Invited paper at SPIE Wavelet XI, San Diego, July 31-August 4, 2005

Monochrome and Color Image Denoising Using Neighboring Dependency and Data Correlation

Dongwook Cho and Tien D. Bui

Dept. of Computer Science & Software Engineering, Concordia University,
1455 De Maisonneuve Blvd. West, Montreal, Quebec, H3G 1M8, Canada

ABSTRACT

In this paper, two approaches for image denoising that take advantages of neighboring dependency in the wavelet domain are studied. The first approach is to take into account the higher order statistical coupling between neighboring wavelet coefficients and their corresponding coefficients in the parent level. The second is based on multivariate statistical modeling. The estimation of the clean coefficients is obtained by a general rule using Bayesian approach. Various estimation expressions can be obtained by *a priori* probability distribution, called multivariate generalized Gaussian distribution (MGGD). The experimental results show that both of our methods give comparatively higher peak signal to noise ratio (PSNR) as well as little visual artifact for monochrome images. In addition, we extend our approaches to a denoising algorithm for color image that has multiple color components. The proposed color denoising algorithm is a framework to consider the correlations between color components yet using the existing monochrome denoising method without modification. Denoising results in this framework give noticeable better improvement than in the case when the correlation between color components is not considered.

Keywords: Denoising, wavelet, color image, correlation, neighboring dependency

1. INTRODUCTION

For the last decade, various denoising approaches using wavelet transform have been proposed and proved to be efficient for images as well as for signals. They have shown that denoising using wavelet transforms produces superb results. This is because wavelet transform has the compaction property of having only a small number of significant coefficients and a large number of detailed coefficients. Therefore, it is possible to suppress the noise in the wavelet domain by killing the detailed coefficients that represent the detailed information as well as the noise.

Wavelet coefficients are not strongly correlated, but they still have dependency on each other. So many of the recent works have taken into account this dependency in order to obtain better coefficient estimate. Cai and Silverman[1] proposed a simple and effective approach for signal denoising by incorporating the neighboring coefficients. Chen and Bui[2] designed new neighboring threshold for multiwavelet based on *NeighBlock*. Their method considers both multiwavelet properties and neighboring dependency. Mihcak et al.[3] proposed a local variance estimator to get a locally-adaptive shrinkage value. Malfait and Roose proposed an image denoising algorithm using Markov random field image model as *a priori*.[4] Also Pizurica et al.[5] considered a joint inter- and intra-scale statistical model and improved the approach by Malfait and Roose. A parent coefficient in the coarser level was also considered to estimate a threshold by Sendur and Selesnick.[6] They obtained better results when they applied the local variance together with the dual-tree complex wavelet transform (DT CWT).[7] DT CWT provides better *shift-invariant* features and *directional selectivity* than the usual separable wavelet transform.[8] Portilla et al.[9] presented an image denoising algorithm which is based on a Gaussian scale mixture (GSM) model using an overcomplete multiscale oriented basis. They define a vector using neighboring coefficients and obtain an accurate estimate by the vector operations. All these works show that incorporating different information like neighbors and parents is helpful to remove noise and preserve details for natural image denoising.

Further author information: (Send correspondence to T.D.B.)
T.D.B.: E-mail: bui@cse.concordia.ca, Telephone: 1-514-848-2424 ext. 3014
D.C.: E-mail: d_cho@cse.concordia.ca

Most of the image denoising approaches we have mentioned have been designed for monochrome images. However more common image type in practice is color image composed of multiple color components. Only a few color image denoising algorithms using wavelet transform such as Refs. 10 and 11 have been developed. Unlike monochrome images, color images can be expressed as multiple components of monochrome images or a set of pixels represented by vectors. Therefore it is possible to extend most of the existing denoising approaches to color image straightforwardly by denoising each color component independently. In this case, correlation and dependency between color components are ignored. In order to estimate denoised color pixels accurately, it is necessary to take advantage of correlation or dependency of color components. In the spatial domain, there have been some works which consider a pixel as a vector of colors.[12,13] In this case, it is possible to develop a rule to estimate a correlated vector by averaging or utilizing order statistics between the vectors (or pixels).

In this paper, we discuss denoising approach for both monochrome and color images. We introduce two efficient methods we have proposed for monochrome image denoising. The first approach is to take advantage of the higher order statistical coupling between neighboring wavelet coefficients and their corresponding coefficients in the parent level. The other is based on multivariate statistical modeling. Here the clean coefficients are estimated by a general rule using Bayesian approach. Various estimation expressions can be obtained by *a priori* probability distribution called multivariate generalized Gaussian distribution (MGGD). The method can take into account various related information. Besides, we have developed the general framework which utilizes color components in the wavelet domain by decorrelating them statistically. In the decorrelated space, the sample space is randomly distributed and this results in more robust estimation for clean image.

The organization of this paper is as follows. In Sect. 2, we study two different approaches for monochrome image denoising. These methods have been motivated by how to efficiently incorporate and utilize neighboring wavelet coefficients. In Sect. 3 color denoising method is proposed based on one of our monochrome image denoising methods Experimental results are shown and compared with the existing methods in Sect. 4. And finally we give the conclusion and future work to be done in Sect. 5.

2. TWO APPROACHES

In this section, we briefly study and compare two different image denoising methods. In both methods, neighboring coefficients of given noisy image are utilized to estimate those of clean image accurately. However they have different approaches to obtain the clean estimates. The first method utilizes the L^p-norm of a vector composed of all the related neighboring coefficients by comparing it with universal threshold.[14] On the other hand, the second method generally estimates the clean coefficients using Bayesian statistics based on our multivariate *a priori* model.

We define some common notations first. Let A be a clean natural image with size $N \times N$, B a noisy image which can be expressed as $B = A + \sigma C$, and C zero-mean Gaussian white noise, which is $C \sim N(0,1)$. σ^2 is noise variance. After performing multiresolution wavelet decomposition on B, we get the wavelet coefficient $y_{j,k}$, which is the k-th wavelet coefficient in j-th level for B. Due to the linearity of the wavelet transform, we have:

$$y_{j,k} = x_{j,k} + \sigma z_{j,k}, \tag{1}$$

where $x_{j,k}$ and $z_{j,k}$ are the wavelet coefficients of A and C respectively in the same location as $y_{j,k}$.

2.1. *NeighLevel* : a simple and efficient wavelet denoising method

In the wavelet domain, the strong dependency between parent and child coefficients has been widely realized in image coding and denoising since *zerotrees* were introduced by Shapiro.[15] Among them, Cai and Silverman[1] proposed a simple and effective approach for a signal denoising. The method, called *NeighBlock*, takes the neighboring coefficients into account and obtains a threshold by comparing the sum of squared neighboring coefficients with Donoho's universal threshold. In addition, parent and child coefficients have inter-dependency similar to neighbors. Therefore, if we can properly utilize neighbors spread both vertically and horizontally as shown in Fig. 1, a better performance can be expected.

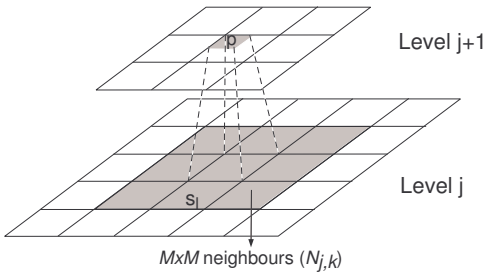


Figure 1. Choosing context for *NeighLevel*.

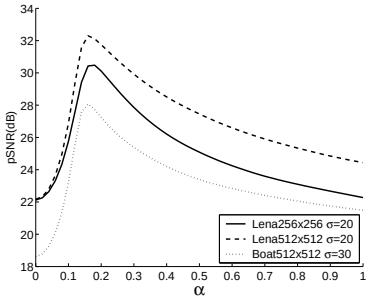


Figure 2. Performance changes as function θ in $\lambda^* = \theta\lambda$ for different kinds and sizes of images when $NeighLevel$ ($M = 5$) is applied using DT CWT.

Based on these ideas, we proposed an efficient image denoising approach called *NeighLevel* for a monochrome image in Ref. 16. This can be briefly described in the following equation:

$$\hat{x}_{j,k} = y_{j,k}\left[1 - (M^2+1)\lambda^{*2}\Big/\left(\sum_{s_l \in N_{j,k}} s_l^2 + p^2\right)\right]_+$$

or in general,

$$\hat{x}_{j,k} = y_{j,k}\left(1 - d\left|\frac{\lambda^*}{\|\mathbf{y}\|_r}\right|^r\right)_+, \tag{2}$$

where s_l denotes the coefficient to be thresholded and its neighbors in an $M \times M$ window, and p is a corresponding parent of the coefficient to be thresholded, which is a coefficient matched in the coarser level (see Fig. 1). $\mathbf{y}$ is a d-dimensional vector composed of all the related coefficients including the coefficient to be thresholded, neighbors and parent and $\|\mathbf{y}\|_r$ denotes L^r-norm (we mostly set r to 2). $\lambda^* = \theta\lambda$, where λ is the universal threshold $\lambda = \sqrt{2\sigma^2 \log N^2}$. θ is given as a parameter which satisfies $0 \le \theta \le 1$.

In Eq. (2), it should be noted that a normalized factor, d or M^2+1, is used which is the number of correlated elements in the context. By this rule, the effect of the local variance from the parent level is considered as well as from the current level.

In this method, it is important to choose an appropriate paramter θ. The universal threshold is designed for smoothness rather than for minimizing the errors. So λ is more meaningful when the signal is sufficiently smooth or the length of the signal is close to infinity. Natural images, however, are usually neither sufficiently smooth nor composed of infinite number of pixels. In fact, if we suppose that an optimal threshold that minimizes mean square error (MSE) (or maximizes peak signal to noise ratio (PSNR)), θ is always much less than 1.0 for natural images as shown in Fig. 2. Especially we got very similar θ value for different kinds and sizes of images when we applied soft thresholding rule. It might vary depending on the wavelet filter, but the appropriate range is similar for different images and noise level we have tested in our experiments.

2.2. Bayesian Estimation for Multivariate Statistical Model

Another method we are going to present is to use the statistical model of clean wavelet coefficients in addition to taking advantage of neighboring coefficients. This drives us to use multivariate statistical model. In Ref. 17, we have proposed the general estimation rule in the wavelet domain to obtain the denoised coefficients from the noisy image based on the multivariate statistical theory and Bayesian estimator. We briefly introduce the estimator in the following.

Let $\mathbf{x}$ be a d-dimensional wavelet coefficient vector, $\mathbf{x} = (x_1, x_2, \cdots, x_d)^t$, where x_1 is the wavelet coefficient under consideration and x_i ($i = 2, \cdots, d$) are the related coefficients to be taken into consideration, e.g. neighbors, parent and offsprings. Here for simplicity, we replace the double subscripts in $x_{j,k}$ by a single subscript x_i. The

corresponding vectors $\mathbf{y}$ and $\mathbf{z}$ can be similarly defined for the noisy image B and the noise C. We assume that x_i, y_i and z_i correspond to each other in both decomposition level and location. Therefore,

$$\mathbf{y} = \mathbf{x} + \sigma\mathbf{z}. \tag{3}$$

For the sake of simplicity, we omit subscripts j, k in Eq. (3) and the rest of the paper.

Our concern lies mainly in obtaining the estimate of a clean wavelet coefficient vector, $\hat{\mathbf{x}}$. $\hat{\mathbf{x}}$ should be obtained only from $\mathbf{y}$, a wavelet coefficient vector of the noisy image B. One of the ways to estimate $\hat{\mathbf{x}}$ is to use MAP estimator to maximize $p(\mathbf{x}|\mathbf{y})$. MAP estimator for $\hat{\mathbf{x}}$ can be obtained as follows:

$$\begin{aligned} \hat{\mathbf{x}} &= \arg\max_{\mathbf{x}\in\mathbb{R}^d} [\ln p(\mathbf{y}|\mathbf{x}) + \ln p(\mathbf{x})] \\ &= \arg\max_{\mathbf{x}\in\mathbb{R}^d} F(\mathbf{x}), \end{aligned} \tag{4}$$

where $F(\mathbf{x})$ represents the term inside arg max. This means that the optimal value $\hat{\mathbf{x}}$ with minimum probability error can be estimated by $p(\mathbf{y}|\mathbf{x})$ and $p(\mathbf{x})$.

From Eq. (3), $p(\mathbf{y}|\mathbf{x})$ is the multivariate Gaussian distribution with $N(\mathbf{0}, \Sigma_z = \sigma^2\mathbf{I})$ since Gaussian noise is independently and identically distributed for each element of the vector. Hence,

$$\ln p(\mathbf{y}|\mathbf{x}) = -\frac{d}{2}\ln(2\pi\sigma^2) - \frac{(\mathbf{y}-\mathbf{x})^t(\mathbf{y}-\mathbf{x})}{2\sigma^2}. \tag{5}$$

We assume that $p(\mathbf{x})$ is known. $p(\mathbf{x})$ might vary depending on the type of sample images. Also suppose that $g(\mathbf{x}) = \ln p(\mathbf{x})$ and there exists $\hat{\mathbf{x}}$ which satisfies $F(\hat{\mathbf{x}}) > \lim_{x_i\to\pm\infty} F(\mathbf{x})$. From Eqs. (4) and (5), Eq. (4) is equivalent to the solution of the following equation:

$$\begin{aligned} \nabla F(\hat{\mathbf{x}}) = -\frac{\hat{\mathbf{x}}-\mathbf{y}}{\sigma^2} + \nabla g(\hat{\mathbf{x}}) = 0 \\ \Leftrightarrow \hat{\mathbf{x}} = \mathbf{y} + \sigma^2\nabla g(\hat{\mathbf{x}}). \end{aligned} \tag{6}$$

Therefore the estimate of $\mathbf{x}$ highly depends on the probability density of clean wavelet coefficients, $p(\mathbf{x})$.

The existing models for wavelet denoising are usually based on univariate statistical model whereas $p(\mathbf{x})$ is a multivariate pdf in our model. There are several multivariate functions which are symmetric spherically like multivariate Gaussian model. In our paper, we use extended generalized Gaussian distribution (GGD) model[18] for its simple form and to achieve good fitting results. We call this model multivariate generalized Gaussian distribution (MGGD):

$$p(\mathbf{x}) = \gamma\exp\left\{-\left(\frac{(\mathbf{x}-\mu)^t\Sigma_{\mathbf{x}}^{-1}(\mathbf{x}-\mu)}{\alpha}\right)^{\beta}\right\}, \tag{7}$$

where α and β are parameters which can represent the spherical shape of the model and γ indicates a normalized constant defined by α, β and the covariance matrix $\Sigma_{\mathbf{x}}$.

When the dimension of $\mathbf{x}$ is one (scalar), the MGGD is still applicable and is denoted by UGGD. MGGD is a particular case of the v-spherical distribution defined by Fernández.[19] Using MGGD model, we can derive more specific forms of Eq. (6). Since we can assume that $\mu = \mathbf{0}$,

$$\nabla g(\mathbf{x}) = -\frac{2\beta}{\alpha^{\beta}}(\mathbf{x}^t\Sigma_{\mathbf{x}}^{-1}\mathbf{x})^{\beta-1}\Sigma_{\mathbf{x}}^{-1}\mathbf{x}. \tag{8}$$

From Eqs. (6) and (8),

$$\hat{\mathbf{x}} = \left(\Sigma_{\hat{\mathbf{x}}} + \frac{2\sigma^2\beta}{\alpha^{\beta}}(\hat{\mathbf{x}}^t\Sigma_{\hat{\mathbf{x}}}^{-1}\hat{\mathbf{x}})^{\beta-1}I\right)^{-1}\Sigma_{\hat{\mathbf{x}}}\mathbf{y}. \tag{9}$$

To simplify Eq. (9), we define $q(\hat{\mathbf{x}}) = \hat{\mathbf{x}}^t \Sigma_{\hat{\mathbf{x}}}^{-1} \hat{\mathbf{x}}$. Hence :

$$q(\hat{\mathbf{x}}) = \mathbf{y}^t \left(\Sigma_{\hat{\mathbf{x}}} + \frac{2\sigma^2 \beta \{q(\hat{\mathbf{x}})\}^{\beta-1}}{\alpha^\beta} I \right)^{-2} \Sigma_{\hat{\mathbf{x}}} \mathbf{y}. \tag{10}$$

Eqs. (9) and (10) allow us to solve for $\hat{\mathbf{x}}$.

However, there is no general solution for Eq. (10). To overcome this problem, we can define a particular condition for α, β and $\Sigma_{\hat{\mathbf{x}}}$ or use a numerical method. In our case, we simply use Newton's method.

3. COLOR IMAGE DENOISING

Unlike monochrome images, a color image can be described as a set of multiple image components. In this paper, the color image representation is based on the trichromatic theory, which separates the color image as three color components, i.e. red, green, and blue (RGB color space). Therefore we assume that a given noisy image includes randomly distributed Gaussian additive noise for each RGB channel. RGB color space models a physical color element, so it does not consider human visual system (HVS). Various color spaces which consider HVS (e.g. HSI, YIQ, L*a*b, etc) have been proposed for this purpose.[13] Many of these color spaces can separate two uncorrelated parts, luminance and chrominance* while RGB color components are highly correlated.

The main difficulty in applying the denoising algorithm for monochrome images to color images is the fact that a color image is multi-channel and the channels are correlated to each other. The simplest and straightforward way to apply the denoising algorithms dicussed in Sect. 2 to color images is to consider each color channel as a single monochrome image and denoise each channel separately. As can be expected, this approach does not take advantage of any color information.

Our proposed color denoising scheme is depicted in Fig. 3. The framework estimates clean wavelet coefficients by different analyses which can make the color channels decorrelated and consider HVS. The analyses generate multiple estimates and construct probable ranges of clean coefficients in the color vector space. In the limited ranges of the vector space, we can decide an estimated coefficient vector which has the highest *a posteriori*.

3.1. Color Space Conversion for Human Visual System

Red, green, and blue are primary colors to represent any physical wavelength of visible light. Physically any optical systems including optic nerves of man and charge-coupled devices acquire these primary colors from the reflected light of an object. However, our color *translation* system is different in that luminance information is considered to be more important than hue and saturation. Among many color spaces that consider HVS, we use YUV color conversion from RGB space defined in the following:

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}, \tag{11}$$

where $(R, G, B)^t$ and $(Y, U, V)^t$ denote color vectors for each space. YUV color space can be converted linearly from RGB space. Due to the linear properties of both the wavelet transform and the conversion to YUV color space, the conversion to YUV can therefore be performed in the wavelet domain. We prefer the linear conversion in order to preserve the noise properties. In Eq. (1), we can notice that the noise is still zero-mean Gaussian after a linear transformation. However, the variance for each color channel may vary. In our case, the noise variances for Y, U, V components become

$$\sigma_Y^2 = (0.299\sigma_R)^2 + (0.587\sigma_G)^2 + (0.114\sigma_B)^2$$
$$\sigma_U^2 = (0.596\sigma_R)^2 + (0.275\sigma_G)^2 + (0.321\sigma_B)^2$$
$$\sigma_V^2 = (0.212\sigma_R)^2 + (0.523\sigma_G)^2 + (0.311\sigma_B)^2$$

respectively, where σ_R, σ_G, and σ_B are given noise standard deviations for each RGB channel.

*Luminance part substitutes intensity, lightness, or brightness depending on the color system. Chrominance information includes hue and saturation.

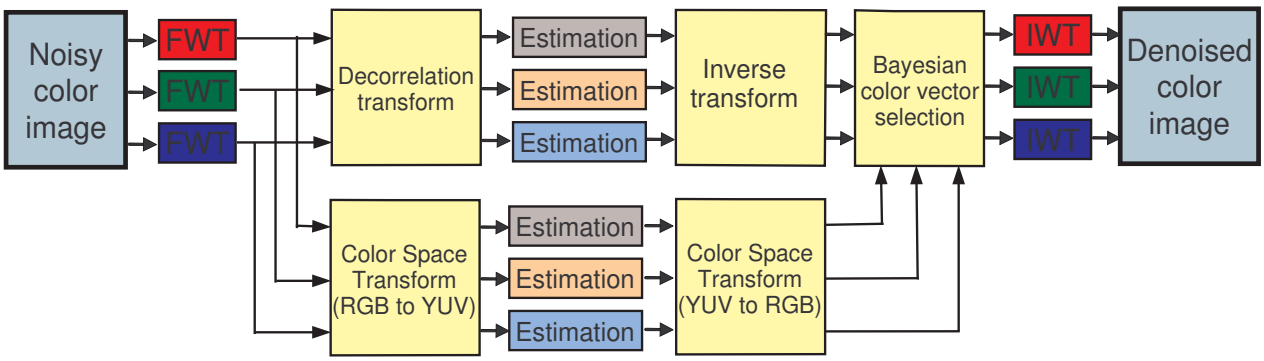


Figure 3. Proposed color image denoising framework using decorrelation-based selective estimation

3.2. Decorrelation Approach in the Wavelet Domain

Since RGB color space is strongly correlated for natural images, one can assume that uncorrelated space can be helpful for image denoising. In fact, one of the important inspirations in denoising methods using wavelet transform is drawn from weak correlation property between wavelet coefficients. Rao and Jones developed denoising method for a multisensor.[20] They show the similarity between Karhunen-Loeve (KL) and wavelet transforms and utilize the uncorrelation properties in spatio-temporal manner. KL transform is able to decorrelate strongly correlated multichannel data perfectly. It has been used for image analysis and coding widely. From Eq. (1), we define one color pixel as a vector, $\mathbf{v} = (v_r, v_g, v_b)^t$. Hence, we can denote a wavelet vector composed of wavelet coefficients from each color component as follows:

$$\mathbf{y} = \mathbf{x} + \Sigma_z^{1/2}\mathbf{z}, \tag{12}$$

where $\Sigma_z^{1/2} = \mathtt{diag}(\sigma_R, \sigma_G, \sigma_B)$. In order to decorrelate the given wavelet vectors, we obtain the KL transform matrix Φ, which satisfies $\Phi^t \Sigma_{\mathbf{y}} \Phi = \Lambda$, where $\Sigma_{\mathbf{y}}$ is a covariance matrix of the given wavelet vectors $\mathbf{y}$, Λ is a diagonal matrix containing the eigenvalues of $\Sigma_{\mathbf{y}}$, and Φ is an eigenmatrix of $\Sigma_{\mathbf{y}}$ expressed as $\Phi = [a_{ij}]$, where a_{ij} denotes the element of i-th row and j-th column in a 3×3 matrix Φ. They are all 3×3 matrices. Then the transformed vectors

$$\mathbf{y}_{KL} = \Phi^t \mathbf{y} \tag{13}$$

become ideally uncorrelated. In this case, our denoising problem is to estimate $\mathbf{x}_{KL} = \Phi^t \mathbf{x}$. From Eqs. (12) and (13), $\mathbf{y}_{KL} = \mathbf{x}_{KL} + \Phi^t \Sigma_z^{1/2} \mathbf{z}$. Therefore,

$$\sigma_j^2 = (a_{1j}\sigma_R)^2 + (a_{2j}\sigma_G)^2 + (a_{3j}\sigma_B)^2,$$

where σ_j denotes noise standard deviation for a transformed channel j and $j = 1, 2, 3$. Since KL transform is linear like YUV color conversion, the noise model in the transformed domain is also Gaussian. Then we can apply a monochrome denoising algorithm without modification for each transformed channel. Since Φ is orthogonal, inverse KL transform can be obtained from Eq. (13) as follows:

$$\mathbf{y} = \Phi \mathbf{y}_{KL} \tag{14}$$

The main advantage of KL transform is the ideal decorrelation depending on the given data. In our framework, KL transform is performed for each subband in the wavelet domain in order to take maximum advantages of multiresolution analysis property of wavelets. We note that any denoising algorithm can be applied to each of the estimation procedures for the denoised wavelet coefficients.

3.3. Bayesian Selection of Candidate Vectors

Previous sections imply that there might be many ways to estimate the clean wavelet coefficients. In this paper we will follow the approach in Sect. 2.2 where we assume that the clean wavelet coefficient is an estimate that maximizes the *a posteriori* $p(\mathbf{x}|\mathbf{y})$. If $p(\mathbf{x}|\mathbf{y})$ can be modeled accurately, we can estimate the clean coefficient in a locally confined vector space with less computational cost. Suppose that the power set $2^D = \{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \cdots, \hat{\mathbf{x}}_m\}$ is constructed from the estimated vector set D obtained by the two different ways at the end of step 4 of the framework in Fig. 3. Then in order to estimate the most probable we evaluate the following equation for each element vector of 2^D:

$$\begin{aligned}\hat{\mathbf{x}} &= \arg\max_{\mathbf{x}\in 2^D} \ln p(\mathbf{y}|\mathbf{x}) + \ln p(\mathbf{x}) \\ &= \arg\max_{\mathbf{x}\in 2^D} \left[-\frac{1}{2}\ln(2\pi)^3|\Sigma_z| + \gamma - \frac{(\mathbf{y}-\mathbf{x})^t\Sigma_z^{-1}(\mathbf{y}-\mathbf{x})}{2\sigma^2} - \left(\frac{(\mathbf{x}-\mu)^t\Sigma_{\mathbf{x}}^{-1}(\mathbf{x}-\mu)}{\alpha}\right)^{\beta}\right]. \end{aligned} \quad (15)$$

4. EXPERIMENTAL RESULTS

In our experiments, we have used 8-bit gray-level images and RGB color images with 512×512 sizes from USC-SIPI image database and other public sources. The noise model we assume is zero-mean additive white Gaussian noise (AWGN). In addition, we also used color images taken by consumer digital cameras to see how the denoising algorithms can enhance the practically used images which include naturally generated noise.

The wavelet filter we have used is dual-tree complex wavelet transform (DT CWT) suggested in Ref. 8. DT CWT has helpful properties for the denoising such as redundancy and directionality. We also have chosen Daubechies' length 8 wavelet filter (DAUB. 8) which is one of the most common mother wavelets for denoising for comparison.

4.1. Evalutions for Monochrome Images

For *NeighLevel*, determining the parameter θ in Eq. (2) is required. We found that the value is located in a particularly narrow range even for diverse types of images with different size and noise level. We set θ to $0.16 \sim 0.19$ depending on the sizes of the neighboring window and parent when we use DT CWT. From our experiments, we found that θ is slightly bigger if the neighboring window becomes smaller.

For the other approach using MGGD model, we need to choose the elements of vector $\mathbf{x}$ which includes the estimated coefficient itself and other closely related coefficients such as neighbors, parent, and others. In addition, the proper parameters for the statistical model are necessary for better estimation. This could be difficult since they are decided case by case empirically depending the type of images, the subband in the wavelet domain and the chosen elements of vector $\mathbf{x}$. For this combination of elements, we select the parameters of MGGD model as $\alpha = 1/6$ and $\beta = 1/2$ for simplicity in our experiments. In our case, we use the following estimation for $\hat{\Sigma}_{\mathbf{x}}$: $\widehat{\Sigma}_{\mathbf{x}} = \Sigma_{\mathbf{y}} - \sigma^2 I$ since the noise is independently distributed. In this case, an $M \times M$ *local window* that surrounds each element of $\mathbf{y}$ is applied. This is on the same basis as the local variance which is empirically used in recent works.[3,7] In our experiments, the local covariance produces higher quality of image when a 7×7 window is applied for each element.

To evaluate and analyze our algorithms, we compared them with the existing effective approaches.[3,7,9,21–26] Denoised images can be compared both visually and numerically. For different noise variances, the measured PSNR values are listed in Tab. 1 for our proposed and other methods. The results are categorized in terms of the type of the wavelet used since denoising results are strongly dependent on the wavelet transforms. We use the experimental results from the original papers and the PSNR table in Ref. 7. GSM results for DAUB. 8 filters are obtained from the software offered by Portilla.[9] A comparison of selected methods is given in Figs. 4 for a 512×512 size *Boat* images. *Wiener* filter[†] is also included since it considers the neighboring dependency and achieves efficacious performance with simple linear filter even if wavelet transforms are not applied. Our

[†]We used `wiener2` function in MATLAB image processing toolbox with a 5×5 neighboring window and unknown noise level.

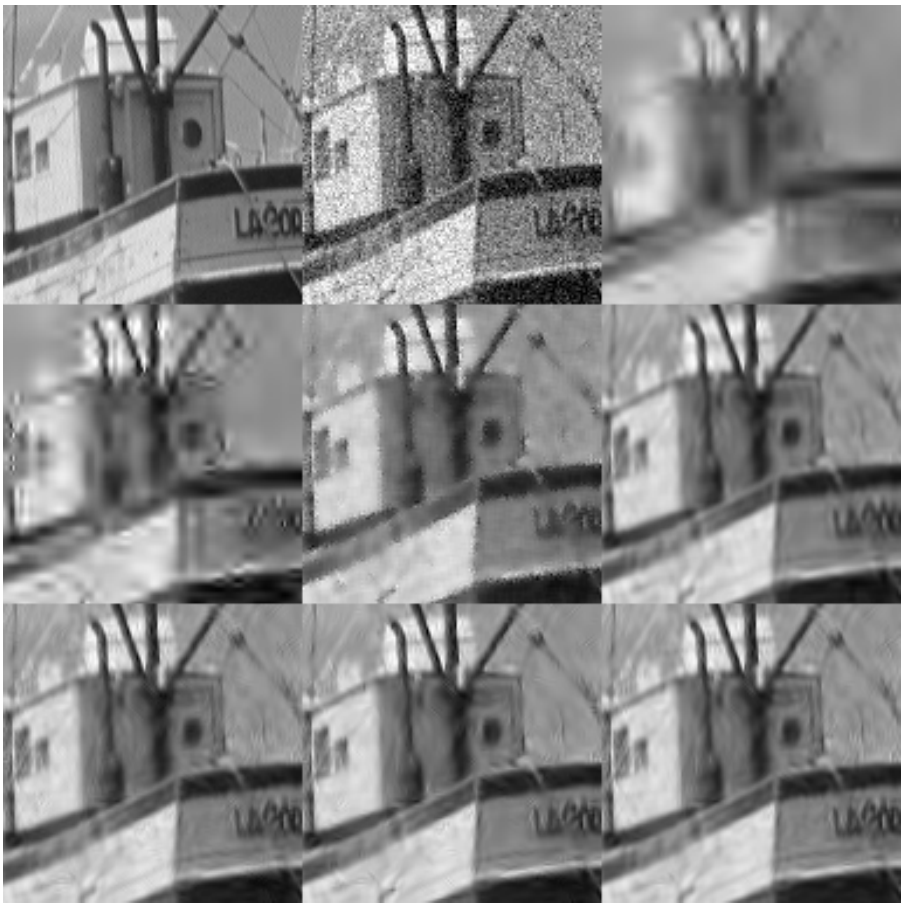


Figure 4. Cropped images (128×128) using proposed algorithms for 512×512 *Boat* image with σ=25 : Original (top-left), Noisy (top-center; 18.60dB), $VisuShrink$ soft (top-right; 24.06dB), $VisuShrink$ hard (middle-left; 25.03dB), $Wiener$ filter (middle-center; 27.22dB), $NeighSure$ (middle-right; 28.48dB), $NeighShrink$ (bottom-left; 28.90dB), $NeighLevel$ (bottom-center; 29.11dB), Multivariate (bottom-right; 29.12dB).

Table 1. Comparison table for proposed and existing methods with different Gaussian noise (*Lena* 512 × 512).

Wavelet	Approach	PSNR(dB) by noise level(σ)				
		10	15	20	25	30
Noisy image		28.12	24.62	22.14	20.16	18.60
–	`Wiener2`	32.67	31.28	30.03	28.85	27.83
DAUB. 8	MGGD	34.55	32.71	31.44	30.46	29.64
	NeighLevel	34.51	32.60	31.30	30.27	29.47
	SureShrink[21]	33.42	31.50	30.17	29.18	28.47
	BiShrink[7]	34.33	32.48	31.16	30.12	29.38
	HMT[22]	33.81	31.73	30.36	29.21	28.32
	LAWMAP[3]	34.24	32.27	30.92	29.90	-
	GSM[9]	34.23	32.35	31.03	30.23	29.21
	BayesShrink[25]	33.29	31.38	30.14	29.19	28.45
DT CWT	MGGD	35.35	33.70	32.46	31.48	30.68
	NeighLevel	35.41	33.72	32.50	31.48	30.70
	BiShrink[7]	35.31	33.64	32.37	31.37	30.51
	CHMT[23]	34.90	-	-	29.90	-
Bi 10/18	MMSE[26]	34.93	33.01	31.69	30.60	-
SI Symm. 8	SI-AdaptShr[24]	-	33.37	32.09	31.07	-
Steer. pyramid	GSM[9]	35.61	33.90	32.66	31.69	-

Matlab implementation takes less than 3 seconds for a 512×512 image with DT CWT on 2.4GHz Pentium IV for *NeighLevel*. Since there is no general solution for Eq. (10), iterative numerical solution is applied for the multivariate approach. It takes about 30 seconds for a 512×512 image with DAUB. 8 on the same system when 10 elements are used. However, when $\hat{\mathbf{x}}$ can be calculated explicitly in Eq. (9) using specific constraints, it only takes less than 3 seconds under the same condition.

4.2. Evalutions for Color Images

In order to evaluate our method, we show the results using *Wiener*, a straightforward extension in RGB space using *NeighLevel*, and our proposed framework described in Sect. 3. Wiener filter is the optimal minimum mean squared error estimator which considers neighboring information. The preliminary estimator for clean wavelet coefficients is *NeighLevel* presented in Sect. 2, in order to investigate how the correlation between color components affects the denoising results. We also tried to implement a method from Ref. 27, which takes advantage of chromatic filters with anisotropic diffusion for luminance channel. Table 2 shows the denoised results with additive white noise $N(0, 30^2)$ in terms of peak signal to noise ratio (PSNR). In Figs. 5 and 6, some of the results are displayed. The compared denoising results in Fig. 5 are interesting since Fig. 5(c) has rainbow-colored artifact while Fig. 5(d) looks more consistent with the white mountains. We also tried to enhance an image without *artificial* Gaussian noise in Fig. 6. As can be seen, our proposed wavelet approach can remove the noise in the image without too much blurring. These results show that we can achieve high-quality color image denoising by decorrelation of color components for most of natural images. However, for images that mainly use primary colors hence are comparatively less correlated such as *Pepper* image, they do not show much difference as can be seen in Table 2.

5. CONCLUSION

In this paper, we have presented gray-level image denoising methods and their application to color image denoising. These approaches show that the dependency between neighboring wavelet coefficients is critical information for image denoising. Also our experiments on color image denoising reflect that the correlation between color components must be maximally utilized for an efficient denoising algorithm.

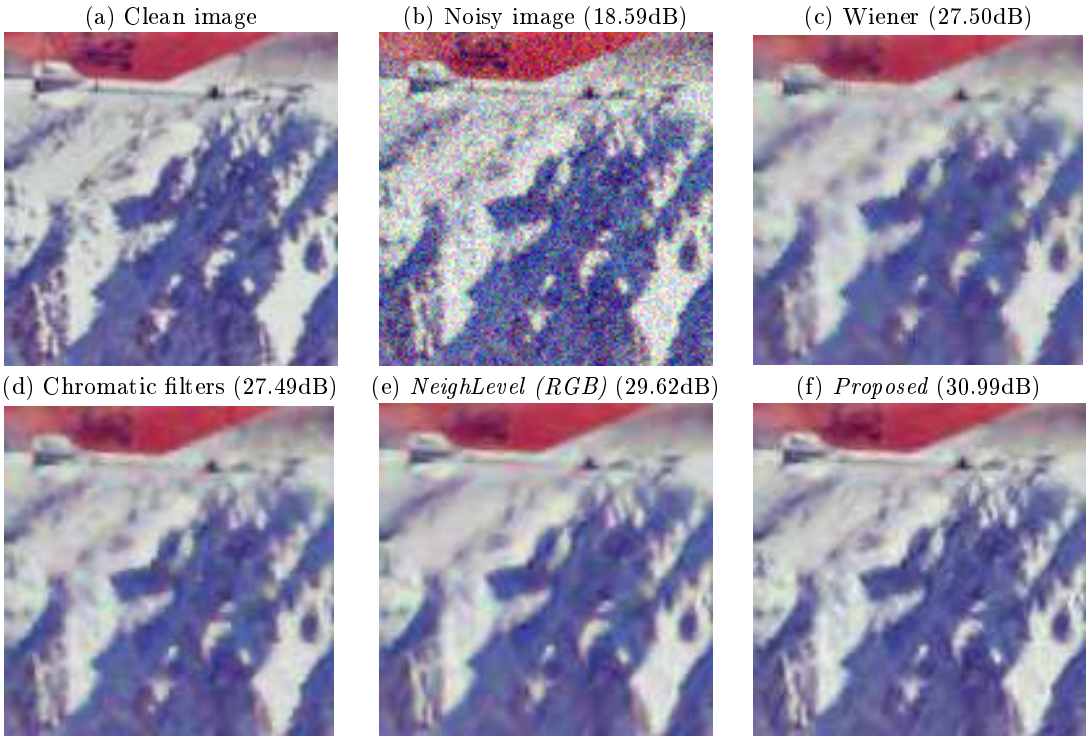
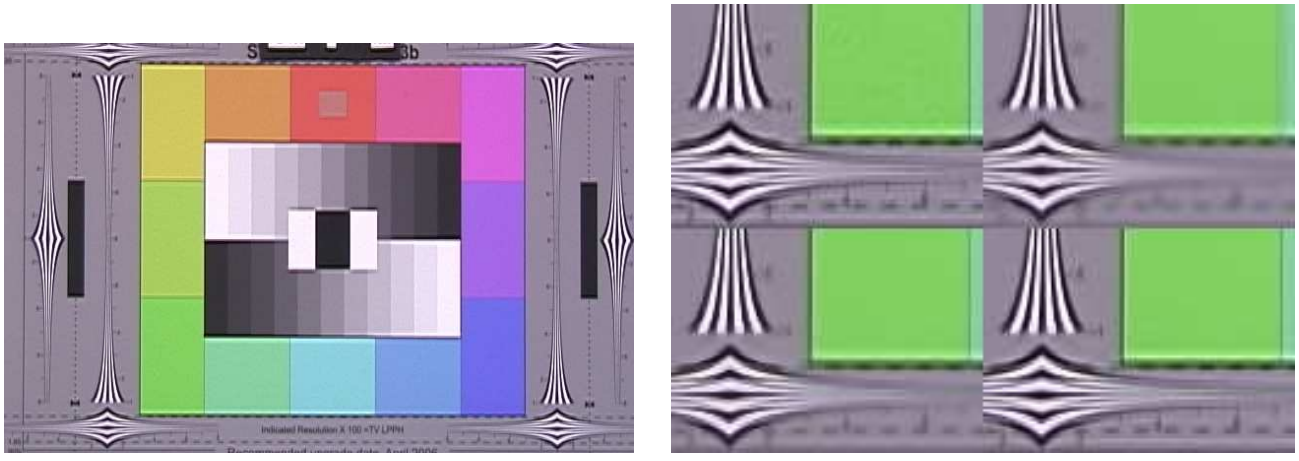


Figure 5. Denoising results for cropped *F-16* image with Gaussian noise $N(0, 30^2)$.



(a) A captured frame using Sony PD-100

(b) original (top-left), Wiener (top-right), chromatic filter[27] (bottom-left), and proposed (bottom-right)

Figure 6. Denoising results of a natually corrupted image frame captured by consumer digital camcoder

Table 2. PSNR values of denoised 512×512 color images with AWGN $N(0, 30^2)$ from USC SIPI database

Image	Noisy image	Wiener filter	Chromatic Filters[27]	*NeighLevel* (RGB)	Proposed
Lena	18.59 dB	27.63 dB	27.84 dB	30.01 dB	31.09 dB
F-16	18.59 dB	27.50 dB	27.49 dB	29.62 dB	30.99 dB
Lake	18.59 dB	25.79 dB	25.94 dB	26.79 dB	27.56 dB
Baboon	18.59 dB	22.74 dB	22.90 dB	23.81 dB	24.98 dB
Pepper	18.59 dB	27.67 dB	27.15 dB	29.48 dB	29.72 dB

ACKNOWLEDGMENTS

This work was supported by research grants from the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

1. T. T. Cai and B. W. Silverman, "Incorporating information on neighbouring coefficients into wavelet estimation," *Sankhya: The Indian Journal of Statistics, Serie A* **63**, pp. 127–148, 2001.
2. G. Y. Chen and T. D. Bui, "Multiwavelet denoising using neighbouring coefficients," *IEEE Signal Processing Letters* **10**(7), pp. 211–214, 2003.
3. M. K. Mihcak, K. R. I. Kozintsev, and P. Moulin, "Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Processing Letters* **6**(12), pp. 300–303, 1999.
4. M. Malfait and D. Roose, "Wavelet-based image denoising using a markov random field a priori model," *IEEE Transactions on Image Processing* **6**, pp. 549–565, Apr. 1997.
5. A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy, "A joint inter- and intrascale statistical model for bayesian wavelet based image denoising," *IEEE Transactions on Image Processing* **11**, pp. 545–557, May 2002.
6. L. Sendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE Transactions on Signal Processing* **50**(11), pp. 2744–2756, 2002.
7. L. Sendur and I. W. Selesnick, "Bivariate shrinkage with local variance estimation," *IEEE Signal Processing Letters* **9**(12), pp. 438–441, 2002.
8. N. G. Kingsbury, "Image processing with complex wavelets," *Phil. Trans. Royal Society London A* **357**, pp. 2543–2560, Sept. 1999.
9. J. Portilla, M. W. V. Strela, and E. P. Simoncelli, "Image denoising using scale mixtures of gaussians in the wavelet domain," *IEEE Transactions on Image Processing* **12**, pp. 1338–1351, Nov. 2003.
10. K. Q. Huang, Z. Y. Wu, G. Fung, and F. H. Y. Chan, "Color image denoising with wavelet thresholding based on human visual system model," *Signal Processing: Image Communications* **20**(2), pp. 115 – 127, 2005.
11. B. A. Thomas and J. J. Rodriguez, "Wavelet-based color image denoising," in *IEEE International Conference on Image Processing (ICIP)*, **2**, pp. 804–807, Sept. 2000.
12. M. J. Cree, "Observations on adaptive vector filters for noise reduction in color images," *IEEE Signal Processing Letters* **11**(2), pp. 140 – 143, 2004.
13. K. Plataniotis and A. N. Venetsanopoulos, *Color Image Processing and Applications*, Springer-Verlag, 2000.
14. D. L. Donoho, "Denoising by soft-thresholding," *IEEE Transactions on Information Theory* **41**, pp. 613–627, 1995.
15. J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing* **41**, pp. 3445–3462, Dec. 1993.
16. D. Cho, T. D. Bui, and G. Y. Chen, "Image denoising based on wavelet shrinkage using neighbour and level dependency." Submitted to *IEE Proceedings on Vision, Image & Signal Processing*, 2004.

17. D. Cho and T. D. Bui, "Multivariate statistical modeling for image denoising using wavelet transforms," *Signal Processing: Image Communications* **20**(1), pp. 77 – 89, 2005.
18. S. G. Mallat, "A theory for multiresolution signal decomposition:the wavelet representation," **11**, pp. 674–693, July 1989.
19. C. Fernández, J.Osiewalski, and M. Steel, "Modelling and inference with v-spherical distributions," *Journal of the American Statistical Association* **90**(432), pp. 1331–1340, 1995.
20. A. M. Rao and D. L. Jones, "A denoising approach to multisensor signal estimation," *IEEE Transactions on Signal Processing* **48**(5), pp. 1225 – 1234, 2000.
21. D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *Journal of the American Statistical Association* **90**(432), pp. 1200–1224, 1995.
22. M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based signal processing using hidden markov models," *IEEE Transactions on Signal Processing* **46**(4), pp. 886–902, 1998.
23. H. Choi, J. K. Romberg, R. G. Baraniuk, and N. G. Kingsbury, "Hidden markov tree modeling of complex wavelet transforms," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, **1**, pp. 133–136, (Istanbul, Turkey), June 2000.
24. S. G. Chang, B. Yu, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Transactions on Image Processing* **9**, pp. 1522–1531, Sept. 2000.
25. S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Transactions on Image Processing* **9**, pp. 1532–1546, Sept. 2000.
26. X. Li and M. T. Orchard, "Spatially adaptive image denosing under overcomplete expansion," in *IEEE International Conference on Image Processing (ICIP)*, **3**, pp. 300–303, Sept. 2000.
27. L. Lucchese and S. K. Mitra, "A new class of chromatic filters for color image processing. theory and applications," *IEEE Transactions on Image Processing* **13**(4), pp. 534–548, 2004.