

Anisotropic Diffusion Filtering Operation and Limitations - Magnetic Resonance Imaging Evaluation

Caio A. Palma, Fábio A.M. Cappabianco, Jaime S. Ide*
Paulo A.V. Miranda**

* *Group of Biomedical Imaging and Signal Research, Institute of
Science and Technology, Universidade Federal de São Paulo, São José
dos Campos, Brazil. cappabianco@unifesp.br*

** *Department of Computer Science, University of São Paulo São
Paulo, Brazil. pmiranda@vision.ime.usp.br*

Abstract: High-frequency noise is present in magnetic resonance images and it is usually removed by a filtering process. The anisotropic diffusion filter (ADF) was proposed to adaptively remove the noise, maintaining the image edges. However, as quantified in this paper for the first time, ADF methods still produce unsatisfactory results. While previous ADF implementations used the gradient of the strongest edges or the standard deviation of a planar region (i.e. without any edges), it is proposed a novel approach with improved parameter estimation based on both edge and planar region, overcoming some of ADF important limitations. Our framework was validated in more than thirty magnetic resonance images leading to more consistent results.

Keywords: Medical imaging and processing; Biomedical and medical image processing and systems

1. INTRODUCTION

Magnetic resonance (MR) images are exposed to Johnson-Nyquist or thermal noise (Johnson [1928]), and different approaches were proposed to adaptively remove noise and preserve edges between objects with distinct intensities. The anisotropic diffusion filter (ADF) (Perona and Malik [1990]) is a popular iterative methodology used to restore MR image intensities (Gerig et al. [1992]). By means of an edge-stopping function (ESF), this filter is capable of smoothing noisy pixels, while preserving the intensity of most edges. Nevertheless, in order to achieve accurate results, the ESF and other parameters of the ADF must be set according to the image nature, the scanned subject, the protocol of acquisition, and the signal-to-noise ratio (SNR), among other factors.

The proposed ADF approaches for SNR automatic estimation are not precise (Kim et al. [2005]). That is because they are based on either the image strongest edges (Perona and Malik [1990], Black et al. [1998], Voci et al. [2004]) or on the standard deviation of a planar region (i.e. an image region without edges) (Tsotsios and Petrou [2013]), making them vulnerable to image under- and over-smoothing.

We present a quantitative analysis describing ADF limitations and a novel framework based on both the strongest edges and on planar regions of the image, in order to set ADF parameters optimally. The evaluation comprises magnetic resonance (MR) images with different acquisition protocols.

* The authors thanks to Fapesp (Jovem Pesquisador 11/08573-4), CAPES, and CNPQ (486988/2013-9) for the financial support.

The remainder of this paper is organized as follows: Section 2 introduces the ADF method, and its extensions; Section 3 analyzes the ADF limitations; Section 4 contains the proposed framework for ADF; experiments and results follow in Section 5; and in Section 6 we state the conclusions.

2. METHODS

Anisotropic diffusion filter (Perona and Malik [1990]) has been successfully employed in the context of image processing to remove high frequency noise while conserving the main edges of existing objects. It was used in the context of MR imaging (Gerig et al. [1992]), and automated in several different ways (Black et al. [1998], Weickert [1998], Voci et al. [2004], Krissian and Aja-Fernandez [2009]).

In its discrete form, ADF is an iterative algorithm that simulates the diffusion process by:

$$I_s^{t+1} \approx I_s^t + \frac{\lambda}{|\eta_s|} \sum_{p \in \eta_s} g(|\nabla I_{s,p}^t|, \gamma) \nabla I_{s,p}^t, \quad (1)$$

where I_s^t is the intensity of a pixel s from image I at instant t , λ is a scalar related to the diffusion rate, γ is a positive constant selected according to the desired smoothing level, η_s stands for the set of adjacent pixels of s , $g(\cdot)$ is an ESF, and $\nabla I_{s,p}^t$ is the magnitude of the image directional gradient from pixel s to p at instant t . The directional gradient $\nabla I_{s,p}^t$ can be approximated by $I_p^t - I_s^t$. To simplify the notation, we will replace $\nabla I_{s,p}^t$ with x whenever pixel information and the iteration number are irrelevant to the context.

In order to achieve a monotonic variation of the intensities at each iteration in the context of 3D images, Gerig et al. [1992] proved that the constant $\lambda/|\eta_s|$ should not exceed $1/7$ if an adjacency composed by the 6-nearest neighbors is used, that is, $\lambda \leq 6/7$. For an adjacency composed by the 26-nearest neighbors we should set $\lambda/|\eta_s| \leq 3/47$, being ¹ $\lambda \leq 78/47$.

Noise pixels are filtered faster than edge pixels for two main reasons. First, opposite direction gradients attenuate or cancel their effects, according to Equation 1. That usually happens to edge pixels, but not to noise pixels. Second, ADF has a cumulative effect by the number of neighbors with high gradient magnitude with the same direction. As noise pixels are randomly spread through the image, mostly their intensity are distinct from their surrounding neighbors. Edge pixels, on the contrary, have a similar intensity to a considerable part of their adjacency.

2.1 Parameter Selection

The ADF requires an edge-stopping function and three parameters: the smoothing constant γ , the number of iterations t_{max} , and the size of the adjacency relation η_s . The size of the adjacency relation reflects a trade-off between precision and computational effort. The other two parameters determine the results in terms of image under- and over-smoothing. In this paper, we are more concerned about the smoothing degree the ADF and hence, we will focus on t_{max} and γ .

With respect to γ , Perona and Malik [1990] suggested to use a noise estimator based on Canny [1986], i.e., 90% of the accumulated histogram of the absolute values of the gradient. Black et al. [1998] proposed the use of the median absolute deviation of the image gradient since it is more robust to the influence of outliers. After that, Voci et al. [2004] reinforced the importance of the idea of estimating γ at each iteration. This strategy reduces the effect of ADF over edges with higher gradient intensity. More recently, Tsiotsios and Petrou [2013] proposed a stopping criteria for the number of iterations based on the contrast of the edges at each iteration, and a novel gradient threshold estimation for parameter γ utilizing the most uniform block of pixels in the image.

All the proposed methods compute γ and t_{max} parameters based on either the gradient of the image, or on the standard deviation of a non-edge region requiring empirical tuning (Kim et al. [2005]).

2.2 Edge-Stopping Function

The edge-stopping function can assume several forms. At first, Perona and Malik [1990] initially proposed the functions:

$$g(x, \gamma) = \exp(-x^2/2\gamma^2), \quad (2)$$

$$g(x, \gamma) = \left[1 + (x/\gamma)^2\right]^{-1}, \quad (3)$$

Black et al. [1998] proposed a nice improvement using Tukeys biweight ESF in Equation 4.

¹ The maximum integration constant Δt for 26-nearest neighbors in the Appendix of Gerig et al. [1992] is incorrect.

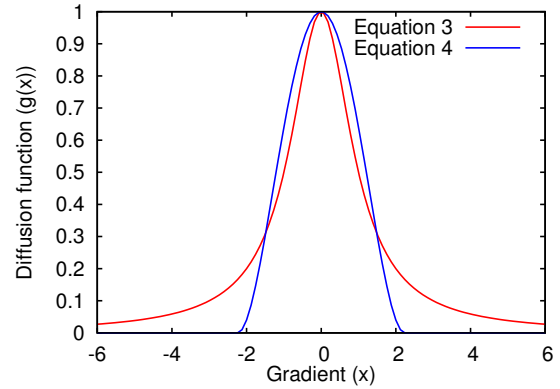


Fig. 1. Curves from Equations 3 and 4 for $\gamma = 1$. We can note that curve from Equation 3 slowly approaching zero, while the curve from Equation 4 presents the desirable cutting behavior.

$$g(x, \gamma) = \begin{cases} [1 - (x^2/5\gamma^2)]^2 & |x| \leq \gamma\sqrt{5}, \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

After a number of iterations of the diffusion filter using the ESF of Equation 4, the filter will stop changing pixel's intensities. Using Equation 3, on the other hand, causes the filter to smooth the image indefinitely. Even though both functions start rejecting gradient intensities greater than a given threshold, the function described by Equation 4 reaches zero for $|x| > \gamma$, while the function from Equation 3 just approaches zero as showed in Figure 1.

3. ANISOTROPIC DIFFUSION FILTER LIMITATIONS

Our first concern is to determine to what extent ADF is capable of removing noise while preserving image edges. To answer this question, we use a MR synthetic image dataset called BrainWeb Phantom² because it contains images with variable noise levels.

Is ADF able to remove all noise, preserving the strongest edges? Figure 2 contains a synthetic input image from BrainWeb Phantoms dataset (BWP) with an induced noise level of 3% and 5% of the maximum image intensity, and the optimal results of ADF after 200 iterations, fixing γ manually at 225 and 250, respectively. For an induced noise level of 3%, the edges were almost completely preserved and noise was almost totally removed. Nevertheless, for an induced noise level of 5%, some edges are blurred while several noise pixels still remain. The conclusion is that for images with noise intensity equal to or higher than 3% of the maximum image intensity, it is not feasible to preserve the edges while completely removing the noise using the ADF.

An effective strategy to avoid smoothing strong edges is to estimate γ at each ADF iteration. Figure 3 shows an axial slice of a synthetic image with noise level of 5%, the result of five iterations of ADF fixing γ at 1143.0, and the result of five iterations of ADF with decreasing γ at 1143.0, 372.0, 170.0, 97.0, and 62.0. Clearly, the edges are more smoothed using fixed γ value.

² <http://brainweb.bic.mni.mcgill.ca/brainweb/>

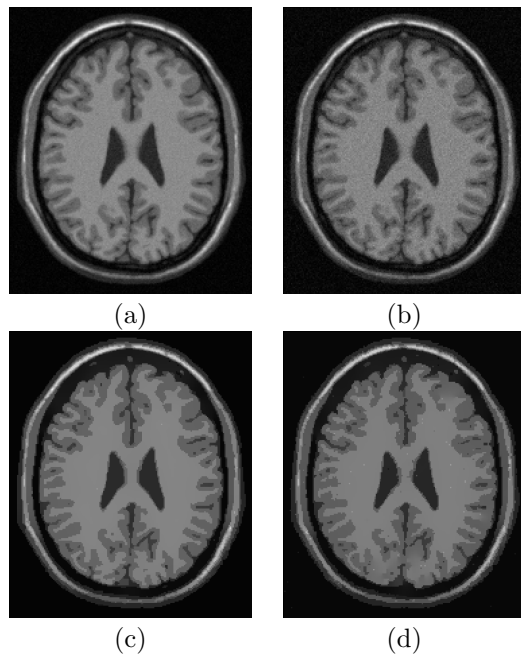


Fig. 2. Axial slice of input synthetic MR images of the human brain exposed to (a) 3% and (b) 5% of noise level; Best filtering results for (c) 3% and (d) 5% of noise level, respectively, after 200 iterations with fixed γ parameter manually set.

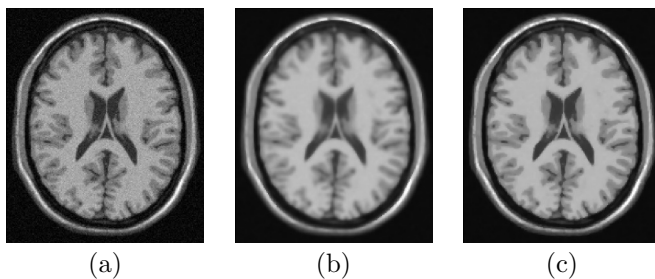


Fig. 3. (a) Axial slice of synthetic MR image of the brain with noise intensity of 5% of the maximum intensity; (b) image filtered with γ fixed at 1143.0 and 5 iterations; (c) image filtered in 5 iterations with decreasing γ .

We have a different scenario for weaker gradient edges. They will be smoothed even applying small γ values. Figure 4 shows how weak borders in the cerebellum and sub-cortical region of the brain are progressively blurred by employing 5 to 400 iterations of ADF with a γ value of 150. As the number of iterations consists in a trade-off between removing the noise and preserving the weak edges and the surrounding structures, it is desirable to execute a low number of ADF iterations (i.e. less than 10) to preserve even the weaker borders of the image.

There is a delicate situation that occurs when the image SNR is very low (i.e. noise level higher than 3%). As it is impossible to remove all noise and preserve even the stronger edges, one must choose between keeping sharper edges unchanged, or removing higher intensity noise. The choice between these two approaches can be made solely based on the initial value of γ parameter. Figure 5 shows

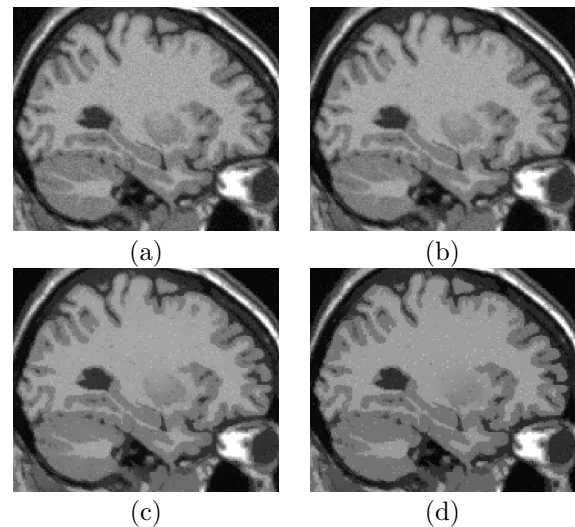


Fig. 4. (a) Sagittal slice of synthetic MR image of the brain with noise intensity of 5% of the maximum intensity, and the resultant images after applying ADF with (b) 5, (c) 50, and (d) 200 iterations, fixing γ at 150.

that both alternatives generate good results, depending on the post-processing objectives.

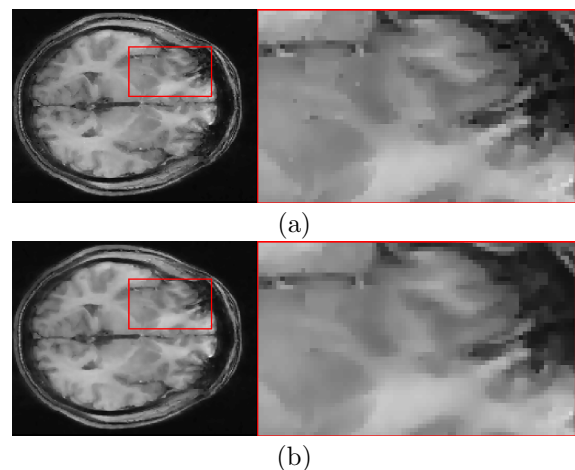


Fig. 5. Image filtered by ADF with (a) conservative and (b) aggressive selection of γ parameter value.

In short, the presented analysis elucidates that ADF should be executed using a low number of iterations, and employing a decreasing γ value. If SNR is too low, one must choose to preserve sharper edges or to remove stronger noise pixels, based on the filtering purposes. These conclusions motivated a novel ADF framework described in the next section.

4. OPERATIONAL ANISOTROPIC DIFFUSION FILTER

We present a novel ADF framework with improved γ parameter estimation, based on both a planar region \mathcal{F} and on a set of the strongest edges of the image, denominated edge region \mathcal{E} . We also propose a novel procedure to delineate the planar region that is robust to partial volume effects (González Ballester et al. [2002]). Edge detection is executed using a Canny proposed methodology in Canny

[1986] with hysteresis of 80% and 90% of the accumulated histogram of the absolute values of the gradient.

The optimal γ parameter is computed according to the desirable conservative or aggressive behavior, and updated at each iteration. The number of iterations t_{max} is also defined dynamically according to the current SNR, as the image is smoothed.

4.1 Planar Region Delineation

We will state here that a pixel belongs to an edge if its intensity is composed by partial volume belonging to more than one object of interest, such as tissues, organs, or the background. We will denominate a set of connected pixels that does not contain any edges as a planar region.

Even though partial volume and planar region delineation are difficult and extensive research topics (Chiverton and Wells [2008]), many proposed high-frequency noise filtering approaches tend to minimize the problem, treating it as a trivial step in their pipeline (Dietrich et al. [2008]). Due to the presence of artifacts or low SNR, it is complicated to estimate a default regular volume, such as a cube or sphere, without the intervention of a specialist. A large regular volume may not even exist depending on the scanned body part, and the acquisition process. In a small volume, for instance, the presence of partial volume pixels in its boundary could strongly influence the SNR estimation. Hence, a general planar region delineation may not be precise even when applying more complex algorithms such as in Rajan et al. [2010], Tsiotsios and Petrou [2013].

We propose the following methodology to delineate a general planar region, robust to partial volume pixels. First, the input image (Figure 6(a)) is completely smoothed, applying to it a median filter with an adaptive radius that takes into consideration the image dimensions and slice thickness (Figure 6(b)). Note that this initial filtering is only used for clustering purposes and disposed. The filter should be strong enough to remove the noise present in the biomedical image of the selected modality. Then, a clustering algorithm is used to separate the image in homogeneous regions. In this paper, we use the Optimum-Path Forest clustering algorithm (Rocha et al. [2008], Cappabianco et al. [2012]). As the median filter blurred image edges, we choose one of the clustered regions (e.g. the largest) and morphologically erode it with a radius larger than the one used by the median filter (Figure 6(c)). We expect that the erosion eliminates all partial volume pixels in the boundary of the cluster. Finally, if the erosion causes disconnection of original cluster, the largest connected component is taken. The standard deviation of the planar region computed for the image illustrated in Figure 6, is reduced from 94.20 to 50.52 after the erosion, a difference of 86%.

4.2 Initial Gamma(γ) Parameter

We start computing $\gamma_{\mathcal{F}}$, the minimum γ value to smooth the majority of the noise pixels in \mathcal{F} , and $\gamma_{\mathcal{E}}$, the maximum γ value that does not smooth most of the edges in \mathcal{E} . Figure 7 contains plots of the standard deviation of \mathcal{F} and \mathcal{E} according to the employed γ parameter for a synthetic MR image with a noise level of 3%. The optimum $\gamma_{\mathcal{F}}$ is

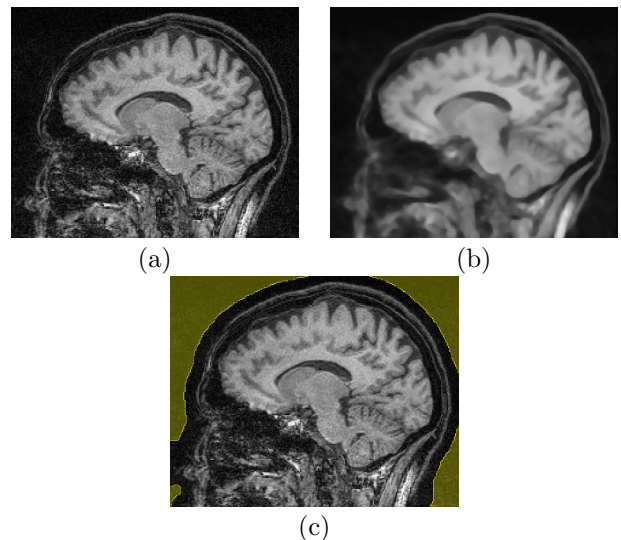


Fig. 6. Planar region delineation methodology illustration on a sagittal slice of a 1.5 Tesla T1-weighted MR image of the brain: (a) the original input image; (b) image over smoothed by median filter; (c) the input image covered by the resultant planar region in yellow.

327.12 and the optimum $\gamma_{\mathcal{E}}$ is 367.29. As happened in this case, whenever $\gamma_{\mathcal{E}} \geq \gamma_{\mathcal{F}}$, any γ within this interval will keep stronger edges and virtually remove all noise.

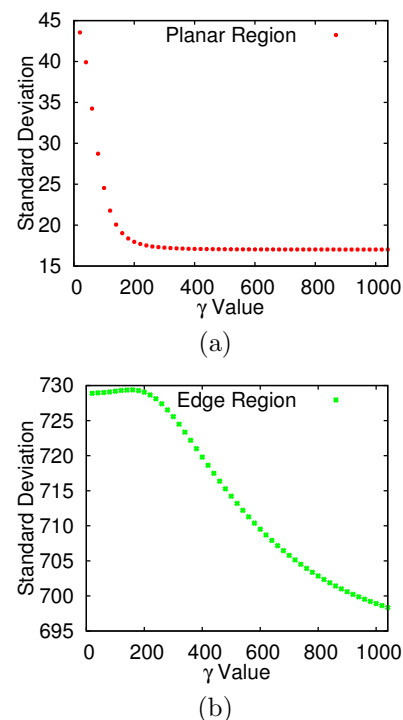


Fig. 7. Plots of the standard deviation of regions (a) \mathcal{F} and (b) \mathcal{E} after the first iteration of ADF versus the employed γ parameter in a synthetic MR image with a noise level 3% of the maximum image intensity.

Figure 8 contains the same plots for a synthetic MR images with noise level of 9%. In this case, the optimum $\gamma_{\mathcal{F}}$ and $\gamma_{\mathcal{E}}$ values are 775.22 and 248.07, respectively. With $\gamma_{\mathcal{E}} < \gamma_{\mathcal{F}}$, one must employ a conservative or aggressive approach due

to the ADF limitations. Therefore, given a conservativeness parameter $c \in [0, 1]$, γ is set by Equation 5 at each iteration:

$$\gamma = \begin{cases} \gamma_{\mathcal{F}}, & \gamma_{\mathcal{F}} \leq \gamma_{\mathcal{E}}, \\ c(\gamma_{\mathcal{F}} - \gamma_{\mathcal{E}}), & \text{otherwise.} \end{cases} \quad (5)$$

The closer we set c to 0, the more conservative the ADF will be.

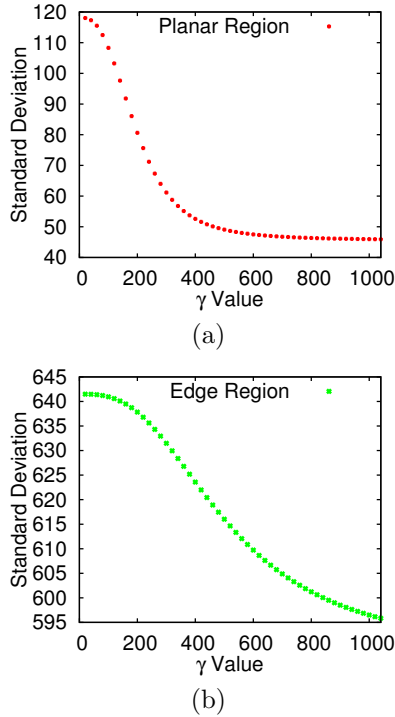


Fig. 8. Plots of the standard deviation of regions (a) \mathcal{F} and (b) \mathcal{E} after the first iteration of ADF versus the employed γ parameter in a synthetic MR images with noise level 9% of the maximum image intensity.

4.3 Iterative Gamma(γ) Updating

The initial γ will be utilized at the first ADF iteration. It would be tempting to utilize the same strategy for the next ADF iterations. Nevertheless, this is a very time consuming procedure. As presented in Section 5, it takes more ADF iterations to compute each of $\gamma_{\mathcal{E}}$ and $\gamma_{\mathcal{F}}$ parameters than to actually filter the image. An alternative is to compute the expected γ reduction based on the highest gradient noise pixels affected in the current iteration.

The maximum gradient value of Equation 1 using Equations 2 to 4 occurs for $x = \pm\gamma$. As mentioned in Section 3, edge pixels probably have more adjacent pixels with similar intensity than noise pixels. We can reasonably state that most of edge pixels have more than 1/4 of its neighbors with similar intensity, and that most of the noise pixels have less than 1/4. Then, we have a gradient reduction of:

$$0.25e^{-0.5}\gamma\lambda, \text{ using ESF of Equation 2,} \quad (6)$$

$$0.125\gamma\lambda, \text{ using ESF of Equation 3,} \quad (7)$$

$$0.16\gamma\lambda, \text{ using ESF of Equation 4.} \quad (8)$$

As $x = \pm\gamma$, these reductions in the gradient are the same reduction employed to γ value of the next iteration. Empirical experiments showed that the estimated γ reduction is very similar to the γ computed from $\gamma_{\mathcal{E}}$ and $\gamma_{\mathcal{F}}$ at each iteration with $c = 1.0$.

4.4 Stopping Criteria

We cannot rely on \mathcal{F} as an stopping criteria, since a more conservative approach will not eliminate all the noise. A very simple solution is to use a metric based on the current γ and the initial $\gamma_{\mathcal{E}}$ values. Despite of the ESF, the proportion $\gamma_{\mathcal{F}}/\gamma_{\mathcal{E}}$ for a synthetic image with no noise is close to 1/7. That means that the image has almost no noise when the current γ is close to $\gamma_{\mathcal{E}}/7$.

In short, we compute the initial \mathcal{F} and \mathcal{E} regions and their respective optimum γ values, $\gamma_{\mathcal{F}}$ and $\gamma_{\mathcal{E}}$, respectively. Then, at each ADF iteration we compute γ by Equation 5 and continue to apply the ADF until $\gamma \leq \gamma_{\mathcal{E}}/7$.

5. EXPERIMENTS

We used two different datasets in our experiments: BWP with noise levels 0%, 1%, 3%, 5%, 7%, and 9% of the maximum image intensity; and a real brain MR image dataset (RBI) containing 19 1.5 Tesla and 10 3.0 Tesla images.

In the worse case when very poor SNRs was present, the search for initial $\gamma_{\mathcal{E}}$ and $\gamma_{\mathcal{F}}$ values took 9 ADF iterations each, and 7 more iterations were required for the actual filtering process, totalling 25 iterations.

All ESF produced satisfactory results under the proposed framework. Nevertheless, we lay emphasis here on the best results obtained while employing Equation 4 due to its faster descent behavior after achieving the maximum intensity.

Images from BWP dataset required more or less ADF iterations based on the noise level. The 9% noise level image is ideal to show the behavior of the ADF with respect to the conservativeness parameter c . Figure 9 shows the filtering results, with several different c values. We can note that for smaller c values, the edges are sharper, but the strongest noise pixels remain.

RBI dataset had distinct results for 1.5 Tesla and 3.0 Tesla images. 3.0 Tesla required just one ADF iteration, while 1.5 Tesla could not be filtered without affecting the image edges. Figure 10 shows the result for a 1.5 Tesla image, with several different c values. We can state the same conclusions pointed by BWP dataset, showing that the proposed framework is appropriated for real MR images of any acquisition protocol.

Comparing to other ADF extensions, the proposed methodology with $c = 1.0$ will generate a very similar result to the work in Tsitsios and Petrou [2013]. The result of employing the methodology in Black et al. [1998] is more blurred than the proposed methodology with $c = 1.0$ (see Figure 11). None of the previous works was able to generate conservative results as proposed here.

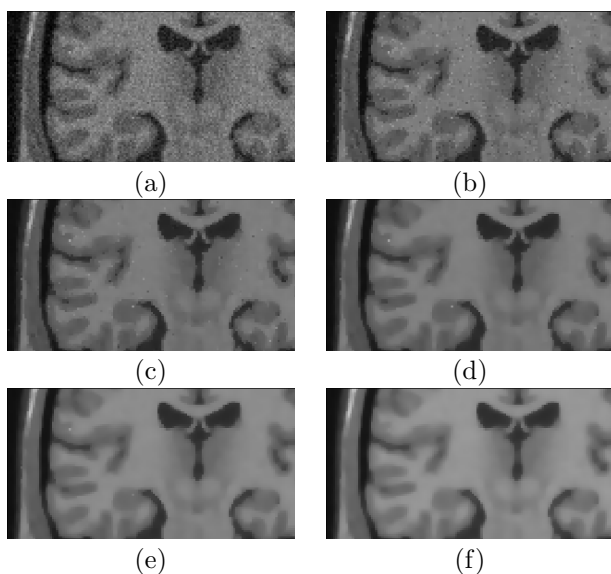


Fig. 9. Axial slice of (a) synthetic MR input image with noise level of 9% of the maximum image intensity and the filtered image by the proposed framework with (b) $c=0.0$, (c) $c=0.25$, (d) $c=0.5$, (e) $c=0.75$, and (f) $c=1.0$.

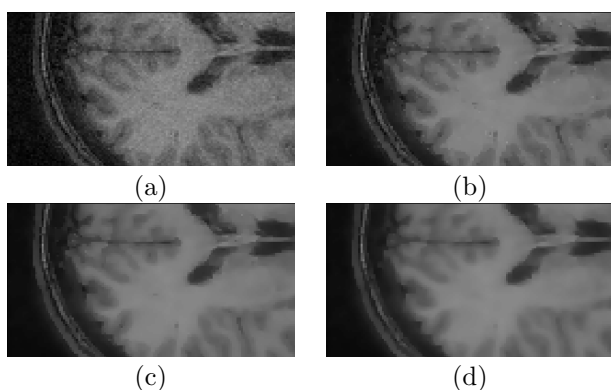


Fig. 10. Axial slice of (a) real MR input image and the filtered image by the proposed framework with (b) $c = 0.0$, (c) $c = 0.5$, and (d) $c = 1.0$.

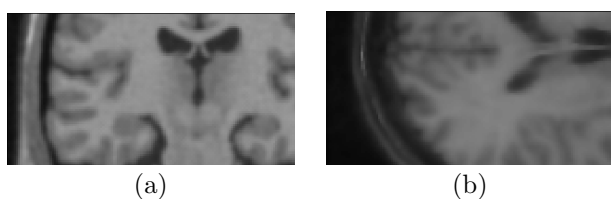


Fig. 11. Results of the methodology in Black et al. [1998] for (a) the synthetic MR image, and (b) the real MR image.

6. CONCLUSION

This paper presented a novel analysis about ADF operation and limitations. It became clear that the ADF significantly blurs the edges while removing an induced noise level greater than 3% of the maximum image intensity. To deal with this issue, a robust framework was proposed, that adjusts all ADF parameters based on statistics derived from the significant edges and from a planar

region, allowing a conservative or an aggressive behavior. The methodology was successfully validated using real and synthetic MR images with different noise levels.

Future works comprises local optimal γ estimation based on local \mathcal{E} and \mathcal{F} regions, and a deeper analysis of the ESF behavior over the proposed methodology.

REFERENCES

- M.J. Black, G. Sapiro, D.H. Marimont, and D. Heeger. Robust anisotropic diffusion. *Image Processing, IEEE Transactions on*, 7(3):421–432, 1998.
- J. Canny. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 1(6):679–698, 1986.
- F.A.M. Cappabianco, A.X. Falcão, C.L. Yasuda, and J.K. Udupa. Brain tissue MR-image segmentation via optimum-path forest clustering. *Computer Vision and Image Understanding*, 116(10):1047–1059, 2012.
- J.P. Chiverton and K. Wells. Adaptive partial volume classification of MRI data. *Physics in Medicine and Biology*, 53(20):5577–5594, 2008.
- O. Dietrich, J.G. Raya, S.B. Reeder, M. Ingrisch, M.F. Reiser, and S.O. Schoenberg. Influence of multichannel combination, parallel imaging and other reconstruction techniques on mri noise characteristics. *Magnetic resonance imaging*, 26(6):754–762, 2008.
- G. Gerig, O. Kubler, R. Kikinis, and F.A. Jolesz. Nonlinear anisotropic filtering of mri data. *Medical Imaging, IEEE Transactions on*, 11(2):221–232, 1992.
- Miguel Ángel González Ballester, Andrew P Zisserman, and Michael Brady. Estimation of the partial volume effect in mri. *Medical Image Analysis*, 6(4):389–405, 2002.
- J. B. Johnson. Thermal Agitation of Electricity in Conductors. *Phys. Rev.*, 32:97–109, Jul 1928.
- H.Y. Kim, J. Giacomantone, and Z.H. Cho. Robust anisotropic diffusion to produce enhanced statistical parametric map from noisy fmri. *Computer Vision and Image Understanding*, 99(3):435–452, 2005.
- K. Krissian and S. Aja-Fernandez. Noise-Driven Anisotropic Diffusion Filtering of MRI. *IEEE Transactions on Image Processing*, 18(10):2265–2274, 2009.
- P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.
- J. Rajan, D. Poot, J. Juntu, and J. Sijbers. Segmentation based noise variance estimation from background mri data. *Image Analysis and Recognition*, pages 62–70, 2010.
- L.M. Rocha, A.X. Falcão, and L.G.P. Meloni. A Robust extension of the mean shift Algorithm using optimum path forest. *International Workshop on Combinatorial Image Analysis*, 1, April 2008.
- Chourmouzios Tsiotsios and Maria Petrou. On the choice of the parameters for anisotropic diffusion in image processing. *Pattern Recognition*, 46(5):1369–1381, 2013.
- F. Voci, S. Eiho, N. Sugimoto, and H. Sekibuchi. Estimating the gradient in the perona-malik equation. *Signal Processing Magazine, IEEE*, 21(3):39–65, 2004.
- J. Weickert. *Anisotropic diffusion in image processing*, volume 1. Teubner Stuttgart, 1998.