Wavelet Transform Domain Filters: A Spatially Selective Noise Filtration Technique

Yansun Xu, John B. Weaver, Dennis M. Healy, Jr., and Jian Lu

Abstract— **Wavelet transforms are multiresolution decompositions that can be used to analyze signals and images. They describe a signal by the power at each scale and position. Edges can be located very effectively in the wavelet transform domain. A spatially selective noise filtration technique based on the direct spatial correlation of the wavelet transform at several adjacent scales is introduced. A high correlation is used to infer that there is a significant feature at the position that should be passed through the filter. We have tested the technique on simulated signals, phantom images, and real MR images. It is found that the technique can reduce noise contents in signals and images by more than 80% while maintaining at least 80% of the value of the gradient at most edges. We did not observe any Gibbs' ringing or significant resolution loss on the filtered images. Artifacts that arose from the filtration are very small and local. The noise filtration technique is quite robust. There are many possible extensions of the technique. We see its applications in spatially dependent noise filtration, edge detection and enhancement, image restoration, and motion artifact removal. We have compared the performance of the technique to that of the Weiner filter and found it to be superior.**

I. Introduction

FOURIER transform domain filters used in signal and image processing involve a tradeoff between the signal-to-noise ratio (SNR) and the spatial resolution of the signal/image processed. Low-pass filters will not only smooth away noise but also blur edges in signals and images; high-pass filters can make edges even sharper and improve the spatial resolution but will also amplify the noisy background [1]. In many medical imaging applications, noise is often removed by averaging over many identical image acquisitions. This leads to a very long imaging time and turns an SNR and resolution tradeoff into an imaging time and image resolution tradeoff. Filtering images with advanced image processing techniques is still often needed after acquisition to mitigate the lack of SNR. Image processing techniques based on statistical models (Bayesian approaches) have shown certain strength at reducing noise, enhancing edges, and removing ringing artifacts in clinical images [2], [3]. We are studying spatial filters in the wavelet transform domain as an alternative to Fourier transform domain filters for medical imaging applications. The preliminary results from these studies show that noise is reduced very effectively by wavelet filters with very little resolution loss; most sharp edges are preserved, and some of them are being enhanced [4]–[6].

Wavelet transforms are multiresolution representations of signals and images. They decompose signals and images into multiscale details [7]–[9]. The basis functions used in wavelet transforms are locally supported; they are nonzero only over part of the domain represented. Sharp transitions in images are preserved and depicted extremely well in wavelet expansions [9]. This special treatment of edges by wavelet transforms is very attractive in image filtering.

The filters we are exploring pass essentially all the signal at large scales. The signal at small scales is passed if it is around an identified edge; it is eliminated as noise if it is not around an identified edge. Because most noise power is confined to small scales, the reduction of signal at small scales reduces noise preferentially. However, to keep edges sharp, small-scale information is required. By passing small-scale data around identified edges, noise is reduced, and the identified edges stay sharp. The key to this technique is to identify edges. Edges are identified as features that have signal peaks across many scales. An edge occurs at a position where there are maxima in the nonorthogonal wavelet transform at several adjacent scales. Direct spatial correlations of the wavelet transform at different scales are used to identify the edges; the small scale data is passed at positions where the correlation is large and suppressed if the correlation is small. The orthogonal wavelet transform used in our previous work [4] is not appropriate because the transformed signal is uncorrelated across scale. The nonorthogonal wavelet introduced by Mallat and Zhong [9] offers much better edge detection because the signal is correlated across scale and the wavelet is an "edge detector."

The signal power at large scales corresponds to that at low frequencies in the Fourier transform; the power at small scales corresponds to that at high frequencies in the Fourier transform [7], [8]. This filter can be seen as a low-pass filter that passes selected high-frequency data. The high-frequency data passed is that which occurs at positions where edges are identified. Compared with the optimum linear filter, the Wiener filter, which is solely adapted to SNR at a single scale, the wavelet filter performs much better in preserving high-frequency data around edges.

Manuscript received March 5, 1992; revised August 2, 1993. This work was supported by a grant from the Whitaker Foundation and by a DARPA contract administered by the AFOSR as AFOSR-90-0292. The associate editor coordinating the review of this paper and approving it for publication was Dr. Michael Unser.

Y. Xu and J. B. Weaver are with the Department of Diagnostic Radiology, Dartmouth-Hitchcock Medical Center, Lebanon, NH 03756 USA.

D. M. Healy, Jr. is with the Department of Mathematics and Computer Science, Dartmouth College, Hanover, NH 03755 USA.

J. Lu is with the Thayer School of Engineering, Dartmouth College, Hanover, NH 03755 USA.

IEEE Log Number 9404455.

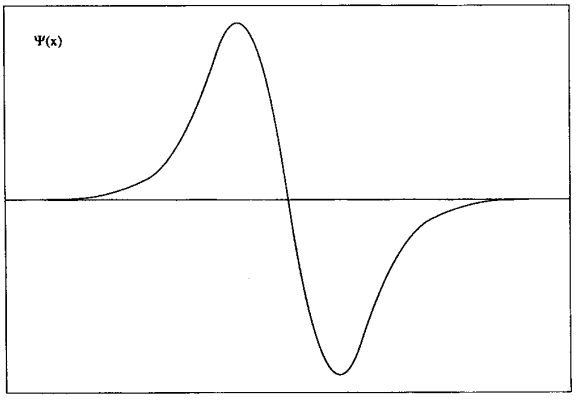


Fig. 1. Plot of a typical quadratic-spline wavelet used here for noise filtration.

II. Wavelet Transforms

Wavelets are families of functions $\Psi_{s,t}(x)$ generated from a single base wavelet $\Psi(x)$ by dilations and translations

$$\Psi_{s,t}(x) = \frac{1}{\sqrt{|s|}} \Psi\left(\frac{x-t}{s}\right) \qquad s \neq 0 \tag{1}$$

where s is the dilation (scale) parameter, and t is the translation parameter. Wavelets must have mean zero, and the useful ones have localized support in both spatial and Fourier domains. There are orthogonal and nonorthogonal wavelet sets that span $L^2(\Re)$. In this paper, we use nonorthogonal wavelets first introduced by Mallat *et al.* [9]; the base wavelet is shown in Fig. 1. It is very close to the derivative of a Gaussian function. The set of $\Psi_{m,n}(x)$ spans $L^2(\Re)$ when $s = 2^m$, $t = n$

$$\Psi_{m,n}(x) = 2^{-m/2}\Psi(2^{-m}(x-n)) \tag{2}$$

where m is the scale index ($m = 0, 1, 2, \ldots$), and n is the translation (spatial) index ($n = \ldots, -2, -1, 0, 1, 2, \ldots$). The discrete wavelet transform $W(m,n)$ of a 1-D function $f(x)$ is defined as the projection of the function onto the wavelet set $\Psi_{m,n}(x)$.

$$W(m,n) = \int_{-\infty}^{+\infty} dx \quad \overline{\Psi_{m,n}(x)} \quad f(x). \tag{3}$$

Since the set of $\Psi_{m,n}(x)$ spans the space containing $f(x)$, the reconstruction of function $f(x)$ from its wavelet transform $W(m,n)$ is possible

$$f(x) = \sum_m \sum_n \Psi'_{m,n}(x) W(m,n) \tag{4}$$

where $\Psi'_{m,n}(x)$ is the normalized dual basis of $\Psi_{m,n}(x)$. For the wavelet expansion we use here, $\Psi' \approx \Psi$.

The wavelet transform $W(m,n)$ gives a scale-space decomposition of signals and, with simple modifications, images. It decomposes the signal into different resolution scales, with m indexing the scale and n indexing position in the original signal space.

In practice, we are concerned with a finite length, discrete (sampled), 1-D data set $\{f(k), k = 1, 2, \ldots, N\}$, and we need appropriate discrete and finite versions of the calculations involved in the wavelet decomposition [9]. In particular, there is a fixed limit to the resolution and, therefore, a lower bound on the scale index m, which we may take as $m = 1$ without loss of generality. It is useful to model this resolution limit by representing the data $f(k)$ as samples of a smoothed, or low-passed, version of a continuous signal:

$$f(k) = \int_{-\infty}^{+\infty} dx \quad \overline{\Phi(x-k)} \quad f(x) \tag{5}$$

with respect to a smoothing or scaling function Φ. Based on this representation of the data, one may compute the wavelet coefficients in (3) by means of a purely discrete algorithm, as detailed in [9]. Beyond these considerations, there is also an effective upper limit on the scale m imposed by the finite length of the signal.

Consequently, the nonorthogonal, discrete, dyadic wavelet coefficients $W(m,n)$ are computed on a 2-D space of $m = 1, 2, \ldots, M-1$ and $n = 1, 2, \ldots, N$, where $M = \log_2 N$, with the remaining information contained in the coarse scale averages $S(M,n) = 2^{-M} \int_{-\infty}^{+\infty} dx\, \overline{\Phi(2^{-M}(x-n))}\ f(x)$. This information determines the signal $f(x)$:

$$f(x) = \sum_{m=1}^{M-1} \sum_{n=1}^{N} \Psi'_{m,n}(x) W(m,n) + \sum_{n=1}^{N} \Phi(2^{-M}(x-n)) S(M,n). \tag{6}$$

The $M \cdot N$ coefficients obviously form an overcomplete representation of the signal. For a data set of $N = 256$ points, M is equal to 8, i.e., there are eight wavelet scales. At each scale, there are 256 data points corresponding to the signal detail projected at that scale. Orthogonal wavelet transforms have fewer coefficients at course scales, which makes correlations across scale difficult. Fig. 2 shows a simulated 1-D data set of 256 points and its discrete, dyadic wavelet transform at all eight scales. In the simulated data, we placed two small "bumps" on top of a large boxcar and added Gaussian distributed white noise. The SNR of the data is about 18 dB. The edges of the large boxcar and two small bumps are well localized in the wavelet transform domain and show up at most wavelet scales (from small to large). Noise power, however, is concentrated only at a few small scales.

III. Filtering Algorithm

Several edge detection and noise reduction techniques based on the approaches of wavelet and subband decompositions have been proposed in recent years [10], [11]. Witkin first introduced the idea of using the scale space correlation of the subband decompositions of a signal to filter noise from the signal [10]. He developed an algorithm to track major edges in a signal from coarse scales to fine scales in the subband decompositions and was able to distinguish major edges from noise background at finer scales. Recently Mallat

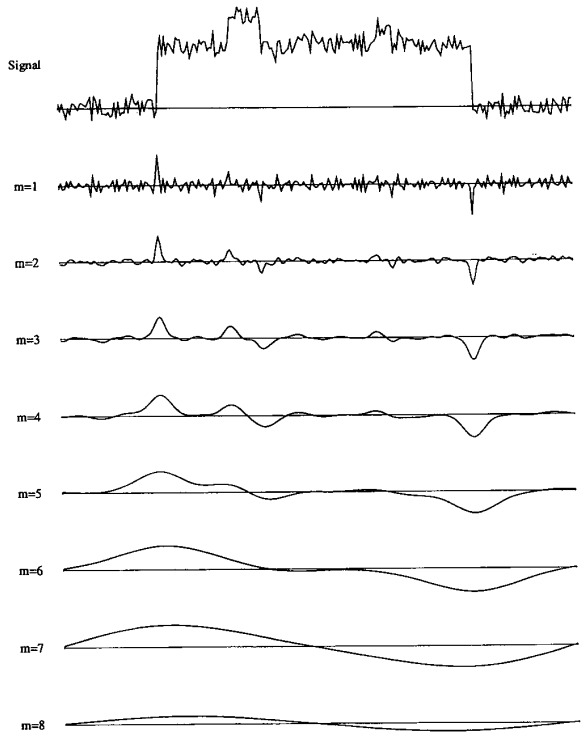


Fig. 2. Simulated 1-D data of 256 points and its discrete dyadic wavelet transform at all eight scales.

et al. introduced the complete signal representation by wavelet transform domain maxima [9] (wavelet transforms are special types of subband decompositions). They were able to distinguish edge maxima from noise maxima by analyzing the singularity properties of wavelet transform domain maxima of a signal across the various scales [11]. Our approach to filtering noise from a signal also relies on the variations in scale of the wavelet transform data of the signal, but rather than detecting edges directly on the wavelet transform data with a complicated algorithm, such as those introduced in [10] and [11], we use the direct multiplication of wavelet transform data at adjacent scales to distinguish important edges from noise and accomplish the task of removing noise from signals. Rosenfeld *et al.* demonstrated that it is very efficient and quite accurate to simply use the direct multiplication of the subband decompositions of an image to locate important edges in digital images [12], [13]. Although it may be slightly less accurate than the techniques in [10] and [11], we believe that our approach is more straightforward, easier to implement, and significantly more robust.

The wavelet transform domain noise filtration technique we are developing is based on the fact that sharp edges have large signal over many wavelet scales, and noise dies out swiftly with increasing scale. We are using the direct spatial correlation $\mathrm{Corr}_l(m,n)$ of wavelet transform contents at several adjacent scales to accurately detect the locations of edges or other significant features.

$$\mathrm{Corr}_l(m,n) = \prod_{i=0}^{l-1} W(m+i,n), \qquad n = 1,2,\ldots,N \quad (7)$$

where l is the number of scales involved in the direct multiplication, $m < M - l + 1$, and M is the total number of scales. The absence of edges or other significant features in a localized region of the signal allows the noisy background to be removed there. The direct spatial correlation of edge-detection data over several scales sharpens and enhances major edges while suppressing noise and small sharp features. The correlation, $\mathrm{Corr}_l(m,n)$ will improve the accuracy of locating important edges in signals and images [12], [13].

Fig. 3 demonstrates the effect of this wavelet filter on the smallest (first) scale of the wavelet transform of the signal shown in Fig. 2. The wavelet transform of the signal at the first scale $W(1,n)$ is shown in Fig. 3(a). Fig. 3(b) gives the direct multiplication of the wavelet transform contents at the first two smallest scales $\mathrm{Corr}_2(1,n) = W(1,n)W(2,n)$. Note that the two edges of the large boxcar in the original data set show up much sharper and stronger in $\mathrm{Corr}_2(1,n)$ than they appear in $W(1,n)$. Furthermore, one may observe that they are much larger in $\mathrm{Corr}_2(1,n)$ than the edges of the two small bumps and the noisy background. Our algorithm uses the strength of these correlated features to discriminate significant signal features from noise, as follows.

First, the power of the $\{\mathrm{Corr}_2(1,n)\}$ data is rescaled to that of the $\{W(1,n)\}$ data. The most important edges (two major edges in Fig. 3) are identified in $W(1,n)$ and $\mathrm{Corr}_2(1,n)$ by comparing the absolute values of $\mathrm{Corr}_2(1,n)$ and $W(1,n)$. An edge is identified at any position n for which $|\mathrm{Corr}_2(1,n)| > |W(1,n)|$. This edge position and its corresponding value $W(1,n)$ are stored. Finally, all the edges identified in this way are extracted from $\mathrm{Corr}_2(1,n)$ and $W(1,n)$ by resetting the values of these signals to 0's at the positions identified. We refer to the remainder of the data points in $W(1,n)$ and $\mathrm{Corr}_2(1,n)$ after the first round of edge extraction as $W'(1,n)$ and $\mathrm{Corr}'_2(1,n)$. By rescaling the power of $\mathrm{Corr}'_2(1,n)$ to that of $W'(1,n)$ and comparing their absolute values, the next most significant edges (edges of the two small bumps in Fig. 3) are extracted from $W(1,n)$ and $\mathrm{Corr}_2(1,n)$. This procedure of power normalization, data value comparison, and edge information extraction can be iterated many times until the power of the unextracted data points in $W(1,n)$ is nearly equal to some reference noise power at the first wavelet scale. In digital image processing, one can often use the background noise at the "dark" (signal-free) regions near the boundaries of an image as the reference noise [4]. Through the wavelet transform of this "reference noise," the average noise power at each wavelet scale can be estimated for a particular image and used as a threshold to terminate the iteration procedure explained above.

All the edge information in the original data that is extracted from $W(1,n)$ during this iteration process is kept in a data vector. We call this data vector $W_{\mathrm{new}}(1,n)$. By replacing $W(1,n)$ with $W_{\mathrm{new}}(1,n)$, we can have a new and spatially filtered first scale wavelet transform data where most of

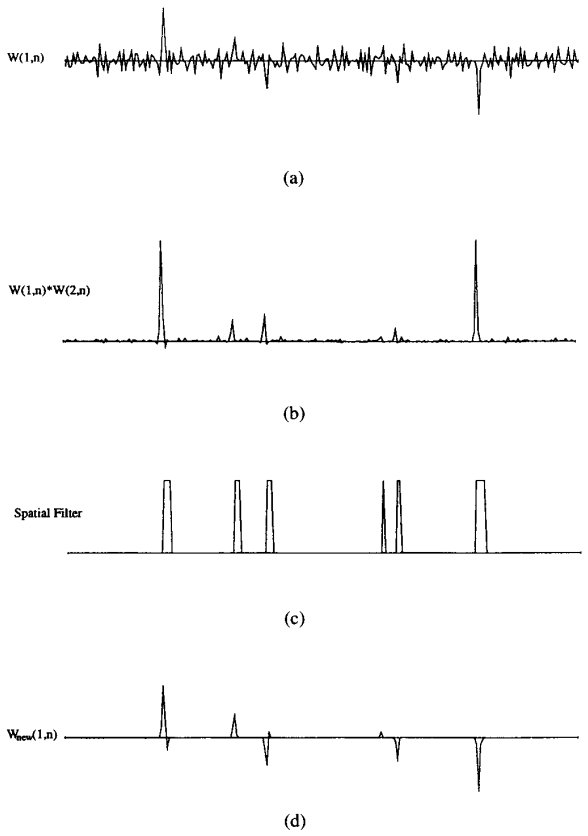


Fig. 3. Graphic illustration of the noise filtration technique based on the direct spatial correlation of wavelet transform data at several adjacent resolution scales: (a) First scale wavelet transform $W(1,n)$ before filtering; (b) direct multiplication of $W(1,n)$ and $W(2,n)$; (c) demonstration of the filtering technique in the form of a spatially selective filter for the first scale; (d) first scale wavelet transform $W_{\text{new}}(1,n)$ after filtering.

TABLE I
ENTIRE FILTERING PROCESS IN THE WAVELET TRANSFORM DOMAIN DESCRIBED BY A "CONTROL FLOW" TYPE OF SCHEME

```
First, Save a copy of W(m,n) to WW(m,n)
Initialize the "spatial filter mask": mask(m,n) to 0's

Loop for each wavelet scale m
{
  Loop for the iteration process
  {
    Compute the power of Corr2(m,n) and W(m,n):
     PCorr(m) = Σ_n Corr2(m,n)^2
     PW(m)    = Σ_n W(m,n)^2
    Re-scale the power of Corr2(m,n) to that of W(m,n):
    Loop for each pixel point n
    {
       new Corr2(m,n) = Corr2(m,n) * sqrt(PW(m)/PCorr(m))
    } end loop n
    Loop for each pixel point n
    {
      Compare pixel values in new Corr2(m,n) and W(m,n):
       if |Corr2(m,n)| > |W(m,n)|
       {
         Extract edge information from W(m,n) and Corr2(m,n),
         and save it in the "spatial filter mask":
          Corr2(m,n) = 0.0
          W(m,n) = 0.0
          mask(m,n) = 1
       } end if
    } end loop n
  } iterate until PW(m) ≤ the noise threshold at scale m
  Apply the "spatial filter mask" to the saved copy, WW(m,n),
  at scale m. Save the filtered data to Wnew(m,n):
    Loop for each pixel point n
    {
      Wnew(m,n) = mask(m,n) • WW(m,n)
    } end loop n
} end loop m
```

the noise is removed and most of the original edges are preserved. Fig. 3(d) shows the spatially filtered first scale wavelet transform data of the signal shown in Fig. 2. This noise filtration and edge extraction technique can be thought of as a spatially dependent filter. This spatial filter is demonstrated by the mask shown in Fig. 3(c); it spatially selects which part of the data to keep (the edges) and which part of the data to eliminate (the noise); the signal is passed where the wavelet transform is highly correlated across scales and suppressed elsewhere.

The procedure can be summarized as a comparison of the normalized $|\text{Corr}_2(m,n)|$ data with $|W(m,n)|$ to guide the extraction of the edge information from $W(m,n)$ iteratively at the mth wavelet scale. By repeating the procedure at every resolution scale, we can acquire all the spatially filtered wavelet transform data $W_{\text{new}}(m,n)$. The entire procedure is illustrated by the "control flow" pseudo-software code in Table I. The reconstruction from $W_{\text{new}}(m,n)$ through the inverse wavelet transform shown in (6) will yield the final filtered signal. The inverse wavelet transform that we implemented in our technique uses a kernel function that is the same as the basis function (the wavelet kernel shown in (2)) used to build the forward wavelet transform [9]. Fig. 4 shows the filtered 1-D simulation data and its spatially filtered wavelet transform data $W_{\text{new}}(m,n)$. Edges of the large boxcar and the higher bump remain as sharp after filtration as they were before filtration. Noise reduction is remarkable; the synthetic noise in the signal has been reduced by 80 to 90%. The SNR is raised from 18 to 26.3 dB. There is no Gibbs' ringing around major edges in the filtered signal; artifacts arising from the filtration are very small and local. We do notice, however, that there is slight degradation in both the edges and the contrast of small features (i.e., the smaller bump in the simulated data). It is difficult for the filter to discriminate between noise and the features that are the same size as the noise.

Fig. 5(a) is an overlay plot of the simulated signal before and after the noise filtration. The edge preserving property of this technique is clearly seen from a close-up view at the larger bump in the signal. Fig. 5(b) gives an overlay plot of a real image line before and after the noise filtration. This image line data is drawn from a fairly noisy magnetic resonance (MR) image of human neck. The intensity gradients at most edges in the filtered line remain at more than 80% of their original values. No large artifacts are observed. A close-up view of the central portion of the line shows that this wavelet domain filter technique protects edges very well.

We should mention that the test data we have shown so far contain mainly very sharp edges (even though some of them are small) and noise. It appears that the wavelet filter works well on those signals. Most clinical images (such as

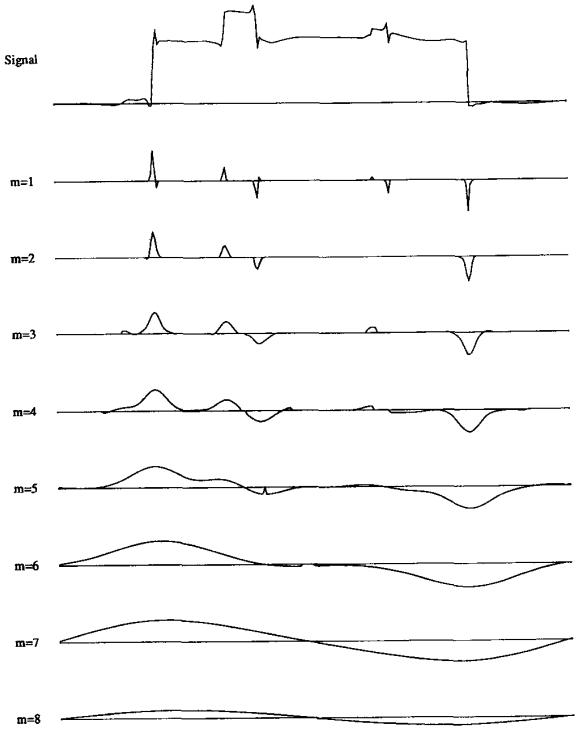
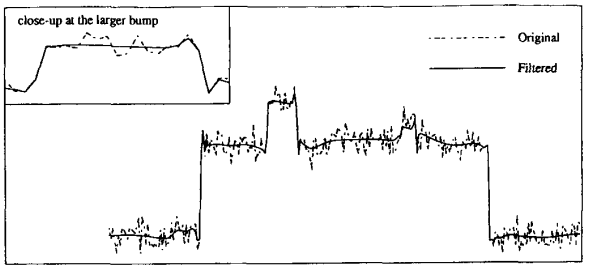
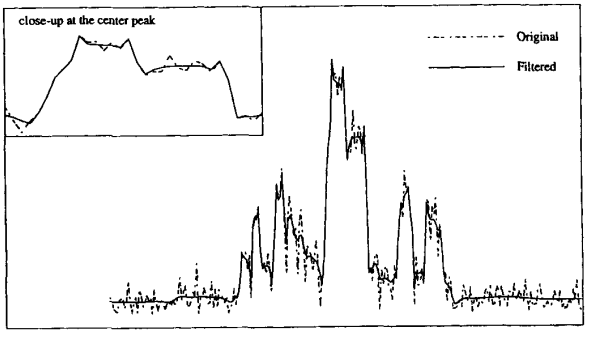


Fig. 4. Same 1-D data and its discrete dyadic wavelet transform shown in Fig. 2 after being processed with the wavelet domain filtering technique.



(a)



(b)

Fig. 5. Plots of the synthetic noisy signal shown in Fig. 2 and the image line #109 from an MR image against their corresponding filtered data: (a) Syntheticsignal (with a closeup at the larger bump). The standard deviation of the synthetic Gaussian white noise is about 10% of the amplitude of the large boxcar; (b) MR image line 109 (with a closeup at the center peak).

computed tomography and MR images) show up as regions of tissues with different intensities and fairly sharp boundaries. We have, however, also tested the filter on some noisy data with rather smooth features. In Fig. 6(a), the original data (dashed line) has a fairly large and smooth "bump" on top of a large boxcar embedded in the noisy background. The solid line is the filtered data. Fig. 6(b) gives the filtered data (solid line) against the original data without the presence of noise (dashed line). Smooth features have significant signal power in the middle size scales. The presence of more signal power in the middle size scales retains slightly more noise power in the same scales. When the spatial filter passes signals in the small neighborhood of a sharp edge, the noise power that is passed along with the signal is not very noticeable. When the filter passes signals over a wide area around a smooth feature, the noise power that is passed along is more noticeable as shown in Fig. 6(b).

IV. Results on 2-D Images

Two-dimensional wavelet transforms used in image analysis and processing are not very different from the 1-D transforms shown in (3) and (4), except that the wavelet kernel (basis function) becomes a 2-D one $\Psi_{m,n_x,n_y}(x,y)$, and the transformed image data will be a 3-D data set $W(m,n_x,n_y)$.

$$W(m,n_x,n_y) = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy \times \overline{\Psi_{m,n_x,n_y}(x,y)} \quad f(x,y) \tag{8}$$

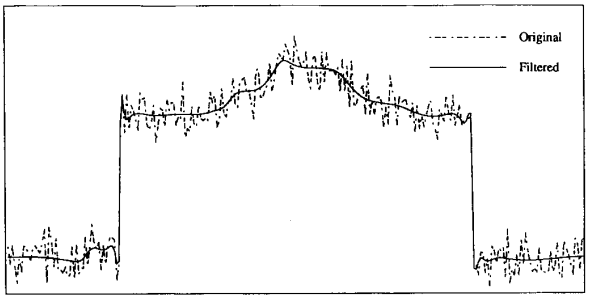
and

$$f(x,y) = \sum_m \sum_{n_x} \sum_{n_y} \Psi'_{m,n_x,n_y}(x,y)\, W(m,n_x,n_y) \tag{9}$$

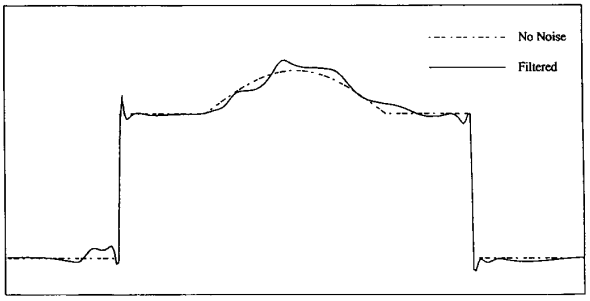
where $\Psi'(x,y)$ is the dual basis of $\Psi(x,y)$, and the scale index m is not greater than $M = \log_2[\min(Nx,Ny)]$ if $n_x = 1,2,\ldots,N_x$ and $n_y = 1,2,\ldots,N_y$. The 2-D wavelet transform kernel we used is a separable kernel function [9]. The 2-D filtering algorithm is built on the direct spatial correlation data

$$\mathrm{Corr}_l(m,n_x,n_y) = \prod_{i=0}^{l-1} W(m+i,n_x,n_y), \qquad 1 \le n_x \le N_x,\ 1 \le n_y \le N_y \tag{10}$$

where l is the number of resolution scales involved in the direct multiplication, and m is the intended processing scale. The iterating process of rescaling the $\mathrm{Corr}_l(m,n_x,n_y)$ data to $W(m,n_x,n_y)$, comparing the absolute values of $\mathrm{Corr}_l(m,n_x,n_y)$ and $W(m,n_x,n_y)$, and extracting edge information from the two data sets will yield the spatially filtered wavelet transform data $W_{\mathrm{new}}(m,n_x,n_y)$. The processed 2-D image data can then be generated through the inverse transform (9).



(a)



(b)

Fig. 6. One-dimensional filtering of a synthetic noisy signal with a large and smooth feature. The standard deviation of the synthetic Gaussian white noise is about 10% of the amplitude of the large boxcar: (a) Original noisy signal against the filtered data; (b) original clean (no noise) synthetic signal against the wavelet filtered noisy data. In the wavelet filter, $l = 2$ (the number of scales to correlate), and the first six wavelet scales are processed.

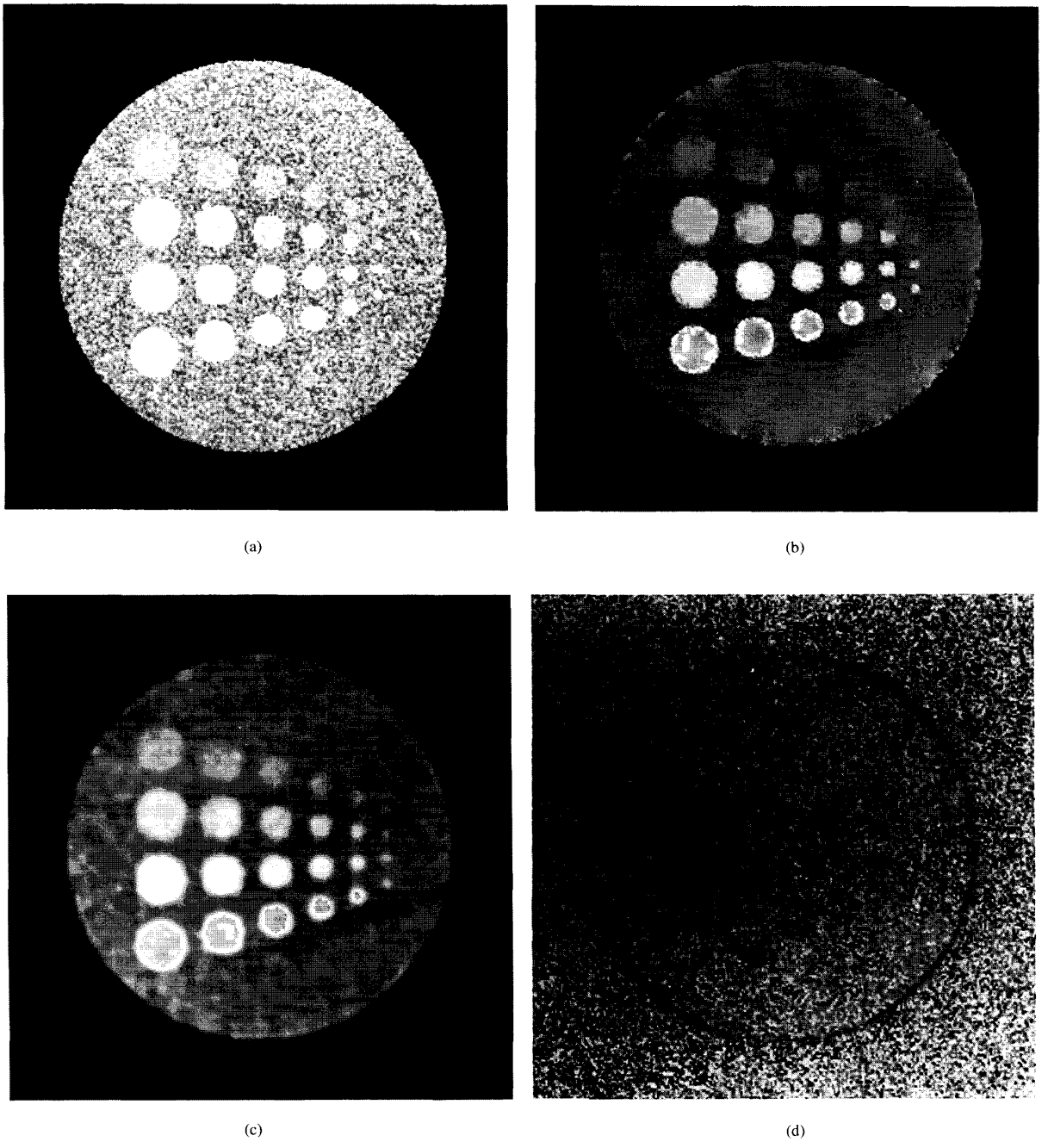
We synthesized a digital image of 256 (Nx) by 256 (Ny) matrix size to test our wavelet transform domain noise filtration technique. In the synthetic image, there are 24 small circles of different sizes and intensities on top of a large circle. We added a certain amount of synthetic Gaussian distributed white noise into the image data. The synthetic image is shown in Fig. 7(a), its average signal to noise ratio (SNR) is about 15 dB. Fig. 7(b) and (c) show filtered images with the wavelet filter and Wiener filter, respectively. In implementing the wavelet filter, we used the wavelet transform data only at two adjacent scales ($l = 2$) in calculating the direct spatial correlation. In addition, we filtered the wavelet transform data at the first six scales (from scale 1 to scale 6). There are eight wavelet scales. The boundaries of most of the small circles and the large circle remain very sharp after the noise filtration. The average SNR is enhanced to more than 25 dB (90% noise reduction) in Fig. 7(b). There is very little reduction of the spatial resolution, and artifacts are small and local. However, we did observe that certain noisy spots were left around the neighborhoods of the sharp edges. The smallest, lowest contrast circles could not be fully recovered from the noisy background. In the Wiener filtered image shown in Fig. 7(c), the amount of noise reduction is not as great as that in Fig. 7(b); the average SNR is 22.6 dB. On the other hand, the loss of image sharpness and ringing at the edges of the phantom are more noticeable. An error image acquired by subtracting the wavelet filtered image data from the original noisy synthetic image data and taking the absolute values is shown in Fig. 7(d).

In our implementation of the Wiener filter, we assumed the noise was white. A noisy image power spectrum was first computed by an averaged periodogram to reduce estimation variance. The noise power was then estimated at the high-frequency band of the noisy image spectrum, and the object spectrum was obtained by subtracting the estimated noise power from the noisy image spectrum. This method is referred to as spectral pruning or subtraction [14], [15]. The amount of spectral subtraction is controlled by the estimated noise power and a multiplicative factor that represents the tradeoff between the noise reduction and the image resolution. The more spectral subtraction is applied, the more noise is removed at the expense of having more blurring and ringing at edges.

We have also tested this wavelet domain filter on clinical images. Fig 8(a) is a real, axial, gradient echo, MR image of a volunteer taken on a General Electric Signa 1.5-T whole-body MR scanner.[1] The noise is not large in this particular image. The average SNR of the image is below 20 dB. The wavelet filtered head image is shown in Fig. 8(b), and the Wiener filtered one is shown in Fig. 8(c). In Fig. 8(b), we can see that most of the low-contrast and fairly small features in the brain tissues are recovered after the noise filtration. The noise reduction is remarkable. The filtered head image looks very clean and sharp. No ringing or other artifacts are present. In Fig. 8(c), small features in the head image are blurred considerably after Wiener filtering.

Figs. 9(a) and 10(a) give two more MR images from the axial head scan of the same volunteer. The noise in these two images is much larger than that in the previous one. The average SNR in these two images is about 12 dB. Many small and low-contrast features in the brain are heavily contaminated by the noise. Figs. 9(b) and 10(b) show the corresponding images after the wavelet domain filtering, and Figs. 9(c) and 10(c) are the ones after the Wiener filtering. At these low SNR's, the Wiener filter performs very poorly. The images are significantly blurred while the noise is still not smoothed as much as desired. The average SNR is raised to about 20.5 dB in the Wiener filtered images. With the wavelet filter, however, most low-contrast, large features are very well recovered as are the high-contrast, small features. However, the small, low-contrast features in the brain tissue, which do not have a spatial correlation $\mathrm{Corr}_l(m, n)$ superior to that of noise, are not well recovered. There is a little reduction in the spatial resolution; however, it does not appear to be a result of ringing artifacts and edge blurring but rather a result of the slight degradation in tissue contrast. Heavy noise has been more effectively removed. The filtered images again are very clean. The average SNR is raised to 22 dB in the wavelet filtered images. An error image acquired by subtracting the wavelet-filtered image data from the original MR image data and taking the absolute values is shown in Fig. 9(d).

[1] (G.E. Medical Systems, Milwaukee, WI).

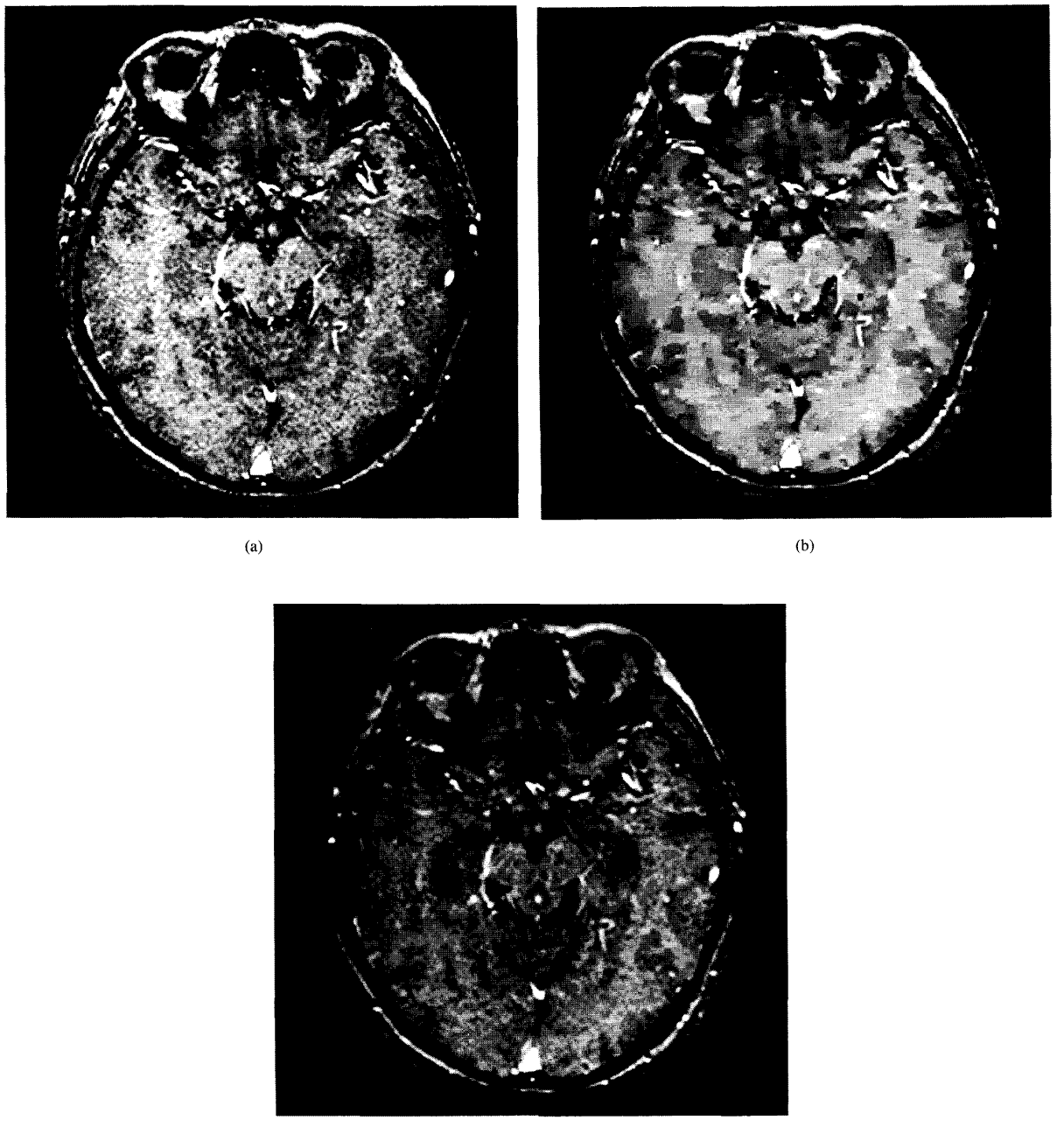


(a) (b) (c) (d)

Fig. 7. Fairly noisy synthetic digital image. The average SNR = 15 dB; the standard deviation of the synthetic Gaussian white noise is about 10% of the amplitude of the large circle image: (a) Before filtering; (b) after 2-D wavelet filtering; (c) after 2-D Wiener filtering; (d) difference between (a) and (b). In the wavelet domain filter, $l = 2$, and the first six wavelet scales are processed.

V. Discussion

The Wiener filter assumes the noise and power spectra of the object *a priori*. In practice, both object and noise spectra are often unknown and have to be estimated from a single noisy observation. The noise spectrum can often be estimated from a signal-free region of the image or in a high-frequency band, but estimating the object spectrum from a noisy observation is not an easy task. In particular, the estimation of the object spectrum in a high-frequency band is hampered by the strong presence of noise. At very low SNR, the Wiener filter tends to close its high-frequency passing band; it essentially rejects almost all



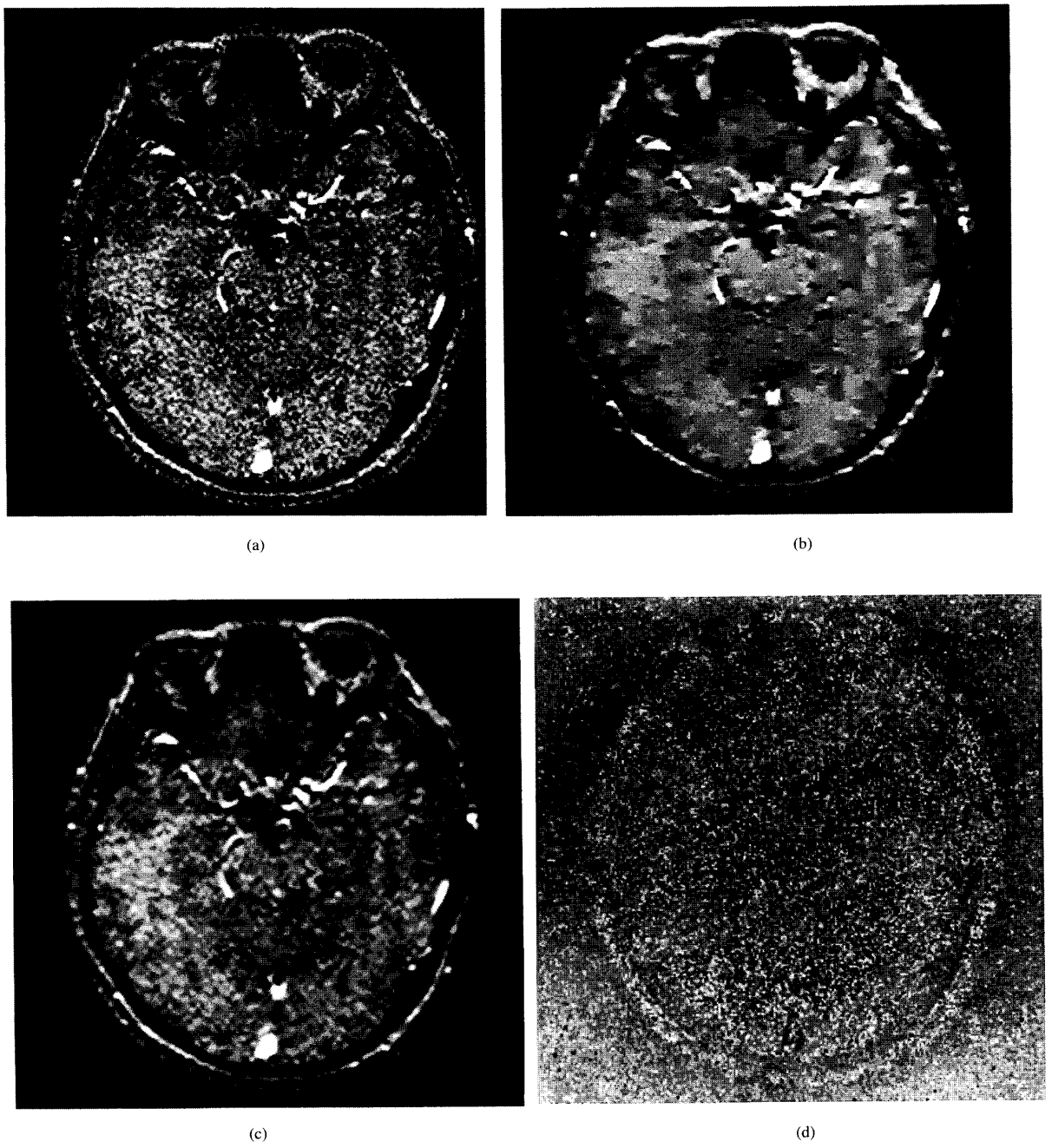
(a) (b)

(c)

Fig. 8. MR image from an axial head scan of a volunteer (SNR = 18 dB): (a) Before 2-D filtering; (b) after 2-D wavelet filtering; (c) after 2-D Wiener filtering. In the wavelet domain filter, $l = 2$, and the first six wavelet scales are processed.

high-frequency components, including those from important edge features.

The wavelet domain filter is superior to the Wiener filter because of its edge and feature-sensitive selectivity in passing certain high-frequency data. The direct spatial correlation of wavelet transform contents at several adjacent scales obviously enhanced major edges in the wavelet transform domain while the noise and small features were suppressed. The tradeoff with this filter is between noise suppression and small feature retention. Features that were large compared with the noise were easily retained and were not degraded. However, the features that were the same size as the noise were also suppressed because they were not distinguished from the noise. Rescaling the power of the direct multiplication values to that

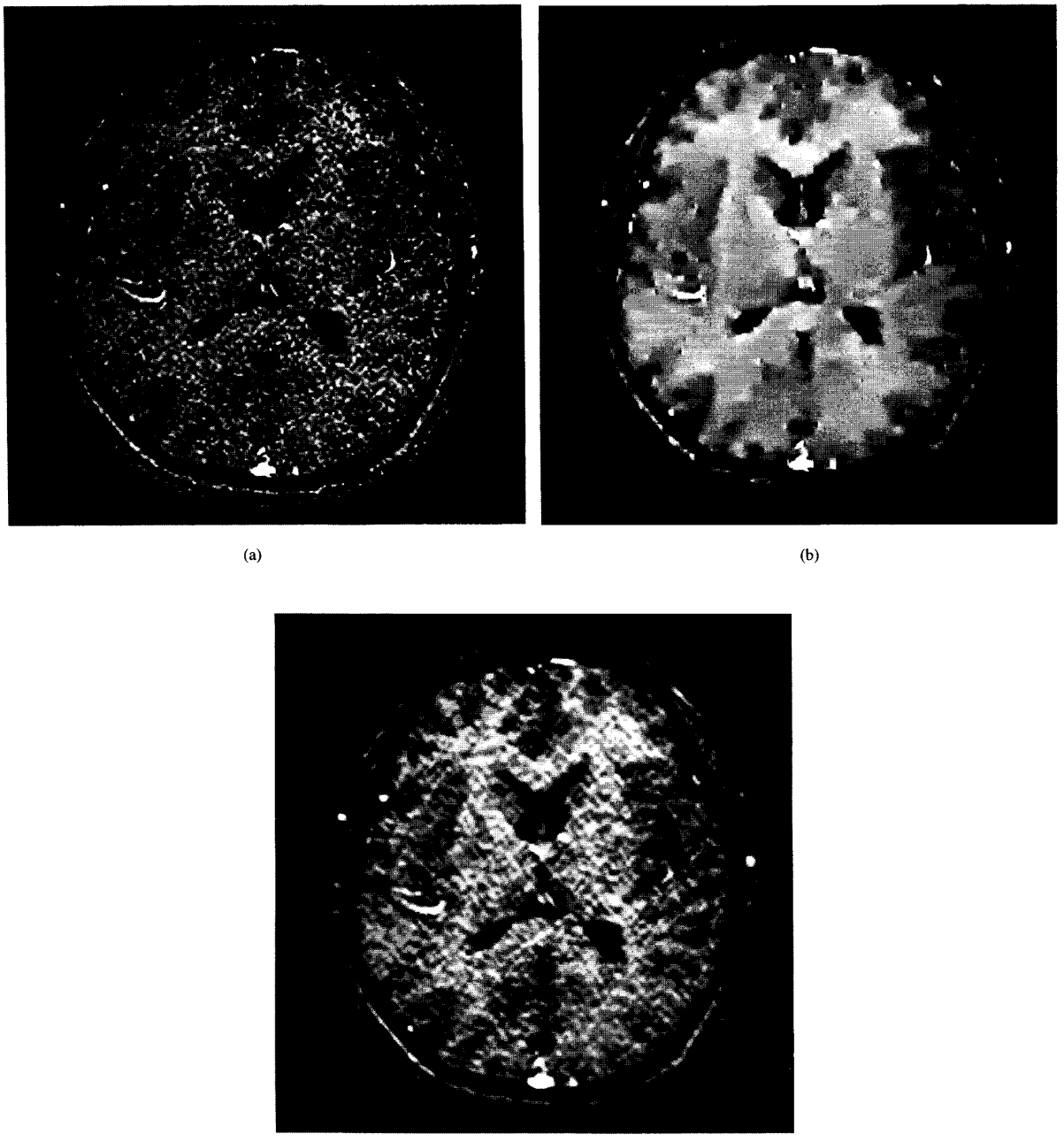


(a) (b) (c) (d)

Fig. 9. Much noisier MR head image of the same volunteer (SNR = 12 dB): (a) Before 2-D filtering; (b) after 2-D wavelet filtering; (c) after 2-D Wiener filtering; (d) difference between (a) and (b). In the wavelet domain filter, $l = 2$, and the first six wavelet scales are processed.

of the wavelet transform data at the smallest scale among the scales involved in the multiplication apparently transferred energy from the noisy background and small features to the major edge contents in the wavelet transformed data.

We found, from our experiments, that the direct spatial correlation involving only two or three adjacent wavelet scales (l =2 or 3) yielded the best filtering results. If we used more wavelet scales in the spatial correlation operation, we did not achieve better edge enhancement; in fact, many times, the situation became worse. We believe this is because the center locations of many types of edges do not occur at the same position over a wide range of scales in the wavelet transform domain; the shape of the edge determines where the maxima in the wavelet transform domain occur at each scale. For a



(a) (b)

(c)

Fig. 10. Another very noisy MR head image of the same volunteer (SNR = 12 dB): (a) Before 2-D filtering; (b) after 2-D wavelet filtering; (c) after 2-D Wiener filtering. In the wavelet domain filter, $l = 2$, and the first six wavelet scales are processed.

perfectly sharp edge the maxima at each scale occurs at the same position, but for other edges, the maxima can move. In addition, the position of the maxima can move if other edges interfere. It was also found that the more wavelet scales are involved in the spatial correlation operation, the more and stronger noise events were passed in the neighborhood of the sharp edges.

In the filtering algorithm, we used the "dark" (signal-free) regions at the boundaries of each image to estimate the noise power at each wavelet scale. This reference noise power serves as a threshold to terminate the edge extraction iteration. We found that the filtering results are not very sensitive to the slightly inaccurate selection of this reference noise power. We inserted an adjusting factor to the estimated reference noise

power. The effect of the filtering did not deviate very much (no more than 5% variation on the amount of resolution reduction in images) when the adjusting multiplicative factor varied from 1.0 to 1.3. However, we found that the adjusting factor was useful because the noise near the boundaries of an MR image is a bit lower than that near the center of the image. We have used a factor of around 1.3 to adjust the reference noise power estimated near the boundaries of MR images before using it as a threshold. This new threshold yields the best wavelet domain filtering results in reducing noise from MR images.

The computation time spent on filtering noise from 2-D images in the wavelet transform domain is not significant. The part of applying the 2-D filtering algorithm is much less time-consuming than the 2-D forward and inverse wavelet transforms. In the entire filtering process, 90% of the total processing time is spent on the 2-D wavelet transformations. However, it has been estimated that parallelization of the wavelet transform algorithm can increase the processing speed by three orders of magnitude [16].

VI. Conclusion

We have introduced an effective wavelet transform domain noise filtration technique. This filter preserves edges and removes noise. Noise is preferentially removed from the wavelet transform data at a given scale by comparing the data at that scale to the correlation of the data at that scale with those at larger scales. Features are identified and retained because they are strongly correlated across scale in the wavelet transform domain. Noise is identified and removed because it is poorly correlated across scale in the wavelet transform domain. Features remain relatively undistorted because they are very well localized in space in the wavelet transform domain; therefore, edges remain sharp after filtration. In comparison with the Wiener filter, the wavelet domain filter has edge and feature-sensitive selectivity in passing high-frequency data at very low SNR, whereas the Wiener filter tends to close its high-frequency passing band exclusively. It appears that the Wiener filter can improve SNR almost as much as the wavelet domain filter does. However, the mean square error is not a perfect criterion to measure the success in filtering noise from digital images. The visual appearance of the wavelet-filtered images is far superior to that of the Wiener-filtered images.

Our results show that this new filtering technique can remove more than 80% of the noise from the tested images and maintain all the edge gradients to more than 80% of their original values. There is no equivalent to Gibbs' ringing at edges after filtration. The noise filtration is anisotropic in both spatial and Fourier domains. Any artifacts present are always small and local. The loss of spatial resolution is almost unnoticeable.

However, this filtering technique does present a tradeoff between the SNR and the presence of small, low-contrast features in the filtered image; features that are the same size as noise are suppressed because they are not distinguished from the noise. We also noticed that the noise near a sharp edge is not removed as effectively as the noise in the smooth regions of an image.

We see several possible improvements of this technique. First, we are currently using the noisy background at the near-boundary regions of the image to estimate the noise power at each wavelet scale. This reference noise power serves as a threshold to terminate the iteration process employed by the filtering technique. We can extend the technique to allow users to select and analyze any noisy regions in the image to establish this reference noise power so that more effective noise filtering can be achieved. Second, spatial continuity of edges can be used to identify weak but long edges in large noise. Spatial continuity of edges will help us effectively remove noise near sharp edges. More importantly, it will help discriminate very weak, low-contrast features from the noisy background. This is a very powerful extension if it can be made to work effectively.

There are many possible applications of this filter. The spatially selective nature of this filter makes it useful in spatially dependent noise filtration, which is impossible for Fourier domain filters. Its good localization and protection of sharp edges will allow the wavelet filters to be very competitive in edge detection, pattern recognition, and computer vision as well as image enhancement and restoration. This wavelet transform domain filtering technique may become a very robust post-processing tool in many imaging applications, e.g., in MR angiography to help to identify the conspicuous blood vessel boundaries.

References

[1] E. R. McVeigh, R. M. Henkelman, and M. J. Bronskill, "Noise and filtration in magnetic resonance imaging," *Med. Phys.*, vol. 12, no. 5, pp. 586–591, 1985.

[2] H. Hart and Z. Liang, "Bayesian image processing in two dimensions," *IEEE Trans. Med. Imaging*, vol. 6, pp. 201–208, 1988.

[3] X. Hu, V. Johnson, W. H. Wong, and C.-T. Chen, "Bayesian image processing in magnetic resonance imaging," *Magn. Reson. Imaging*, vol. 9, pp. 611–620, 1991.

[4] J. B. Weaver, Y. Xu, D. M. Healy, Jr., and L. D. Cromwell, "Filtering noise from images with wavelet transforms," *Magn. Reson. Med.* vol. 21, no. 2, pp. 288–295, 1991.

[5] Y. Xu, J. Lu, D. M. Healy, Jr., and J. B. Weaver, "Filtering MR images with wavelet transforms," in *Proc. 10th Ann. Mtg. SMRM* (San Francisco), Aug. 1991, p. 199.

[6] J. Lu, Y. Xu, J. B. Weaver, and D. M. Healy, Jr., "Noise reduction by constrained reconstructions in the wavelet transform domain," in *Proc. 7th IEEE SPS MDSP Workshop* (Lake Placid, NY), Sept. 1991, p. 1.9

[7] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 41, pp. 909–996, Nov. 1988.

[8] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 11, no. 7, pp. 674–693, July 1989.

[9] S. G. Mallat and S. Zhong, "Complete signal representation with multiscale edges," *NYU Tech. Rep. No. 483*, Dec. 1989.

[10] A. Witkin, "Scale space filtering," in *Proc. 8th Int. Joint Conf. Artificial Intell.*, 1983.

[11] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 617–643, Mar. 1992.

[12] A. Rosenfeld, "A nonlinear edge detection technique," *Proc. IEEE (Lett.)*, vol. 58, pp. 814–816, May 1970.

[13] A. Rosenfeld and M. Thurston, "Edge and curve detection for visual scene analysis," *IEEE Trans. Comput.*, vol. C-20, no. 5, pp. 562–569, May 1971.

[14] F. M. Wahl, *Digital Image Processing*. Norwood, MA: Artech House, 1987.

[15] J. S. Lim, *Two-Dimension Signal and Image Processing*. Englewood Cliffs, NJ: Prentice Hall, 1990.

[16] J. Lu, "Parallel computation of 2-D wavelet transforms for image processing and computer vision," in *Proc. 7th IEEE SPS MDSP Workshop* (Lake Placid, NY), Sept. 1991, p. 10.5

Yansun Xu received the B.S. degree in radio electronics/radio spectroscopy from Peking University, Beijing, China, in 1984 and the Ph.D. degree in radiation physics from Dartmouth College, Hanover, NH, in 1990.

He was with the Department of Diagnostic Radiology of Dartmouth-Hitchcock Medical Center, Lebanon, NH, from 1990 to 1992, where he did research work on developing novel magnetic resonance imaging techniques and new image processing algorithms with the use of wavelet transforms. He has been with the NMR Division, Picker International, Inc., Highland Heights, OH, since 1992, where he is currently a Principal Engineer, developing imaging software for new magnetic resonance imagers. His current research interests are in the areas of applications of wavelet transform analyses in medical image processing and the development of new medical imaging techniques.

John B. Weaver received the B.S. degree in engineering physics from the University of Arizona in 1977 and the Ph.D. degree in biophysics from the University of Virginia in 1983.

From 1983 until 1985, he worked at Siemens Medical Systems in the position of MRI applications scientist. He came to Dartmouth in 1985, where he is an Associate Professor in the Department of Radiology at the Dartmouth-Hitchcock Medical Center and an adjunct assistant Professor at the Thayer School of Engineering at Dartmouth College. His research interests include the application of signal processing to both magnetic resonance imaging and screening mammography, as well as fast magnetic resonance imaging of susceptibility effects and its use in studying brain function.

Dennis M. Healy, Jr. received the B.A. degree in mathematics and physics in 1980 and the Ph.D. degree in mathematics in 1986 from the University of California, San Diego.

He is presently an Associate Professor with the Department of Mathematics and Computer Science at Dartmouth College. His research interests include phase space transforms, signal processing, harmonic analysis on symmetric spaces, and fast transform algorithms for nonabelian groups and their homogeneous spaces.

Jian Lu received the B.E. degree in mechanical engineering from Zhejiang University, Hangzhou, China, in 1984, the M.S.E.E. degree from the China Academy of Railway Sciences in 1987, and the Ph.D. degree in electrical engineering from Dartmouth College in 1993.

He is presently a Postdoctoral Research Associate in the Mathematics and Computer Science Department at Dartmouth College, where he does research in wavelet-based signal and image processing. His current research interests include nonuniform sampling and representation of signals, multiscale signal and image processing, image compression, and medical imaging.