Denoising by Spatial Correlation Thresholding

Lei Zhang and Paul Bao

***Abstract*—This paper presents a spatial-correlation thresholding scheme for noise reduction by wavelet transform. Observing that edge structures are of high magnitude across wavelet scales but noise decays rapidly, we multiply two adjacent wavelet scales to form a spatial-correlation function to enhance significant structures and dilute noise. Dissimilar to the traditional thresholding schemes that apply threshold to the wavelet coefficients, the proposed scheme applies threshold directly to the scale correlation. A robust threshold is presented and experiments show that the proposed scheme outperforms the traditional thresholding methods.**

***Index Terms*—Image denoising, spatial correlation, threshold, wavelet transform.**

I. Introduction

DENOISING is essential in image analysis. Wavelet transform (WT) [1]–[3] -based schemes have proved to be effective, especially the nonlinear threshold-based denoising schemes [6]–[9]. In these approaches, a threshold is preset to determine if a wavelet coefficient should be preserved (shrunk) or eliminated.

Donoho proposed a *wavelet shrinkage* method [6]. The *soft* threshold $t = \sigma\sqrt{2\log N}$ presented in his method proved to be *smooth* and *adaptive* in minimax sense, where σ is the standard deviation of the additive noise and N is the signal length. The word *soft* implies that it shrinks the input w to zero by amount t, i.e., $\eta_t(w) = \text{sgn}(w) \cdot \max(0, |w| - t)$. Following Donoho's pioneering work, some new *soft* thresholds were presented [7]–[9]. Pan *et al.* [8] proposed a *hard* threshold $t(j) = c\sigma_j$ for nonorthogonal wavelet transforms, where σ_j is the standard deviation of noise at the jth scale and c is a constant between 3 and 4. The word *hard* implies that the input w is preserved if it is greater than the threshold, otherwise set to zero: $\eta_t(w) = w \cdot \mathbf{1}\{|w| > t\}$. Generally, soft threshold yields smaller risk in minimax sense but may *over-smooth* the images.

In wavelet domain, the edge structures will evolve with observable magnitudes across scales while noise decreases rapidly [1]. Based on this observation, Xu *et al.* [5] presented a spatially selected filtering technique, where the adjacent scales of WT are multiplied and an iterative selection is utilized to identify the edge structures. The scheme was improved in [8]. In [4], Sadler *et al.* employed the multiscale products of WT for step detection and estimation.

Thresholding is efficient but does not take advantage of the correlation information across wavelet scales, while the latter is exploited by the spatial selective technique. To combine the merits of the two techniques, this paper presents a spatial-correlation thresholding scheme where the adjacent wavelet scales are multiplied to amplify edge structures and dilute noise. Then a threshold is determined and applied to the multiscale products to identify significant structures. The spatial-correlation threshold will better distinguish edge structures from noise than the traditional threshold imposed directly on wavelet coefficients.

II. Wavelet Spatial Correlation

Suppose function $\psi(x)$ is a mother wavelet. Let $\psi_j(x) = 2^{-j}\psi(2^{-j}x)$ be the dilation of mother wavelet $\psi(x)$ on dyadic sequence $\{2^j\}_{j\in Z}$. The dyadic wavelet transform (DWT) of $f(x)$ at scale 2^j and position x is

$$W_j f(x) = f * \psi_j(x) \tag{1}$$

where * denotes convolution operation. In the case of images, two wavelets $\psi^1(x,y)$ and $\psi^2(x,y)$ should be defined. Let $\psi_j^i(x) = 2^{-2j}\psi^i(2^{-j}x, 2^{-j}y)$ be the dilation of $\psi^i(x,y), i = 1, 2$. The DWT of $f(x,y)$ at scale 2^j and position (x,y) has two components

$$W_j^i f(x,y) = f * \psi_j^i(x,y), \quad i = 1, 2. \tag{2}$$

The fast algorithm of 2-D DWT is illustrated in Fig. 1 [2].

Singularities and noise have different evolution across wavelet scales [1]. In the WT domain, edge structures will evolve with considerable peaks along scales but noise will deteriorate rapidly. Multiplying the adjacent wavelet subbands would enhance edges while diluting noise. An edge structure may be centered with relative shifts at different scales, which suggests that the adjacent wavelet scales should be multiplied with some relative translations to maximize the product. In practice, it is sufficient to amplify the edge structures by employing two adjacent scales. We define the spatial-correlation function as

$$P_j^i\left(n, m, \tau_{j,n}^i, \tau_{j,m}^i\right) = W_j^i f(n,m) \cdot W_{j+1}^i f\left(n + \tau_{j,n}^i, m + \tau_{j,m}^i\right), \quad i = 1, 2 \tag{3}$$

where the shifts $\tau_{j,n}^i, \tau_{j,m}^i$ satisfy

$$R_{j+1,j}^i\left(\tau_{j,n}^i, \tau_{j,m}^i\right) = E\left[P_j^i\left(n, m, \tau_{j,n}^i, \tau_{j,m}^i\right)\right] = \max, \quad i = 1, 2. \tag{4}$$

Manuscript received May 7, 2001; revised December 1, 2002. This paper was supported in part by Grant CUHK5982/00E of the Research Grant Council of the Hong Kong Special Administrative Region. This paper was recommended by Associate Editor R. Stevenson.

L. Zhang is with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4L8, Canada (e-mail: johnray@mail.ece.mcmaster.ca).

P. Bao is with the Department of Information Engineering, The Chinese University of Hong Kong, Shatin, New Territory, Hong Kong (e-mail: paulbao@ie.cuhk.edu.hk).

Digital Object Identifier 10.1109/TCSVT.2003.813426

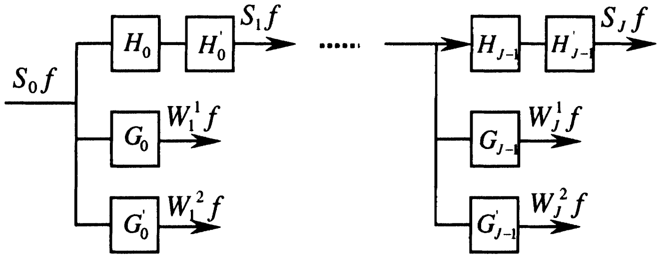


Fig. 1. 2-D DWT structure, where filter $H_i(G_i)$ is the 2^i dilation of $H_0(G_0)$ (putting $2^i - 1$ zeros between each of coefficients of $H_0(G_0)$) and $H_i'(G_i')$ is the transition of $H_i(G_i)$.

Shifts $\tau_{j,n}^i$ and $\tau_{j,m}^i$ depend on the wavelet basis, i.e., the filters H_i and G_i in Fig. 1. $W_j^1 f$ can be written as $W_j^1 f = S_0 f * F_j^1$, where the filter F_j^1 is

$$F_j^1 = H_0 * H_0' * H_1 * H_1' * \cdots * H_{j-2} * H_{j-2}' * G_{j-1}. \quad (5)$$

Similarly, $W_j^2 f = S_0 f * F_j^2$ with

$$F_j^2 = H_0 * H_0' * H_1 * H_1' * \cdots * H_{j-2} * H_{j-2}' * G_{j-1}'. \quad (6)$$

Mathematically, we can obtain

$$R_{j+1,j}^i \left(\tau_{j,n}^i, \tau_{j,m}^i\right) = \left(R_{S_0 f} * R_{F_{j+1}^i, F_j^i}\right) \left(\tau_{j,n}^i, \tau_{j,m}^i\right), \quad i = 1, 2 \quad (7)$$

where

$$R_{S_0 f}\left(\tau_{j,n}^i, \tau_{j,m}^i\right) = E\left[S_0 f(n,m) S_0 f\left(n + \tau_{j,n}^i, m + \tau_{j,m}^i\right)\right] \quad (8)$$

$$R_{F_{j+1}^i, F_j^i}\left(\tau_{j,n}^i, \tau_{j,m}^i\right) = F_{j+1}^i\left(\tau_{j,n}^i, \tau_{j,m}^i\right) * F_j^i\left(-\tau_{j,n}^i, -\tau_{j,m}^i\right). \quad (9)$$

The self-correlation function $R_{S_0 f}$ is symmetric with respect to Z-coordinate and maximizes at the origin (0, 0). The Mallat wavelet $\psi(x)$ [2] used in this paper is an anti-symmetrical smooth quadratic spline, and thus $R_{F_{j+1}^i, F_j^i}$ will be smooth and symmetrical. Convoluting $R_{S_0 f}$ with $R_{F_{j+1}^i, F_j^i}$ will smooth $R_{F_{j+1}^i, F_j^i}$. The silhouette of $R_{j+1,j}^i$ is similar to that of $R_{F_{j+1}^i, F_j^i}$ and they maximize at the same position. Thus $\tau_{j,n}^i$ and $\tau_{j,m}^i$ can be determined by the maxima of $R_{F_{j+1}^i, F_j^i}$. We have

$$\tau_{j,n}^1 = 2^{j-1}, \quad \tau_{j,m}^1 = 2^{j-2}, \quad \left(\tau_{1,m}^1 = 0\right), \quad \tau_{j,n}^2 = 2^{j-2}, \quad \left(\tau_{1,n}^2 = 0\right), \quad \tau_{j,m}^2 = 2^{j-1}. \quad (10)$$

III. The Thresholding Scheme

A. Algorithm Description

Suppose $f = g + \varepsilon$ is the observation of the original image g corrupted by *Gaussian* white noise $\varepsilon \sim N(0, \sigma^2)$. The denoising aims at estimating an image $\hat{g}$ from f. Wavelet-based thresholding schemes have proved to be effective [6]–[9]. Nonsignificant wavelet coefficients below a preset threshold are discarded as noise and the image is restored by the remaining coefficients. Most schemes apply threshold t directly to wavelet coefficients. If t is relatively large, some edge structures may be suppressed as noise. Contrarily, if t is relatively small, many noisy pixels would be undesirably preserved. These thresholds make no use of the correlation information distributed between wavelet scales. In the spatial correlation P_j^i, image edges would be strengthened and noise would be diluted, which leads to an effective discrimination between edges and noise.

In this paper, a new denoising method, the *spatial-correlation thresholding*, is proposed. A threshold $t_{\rm sc}^i(j)$ is applied to P_j^i to identify the significant structures. $t_{\rm sc}^i(j)$ is a scale-dependent *hard* threshold. The algorithm is summarized as follows.

1) Transform input $S_0 f$ into J scales to get $S_J f$ and $W_j f, j = 1, 2, \ldots, J$.
2) Calculate the spatial correlation P_j^i and apply the $hard$ threshold $t_{\rm sc}^i(j)$ to it

$$\hat{W}_j^i f(n,m) = \begin{cases} W_j^i f(n,m) & P_j^i\left(n, m, \tau_{j,n}^i, \tau_{j,m}^i\right) \geq t_{\rm sc}^i(j) \\ 0 & P_j^i\left(n, m, \tau_{j,n}^i, \tau_{j,m}^i\right) < t_{\rm sc}^i(j) \end{cases}, \quad i = 1, 2. \quad (11)$$

3) Recover the estimated image $\hat{g}$ from $S_J f$ and $\hat{W}_j f, j = 1, 2, \ldots, J$.

B. Determination of the Threshold

Referring to Fig. 1, suppose the input is *Gaussian* white noise $S_0 \varepsilon \sim N(0, \sigma^2)$ and its DWT is $W_j^i \varepsilon = S_0 \varepsilon * F_j^i, i = 1, 2$. Due to the commutativity of the convolution operation, filters F_j^1 and F_j^2 can be written as

$$F_j^1 = F_{j,n}^1 * F_{j,m}^1 \quad \text{and} \quad F_j^2 = F_{j,n}^2 * F_{j,m}^2 \quad (12)$$

where

$$\begin{aligned} F_{j,n}^1 &= H_0 * H_1 * \cdots * H_{j-2} * G_{j-1} \\ F_{j,m}^1 &= H_0' * H_1' * \cdots H_{j-2}' \\ F_{j,n}^2 &= H_0 * H_1 * \cdots * H_{j-2}, \\ F_{j,m}^2 &= H_0' * H_1' * \cdots * H_{j-2}' * G_{j-1}'. \end{aligned} \quad (13)$$

$W_j^i \varepsilon$ is *Gaussian* colored noise: $W_j^i \varepsilon \sim N(0, (\sigma_j^i)^2)$, where

$$\sigma_j^i = \left\|F_{j,n}^i\right\| \left\|F_{j,m}^i\right\| \sigma \quad (14)$$

and $\|\cdot\|$ denotes the norm of a vector $S(n) \in l^2(n)$.

The spatial correlation of $W_j^i \varepsilon$ is

$$P_j^i \varepsilon\left(n, m, \tau_{j,n}^i, \tau_{j,m}^i\right) = W_j^i \varepsilon(n,m) \cdot W_{j+1}^i \varepsilon\left(n + \tau_{j,n}^i, m + \tau_{j,m}^i\right) \quad (15)$$

Let $Y_j^i(n,m) = P_j^i \varepsilon(n, m, \tau_{j,n}^i, \tau_{j,m}^i)$. Normalizing $W_j^i \varepsilon$ as

$$\bar{X}_j^i = W_j^i \varepsilon / \sigma_j^i \quad (16)$$

we have $\bar{X}_j^i \sim N(0, 1)$. Then Y_j^i is normalized as $\bar{Y}_j^i = \bar{X}_j^i \cdot \bar{X}_j^i$. Letting

$$\bar{Y}_{j,+}^i = \left(\bar{X}_j^i + \bar{X}_j^i\right)/2 \quad \text{and} \quad \bar{Y}_{j,-}^i = \left(\bar{X}_j^i - \bar{X}_j^i\right)/2 \quad (17)$$

we have

$$\bar{Y}_j^i = \left(\bar{Y}_{j,+}^i\right)^2 - \left(\bar{Y}_{j,-}^i\right)^2. \quad (18)$$

$\bar{Y}^i_{j,+}$ and $\bar{Y}^i_{j,-}$ are *Gaussian* distributed: $\bar{Y}^i_{j,+} \sim N(0, \sigma^2_{\bar{Y}^i_{j,+}})$ and $\bar{Y}^i_{j,-} \sim N(0, \sigma^2_{\bar{Y}^i_{j,-}})$, where

$$\sigma_{\bar{Y}^i_{j,+}} = \frac{1}{2}\sqrt{\sum_k \left(\frac{F^i_{j,n}(k)}{\|F^i_{j,n}\|} + \frac{F^i_{j+1,n}\left(k+\tau^i_{j,n}\right)}{\|F^i_{j+1,n}\|}\right)^2} \cdot \sqrt{\sum_l \left(\frac{F^i_{j,m}(l)}{\|F^i_{j,m}\|} + \frac{F^i_{j+1,m}\left(l+\tau^i_{j+1,m}\right)}{\|F^i_{j+1,m}\|}\right)^2}$$

$$\sigma_{\bar{Y}^i_{j,-}} = \frac{1}{2}\sqrt{\sum_k \left(\frac{F^i_{j,n}(k)}{\|F^i_{j,n}\|} - \frac{F^i_{j+1,n}\left(k+\tau^i_{j,n}\right)}{\|F^i_{j+1,n}\|}\right)^2} \cdot \sqrt{\sum_l \left(\frac{F^i_{j,m}(l)}{\|F^i_{j,m}\|} - \frac{F^i_{j+1,m}\left(l+\tau^i_{j+1,m}\right)}{\|F^i_{j+1,m}\|}\right)^2}. \quad (19)$$

Due to the strong correlation between $F^i_j(n,m)$ and $F^i_{j+1}(n+\tau^i_{j+1,n}, m+\tau^i_{j+1,m}), \sigma_{\bar{Y}^i_{j,+}}$ will be greater than $\sigma_{\bar{Y}^i_{j,-}}$. For the dyadic Mallat wavelet, the ratios of $\sigma^2_{\bar{Y}^i_{j,+}}$ to $\sigma^2_{\bar{Y}^i_{j,-}}$ at the first three scales are $\sigma^2_{\bar{Y}^i_{j,+}}/\sigma^2_{\bar{Y}^i_{j,-}} = 2.12, 3.45, 3.95$.

Denote by $t^i_{\mathrm{sc}}(j)$ the threshold applied to Y^i_j. For the de-noising purpose, it is expected that $t^i_{\mathrm{sc}}(j)$ could suppress almost all the values in Y^i_j, i.e., $P(y^i_j < t^i_{\mathrm{sc}}(j)) \to 1$. Letting $\bar{t}^i_{\mathrm{sc}}(j) = t^i_{\mathrm{sc}}(j)/(\sigma^i_j \sigma^i_{j+1})$, we have

$$\begin{aligned} P\left(y^i_j < t^i_{\mathrm{sc}}(j)\right) &= P\left(\bar{y}^i_j < \bar{t}^i_j(j)\right) \\ &= P\left(\left(\bar{y}^i_{j,+}\right)^2 < \bar{t}^i_j(j) + \left(\bar{y}^i_{j,-}\right)^2\right) \\ &\geq P\left(\left(\bar{y}^i_{j,+}\right)^2 < \bar{t}^i_j(j)\right) \\ &= P\left(|\bar{y}^i_{j,+}| < \sqrt{\bar{t}^i_j(j)}\right). \end{aligned}$$

Because $\bar{Y}^i_{j,+}$ is Gaussian distributed, $\sqrt{\bar{t}^i_{\mathrm{sc}}(j)} \geq 3\sigma_{\bar{Y}^i_{j,+}}$ will lead to

$$\begin{aligned} P&\left(y^i_j < t^i_{\mathrm{sc}}(j) \,\middle|\, t^i_{\mathrm{sc}}(j) \geq 9\sigma^i_j \sigma^i_{j+1} \sigma^2_{\bar{Y}^i_{j,+}}\right) \\ &\geq P\left(\left|\bar{y}^i_{j,+}\right| < \sqrt{\bar{t}^i_{\mathrm{sc}}(j)} \,\middle|\, \bar{t}^i_{\mathrm{sc}}(j) \geq 9\sigma^2_{\bar{Y}^i_{j,+}}\right) > 0.9973 \to 1. \end{aligned}$$

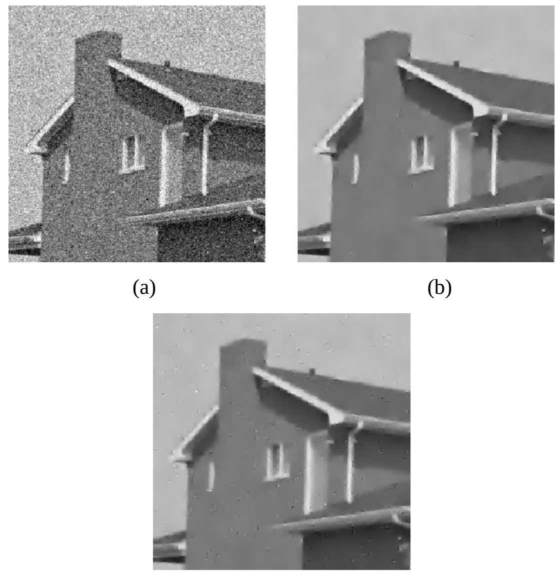
In applications, the input is $S_0 f = S_0 g + S_0 \varepsilon$, where $S_0 g$ is the original image. Due to the linearity of WT, $W^i_j f = W^i_j g + W^i_j \varepsilon$. At the fine scales, $W^i_j \varepsilon$ will be predominant in $W^i_j f$ except for some significant structures to be preserved. To ensure an effective de-noising, $t^i_{\mathrm{sc}}(j)$ should not be set too high. Since the difference between edges and noise is fairly amplified in $P^i_j f$, threshold $t^i_{\mathrm{sc}}(j)$ would be more effective in discriminating edges from noise compared with the traditional thresholding. In our experiments, a setting of $t^i_{\mathrm{sc}}(j) = c \cdot \sigma^i_j \sigma^i_{j+1} \sigma^2_{\bar{Y}^i_{j,n}}$ with $c =$ 10–12 will effectively remove noise while preserving edges.

IV. Experiments

This section illustrates the performance of the proposed spatial-correlation thresholding scheme applying on some benchmark images. The additive noise is assumed zero mean Gaussian white. The *hard* thresholding scheme in [8] is used for comparison, which is described as follows:

(a) (b)

Fig. 2. Test images. (a) *House*. (b) *Cameraman*.



(c)

Fig. 3. (a) Noisy *House* ($\mathrm{SNR} = 14.23$ dB). (b) By spatial-correlation thresholding ($\mathrm{SNR} = 24.96$ dB). (c) By traditional thresholding ($\mathrm{SNR} = 23.87$ dB).

$$h\hat{W}^i_j f(n,m), \\ = \begin{cases} W^i_j f(n,m) & \left|W^i_j f(n,m)\right| \geq t^i(j) \\ 0, & \left|W^i_j f(n,m)\right| < t^i(j) \end{cases}, \quad i = 1,2 \quad (20)$$

where threshold $t^i(j) = c\sigma^i_j$ with $c \in [3,4]$.

Two test images *House* and *Cameraman* are shown in Fig. 2. Fig. 3(a) is the noisy *House* with $\mathrm{SNR} = 14.23$ dB. The de-noised image by the proposed scheme, illustrated in Fig. 3(b), achieves a SNR of 24.96 dB. The traditional hard thresholding result is shown in Fig. 3(c) with $\mathrm{SNR} = 23.87$ dB. Our scheme achieves a higher SNR. Noticeably, the image in Fig. 3(c) is little *over smoothed*. For better illustration, a zoom-in of the "*window*" in *House* is shown in Fig. 4, in which we can find that the *window* edges are better preserved by the spatial-correlation thresholding while they are obviously blurred by the traditional method. Furthermore, it should be noticed that in Fig. 4(c) some *stings* generated by noise are identified in error by the traditional thresholding.

For *Cameraman*, the results are similar to that of *House*. Noisy *Cameraman* of $\mathrm{SNR} = 9.49$ dB is shown in Fig. 5(a) and the recovered images are shown in Fig. 5(b) and (c), with

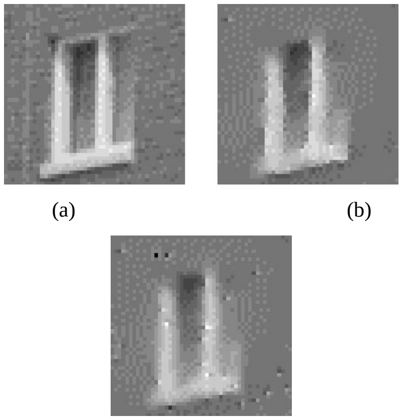


Fig. 4. Zoom-in of the "*window*". (a) Original. (b) By spatial-correlation thresholding. (c) By traditional thresholding.

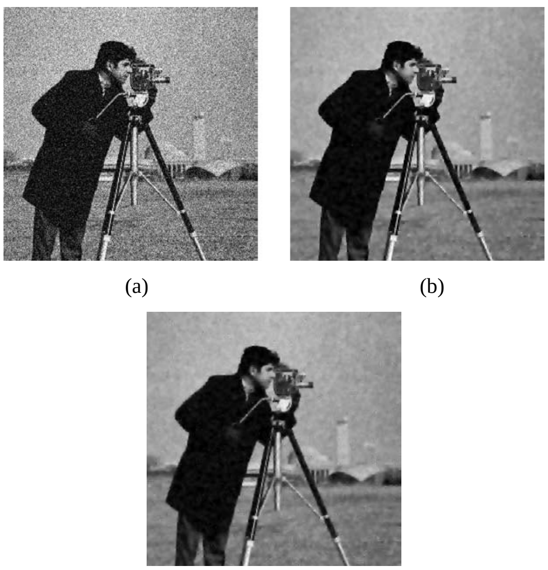


Fig. 5. (a) Noisy *Cameraman* ($\mathrm{SNR} = 9.49$ dB). (b) By spatial-correlation thresholding ($\mathrm{SNR} = 21.13$ dB). (c) By traditional thresholding ($\mathrm{SNR} = 19.73$ dB).

$\mathrm{SNR} = 21.13$ dB and $\mathrm{SNR} = 19.73$ dB, respectively. The proposed scheme outperforms the traditional scheme in both the SNR comparison and visual perception.

V. Conclusion

In this paper, a wavelet-threshold-based denoising scheme is presented. Unlike many popular schemes that directly threshold the wavelet coefficients, the proposed method multiplies adjacent wavelet scales to amplify instantaneous structures and then applies thresholding to the multiplication. It is shown that the proposed threshold can distinguish edges from noise more effectively and achieve better results in the SNR measurement and the visual perception compared with the traditional thresholding schemes.

References

[1] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Trans. Inform. Theory*, vol. 32, pp. 617–643, Mar. 1992.
[2] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, pp. 710–732, July 1992.
[3] M. Vetterli and C. Herley, "Wavelet and filter banks: Theory and design," *IEEE Trans. Signal Processing*, vol. 40, pp. 2207–2232, Sept. 1992.
[4] B. M. Sadler and A. Swami, "Analysis of multiscale products for step detection and estimation," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1043–1051, Apr. 1999.
[5] Y. Xu *et al.*, "Wavelet transform domain filters: A spatially selective noise filtration technique," *IEEE Trans. Image Processing*, vol. 3, pp. 747–758, Nov. 1994.
[6] D. L. Donoho, "De-noising by soft thresholding," *IEEE Trans. Inform. Theory*, vol. 41, pp. 613–627, May 1995.
[7] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *J. Amer. Statist. Assoc.*, vol. 90, pp. 1200–1224, Dec. 1995.
[8] Q. Pan *et al.*, "Two denoising methods by wavelet transform," *IEEE Trans. Signal Processing*, vol. 47, pp. 3401–3406, Dec. 1999.
[9] S. G. Chang *et al.*, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. Image Processing*, vol. 9, pp. 1532–1546, Sept. 2000.