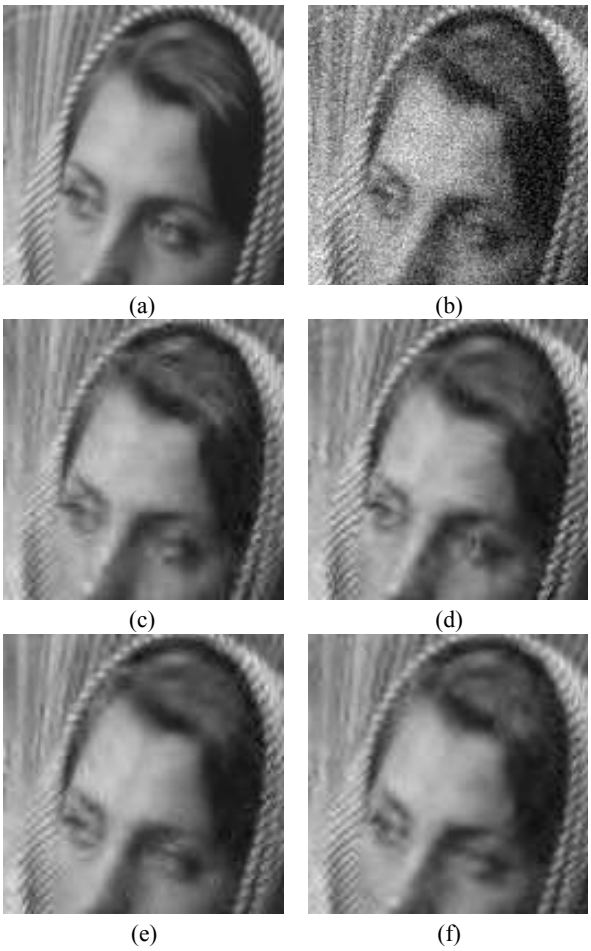

Fig. 2. Comparison of the denoised images obtained from the different locally adaptive DWT-based denoising methods on *Barbara* with $\sigma_{\varepsilon} = 20$. (a) Original image, (b) noisy image and denoised images using (c) the BivariateShrink method [10], (d) the ProbShrink method [17], (e) the GSM method [3], and (f) the proposed method.

Table 4. Output PSNR for different locally adaptive denoising methods that use redundant representation of the DWT

Denoising algorithms	Noise standard deviation σ_{ε}			
	10	15	20	25
	Lena			
ProbShrink (SI-DWT) [17]	35.24	33.46	32.20	31.21
BivariateShrink (DT-CWT) [10]	35.34	33.67	32.40	31.40
Proposed method (DT-CWT)	35.49⋆	33.75⋆	32.45⋆	31.43⋆
GSM (SP) [3]	35.59	33.86	32.61	31.61
	Barbara			
ProbShrink (SI-DWT) [17]	33.46	31.21	29.60	28.32
BivariateShrink (DT-CWT) [10]	33.35	31.31	29.80	28.61
Proposed method (DT-CWT)	34.01⋆	31.77⋆	30.20⋆	28.98⋆
GSM (SP) [3]	34.10	31.88	30.31	29.10
	Boat			
ProbShrink (SI-DWT) [17]	33.26	31.32	29.94	28.89
BivariateShrink (DT-CWT) [10]	33.10	31.36	30.08	29.06
Proposed method (DT-CWT)	33.42⋆	31.47⋆	30.10⋆	29.07⋆
GSM (SP) [3]	33.54	31.60	30.25	29.20
	Bridge			
ProbShrink (SI-DWT) [17]	30.62	28.23	26.75	25.72
BivariateShrink (DT-CWT) [10]	29.90	28.00	26.72	25.78
Proposed method (DT-CWT)	30.96⋆	28.59⋆	27.08⋆	26.02⋆
GSM (SP) [3]	30.84	28.46	26.99	25.96

[2] N. G. Kingsbury, "Complex wavelets for shift invariance analysis and filtering of signals," *Applied Comput. Harmon. Anal.*, vol. 10, no. 3, pp. 234–253, 2001.

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

[4] A. Srivastava, A. B. Lee, E. P. Simoncelli, and S. C. Zhu, "On advances in statistical modeling of natural images," *J. Math. Imag. Vision*, vol. 18, pp. 17–33, 2003.

[5] S. G. Chang, B. Yu, and M. Vettereli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Trans. Image Processing*, vol. 9, no. 9, pp. 1522–1531, 2000.

[6] A. Achim and E. E. Kuruoğlu, "Image denoising using bivariate α-stable distributions in the complex wavelet domain," *IEEE Signal Processing Lett.*, vol. 12, no. 1, pp. 17–20, 2005.

[7] S. Solbo and T. Eltoft, "Homomorphic wavelet-based statistical despeckling of SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 4, pp. 711–721, 2004.

[8] S. M. M. Rahman, M. O. Ahmad, and M. N. S. Swamy, "Wavelet-domain image denoising algorithm using series expansion of coefficient P.D.F. in terms of Hermite polynomials," in *Proc. 3rd. Int. IEEE-NEWCAS Conf.*, Québec, Canada, 2005, pp. 271–275.

[9] J. Liu and P. Moulin, "Information-theoretic analysis of interscale and intrascale dependencies between image wavelet coefficients," *IEEE Trans. Image Processing*, vol. 10, no. 11, pp. 1647–1658, 2001.

[10] L. Şendur and I. W. Selesnick, "Bivariate shrinkage with local variance estimation," *IEEE Signal Processing Lett.*, vol. 9, no. 12, pp. 438–441, 2002.

[11] M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics, Vol. 1, Distribution Theory*, Charles Griffin, London, 4th edition, 1977.

[12] D. E. Barton and K. E. Dennis, "The conditions under which Gram-Charlier and Edgeworth curves are positive definite and unimodal," *Biometrika*, vol. 39, no. 3/4, pp. 425–427, 1952.

[13] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *Journal of the American Statistical Assoc.*, vol. 90, no. 432, pp. 1200–1224, 1995.

[14] A. Hyverinen, "Sparse code shrinkage: Denoising of nongaussian data by maximum likelihood estimation," *Neural Computation*, vol. 11, no. 7, pp. 1739–1768, 1999.

[15] M. K. Mihçak, I. Kozintsev, K. Ramchandran, and P. Moulin, "Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Processing Lett.*, vol. 6, no. 12, pp. 300–303, 1999.

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

[17] A. Pizurica and W. Philips, "Estimating the probability of the presence of a signal of interest in multiresolution single- and multiband image denoising," *IEEE Trans. Image Processing*, vol. 15, no. 3, pp. 654–665, 2006.