Automatic Parameter Optimization for De-noising MR Data

Joaquín Castellanos[1], Karl Rohr[2], Thomas Tolxdorff[3], and Gudrun Wagenknecht[1]

[1] Central Institute for Electronics, Research Center Jülich, Germany
j.castellanos@fz-juelich.de
[2] Dept. Intelligent Bioinformatics Systems IPMB, University of Heidelberg and DKFZ Heidelberg, Germany
[3] Institute of Medical Informatics, Biostatistics and Epidemiology, Charité - University Medicine Berlin, Germany

Abstract. This paper describes an automatic parameter optimization method for anisotropic diffusion filters used to de-noise 2D and 3D MR images. The filtering process is integrated into a closed-loop system where image improvement is monitored indirectly by comparing the characteristics of the suppressed noise with those of the assumed noise model at the optimal point. In order to verify the performance of this approach, experimental results obtained with this method are presented together with the results obtained by median and k-nearest neighbor filters.

1 Introduction

High-resolution MR images are often affected by noise, causing undesired intensity overlapping of represented tissues, making its posterior segmentation and classification difficult. Traditional linear filters, such as mean or Gaussian filters, commonly used to reduce the noise, do not consider the boundaries originated from regions with different intensities, producing smoothing of these edges and suppression of sharp details. As a result, the produced images are blurred and diffuse.

Anisotropic diffusion filters overcome these shortcomings by adjusting the smoothing (diffusion) strength to the boundaries, thus reducing the noise while preserving edges. The anisotropic diffusion approach arose from the use of the Gaussian filter in multi-scale image analysis [1]. Perona and Malik [2] modified the isotropic diffusion equation (Eq. 1) by making the diffusion coefficient term $c(\bar{x},t)$ a function of the magnitude of the gradient of the image intensity,

$$\frac{\partial}{\partial t} I(\bar{x},t) = div\left(c\left(\bar{x},t\right)\nabla I(\bar{x},t)\right) \tag{1}$$

where $I(\bar{x},t)$ stands for the processed image at time t, $\bar{x} = (x,y,z)$ the space coordinates, t the iteration step (time) and ∇I the image gradient.

The diffusion coefficient was defined as a monotonically decreasing function $c(\bar{x},t)=f(|\nabla I(\bar{x},t)|)$ of the gradient, which becomes small when the magnitude of

J. Duncan and G. Gerig (Eds.): MICCAI 2005, LNCS 3750, pp. 320–327, 2005.

the gradient is large and approaches one when the gradient is close to zero. Perona and Malik [2] proposed two such diffusion functions, PMAD1 and PMAD2.

$$c_1(\bar{x},t) = \exp - \left(\frac{|\nabla I(\bar{x},t)|}{k} \right)^2 \qquad c_2(\bar{x},t) = \frac{1}{1 + \left(\frac{|\nabla I(\bar{x},t)|}{k} \right)^2} \tag{2}$$

Gerig et al. [3] introduced a discrete anisotropic (non-linear) diffusion algorithm for de-noising MR images. Other diffusion functions were reported by Black et al. [4] and Weickert [5]. Weeratunga et al. [6] assessed the de-noising performance of several diffusion functions using medical and non-medical images. Suri and Wu [7] give an overview of current trends and outlook on future development of the anisotropic diffusion field.

The main parameters which control the behavior of the smoothing process in anisotropic diffusion are the *number of iterations (it)* and the *diffusion factor* k (Eq. 2), which determines the level of gradient intensity where diffusion is at its maximum. For de-noising applications, the *diffusion factor* needs to be adjusted according to the noise level. The noise is usually estimated with some statistical methods that determine global characteristics (e.g. Black et al. [4] used the median absolute deviation), or by hand-picking some homogeneous areas and measuring the local variance. The *number of iterations* determines how many times the smoothing process is repeated. This parameter is often adjusted manually but it can also be done using an auto-stop criterion. In the latter case, the program can consider the number of pixel (voxel) modifications which occurred between the last two iterations to stop execution [4]. Either way, selecting an appropriate set of parameters is generally quite complicated and time-consuming.

In this work, an iterative method is presented that automatically adjusts these two main parameters. In contrast to previous approaches the estimation of the noise level is only used to determine the initial value of k. The optimization of the parameter k is driven by the feedback output from an evaluation method until a maximum response is reached. The novel evaluation method estimates indirectly the improvement of the image by analyzing the suppressed information and comparing its characteristics with those expected at the optimum. This is repeated several times for different values of *it*. The best combination of the two parameters, according to the evaluation method response, is then selected to finally process the image. Figure 1 shows a diagram of the method described here. The following sections will explain in detail our method and will present achieved experimental results.

2 Method

The three basic modules of the automatic iterative system proposed here are the *de-noising filters*, the *evaluation method* and the *adjustment rules* (Fig. 1). The *de-noising filters* module contains several anisotropic diffusion functions for data processing (e.g., PMAD2) as well as a set of anisotropic diffusion filters modeled

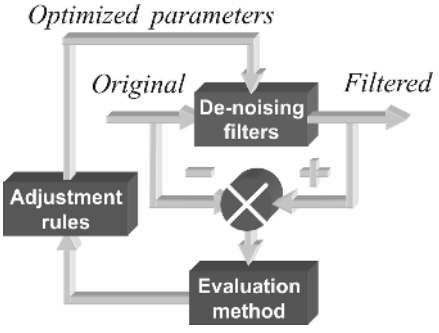

Fig. 1. Diagram of the automatic anisotropic filter system

after Nordström's [8] biased anisotropic formulation (e.g., PMAD2_bias). All the filters use a regularized (smoothed) version of the gradient to estimate the position of the edges [5] which should not be smoothed by the anisotropic filtering process.

The key component of the system, the *evaluation module*, gives feedback on image improvement or degradation during processing. Unlike other techniques, such as *image compression*, de-noising techniques do not have access to uncorrupted reference images to minimize the error between the reference and the processed image. Also, since this method is used for pre-processing, the only information available to it is that contained in the source image and that obtained during processing, so that no a priori anatomical knowledge is used to process the image. These conditions were set to keep the method as flexible and independent as possible.

MR images can be seen as the combination of the intensity information of the examined tissues and the noise generated during the measurement. After processing an MR image with an ideal filter configured using ideal parameters, the processed image would be perfectly clean of noise and only contain tissue information. Hence, the residual image, obtained by subtracting the source image from the processed one (Fig. 2), would consist only of the noise of the source image. The evaluation method takes advantage of this residual information to analyze the characteristics of the suppressed noise.

The characteristics of the noise for magnitude MR images are sufficiently known and therefore used as reference. In this case, the noise has a Rice distribution and its strength is homogeneous across the entire data set [9]. The closer the parametrization of k and it is to optimum, the more the residual image will approximate the characteristics of the initial noise. The images on the bottom row of Figure 2b were obtained by processing the image using successively increasing k values. The great variation in texture between the left and the right image suggest that either the source image needed more filtering or it was strongly smoothed and some anatomical structure have started to emerge in the residual image. The image in the middle was obtained using near-optimum parameters.

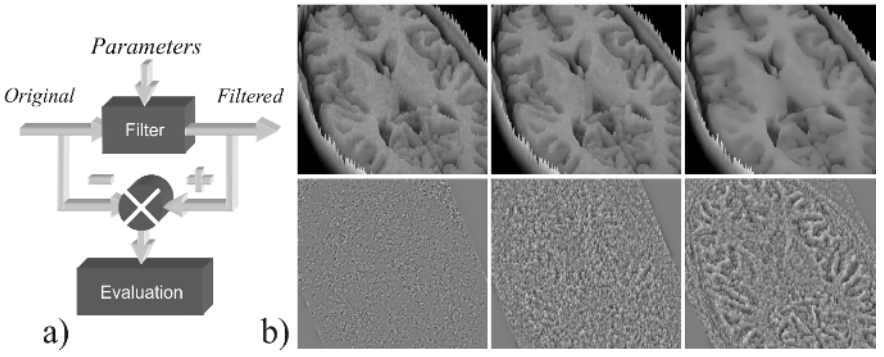

Fig. 2. Residual information used to monitor the noise reduction; a) diagram; b) processed (top) and residual images (bottom). The three pairs correspond to slightly smoothed (left), near optimally smoothed (middle) and heavily smoothed MR images (right).

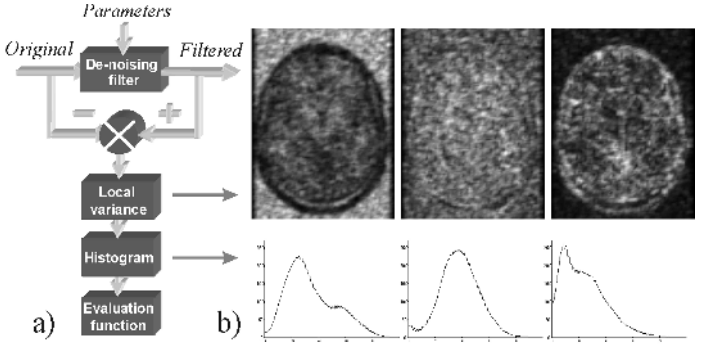

Fig. 3. a) Diagram of the evaluation method; b) results of the *local variance* (top) and from the *histogram* (bottom) modules. The three images correspond to slightly smoothed (left), near optimally smoothed (middle) and heavily smoothed MR images (right).

The *evaluation method* consists of three modules (Fig. 3a). The first one is a *local variance* operator which produces a picture of the noise, measuring the variance within a 3x3 (3x3x3) local region every third pixel (voxel). The local variance image is normalized to prevent bias during subsequent operations. The second module is a *histogram* which extracts the distribution information of the variance image. The results are smoothed with a low-pass filter to prevent discontinuities. The third module is an *evaluation function* of the histogram results which considers the maximum height, the width and the symmetry of the histogram to produce a noise reduction index. The *evaluation function* formula is shown in Equation 3.

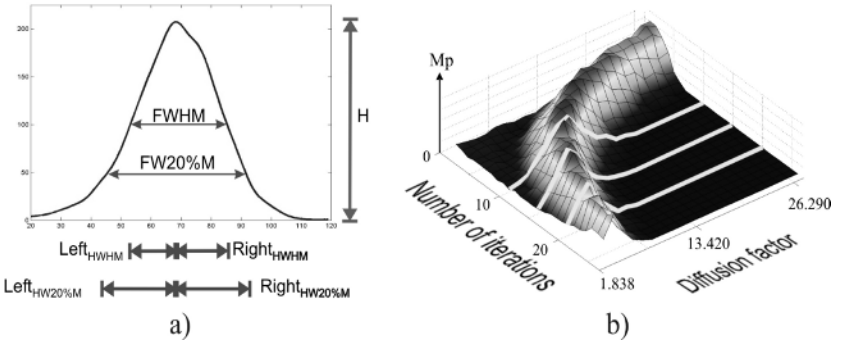

Fig. 4. a) Histogram characteristics considered by the *evaluation function* module; b) results of the *evaluation function*

Its first term H represents the maximum value of the histogram. The inverse of the Full-Width at Half-Maximum (FWHM) and the inverse of the Full-Width at 20%-Maximum (FW20%M) terms are indicators of the variance dispersion, and the two exponential terms are functions of the histogram symmetry based on the right and left Half-Widths at 50% (HWHM) and at 20% (HW20%M) of the maximum (Fig. 4a). This *evaluation function* yields large values when the histogram function is close to a large and narrow Gaussian-type curve, reflecting a homogeneous distribution of the local variance values. Figure 4b shows the results of the *evaluation function* after evaluating part of the parameters interval (it=1 to 25, k=1.838 to 26.290).

$$Mp = H * \left(\frac{1}{\text{FWHM}}\right) * \left(\frac{1}{\text{FW20\%M}}\right) * \left(1 - \exp\left(-\frac{(\text{Symm}_{\text{FWHM}})^3}{0.2}\right)\right) * \left(1 - \exp\left(-\frac{(\text{Symm}_{\text{FW20\%M}})^3}{0.2}\right)\right) \tag{3}$$

where:

$$\text{Symm}_{\text{FWHM}} = \frac{\text{Left}_{\text{HWHM}}}{\text{Right}_{\text{HWHM}}} \quad if \quad \text{Left}_{\text{HWHM}} \leq \text{Right}_{\text{HWHM}}$$

$$\text{Symm}_{\text{FWHM}} = \frac{\text{Right}_{\text{HWHM}}}{\text{Left}_{\text{HWHM}}} \quad if \quad \text{Left}_{\text{HWHM}} > \text{Right}_{\text{HWHM}}$$

$$\text{Symm}_{\text{FW20\%M}} = \frac{\text{Left}_{\text{HW20\%M}}}{\text{Right}_{\text{HW20\%M}}} \quad if \quad \text{Left}_{\text{HW20\%M}} \leq \text{Right}_{\text{HW20\%M}}$$

$$\text{Symm}_{\text{FW20\%M}} = \frac{\text{Right}_{\text{HW20\%M}}}{\text{Left}_{\text{HW20\%M}}} \quad if \quad \text{Left}_{\text{HW20\%M}} > \text{Right}_{\text{HW20\%M}}$$

The pairs *diffusion factor-number of iterations* corresponding to the ridge values are considered to be close to optimum parameter configurations because the response corresponds to a homogeneous distribution of the *local variance*.

The *adjustment rules* module was implemented to avoid evaluating each combination of parameters on the surface while searching for the optimum. This search is greatly simplified if each parameter is analyzed independently. It is more convenient if the continuous variable k is optimized first while the discrete variable *it* is kept constant (represented as white lines in Fig. 4b). The optimum k value of each sample *it* is obtained through a successive approximation scheme which determines the new k based on its current and previous values and on the corresponding results produced by the *evaluation function*. This optimization is repeated several times with different *it* values. From the optimum k values obtained, the median k value and its respective *number of iterations* are used for the final filtering of the image.

3 Results

In order to evaluate the method proposed here, several real and simulated data sets were processed (Fig. 5). For evaluation, three corrupted 3D data sets with increasing noise intensity were generated. These data sets represent different overlapping intensity levels between the tissue types cerebrospinal fluid, gray and white matter. The reference image, taken from the Montréal Neurological Institute (MNI) database [10], was an averaged T1-weighted image of 27 scans of the same individual. Rician noise was added following the equation:

$$I = \sqrt{((I_0 + n1(\sigma))^2 + (n2(\sigma))^2)} \tag{4}$$

where I_0 is the original image and $n1(\sigma)$ and $n2(\sigma)$ are two independent 3D images with zero-mean Gaussian-distributed noise. The standard deviations used to produce three noisy data sets were σ=9.16, 13.75 and 18.33.

These data sets were processed with the automatic method using the second Perona-Malik function PMAD2 (c_2 in Eq. 2) and its biased implementation PMAD2_bias [8]. The same data sets were also processed using a median filter (1 iteration) and a k-nearest neighbor (kNN) filter with k=14 (3 iterations). In all cases, the data were processed using a 26 neighborhood. The PMAD2 filter approximated the original image quite well, although it failed to reduce some speckle noise (Fig. 5e). The kNN filter also gave good results (Fig. 5f), although not as smooth as those of the anisotropic filter.

The experimental results were evaluated together with the corrupted data using the original MNI data set as reference. The evaluation was done using the mean-absolute error (MAE), the root-mean-square error (RMSE), the signal-to-noise ratio (SNR), the peak-signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) [11]. Figure 6 summarizes the obtained results. As can be seen there, the automatic parameterization of the second Perona-Malik function (PMAD2) gave the lowest errors (MAE and RMSE) and the greatest ratios (SNR, PSNR and SSIM). The biased implementation of the same filter (PMAD2_bias) gave comparable results to the k-nearest neighbor filter. Results obtained using the median filter were consistently inferior. The computing time for the MNI data set (181x217x181) was 47 min using a 2.6 GHz Pentium-4

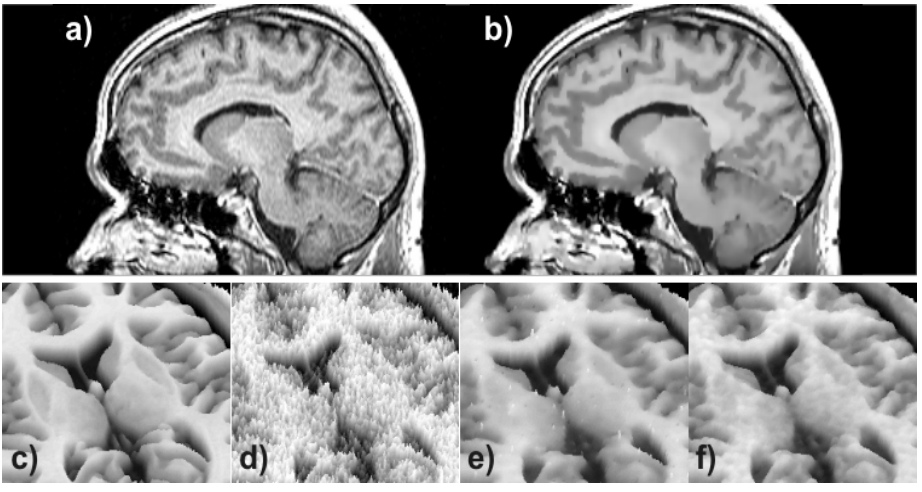

Fig. 5. a) Real MR image; b) after automatic filtering; c) reference MNI image; d) MNI image corrupted with Rician noise (σ=18.33); e) results from the anisotropic filter (PMAD2) using the parameters obtained with the automatic method (it=10, k=10.94); f) results from the k-nearest neighbor (kNN) filter

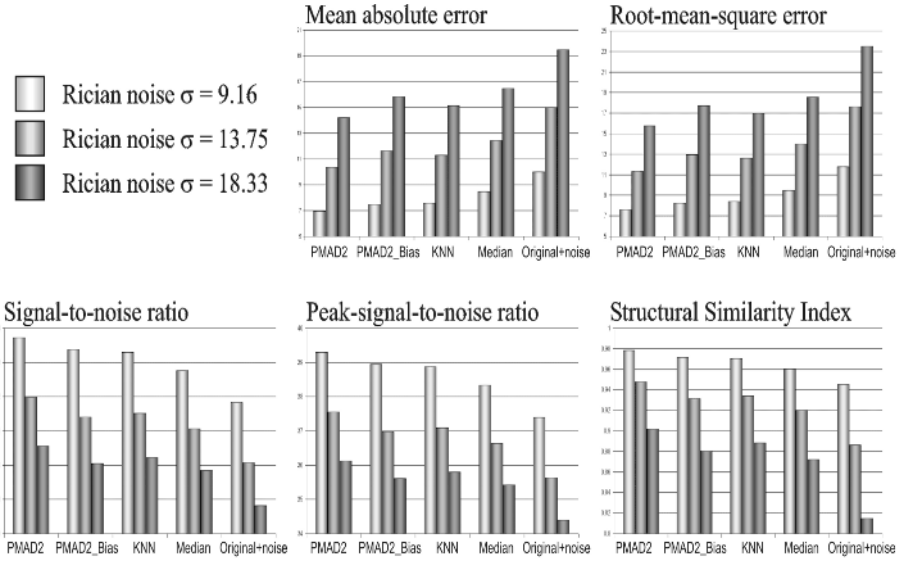

Fig. 6. Experimental results obtained comparing the original MNI data set with the corrupted and processed images

CPU. Each iteration took 1.8 seconds during the iterative optimization (using only 10 transaxial layers) and 33 seconds during the final filtering.

4 Discussion

The proposed evaluation function used to evaluate the filtering results is based on the characteristics of the expected noise model and therefore enables the implementation of a closed-loop system to automatically optimize the diffusion filter parameters. The obtained results, when compared to those obtained with median and k-nearest neighbor filters, indicate that our method is not only viable but also produces better results. In future work, we intend to incorporate adaptive versions of the diffusion filters into the *de-noising filters* module. These filters will locally adjust the global *diffusion factor* value according to the time (number of filter iterations) and to the local homogeneity of the image. In addition, we plan to optimize the behavior of the evaluation method according to the Rician noise model. We expect that these measures further increase the robustness and performance of the method.

Acknowledgment. The authors want to thank the German academic exchange service (DAAD) for financial support (reference no. A/02/11312). Thanks are also due to Alejandro Rodón for language editing.

References

1. Witkin, A.: Scale-space filtering. In: In Int. Join Conf. Artificial Intelligence. (1983) 1019–1022
2. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. IEEE Trans. In Pattern Analysis and Machine Intelligence **12** (1990) 629–639
3. Gerig, G., Kikinis, R., et al: Nonlinear anisotropic filtering of mri data. IEEE Transactions on Medical Imaging **11** (1992) 221–232
4. Black, M., Sapiro, G., et al: Robust anisotropic diffusion. IEEE Trans. Image Processing **7** (1998) 421–432
5. Weickert, J.: Anisotropic Diffusion in Image Processing. ECMI. Teubner (1998)
6. Weeratunga, S., Kamath, C.: A comparison of pde-based non-linear anisotropic diffusion technologies for image denoising. In: Image Processing: Algorithms and Systems. Algorithms and Systems II (2003)
7. Suri, J.S., Wu, D., et al: A comparison of state-of-the-art diffusion imaging technologies for smoothing medical/non-medical image data. In: Pattern Recognition. Algorithms and Systems II (2002)
8. Nordström, N.: Biased anisotropic diffusion - a unified regularization and diffusion approach to edge detection. In: 1st European Conf. on Comp. Vision. (1990) 18–27
9. Sijbers, J.: Signal and Noise Estimation from MR Images. Phd. thesis, University of Antwerp (1998)
10. Collins, D., Neelin, P., Evans, A., et al: Automatic 3d registration of mr volumetric data in standardized talairach space. J Comput Assist Tomogr **18** (1994) 192–205
11. Wang, Z., Bovik, A., Sheikh, H.: Structural similarity based image quality assessment. In: Digital Video Image Quality and Perceptual Coding. Marcel Dekker series in Signal Processing and Communications (2004)