

10-1-2017

# A Multi-Resolution and Adaptive 3-D Image Denoising Framework with Applications in Medical Imaging

Partha Sarathi Mukherjee  
*Boise State University*

Noname manuscript No.  
(will be inserted by the editor)

# A Multi-resolution and Adaptive 3-D Image Denoising Framework with Applications in Medical Imaging

Partha Sarathi Mukherjee

Received: date / Accepted: date

**Abstract** Due to recent increase in the usage of 3-D magnetic resonance images (MRI) and analysis of functional magnetic resonance images (fMRI), research on 3-D image processing becomes important. Observed 3-D images often contain noise which should be removed in such a way that important image features, e.g., edges, edge structures, and other image details should be preserved, so that subsequent image analyses are reliable. Most image denoising methods in the literature are for 2-D images. However, their direct generalizations to 3-D images can not preserve complicated edge structures well. Because, the edge structures in a 3-D edge surface can be much more complicated than the edge structures in a 2-D edge curve. Moreover, the amount of smoothing should be determined locally, depending on local image features and local signal to noise ratio, which is much more challenging in 3-D images due to large number of voxels. This paper proposes an efficient 3-D image denoising procedure based on local clustering of the voxels. This method provides a framework for determining the size of bandwidth and the amount of smoothing locally by empirical procedures. Numerical studies and a real MRI denoising show that it works well in many medical image denoising problems.

**Keywords** 3-D image denoising · Adaptive smoothing · Local clustering · Local smoothing · Multi-resolution · Non-parametric regression

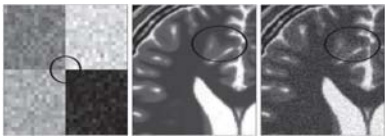
## 1 Introduction

Over the last few years, 3-D magnetic resonance images (MRI) of various body parts (e.g., brain) are being used extensively in medical diagnosis. However, those 3-D images often contain noise due to hardware imperfections and other reasons. Efficient noise removal is necessary for subsequent image analyses to be reliable. The focus of this paper is to remove noise from those 3-D images in such a way that important image features, e.g., edges, edge structures, and other image details are preserved well.

Most image denoising methods in the literature are for 2-D images. For example, the methods based on Markov random field (MRF) modeling, (e.g., Geman and Geman 1984, Besag 1986, Godtlielsen and Sebastiani 1994), local median and other robust filtering (e.g., Sun et al. 1994, Hillebrand and Müller 2007), bilateral filtering (e.g., Chu et al. 1998, Tomasi and Manduchi 1998), adaptive smoothing (e.g., Polzehl and Spokoiny 2000, Takeda et al. 2007), diffusion filtering (e.g., Perona and Malik 1990, Barash 2002), minimization of total variation (e.g., Rudin et al. 1992), wavelet transformation (e.g., Chang et al. 2000, Portilla et al. 2003, Om and Biswas 2015), jump curve/surface estimation (e.g., Qiu 1998, Gijbels et al. 2006, Qiu and Mukherjee 2010), modified non-local means (e.g., Kumar 2013), and many more (e.g., Arivazhagan et al. 2015). See Qiu (2005, 2007) and Katkovnik et al. (2006) for detailed information on this topic.

Many of those 2-D image denoising methods can be generalized for denoising 3-D images, and some of those methods are in the literature as well. For instance, 3-D image denoising based on minimization of Total Variation (TV) is popular in the computer sciences literature (e.g., Keeling 2003, Wang and Zhou 2006). However, the

abilities of those methods to preserve important edge structures and image details are limited. This is because 3-D edge surfaces can have more complicated structures than 2-D edge curves. Examples of complicated edge structures in 3-D images include the structures around a point where many edge surfaces meet (e.g., Figure 1, left panel), vertex of an irregularly shaped cone, vertex of a two-sided cone, and many more. Generalized version of many 2-D image denoising methods can not preserve those structures because the methods are not designed to do so.



**Fig. 1** One cross section from each of three 3-D images: one artificial, one noiseless  $T_2$  phantom, and its noisy version corrupted by Gaussian noise.

In the literature, there are some image denoising methods that are primarily designed to analyze 3-D images. Examples include non-local means algorithm (e.g., Buades et al. 2005, Coupe 2008), 3-D wavelet transformations (e.g., Hostalkova et al. 2007), distance weighted Weiner filtering (e.g., Lu et al. 2001), methods based on jump regression analysis (e.g., Mukherjee and Qiu, 2011, Qiu and Mukherjee 2012), and so forth. Many of those methods work well where the edge structure is relatively simple, e.g., the curvature of the edge surface is small, and a small neighborhood around a voxel contains at most two image regions. However, those methods blur the edge structures around the marked point in the left panel of Figure 1. Medical images often contain complicated structures and fine image details, e.g., the distribution of the gray matter and the white matter in a 3-D brain MRI image provides complicated structures in many places (Figure 1, middle and right panels). Recently, Mukherjee and Qiu (2015) propose an image denoising method based on local pixel/voxel clustering based on their intensity values. However, one major disadvantage of this method is that the smoothing extent is not locally adaptive, and hence fails to efficiently preserve many fine details of the image objects. Therefore, it is imperative that we need to construct 3-D image denoising methods that can perform this task well. This paper aims to provide one such method.

This paper proposes a novel 3-D image denoising method that can preserve complicated edge structures, even if more than two image regions form those structures. Moreover, this method uses multi-resolution technique in the sense that it selects the size of the bandwidth parameter locally, by a data-driven approach. In the background of the image, or around the places where the edge structures are simple, larger bandwidth

should be more suitable to remove noise. However, in places where the edge structure is complicated, or there are lot of image details, smaller bandwidth should be more suitable to reduce the amount of blur. The proposed method based on multi-resolution technique performs that by a local data driven approach. Also, the proposed method is adaptive, in the sense that the smoothing parameter is also selected locally by an empirical procedure. Another major advantage of the proposed method is that it selects the size of the bandwidth and the smoothing parameter locally by a non-iterative procedure, unlike the method in Polzehl and Spokoiny (2000), and thus the computation cost is checked at an acceptable level, which is very important specially in 3-D image denoising.

The proposed method has two major parts. The first part is a pilot screening of the whole image. In this part, overall noise level in the whole image is estimated and the complications of the edge structures are estimated. In the second part, bandwidth parameter is chosen locally from the estimations in the pilot screening, and then the image intensities are clustered into several groups based on their numerical values. Finally, the image intensity at a given voxel is estimated by appropriately smoothing the image intensities of the particular group that contains the given voxel.

The remaining part of the paper is organized as follows. The description of the proposed methodology is provided in Section 2. Section 3 presents numerical studies of the proposed method in comparison with a few state-of-the-art denoising techniques. Section 4 compares the performance of the proposed method on a real MRI with a few other competing methods. Section 5 provides some concluding remarks.

## 2 Proposed methodology

### 2.1 The underlying regression model

A monochrome image can be regarded as an image intensity surface that is usually discontinuous at the boundaries of the image objects (Qiu 2005). Under the jump regression model framework, (Mukherjee and Qiu 2011) suppose that a 3-D image follows the regression model

$$\xi_{ijk} = f(x_i, y_j, z_k) + \varepsilon_{ijk}, \quad \text{for } i, j, k = 1, 2, \dots, n, \quad (1)$$

where  $\{(x_i, y_j, z_k) = (i/n, j/n, k/n), i, j, k = 1, 2, \dots, n\}$  are equally spaced design points or voxels in the design space  $\Omega = [0, 1] \times [0, 1] \times [0, 1]$ ,  $\{\varepsilon_{ijk}\}$  are i.i.d. random errors, which may or may not be Gaussian, but with mean 0 and unknown variance  $\sigma^2$ ,  $f(x, y, z)$  is an unknown regression function denoting the image intensity

function, and  $N = n^3$  is the sample size. We further assume that there exists a partition  $\{A_l, l = 1, 2, \dots, s\}$  of the design space  $\Omega$  such that: (i) each  $A_l$  is a connected region in  $\Omega$ ; (ii)  $\bigcup_{l=1}^s A_l = \Omega$ ; (iii)  $f(x, y, z)$  is continuous in  $A_l \setminus \partial A_l$ , for  $l = 1, 2, \dots, s$ , where  $\partial A_l$  is the boundary point set of  $A_l$ , and (iv) there exist at most finite number of line segments  $\{\ell_l, l = 1, 2, \dots, s^*\}$  in  $[\bigcup_{l=1}^s \partial A_l] \cap \Omega$  such that for each line segment  $\ell_l$  there are  $A_{l_1}, A_{l_2} \in \{A_l, l = 1, 2, \dots, s\}$  satisfying  $\ell_l \subseteq \partial A_{l_1} \cap \partial A_{l_2}$  and for any  $(x^*, y^*, z^*) \in \ell_l$

$$\lim_{(x,y,z) \rightarrow (x^*, y^*, z^*), (x,y,z) \in A_{l_1}} f(x, y, z) = \lim_{(x,y,z) \rightarrow (x^*, y^*, z^*), (x,y,z) \in A_{l_2}} f(x, y, z).$$

Then, we call  $D := [\bigcup_{l=1}^s \partial A_l] \cap \Omega$  the *jump location surfaces (JLSs)* of  $f(x, y, z)$ . Obviously, JLSs describe the places where  $f$  has jumps. In image processing literature, they are called edge or jump surfaces.

## 2.2 Pilot screening

This stage of the proposed method serves two purposes. Firstly, it finds a rough estimate of  $\sigma$ , and secondly, it estimates the number of  $A_l$ 's that intersect a small neighborhood around each voxel. In the second stage of the proposed method, these information are used to select the bandwidth and smoothing parameters locally.

### 2.2.1 Estimation of $\sigma$

To get a rough estimate of  $\sigma$ , we can first apply a standard image denoising filter to get a denoised image and the residual image. Since we only want a rough estimate of  $\sigma$ , we prefer to choose a simple denoising filter so that the computation is fast. The performance of the proposed method is quite robust to the choice of the denoising filter. In this paper, we suggest using a local constant kernel (LCK) smoothing. The procedure is described below.

At a given voxel  $(x, y, z) \in \Omega$ , let us consider its spherical neighborhood  $O_{h_{ps}}(x, y, z)$  where the radius  $h_{ps}$  is a bandwidth parameter for pilot local smoothing. Since we want a rough estimate of  $\sigma$ ,  $h_{ps} = 1.0/n$  serves the purpose. To keep the procedure simple, this value of  $h_{ps}$  is used in all numerical studies in this paper. In  $O^*(x, y, z)$ , the local constant kernel (LCK) estimator of  $f(x, y, z)$  is obtained by

$$\hat{a}(x, y, z) = \sum_{(x_i, y_j, z_k) \in O^*(x, y, z)} \xi_{ijk} K \left( \frac{x_i - x}{h_{ps}}, \frac{y_j - y}{h_{ps}}, \frac{z_k - z}{h_{ps}} \right) \quad (2)$$

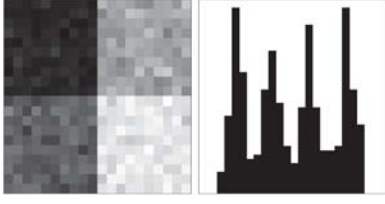
where  $K$  is a 3-D kernel density function defined in a unit ball. For simplicity, we use 3-D Gaussian kernel with variance 1.0 in all numerical studies in this paper.

Once we have the LCK estimator of  $f(x, y, z)$  for all voxels, we estimate  $\sigma$  by

$$\hat{\sigma} = \sqrt{\frac{1}{N} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n (\xi_{ijk} - \hat{a}(x_i, y_j, z_k))^2}. \quad (3)$$

### 2.2.2 Estimation of the number of $A_l$ 's that intersect a specified neighborhood around each voxel

At a given voxel  $(x, y, z) \in \Omega$ , let us consider its spherical neighborhood  $O_{h_{pc}}(x, y, z)$  where the radius  $h_{pc}$  is a bandwidth parameter for pilot local clustering. Again for simplicity, we suggest using a pre-fixed value  $h_{pc} = 3.0/n$ . Next, we use observed image intensity values within this neighborhood to estimate the number of  $A_l$ 's, say  $C_{h_{pc}}(x, y, z)$ , that intersect  $O_{h_{pc}}(x, y, z)$ . If we have large number of voxels in  $O_{h_{pc}}(x, y, z)$ , then the voxels within that neighborhood can be clustered into  $C_{h_{pc}}(x, y, z)$  number of well-separated groups, based on their observed image intensity values. This can be accomplished by a standard clustering algorithm. However, since the image intensities are scalar in this case, we can estimate  $C_{h_{pc}}(x, y, z)$  by a computationally simple algorithm. If we estimate the probability density function of the image intensity values in  $O_{h_{pc}}(x, y, z)$ , then it should have  $C_{h_{pc}}(x, y, z)$  number of local peaks, or local maxima. Moreover, the local minima separate the clusters from one another. One simple way to estimate the probability density function is to use histogram. We construct the histogram of  $\xi_{ijk}$ 's in  $O_{h_{pc}}(x, y, z)$ , with  $B$  number of bins, and find local minima by the following search algorithm. Suppose,  $H_l, l = 1, 2, \dots, B$  are the heights of the bins. Then, if  $H_{l-1} \geq H_l, H_{l+1} \geq H_l$ , and at least one of  $H_{l-1}$  and  $H_{l+1}$  is non-zero, then,  $H_l$  is a local minima. We choose  $B$  as a procedure parameter. Therefore,  $C_{h_{pc}}(x, y, z)$  can be estimated by the number of local minima minus one. Note that if only one  $A_l$  intersects  $O_{h_{pc}}(x, y, z)$ , then the local minima should be two. Moreover, the locations of the local minima define the clusters of voxels. To demonstrate this procedure, we artificially create a neighborhood where four  $A_l$  intersect. One such example of a 2-D image neighborhood corrupted with Gaussian noise is presented in the left panel of Figure 2. In an actual 3-D image neighborhood, each cross-section usually have much lower resolution, but due to an additional dimension, we have about similar number of voxels as in the left panel of Figure 2. The right panel of shows the histogram of the noisy image intensity values where we can see four peaks or local maxima.



**Fig. 2** Left panel: A 2-D image neighborhood containing four regions. Right panel: A histogram of the noisy image intensity values of the 2-D image neighborhood from the left panel.

### 2.3 A multi-resolution and adaptive smoothing

At a given voxel  $(x, y, z) \in \Omega$ , we use estimated  $C_{h_{pc}}(x, y, z)$  to select the size of bandwidth, i.e., the radius of the spherical neighborhood  $h(x, y, z)$ . If  $C_{h_{pc}}(x, y, z)$  is small, then the edge structures are simple in  $O_{h_{pc}}(x, y, z)$ , and so  $h(x, y, z)$  should be large. Conversely, if  $C_{h_{pc}}(x, y, z)$  is large, then the edge structures are complicated in  $O_{h_{pc}}(x, y, z)$ , and so  $h(x, y, z)$  should be small. Therefore, one reasonable approach to select  $h(x, y, z)$  is using a decreasing function of  $\hat{C}_{h_{pc}}(x, y, z)$ . In this paper, we suggest using the following function

$$h(x, y, z) = \frac{1}{n} \left( \max \left( 1.0, \left( M \exp \left( -\frac{1}{M} (\hat{C}_{h_{pc}}(x, y, z) - 1) \right) \right) \right) \right), \quad (4)$$

where the positive number  $M$  is a global procedure parameter controlling the maximum possible neighborhood size. Next, we consider the spherical neighborhood  $O_{h(x,y,z)}(x, y, z)$  of radius  $h(x, y, z)$  around the voxel  $(x, y, z)$  and estimate  $C_{h(x,y,z)}(x, y, z)$ , i.e., the number of  $\Lambda_i$ 's that intersect  $O_{h(x,y,z)}(x, y, z)$  by the procedure described in Section 2.2.2. Note that  $\hat{C}_{h(x,y,z)}(x, y, z)$  can be different from  $\hat{C}_{h_{pc}}(x, y, z)$ , because  $h(x, y, z)$  can be different from  $h_{pc}$ . However, the number of histogram bins  $B$  remains same as the pilot screening stage as  $B$  is used as a procedure parameter. While the performance of the proposed method is dependent on how we select local bandwidth, the function provided in (4) performs well in comparison with many other functions.

Next, we identify the cluster, say  $G(x, y, z)$  that contains the voxel  $(x, y, z)$ . One major advantage of the local clustering technique is that  $G(x, y, z)$  can be a union of disconnected regions, or a union of disconnected or intersecting lines. The true image intensity  $f(x, y, z)$  can be estimated by a weighted average of all  $\xi_{ij}$  in  $G(x, y, z)$ . One similarity measure can be quantified by considering small neighborhoods of size  $\tilde{h}$  (usually smaller than  $h(x, y, z)$ ) around the two voxels  $(x_i, y_j, z_k)$  and  $(x, y, z)$ , and then calculating the  $L_2$  distance of the observed intensity values in those neighborhoods. In this paper, we choose similarity measure  $\tilde{W}_{ijk} = \exp \left( -\frac{\|\tilde{O}(x_i, y_j, z_k) - \tilde{O}(x, y, z)\|_2^2}{2T_L(x, y, z)\hat{\sigma}_L^2(x, y, z)|\tilde{O}(x_i, y_j, z_k)|} \right)$ , where  $\|\tilde{O}(x_i, y_j, z_k) - \tilde{O}(x, y, z)\|_2$  is the  $L_2$  distance of the observed intensity values in the spherical neighborhoods of radius  $\tilde{h}$  around  $(x, y, z)$  and  $(x_i, y_j, z_k)$ ,  $T_L(x, y, z) > 0$  is a

local tuning parameter controlling the smoothness of the denoising procedure, and  $\hat{\sigma}_L(x, y, z)$  is the sample standard deviation of  $\xi_{ijk}$  in  $G(x, y, z)$ . Note that if  $T_L(x, y, z)$  is small, then  $\tilde{W}_{ijk}$  values are small, and if  $T_L$  is large, then  $\tilde{W}_{ijk}$  values are large. Therefore,  $T_L(x, y, z)$  is a smoothing parameter controlling the smoothness locally. If  $\hat{\sigma}_L(x, y, z)$  is large compared to  $\hat{\sigma}$ , then it is likely that there are some fine image details in  $G(x, y, z)$ . Therefore, to preserve those image details,  $T_L(x, y, z)$  should be small. Conversely, if  $\hat{\sigma}_L(x, y, z)$  is small compared to  $\hat{\sigma}$ , then it is likely that there is little image details in  $G(x, y, z)$ , and so  $T_L(x, y, z)$  should be large. For this reason, we can choose  $T_L(x, y, z)$  as a decreasing function of  $\hat{\sigma}_L(x, y, z)/\hat{\sigma}$ . In this paper, we suggest using the following function

$$T_L(x, y, z) = \text{Exp} \left( -\frac{\hat{\sigma}_L(x, y, z)}{\hat{\sigma}} \right), \quad (5)$$

where  $T$  is a global procedure parameter. Here also, the performance of the proposed method is dependent on how we select  $T_L(x, y, z)$ , the function provided in (5) performs well in comparison with many other functions. Then, our proposed estimator of  $f(x, y, z)$  is

$$\hat{f}(x, y, z) = \frac{\sum_{(x_i, y_j, z_k) \in G(x, y, z)} \tilde{W}_{ijk} \xi_{ijk}}{\sum_{(x_i, y_j, z_k) \in G(x, y, z)} \tilde{W}_{ijk}}. \quad (6)$$

If  $\tilde{h}$  is chosen to be too large then a lot of fine details of the image objects will be blurred. Moreover, computation time will increase significantly. Based on numerical simulations,  $\tilde{h} = 1.0/n$  works well in most practical applications. In most images, the performance of the denoising method worsens quite fast with increase of  $\tilde{h}$ . Therefore, in all numerical simulations in this paper, we use  $\tilde{h} = 1.0/n$ . The proposed multi-resolution and adaptive 3-D image denoising is summarized below.

*The proposed multi-resolution and adaptive 3-D image denoising procedure:*

**Part 1: Pilot Screening:**

- Estimate  $\sigma$  by the method as described in Section 2.2.1.
- For voxel  $(x, y, z)$ , estimate  $C_{h_{pc}}(x, y, z)$  by the histogram method described in Section 2.2.2.

**Part 2: Multi-resolution and adaptive smoothing:**

- For voxel  $(x, y, z)$ , determine the local bandwidth parameter  $h(x, y, z)$  by (4).
- Estimate  $\hat{C}_{h(x,y,z)}(x, y, z)$  and find  $G(x, y, z)$  by the histogram method.
- For voxel  $(x, y, z)$ , determine the local smoothing parameter  $T_L(x, y, z)$  by (5).
- For voxel  $(x, y, z)$ , estimate  $f(x, y, z)$  by (6).
- Repeat all steps, except the estimation of  $\sigma$ , for each voxel.

#### 2.4 Selection of procedure parameters:

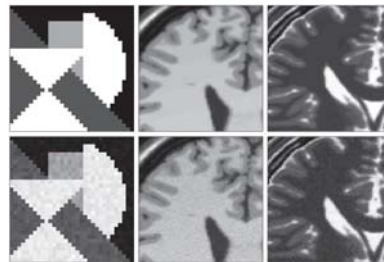
In the proposed denoising method there are three procedure parameters to choose:  $B$ ,  $M$  and  $T$ . In image denoising literature there are several parameter selection approaches such as methods based on (i) minimizing cross-validation, (ii) bootstrap, (iii) visual appearances, (iv) comparing structural similarity index measurements, (v) computationally fast ad-hoc procedures, and so on. In this paper, we suggest using structural similarity index measurement (SSIM) based approach to select the procedure parameters of the proposed method. Wang et al. (2004) introduce this approach to assess image quality. Since our goal is to preserve local image features, this is a reasonable approach. The algorithm goes like this: First, using an arbitrary choice of the procedure parameters we denoise the image and then calculate the SSIM between the denoised image  $\hat{f}$  and the estimated noise  $(\xi - \hat{f})$ . There should be no structural similarity between  $\hat{f}$  and  $(\xi - \hat{f})$  if the performance of the denoising method is good. Therefore, we select the parameter values that minimize SSIM between  $\hat{f}$  and  $(\xi - \hat{f})$ . Here are some guidelines for possible choices of the parameter values:  $B$  values from 2 to 6 are good enough for most images because the number of  $A_i$ 's in a neighborhood is rarely more than 3 or 4. Obviously,  $B$  should be integers only.  $M$  values between 1.8 and 6.0 work for most images as well.  $M$  can be both integers and fractions. Numerical studies in Section 3 and 4 show that the SSIM approach works well in many applications where the amount of noise does not depend heavily on image intensity.

### 3 Numerical Studies

In this section, we present some numerical results concerning the performance of the proposed multi-resolution and adaptive image denoising method, denoted as NEW, in comparison with three state-of-the-art image denoising methods that are widely used in the literature. The three competing methods include the image denoising method based on non-local means algorithm (Buades et al., 2005), denoted as NLM, the edge structure preserving 3-D image denoising method (Mukherjee and Qiu, 2011), denoted as ESPID, a denoising method based on total variation minimization (Rudin et al., 1992), denoted as TV. The NLM method has two bandwidth parameters and another smoothing parameter to choose. The ESPID method has a threshold parameter for edge detection and two bandwidth parameters: one for edge detection and another for smoothing. The TV method has a regularization parameter that controls the amount of smoothing and edge preservation. The proposed method NEW has three parameters to choose:  $B$ , i.e., the num-

ber of bins in the histograms of local image intensity values;  $M$  controlling the maximum possible radius of local neighborhoods; and  $T$ , i.e., the global procedure parameter controlling the local smoothing parameters.

The numerical study presented here includes one artificial image, one  $T_1$ -weighted magnetic resonance image (MRI) phantom and one  $T_2$ -weighted MRI phantom of human brain. The phantom images are collected from 'BrainWeb'. One cross-section from each image is presented in Figure 3. The first row presents noiseless images, and the second row presents their noisy versions corrupted by Gaussian noise. The artificial image has resolution  $32 \times 32 \times 32$ , its image intensity values range from 0.0 to 3.0. This image has many complicated edge structures like intersection of two, three and four edge surfaces, some of which are curved and some are planes. Then, we generate noisy versions of the true artificial image by adding i.i.d. noise from the  $N(0, \sigma^2)$  distribution with  $\sigma = 0.05, 0.10$ , and  $0.15$  representing low, medium and high levels of noise.



**Fig. 3** The first row presents one cross-section from each of the artificial image, and  $T_1$  and  $T_2$  phantom images. The second row presents their noisy versions corrupted by Gaussian noise, when  $\sigma$  is 0.10, 20 and 100, respectively.

Because the methods TV and NLM do not provide data-driven procedures to chose their procedure parameters, to make fair comparisons, we search their procedure parameters by minimizing the estimated MISE value, defined to be the sample mean of integrated square error (cf., Mukherjee and Qiu, 2011) computed from 100 replicated simulations. While the MISE criterion measures the overall performance of an image denoising procedure, it cannot measure how well the edges and other fine details of the image are preserved. To measure the preservation of such fine details of the image, Hall and Qiu (2007) defined a measure of jump size (JS) of an image. Its discretized version for the true image intensity function  $f$  can be written as

$$JS(f) = \frac{1}{(n-2)^3} \sum_{i=2}^{n-1} \sum_{j=2}^{n-1} \sum_{k=2}^{n-1} |f(x'_i, y'_j, z'_k) - f(x''_i, y''_j, z''_k)|,$$

where  $(x'_i, y'_j, z'_k)$  and  $(x''_i, y''_j, z''_k)$  are two immediately neighboring voxels of  $(x_i, y_j, z_k)$  on its two different sides along the estimated gradient direction of  $f$  at  $(x_i, y_j, z_k)$ . Obviously, if  $(x_i, y_j, z_k)$  is an edge voxel,

then  $|f(x'_i, y'_j, z'_k) - f(x''_i, y''_j, z''_k)|$  is close to the jump size of  $f$  at  $(x_i, y_j, z_k)$ . If  $(x_i, y_j, z_k)$  is a continuity voxel of  $f$ , then  $|f(x'_i, y'_j, z'_k) - f(x''_i, y''_j, z''_k)|$  is close to 0. Thus,  $EP(\hat{f}) = |JS(f) - JS(\hat{f})|/JS(f)$  is a reasonable measure of the edge and details preservation for the image denoising method in question. Since we are interested in the gradient directions rather than their magnitudes, a computationally simple  $3 \times 3 \times 3$  Sobel filter (cf., Qiu 2005, Section 4.4.3) is used in this paper when estimating  $f'_x$ ,  $f'_y$  and  $f'_z$ .

In the numerical studies, we also include the proposed method when the parameters are selected by the data driven procedure described at Section 2.4, denoted as NEW-SSIM. We select the parameters of the proposed method that minimizes the average SSIM between  $\hat{f}$  and  $(\xi - \hat{f})$  based on 100 replicated simulations. However, for all other methods, the parameters are selected by minimizing the estimated MISE computed from 100 replicated simulations.

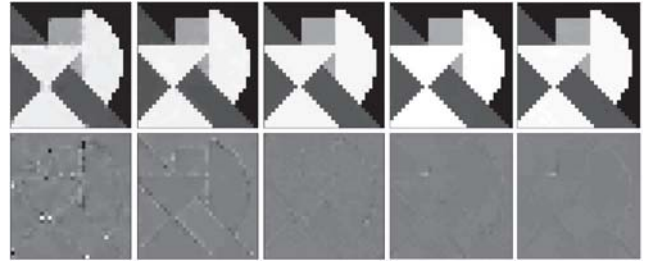
The numerical results for the artificial image is presented in Table 1. When comparing two methods in terms of MISE, if their estimated MISE values are  $MISE_1$  and  $MISE_2$  with standard errors  $SE_1$  and  $SE_2$ , respectively, and if  $MISE_1 < MISE_2$ , then a commonly used practical guideline is that we conclude that method 1 is significantly better than method 2 when  $MISE_2 - MISE_1 > 2(SE_1 + SE_2)$ . Similar comparisons can be made among different methods in terms of EP.

**Table 1** In each entry, the first line presents the estimated MISE value based on 100 simulations and the corresponding standard error (in parenthesis), the second line presents the value of EP and its standard error (in parenthesis), and the third line presents the searched procedure parameter values. This table is about the artificial image shown in Figure 3 when the noise is Gaussian.

	$\sigma = 0.05$	$\sigma = 0.10$	$\sigma = 0.15$
ESPID	0.0112 (0.0003) 0.0820 (0.0018) 0.0375, 12.0, 0.0469	0.0116 (0.0003) 0.1580 (0.0027) 0.0375, 18.0, 0.0469	0.0126 (0.0004) 0.2352 (0.0033) 0.0375, 18.0, 0.0469
TV	0.0005 (0.0000) 0.0363 (0.0007) 32.0	0.0019 (0.0000) 0.0618 (0.0015) 15.0	0.0039 (0.0001) 0.1099 (0.0025) 11.0
NLM	0.0018 (0.0002) 0.0339 (0.0008) 4, 1, 0.10	0.0020 (0.0002) 0.0705 (0.0018) 5, 1, 0.15	0.0025 (0.0002) 0.1018 (0.0024) 7, 1, 0.20
NEW	0.0003 (0.0000) 0.0380 (0.0008) 4.5, 12.0, 6	0.0007 (0.0001) 0.0492 (0.0012) 5.5, 15.0, 6	0.0014 (0.0002) 0.0612 (0.0017) 6.0, 20.0, 6
NEW-SSIM	0.0008 (0.0001) 0.0226 (0.0007) 6.0, 20.0, 6	0.0008 (0.0001) 0.0406 (0.0011) 6.0, 20.0, 6	0.0162 (0.0003) 0.0595 (0.0013) 6.0, 20.0, 2

From Table 1, we see that NEW outperforms its competitors in all cases. In this comparison, ESPID is worst by a large margin because it is designed to preserve only a few types of edge structures. When more than two edge surfaces intersect, ESPID fails to preserve those structures. NEW-SSIM also performs quite well in comparison with estimated MISE-wise optimal performances of other methods, except the case when  $\sigma = 0.05$  when TV performs better than NEW-SSIM. One realization of each of the denoised images by NEW

and its three competitors when  $\sigma = 0.10$  are presented in the first row of Figure 4. Their deviation images, defined as  $\hat{f} - f$ , are presented in the second row of Figure 4. If an image denoising method performs well, then there should not be any non-random pattern in the corresponding deviation image. From Figure 4 we see that NEW and NEW-SSIM indeed performs better than its competitors in this case.



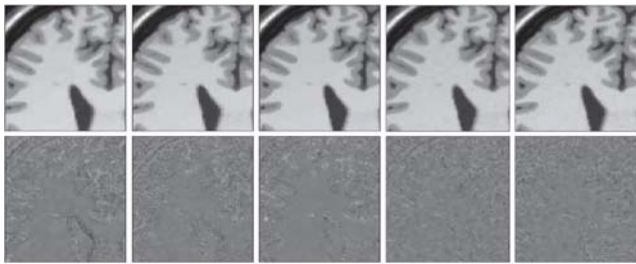
**Fig. 4** The first row shows one cross-section from the denoised artificial images by the methods ESPID, TV, NLM, NEW and NEW-SSIM when the noise is Gaussian with  $\sigma = 0.10$ . The second row shows their deviation images.

Next, we consider one  $T_1$  weighted MRI phantom and one  $T_2$  weighted MRI phantom and focus on one region with many image details from each of those images. Their resolutions are  $90 \times 70 \times 70$ , the image intensity values of the  $T_1$  weighted image range from 0.3 to 1053.2, and the image intensity values of the  $T_2$  weighted image range from 4.5 to 5040.2. We add noise from the distribution  $N(0, \sigma^2)$ , where  $\sigma$  is chosen to be 20 and 50 for the  $T_1$  image, and 100 and 200 for the  $T_2$  image. Then, we apply the four image denoising procedures to these two examples, and their parameters are chosen in the same way as those in the artificial image. The results are presented in Table 2 and Figures 5 and 6.

**Table 2** In each entry, the first line presents the estimated MISE value based on 100 simulations and the corresponding standard error (in parenthesis), the second line presents the value of EP and its standard error (in parenthesis), and the third line presents the searched procedure parameter values. This table is about  $T_1$  and  $T_2$  weighted MRI phantoms shown in Figure 3 when the noise is Gaussian.

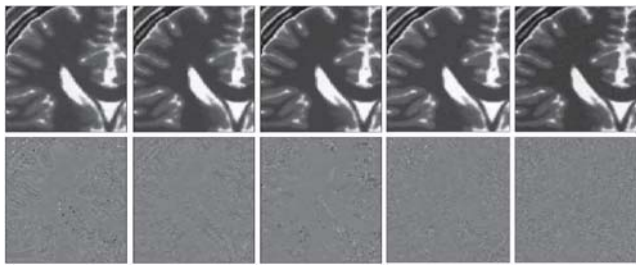
	$T_1, \sigma = 20$	$T_1, \sigma = 50$	$T_2, \sigma = 100$	$T_2, \sigma = 200$
ESPID	197.2 (0.7) 0.0388 (0.0006) 0.0167, 3500, 0.0167	519.4 (2.6) 0.0197 (0.0015) 0.0167, 6000, 0.0200	7246.2 (40.2) 0.0015 (0.0005) 0.0112, 40000, 0.0114	16180.6 (57.3) 0.0530 (0.0009) 0.0167, 35000, 0.0167
TV	170.7 (0.6) 0.0441 (0.0005) 0.105	585.8 (2.0) 0.0905 (0.0012) 0.037	4828.3 (15.4) 0.0353 (0.0004) 0.022	13660.3 (48.3) 0.0641 (0.0008) 0.010
NLM	117.7 (0.4) 0.0449 (0.0006) 3, 1, 20	338.8 (1.7) 0.0548 (0.0011) 3, 1, 40	5730.6 (17.6) 0.0338 (0.0003) 3, 1, 100	9836.4 (44.2) 0.0614 (0.0009) 3, 1, 200
NEW	116.3 (0.4) 0.0249 (0.0006) 1.8, 12.0, 2	380.6 (1.4) 0.0271 (0.0013) 2.0, 12.0, 2	4643.8 (15.1) 0.0009 (0.0004) 2.2, 3.0, 4	10482.2 (39.6) 0.0306 (0.0007) 2.2, 10.0, 2
NEW-SSIM	121.0 (0.4) 0.0271 (0.0006) 2.2, 10.0, 2	411.3 (1.6) 0.0171 (0.0014) 3.0, 3.0, 2	5632.8 (18.8) 0.0359 (0.0006) 2.2, 2.0, 6	16617.4 (56.9) 0.0863 (0.0009) 2.2, 3.0, 6

From Table 2, we see that NEW outperforms its competitors when noise level is low. When noise level



**Fig. 5** The first row shows one cross-section from the denoised  $T_1$  weighted MRI phantom by the methods ESPID, TV, NLM, NEW and NEW-SSIM when the noise is Gaussian with  $\sigma = 20$ . The second row shows their deviation images.

is higher, then NLM is the best in terms of MISE, but NEW still outperforms ESPID and TV. NEW-SSIM performs reasonably well on  $T_1$  phantom, but not on  $T_2$  phantom when noise level is high. From Figures 5 and 6 also, NEW and NEW-SSIM seems to preserve image details better than its competitors when noise level is low.



**Fig. 6** The first row shows one cross-section from the denoised  $T_2$  weighted MRI phantom by the methods ESPID, TV, NLM, NEW and NEW-SSIM when the noise is Gaussian with  $\sigma = 100$ . The second row shows their deviation images.

Next, we consider Rician noise which is commonly observed in magnitude resonance images. The observed image  $Z$  can be described by:

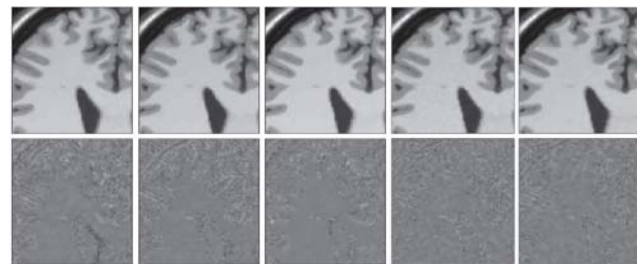
$$Z(x, y, z) = \sqrt{[f(x, y, z) + N_1(x, y, z)]^2 + [N_2(x, y, z)]^2}$$

where  $N_1(x, y, z)$  and  $N_2(x, y, z)$  are two independent random variables with normal distribution  $N(0, \sigma^2)$ . We artificially generate noisy images corrupted by Rician noise where  $\sigma = 0.15$  for the artificial image,  $\sigma = 20$  and 50 for  $T_1$  phantom, and  $\sigma = 100$  and 200 for  $T_2$  phantom. Since conventional denoising methods leave positive bias while estimating the true image intensity function (c.f., Mukherjee and Qiu, 2013), an efficient bias correction procedure (e.g., Gudbjartsson and Patz 1995, Wiest-Daessle (2008), Mukherjee and Qiu 2013) is imperative. In this paper, we use the bias correction method suggested by Mukherjee and Qiu (2013). The performances of the competing methods are presented in Table 3 and Figs 7 and 8. To calculate SSIM between the estimated noise and the denoised image, bias

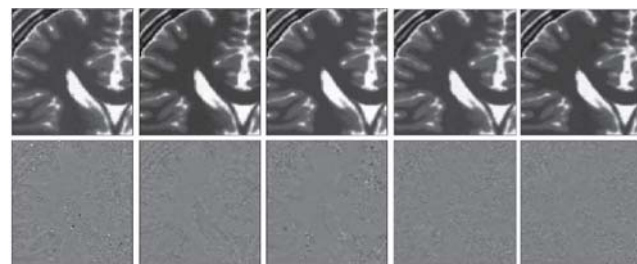
correction method for Rician noise is not used because  $E[Z(x, y, z)] \neq f(x, y, z)$ . In the Rician noise case also, NEW works well on artificial images. In cases of  $T_1$  and  $T_2$  phantom images, NEW works well when the noise level is low. However, NEW-SSIM does not work well on  $T_2$  phantom images when Rician noise level is high.

**Table 3** In each entry, the first line presents the estimated MISE value based on 100 simulations and the corresponding standard error (in parenthesis), the second line presents the value of EP and its standard error (in parenthesis), and the third line presents the searched procedure parameter values. This table is about the Rician noise case.

	Artificial, $\sigma = 0.15$	$T_1, \sigma = 20$	$T_1, \sigma = 50$	$T_2, \sigma = 100$	$T_2, \sigma = 200$
ESPID	0.0135 (0.0004) 0.2006 (0.0039) 0.0375, 15.0, 0.0469	200.1 (0.6) 0.0383 (0.0006) 0.0167, 3500, 0.0167	544.3 (2.1) 0.0292 (0.0013) 0.0167, 6500, 0.200	7352.7 (35.4) 0.0015 (0.0005) 0.0112, 40000, 0.0114	16369.7 (55.5) 0.0485 (0.0009) 0.0167, 28000, 0.0167
TV	0.0050 (0.0001) 0.1358 (0.0030) 13.0	172.9 (0.5) 0.0529 (0.0005) 0.100	601.0 (2.3) 0.0975 (0.0013) 0.036	4876.1 (13.5) 0.0361 (0.0004) 0.022	13752.5 (44.0) 0.0624 (0.0008) 0.010
NLM	0.0027 (0.0002) 0.0760 (0.0021) 7, 1, 0.20	120.4 (0.5) 0.0451 (0.0006) 3, 1, 20	368.9 (1.8) 0.0460 (0.0015) 3, 1, 40	5749.2 (19.4) 0.0340 (0.0004) 3, 1, 100	9988.9 (40.6) 0.0580 (0.0007) 3, 1, 100
NEW	0.0025 (0.0001) 0.1085 (0.0027) 5.0, 5.0, 6	118.6 (0.4) 0.0192 (0.0006) 1.8, 10.0, 2	408.4 (1.5) 0.0143 (0.0013) 2.0, 12.0, 2	4686.4 (14.6) 0.0098 (0.0004) 2.2, 3.0, 4	10674.0 (38.7) 0.0258 (0.0007) 2.2, 10.0, 2
NEW-SSIM	0.0031 (0.0001) 0.0744 (0.0015) 5.0, 20.0, 6	125.0 (0.4) 0.0426 (0.0005) 2.2, 18.0, 2	436.0 (1.6) 0.0093 (0.0012) 2.4, 5.0, 2	4725.9 (13.1) 0.0064 (0.0005) 2.2, 2.0, 4	17225.2 (60.7) 0.0937 (0.0010) 2.2, 3.0, 6



**Fig. 7** The first row shows one cross-section from the denoised  $T_1$  weighted MRI phantom by the methods ESPID, TV, NLM, NEW and NEW-SSIM when the noise is Rician with  $\sigma = 20$ . The second row shows their deviation images.



**Fig. 8** The first row shows one cross-section from the denoised  $T_2$  weighted MRI phantom by the methods ESPID, TV, NLM, NEW and NEW-SSIM when the noise is Rician with  $\sigma = 100$ . The second row shows their deviation images.

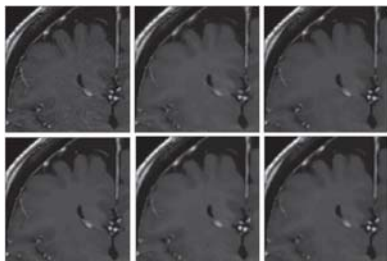
Additional numerical results are provided in “**Supplementary File**”.

#### 4 Real MRI denoising

In this section, we consider a real MRI for which the true image intensity function  $f$  is unknown. The real MRI is collected from <http://www.osirix-viewer.com/datasets/>. The first image in Figure 9 shows one slice of a region of the brain MRI. The selected region has



resolution  $80 \times 80 \times 20$ . All competing denoising methods are applied. Since the true noiseless image intensities are unknown, we can not calculate estimated MISE. Therefore, the parameters are selected based on visual appearances of the denoised images. For the proposed method, we also used the parameter selection based on SSIM. In Figure 9, we see that NEW and NLM perform better than other methods.



**Fig. 9** The first row shows one cross-section from the real MRI, denoised images by the methods ESPID and TV. The second row shows the same cross-section from the denoised images by NLM, NEW and NEW-SSIM.

## 5 Concluding Remarks

We have presented a framework for 3-D image denoising procedure that is multi-scale and adaptive by selecting bandwidth and smoothing parameters locally by data driven methods. This kind of 3-D image denoising procedure can be applied in many fMRI analyses used in many medical studies. However, the parameter selection criterion based on SSIM does not work on images where the noise is heavy and it depends on intensity values e.g., Rician noise. Therefore, future research on this issue of parameter selection need to be conducted. The proposed denoising framework can be successfully applied to diffusion tensor images where image segmentation by edge detection is difficult.

**Acknowledgements** The author thanks the editor, an associate editor and two referees for their valuable comments which greatly improved the quality of this paper.

## References

ARIVAZHAGAN, S., SUGITHA, N., and VIJAY, A. (2015). A novel image denoising scheme based on fusing multiresolution and spatial filters. *Signal, Image and Video Processing*, **9**(4), 885–892.

BARASH, D. (2002). A fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24** 844–847.

BESAG, J. (1986). On the statistical analysis of dirty pictures (with discussion). *Journal of the Royal Statistical Society (Series B)*, **48** 259–302.

BrainWeb: <http://brainweb.bic.mni.mcgill.ca/brainweb/>.

BUADES, B., COLL, B. and MOREL, J.M. (2005). A review of image denoising algorithms, with a new one. *Multiscale Modeling & Simulation*, **4**(2) 490–530.

CHANG, G.S., YU, B. and VETTERLI, M. (2000). Spatially adaptive wavelet thresholding with context modeling for image denoising. *IEEE Transactions on Image Processing*, **9**(9) 1522–1531.

CHU, C.K., GLAD, I.K., GODTLIEBSEN, F. and MARRON, J.S. (1998). Edge-preserving smoothers for image processing (with discussion). *Journal of the American Statistical Association*, **93** 526–556.

COUPE, P., YGER, P., PRIMA, S., HELLIER, P., KERVRANN, C. and BARILLOT, C. (2008). An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. *IEEE Transactions on Medical Imaging*, **27** 425–441.

GEMAN, S. and GEMAN, D. (1984). “Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6** 721–741.

GLIBELS, I., LAMBERT, A. and QIU, P. (2006). Edge-preserving image denoising and estimation of discontinuous surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **28**(7) 1075–1087.

GODTLIEBSEN, F. and SEBASTIANI, G. (1994). Statistical methods for noisy images with discontinuities. *Journal of Applied Statistics*, **21** 459–477.

GUBBJARTSSON, H. and PATZ, S. (1995). The noisy distribution of noisy MRI data. *Magnetic Resonance in Medicine*, **34** 910915.

HALL, P. and QIU, P. (2007). Blind deconvolution and deblurring in image analysis. *Statistica Sinica*, **17** 1483–1509.

HILLEBRAND, M. and MÜLLER, C.H. (2007). Outlier robust corner-preserving methods for reconstructing noisy images. *The Annals of Statistics*, **35** 132–165.

HOSTALKOVA, E., OLDRICH, V. and PROCHAZKA, A. (2007). Multi-dimensional biomedical image de-noising using Haar transform. *Proceedings of the 15th International Conference on Digital Signal Processing*, 175–178, Cardiff, UK.

KATKOVNIK, V., EGIAZARIAN, K., and ASTOLA, J. (2006). *Local Approximation Techniques in Signal and Image Processing*. SPIE Press book. PM 157.

KEELING, S. (2003). Total variation based convex filters for medical imaging. *Applied Mathematics and Computation*, **139**(1), 101–119.

KUMAR, B.K.S. (2013). Image denoising based on non-local means filter and its method noise thresholding. *Signal, Image and Video Processing*, **7**(6), 1211–1227.

LU, H., JUI-HSI, C., HAN, G., LI, L. and LIANG, Z. (2001). A 3D distance-weighted Wiener noise reduction in sinogram space for SPECT imaging. *SPIE proceedings series*, **4320**, 905–913.

MUKHERJEE, P.S. and QIU, P. (2011). 3-D Image Denoising By Local Smoothing And Nonparametric Regression. *Technometrics*, **53**(2) 254–273.

MUKHERJEE, P.S. and QIU, P. (2013). Efficient Bias Correction For Magnetic Resonance Image Denoising. *Statistics in Medicine*, **32**(12) 2079–2096.

MUKHERJEE, P.S. and QIU, P. (2015). Image Denoising by a Local Clustering Framework. *Journal of Computational and Graphical Statistics*, **24**(1) 196–208.

OM, H. and Biswas, M. (2015). A generalized image denoising method using neighbouring wavelet coefficients. *Signal, Image and Video Processing*, **9**(1) 191–200.

PERONA, P. and MALIK, J.(1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12**(7) 629–639.

POLZEHL, J. and SPOKOINY, V.G. (2000). Adaptive weights smoothing with applications to image restoration. *Journal of the Royal Statistical Society (Series B)*, **62** 335–354.

PORTILLA, J., STRELA, V., WAINWRIGHT, M. and SIMONCELLI, E.P. (2003). Image denoising using scale mixtures of gaussians in the wavelet domain. *IEEE Transactions on Image Processing*, **12**(11) 1338–1351.

QIU, P. (1998). Discontinuous regression surfaces fitting. *The Annals of Statistics*, **26** 2218–2245.

QIU, P. (2005). *Image Processing and Jump Regression Analysis*. New York, John Wiley.

QIU, P. (2007). Jump surface estimation, edge detection, and image restoration. *Journal of the American Statistical Association*, **102** 745–756.

QIU, P. and MUKHERJEE, P.S. (2010). Edge Structure Preserving Image Denoising. *Signal Processing*, **90**(10) 2851–2862.

QIU, P. and MUKHERJEE, P.S. (2012). Edge Structure Preserving 3-D Image Denoising By Local Surface Approximation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **34**(8) 1457–1468.

RUDIN, L., OSHER, S. and FATEMI, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D*, **60** 259–268.

SUN, T., GABBOUJ, M. and NEUVO, Y. (1994). Center weighted median filters: some properties and their applications in image processing. *Signal Processing*, **35**(3) 213–229.

TAKEDA, H., FARSIU, S. and MILANFAR, P. (2007). Kernel regression for image processing and reconstruction. *IEEE Transactions on Image Processing*, **16**(2) 349–366.

TOMASI, C. and MANDUCHI, R. (1998). Bilateral filtering for gray and color images. *Proceedings of the 1998 IEEE International Conference on Computer Vision*, 839–846. Bombay, India.

WANG, Y., BOVIK, A.C., SHEIKH, H.R. and SIMONCELLI, E.P. (2004). Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*, **13**(4) 600–612.

WANG, Y. and ZHOU, H. (2006). Total variation wavelet-based medical image denoising. *International Journal of Biomedical Imaging*, **2006** 1–6.

WIEST-DAESSLE, N., PRIMA, S., COUPE, P., MORRISSEY, S. and BARILLOT, C. (2008). Rician noise removal by non-local means filtering for low signal-to-noise ratio MRI: application to DT-MRI. MICCAI. New York/USA; 104117.