Generalized Scale-Based Image Filtering

Andre Souza[a] , Jayaram K. Udupa[a], and Anant Madabhushi[a]
[a]Medical Image Processing Group, Department of Radiology
University of Pennsylvania, Philadelphia, PA 19104.

ABSTRACT

In medical imaging, low signal-to-noise ratio (SNR) and/or contrast-to-noise ratio (CNR) often cause many image processing algorithms to perform poorly. Postacquisition image filtering is an important off-line image processing approach widely employed to enhance the SNR and CNR. A major drawback of many filtering techniques is image degradation by diffusing/blurring edges and/or fine structures. In this paper, we introduce a scale-based filtering method that employs scale-dependent diffusion conductance to perform filtering. This approach utilizes novel object scale information via a concept called generalized ball scale, which imposes no shape, size, or anisotropic constraints unlike previously published ball scale-based filtering strategies. The object scale allows us to better control the filtering process by constraining smoothing in regions with fine details and in the vicinity of boundaries while permitting effective smoothing in the interior of homogeneous regions. Quantitative evaluations based on the Brainweb data sets show superior performance of generalized ball scale-based diffusive filtering over two existing methods, namely, ball scale-based and nonlinear complex diffusion process. Qualitative experiments based on both phantom and patient magnetic resonance images demonstrate that the generalized ball scale-based approach leads to better preserved fine details and edges.

Keywords: Diffusion, image filtering, MR imaging, scale.

1. INTRODUCTION

Noise is ubiquitous in acquired images, especially medical. In many tasks, the utility of an image is determined by how well real intensity interfaces and fine details are preserved in the acquired image. Often the details may be along an edge or a boundary. In addition to influencing diagnostic tasks, noise also affects many image processing and analysis tasks such as segmentation[1, 2], registration[3, 4], and visual rendering[5, 6] that are crucial in many applications.

Methods for improving SNR and CNR can be divided into two categories: those based on image acquisition techniques and methods based on postacquisition image processing. Improving image acquisition techniques usually requires increasing the overall acquisition time, losing spatial resolution, or upgrading of the scanner. Filtering, an off-line image processing approach, is often as effective as improving image acquisition without affecting spatial resolution, and if properly designed, requires less time, and is usually less expensive.

Filtering techniques may be classified into two families: (i) enhancing, wherein wanted (structure) information is enhanced hopefully without affecting unwanted (noise) information, and (ii) suppressing, wherein unwanted information is suppressed hopefully without affecting wanted information. We focus in this paper on filtering as in (ii). Suppressing filtering operations may be further divided into two classes: a) space-invariant filtering[7], b) space-variant filtering[8].

Space-invariant filtering techniques, wherein a spatially independent fixed smoothing operation is carried out over the entire image, often blur important structures along with noise. For overcoming this problem, space-variant filtering techniques, wherein the smoothing operation is modified by local image feature, have been proposed[9-12]. Among these, diffusive filtering methods[12, 11] have become popular. In these methods, image intensity is made to diffuse to

Correspondence to Jayaram K. Udupa: e-mail: {andre, jay, anantm}@mipg.upenn.edu, telephone: (215) 662-6780, fax: (215) 898-9145. This work is supported by DHHS grants NS37172 and EB004395.

Medical Imaging 2005: Image Processing, edited by J. Michael Fitzpatrick, Joseph M. Reinhardt, Proc. of SPIE Vol. 5747 (SPIE, Bellingham, WA, 2005)
1605-7422/05/$15 · doi: 10.1117/12.595784

neighboring pixels in an iterative manner, with the diffusion conductance controlled by intensity gradients. Recently, the use of local scale[13-15] has been suggested for controlling diffusion, with the idea that scale information can exercise finer control on diffusion than plain gradients alone. In [13], local scale is considered the to be smallest scale (Gaussian kernel) at which the gradient measurement is greater than a threshold. In [14], local scale is expressed as a function which is defined as a noise-based measure of the number of zero crossings of second-order derivatives on a fixed-length line segment along a certain direction and the amount of unipolar energy in the directional derivatives. In [15], two scale-based filtering strategies are described, wherein scale at every image element is considered to be the radius of the largest ball centered at the element such that all elements within the ball satisfy a pre-defined homogeneity criterion. We will refer to this scale model as *ball scale* (or *b-scale* for short) from now on. The idea behind b-scale-based filtering is to take into account the local scale information at every spel to control the extent of filtering. The work presented in the present paper is closely related to this idea.

Our proposed filtering method is based on a novel scale idea called, *generalized ball scale* (*g_B-scale* for short)[16]. The g_B-scale at any image element p is considered to be the set of all elements (containing p) within the largest connected set of b-scale regions of a specified minimum size. g_B-scale differs from b-scale and other local morphometric scale models in that it imposes no shape, size, or anisotropic constraints. The paper is organized as follows. In Section 2, we describe our notation, briefly give an overview of anisotropic diffusion theory, and present our g_B-scale-diffusion filtering. In Section 3, qualitative and quantitative evaluations on MRI clinical and phantom images are presented by comparing our method with other diffusion approaches by using a new evaluation strategy. We state our conclusions in Section 4.

2. METHODS

For brevity, we refer to an acquired digital volume image as a *scene* and represent it by a pair $\mathcal{C} = (C, f)$ where $C = \{ c \mid -b_j \le c_j \le b_j$ for some $b \in Z_+^3 \}$, Z_+^3 is the set of 3-tuples of positive integers called *voxels*, f is a function whose domain is C, called the *scene domain*, and whose range is a set of integers [*L*, *H*], and for any $c \in C$, $f(c)$ is referred to as the *intensity* of c. We call $\mathcal{C}$ a *binary scene* if the range of f is {0,1}. A *digital ball* (or simply a *ball*) of radius r centered at any voxel c in C is the set $B_r(c) = \{d \in C \mid \|c - d\| \le r\}$. For any set X, we use the notation $|X|$ to denote its cardinality.

In the rest of this section, we will first outline the non-linear diffusion method[12], and another method based on this called non-linear complex diffusion[17], then describe the b-scale-based and g_B-scale based methods. The abbreviations used are summarized in Table 1.

Table 1. Abbreviations of terms.

Term	Abbreviation
Ball scale	b-scale
Generalized ball scale	g_B-scale
g_B-scale-based diffusion	g_BD
b-scale-based diffusion	bD
Nonlinear complex diffusion	NCD

2.1. Nonlinear anisotropic diffusion

As described in [12], anisotropic diffusion is a locally adaptive smoothing process which reduces blurring near object boundaries. A mathematical formulation of the diffusion process on a vector field $\mathbf{V}$ at a point c in coordinate-free form can be given by

$$\frac{\partial f}{\partial t} = div\,\mathbf{V} = \lim_{\Delta\tau \to 0} \int_{s} \mathbf{V} \cdot d\mathbf{s}, \tag{1}$$

where $\Delta\tau$ is the volume enclosed by the surface s (surrounding c) and $d\mathbf{s} = \mathbf{u}ds$, where $\mathbf{u}$ is a unit vector which is orthogonal and outward-directed with respect to the "infinitesimal" surface element ds. The intensity flow vector field $\mathbf{V}$ controls the diffusion process, and is defined as

$$\mathbf{V} = G\mathbf{F}, \tag{2}$$

where G is the diffusion-conductance function, and $\mathbf{F}$ is the scene intensity gradient vector field. In a linear isotropic diffusion process, G is a constant, and in [12], the authors have argued that such diffusion strategies blur object boundaries and structures. They presented an alternative anisotropic diffusion method in which G varies at each location in the scene as a nonlinear function of the magnitude of the scene intensity gradient so that smoothing within a region with low intensity gradients is encouraged, and smoothing across boundaries, wherein the magnitude of the gradients is much higher, is discouraged.

Nonlinear complex diffusion

The NCD method of [17] presents a complex diffusion-type process which generalizes the nonlinear anisotropic diffusion process by incorporating the free Schrödinger equation. They devised a complex functional form for the diffusion-conductance G, given by

$$G(c) = \frac{e^{i\theta}}{1 + \left(\frac{Im(f_t(c))}{\sigma\theta} \right)^2}, \tag{3}$$

where, for any $c \in C$, $Im(f_t(c))$ is the imaginary value of the intensity function $f_t(c)$ resulting from the diffusion process of NCD at tth iteration, σ is the control parameter that is chosen according to the noise level and edge strength, and θ is the phase angle which should be small ($\theta < 5^\circ$). Diffusion flow magnitude function $|\mathbf{V}|$ has a maximum value at magnitude of gradient $|\mathbf{F}| = \sigma$ and $|\mathbf{V}|$ is monotonically increasing for $|\mathbf{F}| < \sigma$, while it is monotonically decreasing for $|\mathbf{F}| > \sigma$. When σ is large, the generosity of filtering is high and the possibilities of blurring across boundaries increase. On the other hand, for small σ, the generosity of filtering is low and noise still remains after filtering. Because σ is fixed in this method while filtering a given scene as in [12], a fine control on and adaptivity to local region characteristics are missing. How to bring this control and adaptivity is the topic of this paper. This is achieved by using an adaptive homogeneity parameter σ_s in place of σ as a monotonically nondecreasing function of local scale.

2.2. Scale-based diffusive Filtering

Before introducing the details of scale-based diffusion process, we first present the concepts associated with the b-scale and g_B-scale notion[18, 16].

B-scale

For any ball $B_r(c)$ of radius r centered at c, we define a fraction $FO_r(c)$, that indicates the fraction of the set of the voxels in the ball boundary whose intensities are sufficiently uniform with that of c, by

$$FO_r(c) = \frac{\sum_{d \in B_r(c) - B_{r-1}(c)} W_\psi\left(\left| f(c) - f(d) \right|\right)}{\left| B_r(c) - B_{r-1}(c) \right|}, \tag{4}$$

where $W_\psi(x)$ is a fuzzy membership function corresponding to the predicate "x is small". In this paper, a zero mean unnormalized Gaussian function with standard deviation σ_ψ is used for W_ψ. σ_ψ is a homogeneity parameter that is estimated from the given scene as follows [15]. Over the entire scene domain C, local intensity differences $|f(c)-f(d)|$ are computed for all possible pairs (c,d) of voxels such that c and d are 6-adjacent. The upper 10 percentile values of these differences are discarded to account for inter object boundaries. The mean M_h and the standard deviation σ_h of these differences are computed over the lower 90 percentile values. These estimates are then used in setting up the value of σ_ψ as follows:

$$\sigma_\psi = M_h + 3\sigma_h. \tag{5}$$

The rationale for this choice is that, in a normal distribution, three standard deviations on both sides of the mean cover 99.7 percent of the population. The algorithm for *object scale estimation* (OSE) is summarized below.

Algorithm OSE
Input: C, W_ψ, a fixed threshold t_s.
Output: A scale scene $C_S = (C, f_S)$ of the scene C.
begin
for each $c \in C$ *do*
set $r = 1$;
while $FO_r(c) \geq t_s$ *do*
set r to $r+1$;
endwhile;
set $f_S(c)$ to r;
endfor;
end

The algorithm iteratively increases the ball radius r by 1, starting from $r=1$, and checks $FO_r(c)$, the fraction of the object containing c that is contained in the ball boundary. The first time when this fraction falls bellow the tolerance parameter t_s, we consider that the ball enters an object region different from that to which c belongs. Following the recommendation in [18], we have used $t_s = 0.85$. Roughly speaking, $f_S(c)$ is the radius of the largest ball centered at c within which the voxel intensities are sufficiently homogeneous.

g_B-scale

Let $C_S = (C, f_S)$ be the b-scale scene of a given scene $C = (C, f)$ and let $C_S^b = (C, f_S^b)$ be the binary scene resulting from thresholding C_S at a b-scale value $\tau_g > 0$; i.e., for any $c \in C$, $f_S^b(c) = 1$ if $f_S^b(c) \geq \tau_g$, and $f_S^b(c) = 0$, otherwise. Let, for any $c \in C$, $[c]_{\alpha_b}$ denote the α_b-connected component, that contains c, of the voxels with $f_S^b(c) = 1$ in C_S^b, where α_b is a voxel adjacency relation, say 6-adjacency. Then the g_B-scale $G_B(c)$ of c in C is defined as the subset of C such that,

$$G_B(c) = \begin{cases} \{c\}, & \text{if } f_S(c) < \tau_g \\ [c]_{\alpha_b}, & \text{if } f_S(c) \geq \tau_g. \end{cases} \tag{6}$$

The *g_B-scale set* $G_B(C)$ of C is the set $\{G_B(c) \mid c \in C\}$. The g_B-scale regions collectively cover the entire scene domain but also they do not overlap – that is, any two g_B-scale regions are either the same region in C or they are completely disjoint. This implies that, for any voxel $d \in G_B(c)$, there is no need to compute $G_B(d)$.

Algorithm g_BSE
Input: C_S, a threshold τ_g.
Output: $G_B(C)$.
begin

threshold C_S at τ_g and create a binary scene C_S^b;

track and output each α_b-connected component of the 1-valued voxels of C_S^b;

output each 0-valued voxel c as a separate g_B-scale region $\{c\}$;

end

The algorithm for computing g_B-scale essentially calls a (hard) connected component labeling algorithm to track and output each connected component of voxels with value 1 in the binary scene C_S^b resulting from thresholding C_S at τ_g. We have used in this paper 6-adjacency for α_b component in computing $G_B(C)$.

b-scale-based diffusion

As described in [15], the b-scale-based diffusion (bD) method is iterative, and so let t denote the iteration number. Let $C_t = (C, f_t)$ denote the scene resulting from the diffusion process at tth iteration. For any voxels $c, d \in C$, such that $c \neq d$ and c, d are 6-adjacent, let $\mathbf{D}(c,d)$ denote the unit vector along the direction from voxel c toward voxel d. Let $\mathbf{F}_t(c,d)$ represent the component of the intensity gradient vector along $\mathbf{D}(c,d)$, given by

$$\mathbf{F}_t(c,d) = \frac{f_t(d) - f_t(c)}{\sqrt{\sum_{i=1}^{3} \frac{v_i^2 (d_i - c_i)^2}{\min_j [v_j^2]}}} \mathbf{D}(c,d). \tag{7}$$

Note that resolution anisotropy is taken into account in (7), where $v = (v_1, v_2, v_3)$ is the voxel size.

The intensity flow from voxel c to voxel d is affected by the *effective scale* $r_{eff}(c,d)$, which is defined as

$$r_{eff}(c,d) = \min[f_s(c), f_s(d), f_s(e)], \tag{8}$$

where $e \in C$ is the neighboring voxel of c just opposite d, defined by $e = c - (d - c)$. (When c is in the border of the scene domain, we let $e = c$.) The diffusion conductance function for the flow from c to d at the tth iteration via this definition of effective scale, is defined as follows

$$G_t(c,d) = e^{-\frac{|\mathbf{F}_t(c,d)|^2}{2\{\sigma_s(c,d)\}^2}}, \tag{9}$$

where $\sigma_s(c,d)$ is the b-scale-adaptive region-homogeneity parameter used for controlling the intensity flow from d to c, given by

$$\sigma_s(c,d) = \frac{\sigma_\psi \{1 + r_{eff}(c,d)\}}{1 + r_{\mathrm{MAX}}}. \tag{10}$$

Throughout the experiments the maximum ball radius r_{MAX} in the b-scale scene C_S of C is set to 12. σ_ψ is the homogeneity parameter defined in (5). Intensity flow vector $\mathbf{V}_t(c,d)$ from voxel c to voxel d at the tth iteration is now defined by

$$\mathbf{V}_t(c,d) = -G_t(c,d)\mathbf{F}_t(c,d). \tag{11}$$

Then the iterative process is defined as follows

$$f_t(c) = \begin{cases} f(c), & \text{for } t = 0 \\ f_{t-1}(c) - K_D \sum_{d \in C} \mathbf{V}_{t\text{-}1}(c,d) \cdot \mathbf{D}(c,d), & t > 0, \end{cases} \tag{12}$$

where K_D is the integration constant that must be adjusted according to the adjacency criterion. We use $K_D = 1/7$, as indicated in [15, 11]. The flow direction between any voxels $c, d \in C$ is always such that it tries to reduce the gradient between them. Further, this diffusion process described by (7)-(12) is both nonlinear and anisotropic.

g_B-scale-based diffusion

Let $C_d = (C, f_d)$ be the distance scene corresponding to C, where for any $c \in C$, $f_d(c)$ is the distance of c from the boundary of $G_B(c)$. In g_B-scale-based diffusion, the adaptive region-homogeneity parameter $\sigma_s(c,d)$ is redefined by using g_B-scale morphometric information, as follows.

$$\sigma_s = \begin{cases} \dfrac{\sigma_\psi\{1 + d_{eff}(c,d)\}}{1 + d_{\text{MAX}}}, & \text{if } |G_B(c)| > 1 \\ \dfrac{\sigma_\psi\{1 + r_{eff}(c,d)\}}{1 + r_{\text{MAX}}}, & \text{otherwise,} \end{cases} \tag{13}$$

where $d_{eff}(c,d)$ is the *effective distance* give by

$$d_{eff}(c,d) = \min[\, f_d(c),\, f_d(d),\, f_d(e)\,]. \tag{14}$$

d_{MAX} is the maximum distance of any voxel in $G_B(c)$ from the boundary of $G_B(c)$, $e \in C$ is the neighboring voxel of c just opposite d. Now the process of g_B-scale-based diffusion is completely described by (7)-(9), (11)-(14). Intuitively, in this process, diffusion at a voxel is finely controlled by its distance from the boundary of its g_B-scale region. The diffusion rate is generally higher in larger regions than in smaller regions. Further, within a given g_B-scale region, the diffusion rate is greater in the center of the region than in its border. We note that, in this process, diffusion takes place along boundaries in a uniform manner, no matter what their shape is, but not across them.

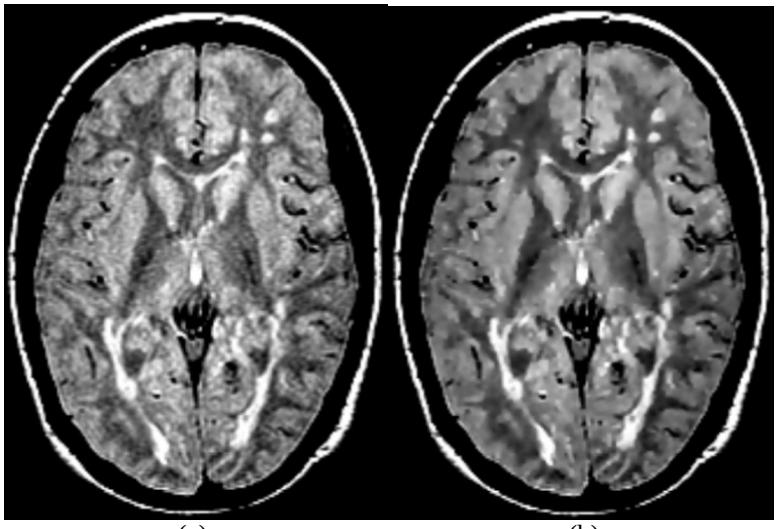
3. RESULTS

In this section, we present the results of qualitative and quantitative experimental evaluation by comparing the three methods (bD, NCD, and g_BD). For the NCD method, we have used $\sigma = \sigma_\psi$ and $\theta = \pi/60$ in (3). Qualitative evaluation is shown on a proton density (PD)-weighted 3D MRI scene of the head of a multiple sclerosis patient. A quantitative evaluation is carried out by utilizing 45 brain phantom 3D scene data sets obtained from the Montreal Neurological Institute Brainweb*. The database consists of scenes with three levels of noise (3%, 7%, and 9%), three protocols (PD, T1, T2), and five slice thickness values (1mm, 3 mm, 5 mm, 7 mm, and 9 mm).

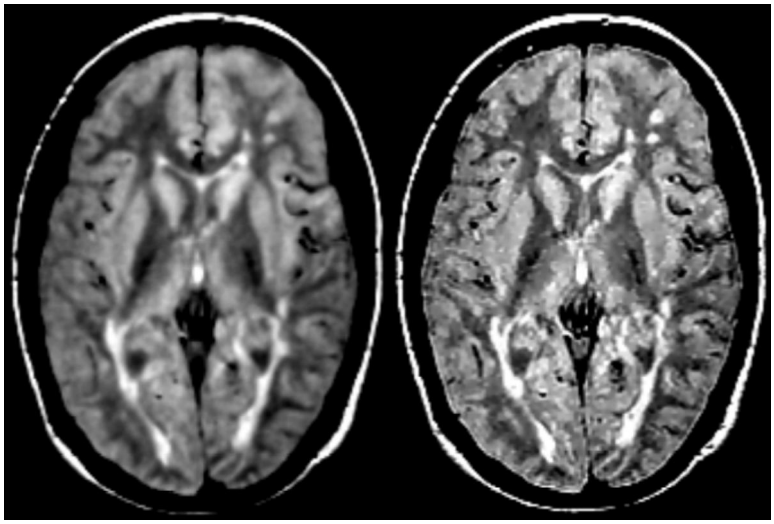
* http://www.bic.mni.mcgill.ca/brainweb

3.1 Qualitative Evaluation

In Figures 1 and 2, we show 3D comparative results among the three methods on an axial PD scene of the head of a multiple sclerosis (MS) patient. The size of the scene domain is $256 \times 256 \times 45$ with a voxel size $v(0.86, 0.86, 3.0)$ mm . Note that the out-of-plane resolution is almost 3.5 times coarser than the in-plane resolution. All processing has been done in the three-dimensional space. Figure 1 displays the results in one slice after applying the methods for three iterations. In Figure 2, the superior portion of the brain from Figure 1 is zoomed up to allow a closer scrutiny of the images. The bD and g_BD methods preserve fine details and edges during the diffusion process better than NCD method. Further, some MS lesions (hyper intense blobs) appear blurred in the filtered scene resulting from NCD.



(a) (b)



(c) (d)

Figure 1. Comparison among the three methods on a 3D MRI scene of the head of a multiple sclerosis patient. (a) A slice of the original scene, (b) after applying the bD method, (c) after applying the NCD method, and (d) after applying g_BD method.

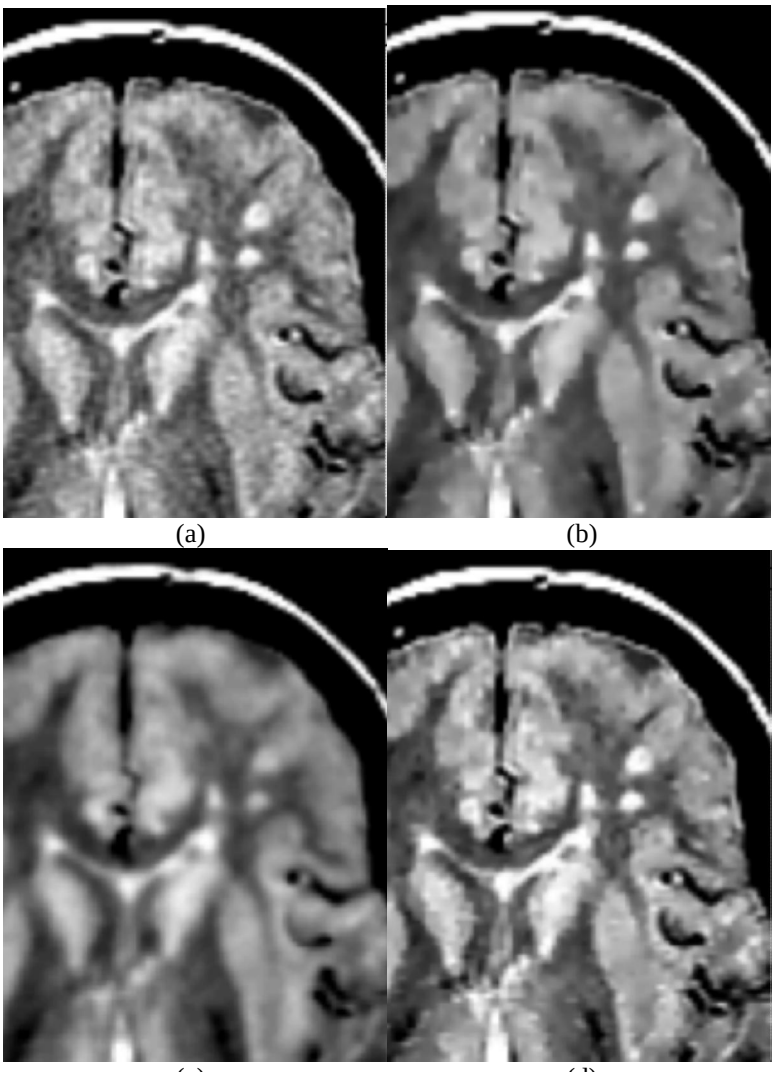


Figure 2. Zooming in the superior portion of the brain from Figure 1. (a) A slice of the original scene, (b) after applying the bD method, (c) after applying the NCD method, and (d) after applying g_BD method.

3.2 Quantitative Evaluation

For quantitative comparison among the methods, we carried out our experiments on 45 MNI phantom scenes corresponding to (i) three levels of noise (3%, 7%, and 9%), (ii) three protocols (PD, T1, and T2) and (iii) five slice thicknesses (1mm, 3 mm, 5 mm, 7 mm, and 9 mm). Let $\{C_{xi} \mid C_{xi} = (C, f_{xi}), 1 \le i \le 45\}$ denote the set of scenes produced after applying the filtering method $x \in$ {bD, NCD, g_BD} to the set of 45 phantom scenes. For the white matter (WM) regions, we compute *residual noise* RN_{xi}^{WM} as the standard deviation of voxel intensities $f_{xi}(c)$ in C_{xi} within the WM regions in a filtered scene resulting from method x for the phantom scene i, and *relative contrast* RC_{xi}^{WM} over the regions in the vicinity of the WM boundary regions, as follows

$$RC_{xi}^{\mathrm{WM}} = \frac{|M_{xi}^{\mathrm{OWM}} - M_{xi}^{\mathrm{BWM}}|}{\sqrt{\sigma_{\mathrm{xi}}^{\mathrm{OWM}} \sigma_{\mathrm{xi}}^{\mathrm{BWM}}}}, \tag{15}$$

where M_{xi}^{OWM} and $\sigma_{xi}^{\mathrm{OWM}}$ denote the mean and standard deviation, respectively, of voxels intensities $f_{xi}(c)$ in C_{xi} within WM region, and M_{xi}^{BWM} and $\sigma_{xi}^{\mathrm{BWM}}$ denote similar entities outside WM region in their immediate vicinity. In Figure 3, we demonstrate the behavior of each method by plotting how the RN and RC values vary as the number of iterations from 1 to 200 for the different protocols at three levels of noise averaged by five slice thickness values. The values of RN_{xi}^{WM} and RC_{xi}^{WM} were normalized by their maximum in each method. Table 2 shows the *area under the curve* (AUC) values for each curve. We note that, in assessing the performance of suppressing filter, the trade-off that exists between RN and RC should be analyzed. This is akin to the trade-off issue between false positives and false negatives arising in image segmentation and object detection tasks. In analogy with ROC curves for the latter task, we call the curves depicting the trade-off behavior between RN and RC *filter operating characteristic* (FOC) curves. It is usually easy to reduce RN simply by increasing the number of iterations or by changing the parameter that controls the conductance function. However, it is challenging to reduce RN and at the same time increase RC. The FOC curve depicts how well this trade-off is handled by a filtering method. A higher value of AUC for the method indicates more accurate and effective filtering. From Table 2, we observe that, for every protocol and each level of noise, AUC for scale-based diffusive filtering (bD and g_BD) is higher than that for the NCD method and that the g_BD method always outperforms the bD method.

Table 2. AUC values for the curves shown in Figure 3

Protocol	Noise level	bD	NCD	g_BD
T1	3%	0.8713	0.7289	0.9154
	7%	0.8596	0.7509	0.9061
	9%	0.8636	0.7551	0.9025
T2	3%	0.8100	0.6885	0.8837
	7%	0.8187	0.7811	0.8688
	9%	0.8335	0.7918	0.9311
PD	3%	0.7875	0.6908	0.9072
	7%	0.8166	0.7771	0.8601
	9%	0.8716	0.8398	0.8969

4. DISCUSSION AND CONCLUSION

In this paper, we have presented a new scale-based diffusive filtering (g_BD method) wherein a fine control on and adaptivity to local region characteristics is incorporated by using a local morphometric scale model, called g_B-scale. Further, unlike other local morphometric scale models, g_B-scale imposes no shape, size or anisotropic constraints. We demonstrate qualitatively and quantitatively that the g_BD method outperforms both bD and NCD methods. The filtered scenes produced by the g_BD method have fine details and structures better preserved while suppressing noise. We have introduced a new strategy for the evaluation of filtering methods by utilizing FOC curves and the area under the FOC curve. We believe that by studying the dynamics of the filtering method via FOC, in the same manner as described in [19] via delineation operating characteristics (DOC), one can devise techniques for the optimal selection of parameters and stopping criteria. On a Pentium III (450 MHz and 256 MB RAM) PC, the g_BD method performs as quickly as the bD method, on average, under 1 minute for three iterations for a $256\times256\times45$ scene.

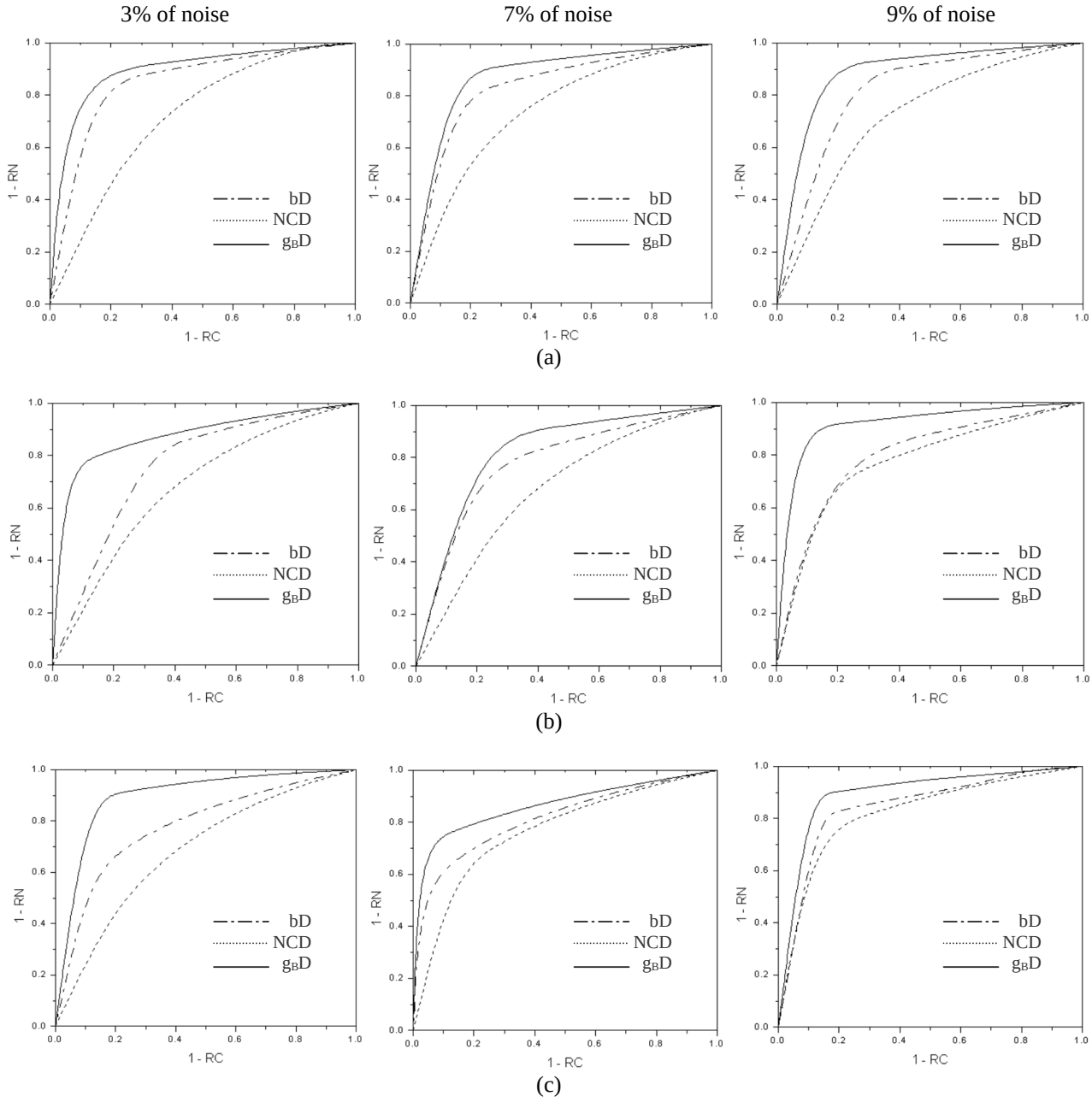


Figure 3. Curves for 45 MR phantom scenes depicting the variation of RN and RC as the number of iterations is changed from 1 to 200 for (a) T1, (b) T2 , and (c) PD protocols.

REFERENCES

1. P. K. Saha and J. K. Udupa, Optimum image thresholding via class uncertainty and region homogeneity, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 23, pp. 689–706, 2001.

2. N. R. Pal and S. K. Pal, A review of image segmentation techniques, *Pattern Recognition*, vol. 26, no. 9, pp. 1277-1294, 1993.

3. W. M. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, Multi-modal image registration by maximization of mutual information, *Med. Image Anal.*, vol. 1, pp. 35-51, 1996.

4. H. Lester and S. R. Arridge, A survey of hierarchical nonlinear medical image registration, *Pattern Recognition*, vol. 32, pp. 129-149, 1999.

5. J. K. Udupa and G. Herman, Eds., *3D Imaging in Medicine*. Boca Raton, FL: CRC Press, 1991.

6. K. Höhne, H. Fuchs, and S. Pizer, Eds., *3D Imaging in Medicine: Algorithms, Systems, Applications*, Berlin, Germany: Springer-Verlag, 1990.

7. A. Rosenfeld and A. C. Kak, *Digital Picture Processing*. Orlando, FL: Academic, 1982.

8. J. S. Lee, Digital image enhancement and noise filtering by use of local statistics, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-2 , pp. 165-168, 1980.

9. C. B. Ahn, Y. C. Song, and D. J. Park, Adaptive template for signal to noise ratio enhancement in magnetic resonance imaging, *IEEE Trans Med Imaging*, vol. 18, pp. 549-556, 1999.

10. J. Weikert *et al*., Theoretical foundations of anisotropic diffusion in image processing, *Theoretical Foundations of Computer Vision. Computing*, pp. 221-236, suppl. 11, 1996.

11. G. Gerig, O. Kübler, R. Kikinis, and F. A. Jolesz, Nonlinear anisotropic filtering of MRI data, *IEEE Trans Med Imaging*, vol. 11, pp. 221-232, 1992.

12. P. Perona and J. Malik, scale-space and edge detection using anisotropic diffusion, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12 , pp. 629-639, 1990.

13. J. H. Elder and S. W. Zucker, Local scale control for edge detection and blur estimation, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 20 , pp. 699-716, 1998.

14. P. Liang and Y. F. Wang, Local scale controlled anisotropic diffusion with local noise estimate for image smoothing and edge detection, *in Proc. Int. Conf. Computer Vision*, Bombay, India, pp. 193-200, 1998.

15. P. K. Saha and J. K. Udupa, Scale-based diffusive image filtering preserving boundary sharpness and fine structures, *IEEE Trans Med Imaging*, vol. 20, pp. 221-232, 2001.

16. A. Madabhushi, J. K. Udupa, and A. Souza, Generalized ball-scale: theory, algorithms, and application to image inhomogeneity correction, *Proc. of SPIE: Medical Imaging 2005* (to appear).

17. G. Gilboa, N. Sochen, and Y. Zeevi, Image enhancement and denoising by complex diffusion processes, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 26 , pp. 1020-1036, 2004.

18. P. K. Saha, J. K. Udupa, and D. Odhner, Scale-based fuzzy connected image segmentation: theory, algorithms, and validation, *CVIU*, vol. 77, pp. 145-174, 2000.

19. J. K. Udupa and Y. Zhuge, Delineation operating characteristic (DOC) curve for assessing the accuracy behavior of segmentation algorithms, *Proc. of SPIE: Medical Imaging*, vol. 5370, pp. 640-647, 2004