IMAGE DENOISING EMPLOYING A MIXTURE OF CIRCULAR SYMMETRIC LAPLACIAN MODELS WITH LOCAL PARAMETERS IN COMPLEX WAVELET DOMAIN

Hossein Rabbani[*], *Mansur Vafadust*[*], *Ivan Selesnick*[**] *and Saeed Gazor*[***]

[*]Department of Biomedical Engineering, Amirkabir University of Technology, Tehran, Iran
[**]Department of Electrical and Computer Engineering, Polytechnic University, Brooklyn, NY, USA
[***]Department of Electrical and Computer Engineering, Queen's University, Kingston, Ontario, Canada

ABSTRACT

In this paper, we present a new image denoising algorithm. We assume a mixture of bivariate circular symmetric Laplacian probability density functions (pdfs) where for each wavelet coefficients may have different local parameter. This pdf characterizes simultaneously 1) the heavy-tailed nature, 2) the interscale dependencies of the wavelet coefficients and also 3) the empirically observed correlation between the coefficient amplitudes. We employ this local bivariate mixture model to derive a Bayesian image denoising technique. This proposed pdf, potentially can fits better the statistical properties of the wavelet coefficients than several other existing models. Our simulation results reveal that the proposed denoising method is among the best reported in the literature. This is justified since the accuracy of the employed distribution for noise-free data determines the denoising performance.

Index Terms— *circular symmetric Laplacian pdf, MAP estimator, mixture model, complex wavelet transforms*

1. INTRODUCTION

The wavelet based image denoising is often performed using Bayesian techniques for estimation of clean coefficients from the noisy data observation [1]-[5]. For example using the maximum a posteriori (MAP) estimator, the solution requires a priori knowledge about the distribution of the wavelet coefficients. For a given distribution, a particular estimator could be obtained (called shrinkage function). For example, the classical soft thresholding is obtained by a Laplacian pdf [2].

The clustering property of the wavelet transforms state that if a particular wavelet coefficient is large/small, then the adjacent coefficients are likely large/small too. This local property is employed in [3] using a Gaussian pdf with local variance. Another property of the wavelet transforms is the compression that is the wavelet coefficients of real-world signals tend to be sparse. In order to take this property into account, in [4] a mixture distribution is proposed. The third property of the wavelet transforms is the persistence, i.e., the large/small values of wavelet coefficients tend to propagate across scales. This property means that the bivariate pdfs, such as circular symmetric Laplacian pdf [2], which exploit the dependency between coefficients, better model the statistical properties of wavelet coefficients in comparison with univariate pdfs. (Since the univariate pdfs assume that coefficients in adjacent scales are independent.) In this paper, we use a mixture of two circular symmetric Laplacian pdfs with local parameters to describe all the above properties. In contrast in [6] a mixture of univariate Laplace pdfs with local parameters is employed to derive the LapMixShrinkL A bivariate mixture model without local parameters is used for developing a denoising algorithm in [7]. This paper combines the ideas in [6] and [7].

The rest of this paper is organized as follows: A brief review on Bayesian denoising is presented in Section 2. We obtain a shrinkage function namely, CShrinkL, using the circular symmetric Laplacian pdf with local variance [2] in Section 2.1. In Section 2.2 the bivariate mixture pdf with local parameters is introduced. In Section 2.3, we obtain the shrinkage function derived from the proposed local circular symmetric Laplacian mixture namely, CsLapMixShrinkL. In Section 3, we use the proposed pdf for wavelet-based denoising of several images corrupted with additive Gaussian noise in various noise levels. The simulation results show that our algorithm achieves better performance visually and in terms of peak-signal-to-noise-ratio (PSNR) compared with several methods: 1) the hard thresholding (HT) [1], 2) CShrinkL [2], 3) alpha-stable based Bayesian processor (WIN-SAR) [5], 4) LapMixShrinkL [6] and 5) CsLapMixShrink [7]. Finally the summarizing remarks are given in Section 4.

2. BAYESIAN DENOISING

Denoising can significantly improve the visual quality of an image. The main sources of noise are arising from the electronic hardware (shot noise) and from the channels during transmission (thermal noise). Most noise sources are modeled by additive white Gaussian noise (AWGN).

Suppose we observe a noisy wavelet coefficient $\bar{y}(k)$, where $\bar{y}(k)=(y_1(k),y_2(k))$. Using a (complex) wavelet transform [2] with subsampeling across scales, we can write

$\overline{y}(k)=\overline{w}(k)+\overline{n}(k)$, where $\overline{w}(k)=(w_1(k),w_2(k))$ represents the noise free image and $\overline{n}=(n_1,n_2)$ represents the AWGN. As illustrated in Fig. 1, $w_2(k)$ represents the parent of $w_1(k)$ (w_2 is the wavelet coefficient at the same spatial position as w_1, and at the next coarser scale). The objective is to estimate the noise-free coefficient, $\overline{w}(k)$. Using MAP criteria, the estimator is defined by:

$$\hat{\overline{w}}(k)=\arg\max_{\overline{w}(k)}\left[p_{\overline{w}(k)|\overline{y}(k)}\left(\overline{w}(k)\,|\,\overline{y}(k)\right)\right]. \quad (1)$$

After some manipulations, this equation is written as:

$$\hat{\overline{w}}(k)=\arg\max_{\overline{w}(k)}\left[p_{\overline{n}}\left(\overline{y}(k)-\overline{w}(k)\right)p_{\overline{w}(k)}\left(\overline{w}(k)\right)\right]. \quad (2)$$

Assuming that the noise samples n_1, n_2 are independent white zero-mean Gaussian with variance σ_n^2, we obtain:

$$\hat{\overline{w}}(k)=\arg\max_{\overline{w}(k)}\Big[-\left(y_1(k)-w_1(k)\right)^2/\sigma_n^2 -\left(y_1(k)-w_1(k)\right)^2/\sigma_n^2+2f\left(w_1(k),w_2(k)\right)\Big], \quad (3)$$

where $f\left(w_1(k),w_2(k)\right)=\log\left(p_{\overline{w}(k)}\left(\overline{w}(k)\right)\right)$.

By setting the partial derivative to zero with respect to $\hat{w}_i(k)$ for $i=1,2$, we obtain:

$$\frac{y_i(k)-\hat{w}_i(k)}{\sigma_n^2}+\frac{\partial f\left(\overline{w}(k)\right)}{\partial \hat{w}_i(k)}=0,\quad \text{for } i=1,2. \quad (4)$$

2.1. CShrinkL function

The solution of (4) depends on the pdf of noise-free wavelet coefficients. In [2], the following circular symmetric Laplacian pdf with local variance is proposed in order to describe that $w_1(k)$ and $w_2(k)$ are uncorrelated while are dependent:

$$p_{\overline{w}(k)}\left(\overline{w}(k)\right)=\frac{3}{2\pi\sigma^2(k)}\exp\left(-\frac{\sqrt{3}}{\sigma(k)}\sqrt{w_1^2(k)+w_2^2(k)}\right). \quad (5)$$

In this case, denoteing $\hat{w}_i(k)$ as the MAP estimate of $w_i(k)$ for $i=1,2$, we obtain following bivariate shrinkage function:

$$\hat{w}_i(k)=\frac{\left(\sqrt{y_1^2(k)+y_2^2(k)}-\sqrt{3}\sigma_n^2/\sigma(k)\right)_+}{\sqrt{y_1^2(k)+y_2^2(k)}}\,y_i(k)\quad i=1,2, \quad (6)$$

where $(a)_+=\max(0,a)$. We define CShrinkL function as:

$$\text{CShrinkL}\left(\overline{g}(k),\tau(k)\right):=\left(1-\tau(k)/\left\|\overline{g}(k)\right\|\right)_+\overline{g}(k),$$

where $\overline{g}(k)=\left(g_1(k),g_2(k)\right)$. We rewrite (6) as:

$$\hat{\overline{w}}(k)=\text{CShrinkL}\left(\overline{y}(k),\sqrt{3}\sigma_n^2/\sigma(k)\right). \quad (7)$$

To implement this estimator, we need to know σ_n and $\sigma(k)$. For each data point, $\overline{y}(k)$, an estimate of $\sigma(k)$ is formed based on a local neighborhood $\mathrm{N}(k)$ as illustrated in Fig. 1. We use a square window $\mathrm{N}(k)$ centered at $\overline{y}(k)$. Then we obtain an empirical estimate for $\sigma(k)$ as:

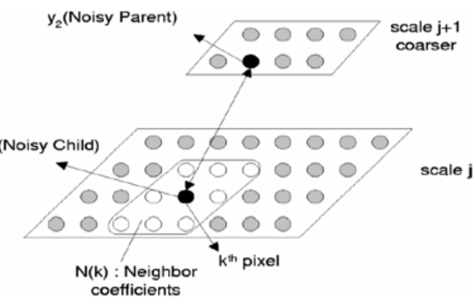

Fig. 1. Illustration of neighborhood [2].

$$\hat{\sigma}^2(k)=\frac{1}{2M}\left(-\sigma_n^2+\sum_{j\in\mathrm{N}(k)}\left(y_1^2(j)+y_2^2(j)\right)\right), \quad (8)$$

where M is the number of coefficients in $\mathrm{N}(k)$.

2.2. Bivariate mixture models with local parameters

We assume a pdf as a mixture of two circular symmetric Laplacian pdfs with local parameters in order to model the distribution of wavelet coefficients of images as follows:

$$\begin{aligned}p_{\overline{w}(k)}\left(\overline{w}(k)\right)&=a(k)\,p_1\left(\overline{w}(k)\right)+\left(1-a(k)\right)p_2\left(\overline{w}(k)\right),\\&=\frac{3a(k)}{2\pi\sigma_1^2(k)}e^{-\frac{\sqrt{3}}{\sigma_1(k)}\left\|\overline{w}(k)\right\|}+\frac{3(1-a(k))}{2\pi\sigma_2^2(k)}e^{-\frac{\sqrt{3}}{\sigma_2(k)}\left\|\overline{w}(k)\right\|},\end{aligned} \quad (9)$$

where $a(k)\in[0,1]$, $\sigma_1(k)$ and $\sigma_2(k)$ are the mixture model parameters. We see that this pdf represents two uncorrelated and independent random variables. Interestingly, because this pdf is mixture, it can better model the heavy-tailed property of wavelet coefficients than single circular symmetric Laplacian distribution.

To characterize the parameters in (9) it is necessary to have the parameters $\sigma_1(k),\sigma_2(k)$ and $a(k)$. For this mixture model, we use a local version of expectation maximization (EM) algorithm, which is an iterative numerical algorithm, to estimate these parameters. This iterative algorithm has two steps. If $S(k,m)$ denotes variable S at point k for iteration m and we start the algorithm with $m=0$ (first iteration), assuming the observed data $\overline{w}(k)$, the E-step calculates the responsibility factors:

$$\begin{aligned}r_1(k,m)&=\frac{a(k,m)p_1(\overline{w}(k),m)}{a(k,m)p_1(\overline{w}(k),m)+(1-a(k,m))p_2(\overline{w}(k),m)},\\r_2(k,m)&=1-r_1(k,m).\end{aligned} \quad (10)$$

The M-step updates the parameters $a(k,m),\sigma_1(k,m)$ and $\sigma_2(k,m)$. $a(k,m)$ is computed by:

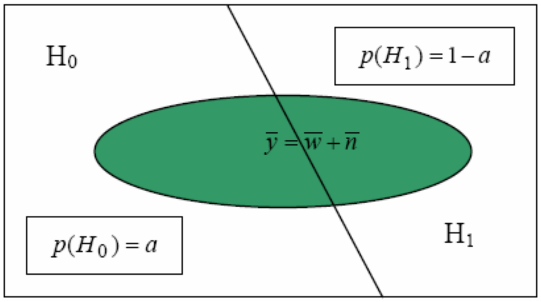

Fig. 2. Sample space based on binary random variable *v*.

$$a(k,m+1)=\frac{1}{M}\sum_{j\in \mathrm{N}(k)} r_1(j,m), \qquad (11)$$

where M is the number of coefficients in square window $\mathrm{N}(k)$ centered at $\bar{w}(k)$ as illustrated in Fig. 1 and the parameters $\sigma_1(k,m)$, $\sigma_2(k,m)$ are computed by:

$$\sigma_i(k,m+1)=\frac{\sqrt{3}}{2}\sum_{j\in N(k)} r_i(j,m)\|\bar{w}(k)\| \Big/ \sum_{j\in N(k)} r_i(j,m), \; i=1,2. \qquad (12)$$

With this new parameters we calculate $p_i(\bar{w}(k),m+1)$ for $i=1,2$. Using (11) we find $r_i(k,m+1)$ for $i=1,2$ by (10) and iterate these steps to converge of the parameters. For many mixture models (such as $\bar{y}$ in next section) a closed form for computing $\sigma_i(k)$ doesn't appear. In these cases, the following formulas produced from bivariate Gaussian mixture pdf can be used to estimate $\sigma_i(k)$:

$$\sigma_i^2(k)\leftarrow \frac{1}{2}\sum_{j\in \mathrm{N}(k)} r_i(j)\|\bar{w}(k)\|^2 \Big/ \sum_{j\in \mathrm{N}(k)} r_i(j), \qquad i=1,2. \qquad (13)$$

2.3. CSLapMixShrink function

This section describes a non-linear shrinkage function for wavelet-based denoising that is derived by assuming that the pdf of the noise-free wavelet coefficients are given by (9). According to Fig. 2, using a binary sample space, we can obtain an estimator for $\bar{w}(k)$ by combining the estimates of $\bar{w}$ under H_0 and H_1 as follows:

$$\hat{\bar{w}}(k) = p_a(\bar{y}(k))\hat{\bar{w}}_1(k) + p_{1-a}(\bar{y}(k))\hat{\bar{w}}_2(k), \qquad (14)$$

where $p(H_0 \mid \bar{y}(k)) := p_a(\bar{y}(k))$ and $p(H_1 \mid \bar{y}(k)) := p_{1-a}(\bar{y}(k))$ are respectively the probability of being in H_0 and H_1, when $\bar{y}(k)$ has been observed. For $i=1,2$ the expression $\hat{\bar{w}}_i(k)$ is an estimate of $\bar{w}(k)$ based on the assumption that $\bar{w}(k)$ was generated under H_{i-1}. If H_{i-1} states that our proposed pdf is circular symmetric Laplacian pdf with variance $\sigma_i(k)$, then we use (7) to get $\hat{\bar{w}}_i(k)$, i.e.:

$$\hat{\bar{w}}(k) = p_a(\bar{y}(k))\,\mathrm{CShrinkL}\left(\bar{y}(k), \sqrt{3}\sigma_n^2/\sigma_1(k)\right) + p_{1-a}(\bar{y}(k))\,\mathrm{CShrinkL}\left(\bar{y}(k), \sqrt{3}\sigma_n^2/\sigma_2(k)\right) \qquad (15)$$

Table I. Summary of the proposed algorithm.

Step 1: Signal transformation of noisy observation
Step 2: Initialization of $\sigma_1(k,m), \sigma_2(k,m), a(k,m)$
Step 3: Calculating of $p_1(\bar{w}(k),m), p_2(\bar{w}(k),m)$
Step 3: Calculating of $r_1(k,m), r_2(k,m)$ from (10)
Step 4: Calculating of $a(k,m+1)$ from (11) and $\sigma_1(k,m+1), \sigma_2(k,m+1)$ from (12)
Step 5: Returning to Step 3 to converge of the parameters
Step 6: Using the obtained parameter in (17) for calculating the denoised coefficients
Step 7: Inverse signal transformation

We use the following Bayes' rule in order to calculate $p_a(\bar{y}(k))$ and $p_{1-a}(\bar{y}(k))$:

$$p_a(\bar{y}(k)) = \frac{a(k)g_1(\bar{y}(k))}{a(k)g_1(\bar{y}(k))+(1-a(k))g_2(\bar{y}(k))}, \qquad (16)$$

$$p_{1-a}(\bar{y}(k)) = 1 - p_a(\bar{y}(k)),$$

where $p(\bar{y}(k) \mid H_{i-1}) := g_i(\bar{y}(k))$ denotes the pdf of $\bar{y}(k)$ according to the assumption that $\bar{w}(k)$ was generated under H_{i-1}. Because $\bar{y}(k)$ is the sum of $\bar{w}(k)$ and the independent Gaussian noise, the pdf of $\bar{y}(k)$ is the 2d convolution of the pdf of $\bar{w}(k)$ and the bivariate Gaussian pdf. After some simplification, we can estimate this pdf as follows:

$$g_i(\bar{y}(k)) = \frac{3}{2\pi\sigma_i^2(k)}\exp\left(-\frac{\|\bar{y}(k)\|^2}{2\sigma_n^2}\right)\left\{1+\frac{1}{\sigma_n\sqrt{2\pi}}\sum_{j=0}^{M-1}\left[\frac{\|\bar{y}(k)\|(\frac{2j}{M}-1)-\frac{\sqrt{3}\sigma_n^2}{\sigma_i(k)})}{\sqrt{jM-M^2}}\,\mathrm{erfcx}\left(-\frac{\|\bar{y}(k)\|(\frac{2j}{M}-1)-\frac{\sqrt{3}\sigma_n^2}{\sigma_i(k)}}{\sigma_n\sqrt{2}}\right)\right]\right\} \qquad (16)$$

where $\mathrm{erfcx}(x)=\exp(x^2)(1-2\int_0^x e^{-t^2}dt/\sqrt{\pi})$. Our numerical experiments show that for $M \geq 10$ the approximation error is less than 1%.

Defining $R=[(1-a)g_2(\bar{y}(k))]/[ag_1(\bar{y}(k))]$, we obtain:

$$\hat{\bar{w}}(k) = \frac{\mathrm{CShrinkL}(\bar{y}(k), \frac{\sqrt{3}\sigma_n^2}{\sigma_1(k)}) + R\,\mathrm{CShrinkL}(\bar{y}(k), \frac{\sqrt{3}\sigma_n^2}{\sigma_2(k)})}{1+R} \qquad (17)$$

The algorithm is summarized in Table I.

3. EXPERIMENTAL RESULTS

In this section, we demonstrate the performance of the proposed algorithm by examples using CSLapMixShrinkL and compare the results with other methods such as HT [1], WIN-SAR [5], CShrinkL [2], LapMixShrinkL [6], CSLap-MixShrink [7].

Fig. 3. From top: CSLapMixShrink function, denoised image with CShrinkL, denoised image with soft thresholding and denoised image with our method (CSLapMixShrinkL).

Fig. 3 illustrates CSLapMixShrink function and denoised images produced from CSLapMixShrink, soft thresholding and CSLapMixShrinkL for a part of *512×512* grayscal Barbara image corrupted with additive Gaussian noise ($\sigma_n=20$). We can also see a comparison between our algorithm with other methods at different additive Gaussian noise levels $\sigma_n=10,\ 15,\ 20,\ 25,\ 30$ for two *512×512* grayscale images, namely, Lena and Boat, with other methods in Table II where the best PSNR performance among algorithms is highlighted by bold number.

4. CONCLUSION AND DISCUSSIONS

In this paper we used a mixture of circular symmetric Laplacian pdfs for noise-free wavelet coefficients and derived CsLapMixShrinkL function for image denoising. Experiment results show that this algorithm mostly outperforms several existing effective techniques in the literature. In terms of PSNR up to 0.5 dB improvement is obtained compared with existing algorithms. As seen from these results, our algorithm outperforms the others in most cases. Visually, the denoised image using the proposed method is more similar to the noise-free image than using other algorithms. For example, the CShrinkL [2] produces images that are softer than original images, i.e., the details of the images are blurred. Since, CSLapMixShrink [7] is not a local algorithm the denoised image using this method has more visual artifacts. In [6], it is discussed that some other methods are more effective than LapMixShrinkL for crowded images. In contrast the proposed algorithm also performs well for crowded images.

Instead of this pdf, one may use other mixture pdfs such as a mixture of bivariate Cauchy pdfs with local parameters in order to derive other shrinkage functions.

Table II. PSNR in dB for Several Denoising Methods.

σ_n	HT	WIN-SAR	CS-ShrinkL	LapMix-ShrinkL	CSLapMi-xShrink	CSLapMi-xShrinkL
Lena						
10	34.23	34.33	35.34	**35.35**	34.99	35.30
15	32.60	32.74	33.67	------	33.34	**33.81**
20	31.30	31.52	32.40	32.31	32.11	**32.55**
25	30.23	30.57	31.40	------	31.09	**31.61**
30	29.40	29.89	30.54	30.58	30.34	**30.94**
Boat						
10	32.40	32.85	33.10	33.28	33.03	**33.23**
15	30.50	30.87	31.36	------	31.27	**31.42**
20	29.27	29.65	30.08	29.88	30.04	**30.19**
25	28.23	28.64	29.06	------	28.96	**29.19**
30	27.44	27.92	28.31	28.03	28.14	**28.47**

5. REFERENCES

[1] D. L. Donoho, "Denoising by soft-thresholding," *IEEE Trans. Inform. Theory,* vol. 41, pp. 613–627, May 1995.

[2] L. Sendur, I. Selesnick, "Bivariate shrinkage with local variance," *IEEE Signal Processing Letters,* 9(12):438-441, Dec. 2002.

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

[4] M. S. Crouse, R. D. Nowak, R. G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," *IEEE Trans. Signal Processing,* vol. 46, no. 4, pp. 886-902, Apr. 1998.

[5] A. Achim, P. Tsakalides, "SAR image denoising via Bayesian wavelet shrinkage based on heavy-tailed modeling," *IEEE Trans. Geosci. Remote Sensing,* vol. 41, no. 8, pp. 1773–1784, Aug. 2003.

[6] H. Rabbani, M. Vafadust, S. Gazor "Image denoising based on a mixture of Laplace distributions with local parameters in complex wavelet domain," *Proc. 13th IEEE International Conference on Image Processing,* Atlanta, October 8-11, 2006.

[7] H. Rabbani, M. Vafadust, S. Gazor "Image denoising based on a mixture of circular symmetric Laplacian models in complex wavelet domain," Accepted in *2007 ISSPA*, Sharjah, Feb.2007.