LOCALLY CONTROLLED REGULARIZED SPATIOTEMPORAL ANISOTROPIC DIFFUSION

Pierre Portejoie, Simon Mure, Hugues Benoit-Cattin, Thomas Grenier

Université de Lyon, CREATIS
CNRS UMR5220 ; Inserm U1044
INSA-Lyon ; Université Lyon 1, France.

ABSTRACT

In this paper, we propose a new anisotropic diffusion formulation allowing non-linear spatiotemporal filtering of image sequences.

We first formulate a multidimensional spatiotemporal diffusion equation based on Barash's iterative form, processing independently both spatial, temporal and intensity dimensions with their own diffusion functions and scale parameters. We then introduce a local regularization term designed to smooth the remaining spike noise.

Experimental results processed on synthetic data and real MR images show that considering the temporal information and the regularization term improves the filtering quality. The method is also shown to be robust to noise, blur and temporal intensity evolution. Results are compared to the BM3D method with MSE and SSIM evaluation metrics.

***Index Terms*—** Image Sequence Filtering, Anisotropic Diffusion, Regularization, MRI

1. INTRODUCTION

Image denoising has been studied for several decades in the context of image processing. It is a basic task aiming at recovering the cleanest image from a corrupted input. A lot of different techniques have been designed for solving this issue [1, 2]. While most of them process 2-dimensional images, other techniques extend their formulation to 3-dimensional data considering time as the third dimension, as in image sequences or video processing.

One of the most basic way to filter corrupted images is obviously to pass it through a gaussian filter, although this leads to blurred data. Anisotropic diffusion is one way to solve this problem : the filtering process is related to the local gradient norm, the lower the gradient norm the higher the impact of the low-pass filter. This is an iterative process without image based stopping criteria. First description of anisotropic diffusion for image filtering was introduced by Perona and Malik [3] but Barash proposed a more robust one in [4], easily extendable to color image filtering, such as bilateral filtering introduced by Tomasi in [5]. Image denoising can also be achieved via patch-based approaches like NL-means, originally proposed by Buades et al. [1].

Most of these methods were extended to handle longitudinal data. Montagnat [6] was the first to propose an anisotropic diffusion application to image sequences filtering. More recently, BM3D and BM4D were applied to image sequences and videos in [7, 8, 9], and also used for image enhancement [10] and restoration [11, 12]. The idea that motion estimation was not needed in such approaches was introduced in [13].

Many other approaches have been proposed for noise removal or image filtering, such as more advanced methods based on MRF [14, 15], collaborative filtering [16], sparse transform domain [17] and total variations [18].

Applied publications involving medical images must also be cited: for MR images, anisotropic diffusion filtering in [19], specific segmentation in [20] and operating before the reconstruction of PET image in [21].

In this paper we focus on longitudinal medical images and consider them as image sequences. We extend Perona and Malik's basic conception of anisotropic diffusion including the temporal dimension in the process. This allows us to gain information from both spatial and temporal dimensions. The discrete iterative solution is based on Barash's discrete form [4], chosen for intuitiveness and stability, described using bandwidth matrices for spatial, intensity and temporal dimensions. One drawback of anisotropic diffusion filtering being its poor performance to remove very noisy points in homogeneous regions we also add a regularization term to the iterative procedure formulation to suppress remaining outlier values.

In the following we first describe the concept and implementation of the proposed filter (Section 2). Section 3 describes the data used for the experiments performed and the corresponding results are presented in Section 4.

This work was funded by CNRS grant PEPS INS2I. This work was performed within the framework of the LABEX PRIMES (ANR-11-LABX-0063) of Universite de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR).

2. METHODS

The proposed approach is based on Perona and Malik's anisotropic diffusion equation, originally designed for 2 dimensional images. This section presents our methodological contribution.

For a given image $I(\mathbf{x})$, the heat anisotropic diffusion equation at a pixel $\mathbf{x}$, using the dynamic parameter $c(\mathbf{x})$, introduced by Perona and Malik [3], is expressed as follows:

$$\frac{\partial I(\mathbf{x})}{\partial t} = c(\mathbf{x})\,\Delta I(\mathbf{x}) + \nabla c(\mathbf{x}) \cdot \nabla I(\mathbf{x}). \tag{1}$$

We chose Perona and Malik's recommended function $c(\mathbf{x})$:

$$c(\mathbf{x}) = \exp\big(-\frac{\|\nabla I(\mathbf{x})\|^2}{K^2}\big) \tag{2}$$

with K a constant allowing to tune the Gaussian width.

Then discretizing (2) using Barash's relationship [4] leads to:

$$I^{k+1}(\mathbf{x}) = \frac{\sum_{i\in N} I^k(\mathbf{x}_i).c(\mathbf{x}_i)}{\sum_{i\in N} c(\mathbf{x}_i)} \tag{3}$$

with N the set of $\mathbf{x}$ neighbors and $I^0(\mathbf{x}) = I(\mathbf{x})$.

The proposed method takes as input an image sequence, where the intensity value of a pixel $\mathbf{x}$ at time t is noted $I(\mathbf{x}, t)$. In the the following expressions, k will denote the iteration number of the filtering process and $\mathcal{N}$ the set containing the spatiotemporal neighbors of $(\mathbf{x}, t)$. Based on such considerations, we extend Barash's form to:

$$I^{k+1}(\mathbf{x}, t) = \frac{\sum_{i\in\mathcal{N}} I^k(\mathbf{x}_i, t_i).C\big(\mathbf{x}_i, t_i, I^k(\mathbf{x}_i, t_i)\big)}{\sum_{i\in\mathcal{N}} C(\mathbf{x}_i, t_i, I^k(\mathbf{x}_i, t_i))} \tag{4}$$

with $C\big(\mathbf{x}_i, t_i, I^k(\mathbf{x}_i, t_i)\big)$ the weighting function defined as:

$$C\big(\mathbf{x}_i, t_i, I^k_{\mathbf{x}_i,t_i}\big) = G_s(\mathbf{x}, \mathbf{x}_i).G_t(t, t_i).G_r(I^k_{\mathbf{x},t}, I^k_{\mathbf{x}_i,t_i}) \tag{5}$$

combining a spatial G_s, a temporal G_t and an intensity G_r weighting functions:

$$G_s(\mathbf{x}, \mathbf{x}_i) = g_s\big(d_s(\mathbf{x} - \mathbf{x}_i)\big) \tag{6}$$

$$G_t(t, t_i) = g_t\big(d_t(t - t_i)\big) \tag{7}$$

$$G_r(I^k_{\mathbf{x},t}, I^k_{\mathbf{x}_i,t_i}) = g_r\big(d_r(I^k_{\mathbf{x},t} - I^k_{\mathbf{x}_i,t_i})\big) \tag{8}$$

where g are positive, summing to one and decreasing functions such as Gaussian or rectangular functions, also known as kernel or profile functions, and where d functions are similarity measures such as Euclidean or Mahalanobis distances. In the case of Mahalanobis distance, scale matrices for spatial, temporal and intensity features must be set. One can note that our $C(\cdot)$ function still allows a pre-normalization of the input data by the scale matrices.

As anisotropic diffusion filtering is not able to remove outlier pixels in homogeneous regions, which intensity deviations to the mean intensities of the regions exceed the bandwidth value, we introduce a regularization term to reduce the remaining noise. This leads to the following equation:

$$\begin{aligned} I^{k+1}(\mathbf{x}) =& (1 - \kappa^k_{\mathbf{x}}) \frac{\sum_{i\in\mathcal{N}} I^k(\mathbf{x}_i, t_i).C\big(\mathbf{x}_i, t_i, I^k(\mathbf{x}_i, t_i)\big)}{\sum_{i\in\mathcal{N}} C(\mathbf{x}_i, t_i, I^k(\mathbf{x}_i, t_i))} \\ &+ \kappa^k_{\mathbf{x}} \frac{\sum_{i\in\mathcal{N}} I^k(\mathbf{x}_i, t_i)}{|\mathcal{N}|} \end{aligned} \tag{9}$$

where $|\mathcal{N}|$ is the cardinality of the set $\mathcal{N}$.

The regularization term is the averaged intensity value of $\mathcal{N}(\mathbf{x}, t)$, the neighborhood of $(\mathbf{x}, t)$, and must not be applied systematically to keep the edge preserving property of the anisotropic diffusion filtering. Thus, this term is weighted by a factor $\kappa^k_{\mathbf{x}} \in [0; 1]$ that will depend on $\mathcal{N}(\mathbf{x}, t)$ intensity values. In order to obtain the desired filtering properties, $\kappa^k_{\mathbf{x}}$ should be close to 1 when the intensity value $I^k(\mathbf{x}, t)$ is considered as an outlier, and close to 0 otherwise. Averaging was chosen for its simplicity and its trivial extension to multidimensional intensity values as for multi-channel image sequences (ie RGB videos).

In this work we propose to use $\kappa^k_{\mathbf{x}} \in \{0; 1\}$ and the zero norm $\|.\|_0$, counting the number of non-zero components in a vector, to measure the neighborhood behavior. The vector $\mathbf{I}^k(\mathbf{x}, t) \in \mathbb{R}^{|\mathcal{N}|}$ is built from all intensity values of $\mathcal{N}(\mathbf{x}, t)$ as: $\mathbf{I}^k(\mathbf{x}, t) = \big[I^k(\mathbf{x}_1, t_1), I^k(\mathbf{x}_2, t_2), \ldots, I^k(\mathbf{x}_{|\mathcal{N}|}, t_{|\mathcal{N}|})\big]$. The vector $\Delta\mathbf{I}^k(\mathbf{x}, t)$ is then defined by the subtraction between $\mathbf{I}^k(\mathbf{x}, t)$ and $I^k(\mathbf{x}, t)$. Using these notations, $\kappa^k_{\mathbf{x}}$ is computed as:

$$\kappa^k_{\mathbf{x}} = \begin{cases} 1 & \text{if } \|\Delta\mathbf{I}^k(\mathbf{x}, t)\|_0 + 1 > \tau \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

with τ a threshold tuning the maximum amount of dissimilar neighbors that are accepted before averaging the intensity of $I^k(\mathbf{x}, t)$ using neighbors of $(\mathbf{x}, t)$ instead of using equation (4).

We remind that to compute $\|.\|_0$, it is recommend to scale by a factor superior to one and floor to integer values each component of $\Delta\mathbf{I}^k(\mathbf{x}, t)$ in order to improve robustness to small variations in homogeneous regions.

In the multidimensional case, when $I^k(\mathbf{x}, t)$ is a vector of values as in color images, the j^{th} component of $\Delta\mathbf{I}^k(\mathbf{x}, t)$ is obtained by flooring d, the Euclidean distance between $I^k(\mathbf{x}_j, t_j)$ and $I^k(\mathbf{x}, t)$, divided by h_τ, the optional scale factor ($h_\tau \geq 1$):

$$d\big(I^k_{\mathbf{x}_j,t_j}, I^k_{\mathbf{x},t}; h_\tau\big) = \frac{1}{h_\tau}\|I^k(\mathbf{x}_j, t_j) - I^k(\mathbf{x}, t)\|. \tag{11}$$

In practice, functions described in equations (6), (7) and (8) should be specified (with their own parameters), the shape and size of the neighborhood $\mathcal{N}$ has to be chosen, then τ and h_τ values must be fixed and so does the number of iterations $\#it$ that the filter will run. Compared to spatial anisotropic diffusion approaches, three tuning parameters are added: τ and h_τ for the regularization term and -at least- one for the control of the time dimension contribution.

3. MATERIALS

The proposed method was evaluated both quantitatively on a synthetic image sequence and qualitatively on a 2D longitudinal MR brain acquisition.

Synthetic data is a $256 \times 256 \times 40$ spatiotemporal sequence (Fig.1), composed of several constant spatiotemporal regions and one intensity evolving region. Then all temporal slices were independently corrupted by a Gaussian blur with standard deviation set to σ_β and an additive Gaussian noise with standard deviation set to σ_η. Three corrupted image sequences were created: CS_1 with $[\sigma_\beta = 2, \sigma_\eta = 10]$, CS_2 with $[\sigma_\beta = 4, \sigma_\eta = 20]$, CS_3 with $[\sigma_\beta = 6, \sigma_\eta = 30]$).

Real longitudinal MR acquisitions are part of a MRI natural history study of multiple sclerosis patients and were acquired over 25 time steps [22]. Data pre-processing was based on the framework proposed in [23]: the data is registered on the first acquisition and the intensities are standardized. Then only one 2D+t section of the MRI data was used to perform the experiments.

To quantitatively assess the quality of our filtering results we chose to compare the obtained images with the original unaltered image by computing the MSE and the SSIM [24]. For real data images, the assessment is only qualitative as no ground truth exists.

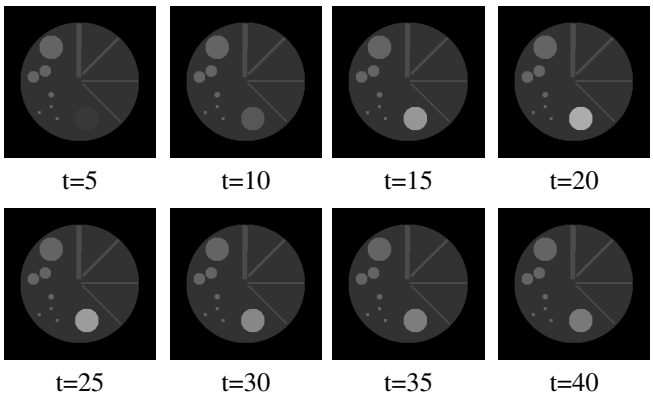


Fig. 1: Synthesized data set used for testing our method before applying Gaussian blur and noise.

Table 1: Parameters used for the experiments.

	h_s	h_t	h_r	$\#it$	$\tau/\lvert\mathcal{N}\rvert$	h_τ
CS_1	1	3	5	10	0.88	2
CS_2	1	3	10	15	0.88	2
CS_3	1	3	25	20	0.88	2
Real Data	1	3	5	10	0.88	2

4. EXPERIMENTS AND RESULTS

In this section, the proposed method is evaluated and compared with another method [16] on synthetic data. Then results of real MRI data processing are exposed.

4.1. Experiments

We will first describe common choice for all our experiments (mainly the functions described in equations (6), (7) and (8)) and then specifiy the experiments.

For all experiments, we use the squared Mahalanobis d_M^2 distance for all distance measurements d_s, d_t and d_r with $\mathbf{H}_s = h_s^2.\mathbb{I}_s$, h_t^2 and $\mathbf{H}_r = h_r^2.\mathbb{I}_r$ the respective scale factors with $\mathbb{I}_s$ and $\mathbb{I}_r$ the corresponding identity matrix of spatial and intensity measure space. The functions $g_s(\cdot)$, $g_t(\cdot)$ and $g_r(\cdot)$ are then all the same profile function $g(u) = exp(-u)$ with the scalar u equal to d_M^2. Such choices allow us to use the same $c(\mathbf{x})$ as (2) generalized to our spatiotemporal approach.

For the corrupted sequences, optimal parameters according to MSE where found using an exhaustive search. According to the value obtained on corrupted sequences, we manually fixed the parameters for real data. All parameters are described in Table 1. We precise that $|\mathcal{N}|$ was fixed according to h_s and h_t parameters so that $|\mathcal{N}| = (2h_s+1) \times (2h_s+1) \times (2h_t+1)$.

4.2. Results

For reading convenience, all data and results are available online[1].

Figure 2 presents results obtained by basic anisotropic diffusion (AD) filtering, by our spatiotemporal anisotropic filtering approach (AD+t) and by our regularized method (RAD+t) described in Section 2. Figure 3 presents results obtained for both real and generated data.

Based on the visual assessment and MSE, we can infer that our proposed filters AD+t and RAD+t perform better than AD for noise removal and contour sharpening. To obtain good filtering results, all the 6 parameters must be adjusted. However from our exhaustive search on this data, it seems that only h_r and $\#it$ should be tuned carefully, as the other ones remain the same for the different images. We also note that noise removal on the first and last few frames of the

[1] http://www.creatis.insa-lyon.fr/%7egrenier/research/PortejoieICIP2015/

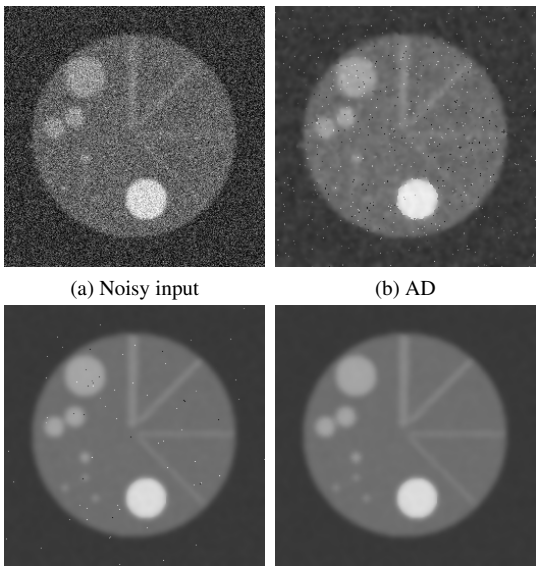
Table 2: Best results of our method (RAD+t) and BM3D filtering according to MSE (smallest is better) and SSIM (highest is better). Optimization was done according to the MSE.

	CS_1	CS_2	CS_3
RAD+t MSE	**25**	**59**	**181**
BM3D MSE	32	64	222
RAD+t SSIM	**0.93**	**0.88**	**0.87**
BM3D SSIM	0.91	0.85	0.85

sequences is negatively impacted by side effects and that this number of frames is linked to h_t.

We quantitatively compared our RAD+t approach to BM3D [16], which is one of the most efficient filtering approach known so far. We applied this algorithm to our data without tuning default parameters as this approach is tuning free. The MSE and SSIM values obtained are presented in Table 2. We observe that our approach is more adapted than BM3D for this particular kind of images sequences.

On a visual level, we can see on the real data that our approach is more efficient for removing noise artifacts and enhancing the quality of the image, we can also note that the details are sharpened.



(a) Noisy input (b) AD

(c) AD+t (d) RAD+t

Fig. 2: Results for the 20th time frame of sequence CS_2.

5. CONCLUSION

In this paper we have proposed a new method to filter image sequences, taking into account both spatial and temporal dimensions in the filtering process. We extended Perona and Malik's anisotropic diffusion filtering works [3] using Barash's form [4] while adding temporal data and a locally controlled regularized term in the process suppressing peak noise in homogeneous regions.

We have shown that our method improves noise removal and preserves region boundaries on image sequences with a time-evolving region. We have also proven the efficiency of using the regularization term for improving denoising.

According to our experiments on synthetic data, we provided results outperforming one of the best state of the art approach applied on this particular application.

The use of three scale parameters to control the anisotropic diffusion allows intuitive tuning of the parameters. However our approach requires a fine tuning of 2 over 6 parameters (three scale parameters, τ and the number of iterations) that we did not describe in this paper.

For future works, we will process video data containing moving objects.

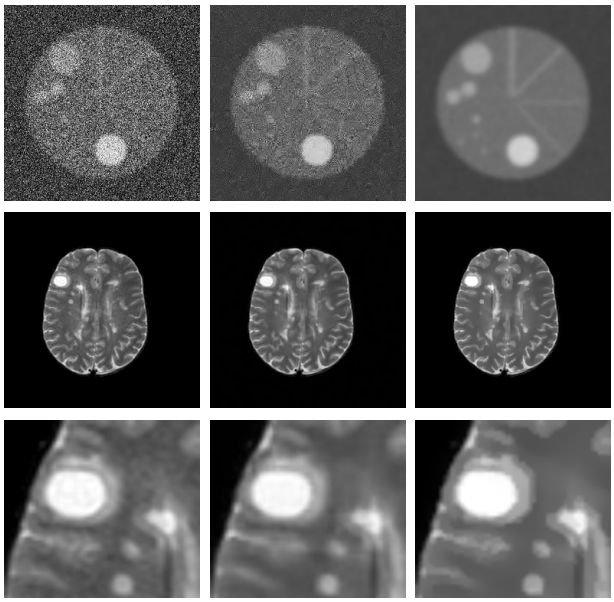


Fig. 3: Input (left) and results for BM3D (center) and RAD+t (right) for 20th time frame of CS_3 (top) and 4th time frame of real MRI data (middle). Bottom presents a detail of a particular region of real data.

6. REFERENCES

[1] A. Buades, B. Coll, and J. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Modeling & Simulation*, vol. 4, no. 2, pp. 490–530, 2005.

[2] V.B. Surya Prasath and D. Vorotnikov, "Weighted and well-balanced anisotropic diffusion scheme for image denoising and restoration," *Nonlinear Analysis: Real World Applications*, vol. 17, no. 0, pp. 33 – 46, 2014.

[3] P. Perona and J. Malik, "Scale-space and edge detection

using anisotropic diffusion," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 12, no. 7, pp. 629–639, Jul 1990.

[4] D. Barash, "Fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 6, pp. 844–847, Jun 2002.

[5] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Computer Vision, 1998. Sixth International Conference on*, Jan 1998, pp. 839–846.

[6] J. Montagnat, M. Sermesant, H. Delingette, G. Malandain, and N. Ayache, "Anisotropic filtering for model-based segmentation of 4D cylindrical echocardiographic images," *Pattern Recognition Letters - Special Issue on Ultrasonic Image Processing and Analysis*, vol. 24, no. 4-5, pp. 815–828, 2003.

[7] H. Malm, M. Oskarsson, E. Warrant, P. Clarberg, J. Hasselgren, and C. Lejdfors, "Adaptive enhancement and noise reduction in very low light-level video," in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, Oct 2007, pp. 1–8.

[8] A. Buades, B. Coll, and J.-M. Morel, "Nonlocal image and movie denoising," *International Journal of Computer Vision*, vol. 76, no. 2, pp. 123–139, 2008.

[9] J. Boulanger, C. Kervrann, P. Bouthemy, P. Elbau, J.-B. Sibarita, and J. Salamero, "Patch-based nonlocal functional for denoising fluorescence microscopy image sequences," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 2, pp. 442–454, Feb 2010.

[10] X. Jiang, H. Yao, S. Zhang, X. Lu, and W. Zeng, "Night video enhancement using improved dark channel prior," in *Image Processing (ICIP), 2013 20th IEEE International Conference on*, Sept 2013, pp. 553–557.

[11] J. Boulanger, C. Kervrann, and P. Bouthemy, "Space-time adaptation for patch-based image sequence restoration," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 6, pp. 1096–1102, June 2007.

[12] X.Li and Y. Zheng, "Patch-based video processing: A variational bayesian approach," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 19, no. 1, pp. 27–40, Jan 2009.

[13] A. Buades, B. Coll, and J.-M Morel, "Denoising image sequences does not require motion estimation," in *Advanced Video and Signal Based Surveillance, 2005. AVSS 2005. IEEE Conference on*, Sept 2005, pp. 70–74.

[14] J. Chen and C.-K. Tang, "Spatio-temporal markov random field for video denoising," in *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, June 2007, pp. 1–8.

[15] M. Maggioni, G. Boracchi, A. Foi, and K. Egiazarian, "Video denoising, deblocking, and enhancement through separable 4-d nonlocal spatiotemporal transforms," *Image Processing, IEEE Transactions on*, vol. 21, no. 9, pp. 3952–3966, Sept 2012.

[16] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *Image Processing, IEEE Transactions on*, vol. 16, no. 8, pp. 2080–2095, Aug 2007.

[17] S. Bhagavathy and J. Llach, "Adaptive spatio-temporal video noise filtering for high quality applications," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, April 2007, vol. 1, pp. I–761–I–764.

[18] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1-2, pp. 89–97, 2004.

[19] L. He and R. Greenshields Ian, "A nonlocal maximum likelihood estimation method for rician noise reduction in mr images," *Medical Imaging, IEEE Transactions on*, vol. 28, no. 2, pp. 165–172, Feb 2009.

[20] E. Ardizzone, R. Pirrone, and O. Gambino, "Automatic segmentation of mr images based on adaptive anisotropic filtering," in *Image Analysis and Processing, 2003.Proceedings. 12th International Conference on*, Sept 2003, pp. 283–288.

[21] J.G. Brankov, M.N. Wernick, Yongyi Yang, and M.V. Narayanan, "Spatially-adaptive temporal smoothing for reconstruction of dynamic and gated image sequences," in *Nuclear Science Symposium Conference Record, 2000 IEEE*, 2000, vol. 2, pp. 15/146–15/150.

[22] C. Guttmann, S. Ahn, L. Hsu, R. Kikinis, and F. Jolesz, "The evolution of multiple sclerosis lesions on serial mr.," *American journal of neuroradiology*, vol. 16, no. 7, pp. 1481–1491, 1995.

[23] D. Meier and C. Guttmann, "Time-series analysis of mri intensity patterns in multiple sclerosis," *NeuroImage*, vol. 20, no. 2, pp. 1193–1209, 2003.

[24] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600 –612, april 2004.