Denoising and Enhancing for MR Images Using NLM and Shock Filter

V. Vanathe¹,*, M. A. Manikandan² and S. Boopathy³

¹,³Department of ECE, Asst. Prof., Dhirajlal Gandhi College of Technology, Salem 636 309, India.
²Department of MCA, Associate Prof., Sona College of Technology, Salem 636 005, India.
e-mail: ¹vanathe@gmail.com

Abstract. In medical field, the processing of medical images are the important to increasing resolution until the recognition is achieved and to remove the noise. The recognition can be achieved by the multiresolution techniques which provides a useful image representation for vision algorithms. After processing the image at a resolution, it is more efficient to analyze only the additional details which are available at a higher resolution level. The removal of noise in medical images is a very challenging issue in the field of medical image processing. Most well known noise reduction methods, which are usually based on the neighborhood of the pixels (NLM Filter) is employed. Certain features in processed image which is better than the original image in appearance is engaged by shock filter. This paper presents an efficient and simple method for denoising and enhancing the medical images. Finally it compares the multiresolution techniques to produce the best denoised MR image using efficient NLM algorithm with enhancement based on the Shock filter in terms of Mean Square Error (MSE), Signal to Noise Ratio (SNR) and Energy values.

Keywords: Multiresolution, Non local means (NLM), Image denoising, Image enhancement, Shock filter, Magnetic resonance imaging (MRI).

1. Introduction

In medical imaging, the Magnetic Resonance Imaging (MRI) is a powerful technique for diagnosis used by physician to detect the structural abnormalities. In general, the visualization of medical image displays a small percentage of the information available and medical images are corrupted by different type of noises. It is very important to obtain precise images to facilitate accurate observations for the given application. A universal property of images is the presence of a peculiar granular pattern of some noise, often referred to as rice or rician noise i.e the MR images are corrupted by noise. The noise in magnetic resonance magnitude images obeys rician distribution. The term rician noise is used to refer to the error between the underlying image intensities and the observed data in the given image. Rician noise is not zero-mean, and the mean depends on the local intensity in the image.

A novel approach based on resolution technique provides a dominant tool for image analysis and it is important for enhancing and denoising the quality of an image. A de-noising should be performed to improve the image quality for more accurate diagnosis. The main objective of image-de-noising techniques is to remove such noises while retaining as much as possible the important signal features. A denoising is the method by Non Local Means for removing the unwanted noise from the image pixels. It compares the weighted average of neighbourhood pixels in an image. Preferably, the resulting denoised image will not contain any noise or added artifacts. Non-Local Means (NLM) filter [9] which are based on neighborhood pixels is in particular interesting for medical imaging. The non-local means method performed exceptionally well compare to other denoising methods. The need of the non-local means algorithm is to accomplish its goals of removing noise and preserving detail.

Shock filters belong to the class of morphological image enhancement methods. Shock filters are based in the idea to apply locally either dilation or an erosion process, depending on whether the pixel belongs to the influence zone of a maximum or a minimum. The use of shock filters as a mean of image enhancement is recommended by the advantages is that it creates segmentation. This paper introduces a noise reduction method from medical images and to enhance the image. The experimental results show the efficacy of the proposed method and it compares the multiresolution techniques with statistical parameters.

*Corresponding author

© Elsevier Publications 2013.
The peak signal-to-noise ratio (PSNR) is the ratio between the maximum possible power of a signal and the power of corrupting noise. PSNR is usually expressed in terms of the logarithmic decibel scale. It is most easily defined via the mean squared error (MSE) and it is given as

$$\text{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - k(i, j)]^2$$  \hspace{1cm} (1)

where, \(I\) and \(k\) are the images.

The PSNR is defined as

$$\text{PSNR} = 10 \log_{10} \left( \frac{A^2}{\text{MSE}} \right) = 20 \log_{10} \left( \frac{A}{(\text{MSE})^{1/2}} \right)$$  \hspace{1cm} (2)

Here \(A\) is the maximum possible pixel value of the image may be 255. When the two images are identical, the MSE will be zero. Energy is used to describe a measure of “information” and corresponds to the mean squared value of the signal. Different kinds of statistical measurement can be used to analysis the performance of the image output. The Peak Signal-to-Noise Ratio (PSNR), Mean Square Error (MSE) and energy are used to evaluate the enhancement performance of various multiresolution techniques

2. Rician Distribution

Magnetic Resonance Imaging (MRI) are corrupted by Rician noise, which arises from complex Gaussian noise in the original frequency domain measurements. MR images are ideal for diagnosing and evaluating many conditions but the presence of rician noise will harm the diagnostic accuracy of an image. MR images are degraded by Rician noise, which arises from complex Gaussian noise in the original frequency domain measurements. If the real and imaginary data with mean values \(A_R\) and \(A_I\) correspondingly, are bankrupt by non zero mean uncorrelated Gaussian noise with standard deviation \(\sigma\), the PDF of the magnitude data will be a Rician distribution. It is given by

$$p(M|A) = \frac{M}{\sigma^2} e^{-\frac{M^2 + A^2}{2\sigma^2}} \frac{I_0 \left( \frac{AM}{\sigma^2} \right)}{\Gamma(1 + v/2)}$$  \hspace{1cm} (3)

where, \(I_0\) is the zeroth order Bessel function of the first kind and \(M\) denotes the pixel variable of the magnitude of an image and \(A\) is given by

$$A = (A_R^2 + A_I^2)^{1/3}$$  \hspace{1cm} (4)

The rician distribution tends to Rayleigh distribution when the SNR goes to zero and at high SNR it tends to Gaussian distribution. The moments of the Rice density function can be expressed systematically as a function of the confluent hypergeometric function.

$$E[M^n] = 2(\sigma^2)^{n/2} \Gamma(1 + n/2) \Gamma(-n/2; 1 - A^2/2\sigma^2)$$  \hspace{1cm} (5)

For even moments, it becomes a simple polynomial in its argument. Particularly, the second moment is given by

$$E[M^2] = 2\sigma^2 + A^2$$  \hspace{1cm} (6)

3. Multiresolution Image Decomposition Technique

Multiresolution techniques are precise functions that decompose the images into several scales i.e hierarchy of scales ranging from the coarsest scale to the finest one.

3.1 Discrete wavelet transform

The Discrete Wavelet Transform (DWT) provides sufficient amount of information and offers significant reduction in computation time. Wavelet coefficients of signal are the projections of the signal onto the multiresolution subspaces. Wavelets are functions generated from one single function (basis function) called the prototype or mother wavelet by dilations (scalings) and translations (shifts) in time (frequency) domain. If the mother wavelet is denoted by \(\Psi(t)\) and other wavelets \(\Psi_{a,b}(t)\) can be represented as

$$\Psi_{a,b}(t) = 1/\sqrt{|a|} \Psi \left( \frac{t - b}{a} \right)$$  \hspace{1cm} (7)
where, $a$ and $b$ are two arbitrary real numbers. The variables $a$ and $b$ represent the parameters for dilations and translations, respectively. The two-dimensional (2-D) DWT represents a real-valued image in terms of shifts and dilations of a lowpass scaling function and bandpass wavelet. The discrete 2-D wavelet functions $\Psi^{j,0}_{ij}$ and scaling functions $\Phi^{j,0}_{ij}$ have several indices — $j$ corresponding to the scale, $o$ corresponding to wavelet orientation (horizontal, vertical, or diagonal) $k,l$ corresponding to the position. In order to keep notation to a minimum, an abstract index $I$ is employed for these indices. Furthermore, 2-D wavelet functions, scaling functions, and images are vectorized by stacking the columns of each, and a single abstract spatial index $m$ is used. The $I^{th}$ scaling coefficient of the image $S$ is computed
\[
C_I = \sum_m \phi_I[m]s[m] 
\] (8)
Similarly the $I^{th}$ wavelet coefficient is computed as
\[
d_I = \sum_m \Psi_I[m]s[m] 
\] (9)

The scaling and wavelet coefficients are collectively denoted by vectors $c$ and $d$ respectively. The reason that the DWT is so enviable is that the wavelet transforms of natural signals and images tend to be very sparse, with a few large scaling and wavelet coefficients dominating the representation. That is, wavelet transforms tend to compress real-world signals.

3.2 Multiwavelet transform

Multi-wavelet transformation is a novel concept of wavelet transformation architecture, which has more than one scaling function $\Phi(t)$ and wavelet function $\Psi(t)$. Multi-wavelet is usually indicated by multi-dimensional vector function. In multiwavelet transform, multiwavelets have two or more scaling functions and mother wavelet for signal representation. The properties of GHM multiwavelet filter are orthogonality, symmetry and compact support. To implement the multiwavelet transform, we require a new filter bank structure where the low pass and high pass filter banks are matrices rather than scalars.

![Figure 1. Two-level decomposition of image sub bands, for scalar wavelets.](image)

Multiwavelet transform domain that there are first and second lowpass coefficient followed by first and second highpass coefficient rather than one lowpass coefficient followed by one highpass coefficient shown in figure 2. And the two level image sub bands for multiwavelet transform is shown in figure 3.

Since the GHM filter has two scaling and two wavelet functions, it has two low pass sub bands and two high pass sub bands in the transform domain. It is easy to see that there are similarities between low frequency sub-images so that it is possible to apply a certain prediction rule to remove the redundancy between the sub bands.

3.3 Laplacian pyramid

The Laplacian pyramid is ubiquitous for decomposing images into multiple scales and is widely used for image analysis. Over-complete decomposition based on difference-of-lowpass filters, the image is recursively decomposed into low-pass and highpass bands. Laplacian pyramids have been used to analyze images at multiple scales for a broad range of applications such as compression and harmonization. The name Laplacian Pyramid is a misnomer; it should be called the Difference of Gaussians Pyramid, since each level (i.e., image) is given roughly by smoothing with two Gaussians of different sizes, then subtracting and sub sampling.

© Elsevier Publications 2013.
Pixels to pixel correlation are first removed by subtracting a low pass filtered copy of the image from the image itself. The difference or error image has low variance and entropy, the low pass filtered image may be represented at reduced sample density. Iteration of the process at appropriately expanded scales generates a pyramid data structure. Let $I$ be the original image and $J$ be the result of applying an appropriate low pass filter to go. The prediction error $E$ is given by $E_1 = I_1 - J_1$. The reduced image $I_1$ is itself lowpass filtered to yield $I_2$ and a second error image is obtained $E_2 = I_2 - J_2$. By these steps we obtain a two dimensional arrays $E_1, E_2, \ldots, E_n$. If we now imagine these arrays stacked one above another, the result is a tapering pyramid data structure shown in figure 4. The value at each node at the pyramid represents the difference between two Gaussian like or related functions convolved with the original image.

The difference between these two functions is similar to the Laplacian operators. The value at each node in the Laplacian pyramid is the difference between the convolutions of two equivalent weighting functions with the original image. Again this is similar to convolving an appropriately scaled Laplacian weighting function with the image. The node value can be obtained directly by applying this operator. The pyramid decompositions are performed on each source image, all these decomposition are integrated to form a composite representation.

The NL-means algorithm replaces the noisy value by a weighted average of all the pixels of the image. The weight of a pixel is significant only if a Gaussian window around it looks like the corresponding Gaussian window around the reference pixel. Thus the non-local means algorithm uses image self-similarity to reduce the noise by averaging similar pixels. This average preserves the integrity of the image but reduces its small fluctuations, which are essentially due to noise.
Each pixel \( p \) of the non-local means denoised image is computed as

\[
\text{NL}(V)(p) = \sum_{q \in V} w(p, q) V(q)
\]  

(10)

where, \( V \) is the noisy image, and weights \( w(p, q) \) meet the following conditions \( 0 \leq w(p, q) \leq 1 \) and \( \sum_{q} w(p, q) = 1 \). Each pixel is a weighted average of all the pixels in the image. The weights are based on the similarity between the neighborhoods of pixels. The weights can then be computed using

\[
w(p, q) = \frac{1}{Z(p)} e^{-\frac{d(p, q)}{h}}
\]  

(11)

\( Z(p) \) is the normalizing constant defined as \( Z(p) = \sum_{q} e^{-\frac{d(p, q)}{h}} \) and \( h \) is the weight-decay control parameter. The non-local means algorithm has three parameters. The first parameter, \( h \), is the weight-decay control parameter which controls where the weights lay on the decaying exponential curve. If \( h \) is set too low, not enough noise will be removed. If \( h \) is set too high, the image will become blurry. When an image contains white noise with a standard deviation of \( \sigma \), \( h \) should be set between 10\( \sigma \) and 15\( \sigma \). The second parameter, \( R_{\text{sim}} \), is the radius of the neighborhoods used to find the similarity between two pixels. If \( R_{\text{sim}} \) is too large, no similar neighborhoods will be found, but if it is too small, too many similar neighborhoods will be found. Common values for \( R_{\text{sim}} \) are 3 and 4 to give neighborhoods of size \( 7 \times 7 \) and \( 9 \times 9 \), respectively.

The third parameter, \( R_{\text{win}} \), is the radius of a search window. Because of the inefficiency of taking the weighted average of every pixel for every pixel, it will be reduced to a weighted average of all pixels in a window. The window is centered at the current pixel being computed. Common values for \( R_{\text{win}} \) are 7 and 9 to give windows of size \( 15 \times 15 \) and \( 19 \times 19 \), respectively. With this change the algorithm will take a weighted average of \( 15^2 \) pixels rather than a weighted average of \( N^2 \) pixels for an \( N \times N \) image.

4. Shock Filter

The basic idea behind shock filters is the process of applying either erosion or dilation in a much localized manner, in order to create a “shock” between two influence zones, one belonging to a maximum and the other to a minimum of the signal. The use of shock filters as a mean of image enhancement is recommended by the advantages this particular method offers:

i. They create strong discontinuities at image edges and
ii. Furthermore, the filtered signal within a region delineated by those edges becomes flat.

In other words, shock filters create segmentation. Due to their discrete mathematical definition they are inherently unstable, meaning that they require special discretization schemes in order to preserve the total variation of the signal. Another property of shock filters is that they satisfy the maximum-minimum principle which states that the range of the filtered image remains within the range of the original image. Another advantage of shock filters over other image enhancement methods, such as Fourier or wavelet-based ones is that phenomena like the Gibbs phenomenon cannot appear.

4.1 Image enhancement through shock filtering

The suggested model is based on the following one-dimensional quasi-linear equation

\[
\frac{\partial u}{\partial t} + a(x) F(u_{xx}, u_x) \frac{\partial}{\partial x} f(u(x, t)) = 0
\]  

(12)

With discontinuous coefficients, where \( u \) is the original signal, \( a \) is a bounded and measurable discontinuous function, and \( F \) and \( f \) are regular (smooth) functions. The function \( F \) plays the same role as in the previous models (i.e., shock localization). The derivatives in its arguments are taken on the original image; as a result, this function is only computed once (at the beginning of the process. The new definition of the shock filter is

\[
\frac{\partial I}{\partial t} = -\text{sign}(I_{\omega \eta}) |I| \Delta I|
\]  

(13)
This particular model uses a different scale for each of the Gaussian kernels it employs: the structure scale $\sigma$ that determines the size of the flow-like patterns and the integration scale $\rho$ which has the role of averaging the orientation information in order to have a more robust orientation estimator. Another improvement in the definition which performs image enhancement and denoising at the same time. It is being more robust to small scale structures or flow like patterns. The main advantage of using shock filter is better contour preservation, better detail preservation, better edge detector etc.

5. Experiments and Results

To test our proposed algorithm we took a Magnetic Resonance Imaging of Human Brain which has size of an medical image is $256 \times 256$. Figure 5 shows the samples of the human brain. In this, the input image does not give the appropriate information for the diagnosing purpose. The rician or rice noise is used to refer to the error between the underlying image intensities and the observed data. The rician noise is added to the input image samples which results in the noisy distributed images shown in figure 6. The results from the experiment using the proposed noise removal method are presented in figure 7. Background part is excluded when our proposed noise removal algorithm is applied in MR image. Figure 7 shows that the rician noise is removed significantly and enhanced using Shock Filter. The accuracy of an image after applying noise removal algorithm, we can observe the image for out looking. It shows, the denoised output images provide more accurate information than the original input image. Performance evaluation involves a qualitative criterion (visual assessment) that reflects the ability of the algorithm to suppress noise while preserving image details.

![Figure 5. Original image.](image1)

![Figure 6. Rician distributed images.](image2)

![Figure 7. MR image after removing noise using non local means filter and enhancing by shock filter.](image3)
6. Conclusion

In this paper, we present a simple and efficient technique to remove noise from the medical images and to determine the pixel value in the noise less image. A multiwavelet system can simultaneously provide perfect reconstruction while preserving length (orthogonality), good performance at the boundaries (via linear-phase symmetry), and a high order of approximation (vanishing moments). The scalar wavelet does not possess all these properties at the same time. But the Laplacian Pyramid decomposition method provides finer performance than the other. With this for the enhancing purpose Shock filter is used which gives the better results.

The experimental results using MATLAB shows that the performance parameters for the Laplacian Pyramid technique are better than the scalar wavelet and multiwavelet. Natural images also have enough redundancy to be re-stored by NL-means. This results open new perspectives on multi-scale image analysis.

References


© Elsevier Publications 2013.

Table 1. Scalar wavelet.

<table>
<thead>
<tr>
<th>Sample Image Name</th>
<th>MSE</th>
<th>PSNR</th>
<th>Energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS</td>
<td>0.0062</td>
<td>70.2921</td>
<td>1.26E+04</td>
</tr>
<tr>
<td>AS1</td>
<td>0.0060</td>
<td>70.7890</td>
<td>1.21E+04</td>
</tr>
<tr>
<td>AS2</td>
<td>0.0058</td>
<td>70.8654</td>
<td>1.29E+04</td>
</tr>
<tr>
<td>AS3</td>
<td>0.0055</td>
<td>70.9777</td>
<td>1.30E+04</td>
</tr>
</tbody>
</table>

Table 2. Multiwavelet.

<table>
<thead>
<tr>
<th>Sample Image Name</th>
<th>MSE</th>
<th>PSNR</th>
<th>Energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS</td>
<td>8.63E−04</td>
<td>78.8235</td>
<td>1.37E+04</td>
</tr>
<tr>
<td>AS1</td>
<td>8.06E−04</td>
<td>79.0689</td>
<td>1.40E+04</td>
</tr>
<tr>
<td>AS2</td>
<td>8.70E−04</td>
<td>78.7338</td>
<td>1.34E+04</td>
</tr>
<tr>
<td>AS3</td>
<td>6.64E−04</td>
<td>79.9059</td>
<td>1.32E+04</td>
</tr>
</tbody>
</table>

Table 3. Laplacian pyramid.

<table>
<thead>
<tr>
<th>Sample Image Name</th>
<th>MSE</th>
<th>PSNR</th>
<th>Energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS</td>
<td>9.53E−05</td>
<td>88.3398</td>
<td>1.45E+04</td>
</tr>
<tr>
<td>AS1</td>
<td>7.36E−05</td>
<td>89.4596</td>
<td>1.41E+04</td>
</tr>
<tr>
<td>AS2</td>
<td>8.20E−05</td>
<td>88.9926</td>
<td>1.38E+04</td>
</tr>
<tr>
<td>AS3</td>
<td>9.81E−05</td>
<td>88.2158</td>
<td>1.44E+04</td>
</tr>
</tbody>
</table>