MRI diffusion-based filtering: A note on performance characterisation

Ovidiu Ghita[*], Kevin Robinson, Michael Lynch and Paul F. Whelan

Vision Systems Group

School of Electronic Engineering

Dublin City University,

Glasnevin, Dublin 9,

Ireland

[*]Corresponding author:

E-mail: ghitao@eeng.dcu.ie

Phone: +353-1-7007637

Fax: +353-1-7005508

MRI diffusion-based filtering: A note on performance characterisation

Ovidiu Ghita, Kevin Robinson, Michael Lynch and Paul F. Whelan

Vision Systems Group

School of Electronic Engineering

Dublin City University, Ireland

Abstract

Frequently MRI data is characterised by a relatively low signal to noise ratio (SNR) or contrast to noise ratio (CNR). When developing automated Computer Assisted Diagnostic (CAD) techniques the errors introduced by the image noise are not acceptable. Thus, to limit these errors, a solution is to filter the data in order to increase the SNR. More importantly, the image filtering technique should be able to reduce the level of noise but not at the expense of feature preservation. In this paper we detail the implementation of a number of 3D diffusion-based filtering techniques and we analyse their performance when they are applied to a large collection of MR datasets of varying type and quality.

Keywords: MRI, feature preserving smoothing, anisotropic diffusion, image segmentation.

1. Introduction

In a survey on image smoothing techniques the approaches encountered may be classified under two broad headings, linear and non-linear [1,2]. Standard linear smoothing techniques based on local averaging or Gaussian weighted spatial operators reduce the level of noise but this is achieved at the expense of poor feature preservation. Consequently, the filtered data appears blurry as step intensity discontinuities such as edges are attenuated. To compensate for these undesirable effects, non-linear techniques have been developed in order to achieve better feature preservation. Among these, the median filter is the simplest operator to remove impulse-like noise [2]. More complex non-linear techniques include statistical approaches based on nonparametric estimation [3,4]. However, while these methods do alleviate somewhat the shortcomings associated with linear techniques, they still perform only

modestly when the data is affected by long tailed noise distributions. To complement these filtering approaches, a number of adaptive techniques have been proposed [2,3,5]. These methods try to achieve the best trade-offs between smoothing efficiency, feature preservation and the generation of artefacts.

Recent developments based on non-linear diffusion [6,7,8,9] alleviate the major limitations associated with conventional linear and non-linear smoothing methods. Diffusion-based filtering was originally developed by Perona and Malik [10] in order to implement an optimal, feature preserving smoothing strategy. Many implementations follow their original approach where the main aim was to improve numerical stability [11,12]. This was advanced by Weikert [13] where he developed a new smoothing algorithm by permitting diffusion along the direction of edges. Gerig et al [6] extended this work to 3D and evaluated its usefulness when applied to medical 2D and 3D datasets. In this paper our aim is to further extend their initial work in presenting the implementation of three diffusion-based smoothing algorithms where a special emphasis has been placed on performance characterisation [14,15]. We have evaluated the algorithms on various MRI datasets and the results are presented and discussed.

This paper is organised as follows. Section 2 describes the extension to 3D of the standard diffusion algorithm. Section 3 details the implementation of an adaptive 3D diffusion-based smoothing strategy. In Section 4 a 3D anisotropic Gaussian filtering technique is introduced. Section 5 presents an extensive performance characterisation of the smoothing strategies described in this paper and Section 6 presents some concluding remarks.

2. Non-linear 3D Diffusion Filtering

In this section we describe the extension to 3D of the smoothing strategy that has been described in Perona and Malik's paper [10]. In their paper smoothing is formulated as a diffusive process where smoothing is performed at intra regions and suppressed at region boundaries.

This non-linear smoothing procedure can be defined in terms of the derivative of the flux function:

$$u_t = div(D(|\nabla u|)\nabla u) \tag{1}$$

where u is the input data, D represents the diffusion function, t indicates the iteration step and *div* is the divergence operator. The smoothing strategy described in Eq. 1 can be translated into an iterative discrete formulation for 3D data as follows:

$$I_{x,y,z}^{t+1} = I_{x,y,z}^{t} + \lambda \sum_{R=1}^{6} [D(\nabla_R I)\nabla_R I]^t \tag{2}$$

where ∇ is the gradient operator than can be defined in a 6 voxel connected neighbourhood (see Eq. 3) and λ is a contrast parameter that takes a value in the range $0<\lambda<0.16$ as suggested in [10]. For a 26 voxel connected neighbourhood the implementation is similar and we just have to adjust the range for the λ parameter accordingly.

$$\begin{aligned} \nabla_1 I_{x,y,z} &= I_{x-1,y,z} - I_{x,y,z} \\ \nabla_2 I_{x,y,z} &= I_{x+1,y,z} - I_{x,y,z} \\ \nabla_3 I_{x,y,z} &= I_{x,y-1,z} - I_{x,y,z} \\ \nabla_4 I_{x,y,z} &= I_{x,y+1,z} - I_{x,y,z} \\ \nabla_5 I_{x,y,z} &= I_{x,y,z-1} - I_{x,y,z} \\ \nabla_6 I_{x,y,z} &= I_{x,y,z+1} - I_{x,y,z} \end{aligned} \tag{3}$$

The diffusion function $D(x)$ should be bounded in the interval $(0\rightarrow1)$ and should have the highest value when the input x has the value zero. These requirements translate to minimal smoothing around boundaries where the gradient has high values. In practice, a large number of functions can be engineered to satisfy these requirements and in our implementation we have used two types of diffusion functions, exponential and reciprocal:

$$D(|\nabla I|) = e^{-\left(\frac{\nabla I}{k}\right)^2} \tag{4}$$

$$D(|\nabla I|) = \frac{1}{1+\left(\frac{\nabla I}{k}\right)^2} \tag{5}$$

where k is the diffusion parameter. The parameter k controls the smoothing level, the smoothing being more pronounced for high values of k. The experimental data indicates that

slightly better results are obtained when the exponential form illustrated in Eq. 4 is used in the expression depicted by Eq. 2.

3. Adaptive 3D Diffusion Smoothing

The adaptive smoothing algorithm implemented in this paper represents the extension to 3D of the algorithm proposed by Chen [16]. Chen [16] demonstrated that the standard diffusion algorithm might not return optimal results when applied to image data defined by a very low SNR. To tackle this limitation he proposed to use two discontinuity measures jointly in order to control the smoothing process. To this end, he used the spatial variance to measure contextual discontinuities and the gradient information as a measure of the local discontinuities.

In order to measure the local discontinuities we calculate the derivatives in a 3x3x3, 13 diagonal derivatives calculated in a 26 voxel-connected neighbourhood.

$$
\begin{aligned}
\nabla_1 I_{x,y,z} &= \left| I_{x+1,y,z} - I_{x-1,y,z} \right| \\
\nabla_2 I_{x,y,z} &= \left| I_{x,y+1,z} - I_{x,y-1,z} \right| \\
\nabla_3 I_{x,y,z} &= \left| I_{x+1,y+1,z} - I_{x-1,y-1,z} \right| \\
&\ldots
\end{aligned}
\tag{6}
$$

where $I_{x,y,z}$ represent the voxel intensity at the position (x,y,z). Then, we can define the local discontinuity measure as:

$$
E_{x,y,z} = \frac{\sum_{i=1}^{13} |\nabla_i|}{13}
\tag{7}
$$

As mentioned earlier the contextual discontinuities are sampled by measuring the spatial variance. The variance is computed in a cubic neighbourhood around the pixel of interest $N_{xyz}(R)$ as follows:

$$
\sigma_{xyz}^2(R) = \frac{\sum_{(i,j,k \in N_{xyz}(R))} \left(I_{i,j,k} - \mu_{xyz}(R) \right)^2}{\left| N_{xyz}(R) \right|}
\tag{8}
$$

where μ_{xyz} is the mean value of the voxels situated in the cubic neighbourhood $N_{xyz}(R)$ and $|N_{xyz}(R)|$ is the number of pixels. Then the spatial variances for the whole volume are scaled between 0 and 1 using the simple transformation illustrated in Eq. 9.

$$\tilde{\sigma}^2_{xyz} = \frac{\sigma^2_{xyz}(R) - \sigma^2_{\min}(R)}{\sigma^2_{\max}(R) - \sigma^2_{\min}(R)} \tag{9}$$

The adaptive smoothing scheme is iterative and is illustrated below:

$$I^{t+1}_{xyz} = I^t_{xyz} + \eta_{xyz} \frac{\sum_{i,j,k \in N_{xyz}(R) \backslash \{(x,y,z)\}} \eta_{ijk} \gamma^t_{ijk} \left(I^t_{ijk} - I^t_{xyz} \right)}{\sum_{i,j,k \in N_{xyz}(R) \backslash \{(x,y,z)\}} \eta_{ijk} \gamma^t_{ijk}} \tag{10}$$

where,

$$\eta_{ijk} = \exp(-\alpha \Phi(\sigma^2_{ijk}(R), \theta_\sigma)) \tag{11}$$

$$\gamma^t_{ijk} = \exp(-E^t_{ijk} / S) \tag{12}$$

$$\Phi(\tilde{\sigma}^2_{xyz}(R), \theta_\sigma) = \begin{cases} 0 & \tilde{\sigma}^2_{xyz}(R) < \theta_\sigma \\ \tilde{\sigma}^2_{xyz}(R) & \tilde{\sigma}^2_{xyz}(R) \geq \theta_\sigma \end{cases} \tag{13}$$

In the formulation illustrated in Eq. 10 the parameters θ_σ and α control the extent to which the contextual discontinuities should be preserved while S is the diffusion parameter that controls the preservation of local discontinuities. It can be seen that γ implements the diffusion function depicted in Eq. 4. For this implementation the parameter $R=3$ (for larger values image details such as thin lines are suppressed) and typical values for controlling parameters are: $\alpha=10$, $\theta_\sigma= (0.05 \rightarrow 0.95)$, $S =(5 \rightarrow 30)$. The algorithm is typically run for 5 to 20 iterations.

4. Anisotropic Gaussian Smoothing

An anisotropic filter based on the familiar Gaussian model has also been implemented in order to provide edge enhancing, directional smoothing. This approach reduces to convolution with a scaled Gaussian mask where the calculation of the kernel's weights becomes the key

issue governing the performance of this smoothing algorithm. By calculating the local gradient vector and favouring smoothing along the edge over smoothing across it, we can achieve boundary-preserving filtering where image regions are smoothed while edges are enhanced.

The weight for a neighbour pixel q can be calculated as a function of the gradient at point p, at the mask origin, and the distance from the origin to the neighbour q. The relationship used in our approach is given in Eq. 14 where $\overline{pq}$ is the vector from the mask centre point p to the neighbour q, ∇u is the gradient vector at position p, λ is the scale parameter that controls the smoothing strength while μ is the shape parameter controlling anisotropy. It can be observed that in the case when $\mu=0$ the anisotropic term $\left(\dfrac{\overline{pq} \cdot \nabla u}{\lambda}\right)^2 \left(2\mu+\mu^2\right)$ is also zero and the smoothing operation reduces to the non-linear isotropic form where smoothing is suppressed at image boundaries (as no directionality is applied).

$$wt(pq, \nabla u) = e^{-\left(\left(\frac{|\overline{pq}||\nabla u|}{\lambda}\right)^2 + \left(\frac{pq \cdot \nabla u}{\lambda}\right)^2 \left(2\mu+\mu^2\right)\right)} \tag{14}$$

The smoothing operation filters the data iteratively with the 3D kernel constructed from Eq. 14. As the number of iterations is increased, more noise and small features are eliminated but even in extreme cases the strong edges in the data are well preserved in both location and strength.

5. Experiments and results

The filtering algorithms described in this paper have been applied to a large number of MR datasets including MRCP, whole body, brain and heart sequences. The aim of these experiments is to conduct a detailed performance characterisation for smoothing algorithms described in this paper in order to produce quantitative results. In our experiments we also included datasets where additional noise had been added to the original sequences.

To evaluate the performance of the smoothing algorithms described in this paper, the first set of experiments were conducted on a synthetic dataset that is defined by a homogenous cubic object with a known greyscale value surrounded by background pixels. To test smoothing algorithms on this artificial dataset is advantageous as the ground truth data is known and the smoothing results are easy to evaluate. We tested the efficiency of the algorithms when the artificial dataset was corrupted with various types of 3D image noise, including Gaussian, Poisson and additive uniformly distributed white noise [1,17]. As quantitative values we have evaluated the local uniformity sampled by the 7x7x7 standard deviation at the location situated at the centre of the cube and the alteration of the greyscale value at the same position when compared with the expected known value. Some experimental results are depicted in Table 1.

Table 1 should be placed here.

In Table 1 the symbols G-15 and G-30 indicate that the synthetic dataset has been corrupted with Gaussian noise (standard deviation 15 and 30 greyscale values). Similarly P-15 and P-30 denote the fact that the test dataset has been corrupted with Poisson noise (distribution 15 and 30 greyscale values) and W-15 and W-30 indicate that the dataset has been corrupted with uniformly distributed white noise (mean deviation 15 and 30 greyscale values).

In order to evaluate globally the noise removal efficiency on real datasets we need to define quantitative measures that indicate the overall performance of the smoothing algorithms that are evaluated. In this regard, we propose to evaluate jointly two quantitative measurements: the smoothness factor that assesses the global uniformity and the edge preservation factor that indicates to what extent the strong edge features are retained and enhanced. To this end, we employed the standard deviation as a measure to evaluate the image local homogeneity. To be statistically relevant [17] the standard deviation should be calculated over a large region but on the other hand the results will be affected by small non-uniformities such as intensity gradients or structural image variations [6]. This requirement is quite difficult to be accomplished if we want to develop an automatic performance characterisation scheme where user intervention is not required. One solution has been advanced by Canny [18] when he decided to select the threshold parameters for an edge detector based on analysis of the cumulative histogram of the gradients. However due to the nature of MR datasets this criteria to identify the gradients generated by noise proved to be inefficient. Thus, in our implementation we have developed an alternative strategy based on observation. In this sense,

we computed the standard deviation for all voxels in the original dataset in a 7x7x7 neighbourhood. These values were sorted with respect to their magnitude and from these values the 25% of the highest values were eliminated, as they are likely to belong to edges and 25% of the lowest values are also eliminated as they are calculated from areas that have no significant texture (such as image regions defined by air). This strategy was applied to select the seed points that belong to image regions defined by a low SNR. Then, the standard deviation for each of the filtered datasets is measured at the same voxel locations (also in a 7x7x7 neighbourhood). To evaluate a quantitative estimation we calculate the RMS value of the standard deviations from the original and smoothed datasets resulting after the application of the smoothing strategies described in previous sections (for details refer to Table 2).

Table 2 should be placed here.

The edge strength is evaluated by plotting the intensity and gradient data at selected locations where edges are located, before and after the application of the smoothing operations. Some graphical results are depicted in Figs 1 to 4. The experimental data illustrated in Figs 1 to 4 indicate that the 3D adaptive smoothing and 3D anisotropic smoothing algorithms perform better than the standard diffusion. The 3D adaptive smoothing algorithm returned better results than the 3D anisotropic when applied to heart, brain and whole body datasets. The 3D anisotropic algorithm performed better when applied to MRCP dataset.

Figure 1 should be placed here
Figure 2 should be placed here
Figure 3 should be placed here
Figure 4 should be placed here

The graphs illustrated in Figs 1 and 4 demonstrate the edge enhancement around image data defined by step-like edges. It can be noticed that the edge localisation is significantly improved. The effect of edge strengthening is even more pronounced for weaker edges in an MRI brain sequence (see Fig. 3) or in image areas affected by a high level of noise, as is the case of the MRCP dataset illustrated in Fig. 2.

The performance of the smoothing algorithms described in this paper is remarkable in discriminating a true edge from image noise (see Fig. 2c). Also notice the improved

performance of the adaptive 3D smoothing algorithm as compared with the performance of the standard diffusion and the 3D anisotropic diffusion algorithms.

In order to emphasise the effectiveness of the smoothing strategies described in this paper we also present the segmentation resulting after the application of a 3D clustering algorithm [19] to the original and smoothed data. Samples of the segmentation results are depicted in Figs 5 to 8.

Figure 5 should be placed here

Figure 6 should be placed here

Figure 7 should be placed here

Figure 8 should be placed here

6. Conclusions

In this paper we have described the implementation of three diffusion-based smoothing schemes and their application to medical 3D data. Our interest has focused on MRI acquisition modalities as MRI datasets are characteristically defined by a low signal to noise ratio (SNR). Hence, our aim was to demonstrate that far superior results are achieved if the MRI data is initially filtered in order to reduce the level of image noise and improve the SNR. In this regard, we have performed a detailed performance characterisation for each smoothing operators evaluated in this paper on both synthetic and real data (including heart, brain, whole body and MRCP image sequences). We conclude that the diffusion-based smoothing techniques offer an efficient approach to noise reduction, and more important this advantage is not achieved at the expense of feature preservation. The experimental data presented and discussed in this paper highlights the ability of the diffusion-based smoothing schemes to distinguish the high gradient image features from the MRI image acquisition noise.

The source code for the 3D smoothing schemes presented in this paper can be downloaded from the following web page: http://www.eeng.dcu.ie/~whelanp/vsg/vsgcode.html.

References

[1] S.J. Orfanidis, Introduction to signal processing, Prentice-Hall; 1996.

[2] M. Sonka, V. Hlavac and R. Boyle, Image processing, analysis and machine vision, 2nd edition, PWS Boston; 1998.

[3] B. Smolka, R. Lukac, A. Chydzinski, K.N. Plataniotis and K. Wojciechovski, "Fast adaptive similarity based impulse noise reduction filter", Real Time Imaging, 9(4): 261-276; 2003.

[4] K. Tang, J. Astola and Y. Neuovo, "Nonlinear multivariate image filtering techniques", IEEE Trans. on Image Processing, 4(6): 788-797; 1995.

[5] D. Comaniciu, P. Meer, "Mean shift: A robust approach toward feature space analysis, IEEE Trans. on Pattern Analysis Machine Intelligence, 24(5): 603-619; 2002.

[6] G. Gerig, R. Kikinis, O. Kubler and F.A. Jolesz, "Nonlinear anisotropic filtering of MRI data", IEEE Trans. on Medical Imaging, 11(2): 221-232; 1992.

[7] 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, 24(4-5): 815-828; 2003.

[8] G.I. Sanchez-Ortiz, D. Rueckert and P. Burger, "Knowledge-based tensor anisotropic diffusion of cardiac magnetic resonance images", Medical Image Analysis, 3(1): 77-101; 1999.

[9] R.T. Whitaker and S.M. Pizer, "A multi-scale approach to nonuniform diffusion", CVGIP: Image Understanding, 57(1): 111-120; 1993.

[10] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion", IEEE Trans. on Pattern Analysis and Machine Intelligence, 12(7): 629-639; 1990.

[11] N. Nordstrom, "Biased anisotropic diffusion. A unified regularization and diffusion approach to edge detection", Image and Vision Computing, 8(4): 318-327; 1990.

[12] J. Weickert, B. M. ter Haar Romeny, M.A. Viergever, "Efficient and reliable schemes for nonlinear diffusion filtering", IEEE Trans. on Image Processing, 7(3): 398-410; 1998.

[13] J. Weickert, "Anisotropic diffusion in image processing", Teubner Verlag, Stuttgart; 1998.

[14] W. Foerstner, "10 pros and cons against performance characterisation of vision algorithms", Proc. ECCV Workshop on Performance Characteristics of Vision Algorithms, Cambridge, UK; 1996.

[15] R.M. Haralick, "Performance characterization in computer vision", CVGIP: Image Understanding, 60(2): 245-249; 1994.

[16] K. Chen, "A feature preserving adaptive smoothing method for early vision", Technical report, National Laboratory of Machine Perception, Centre for Information Science, Peking University, Beijing, China; 1999.

[17] R.O. Duda, P.E. Hart and D.G. Stork, Pattern Classification, 2nd Edition John Wiley and Sons; 2001.

[18] J. Canny, "A computational approach to edge detection", IEEE Trans. on Pattern Analysis and Machine Intelligence, 8(6): 679-698; 1986.

[19] M. Lynch, O. Ghita, P. F. Whelan, "Comparison of 2D and 3D clustering on short axis Magnetic Resonance Images of the left ventricle", CARS 2004 Computer Assisted Radiology and Surgery, June 23 - 26, Chicago, USA; 2004.
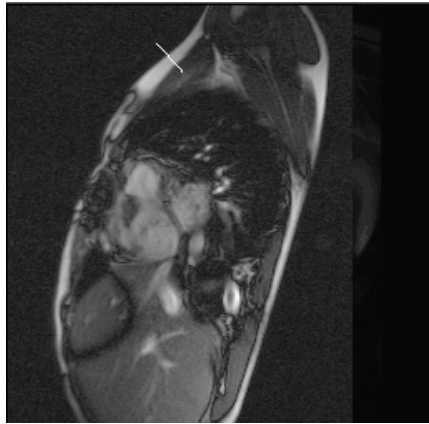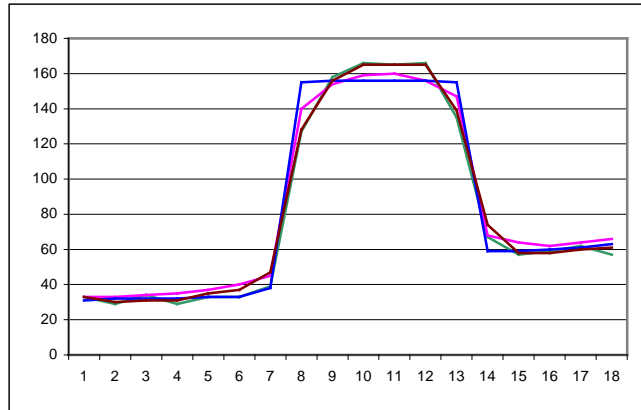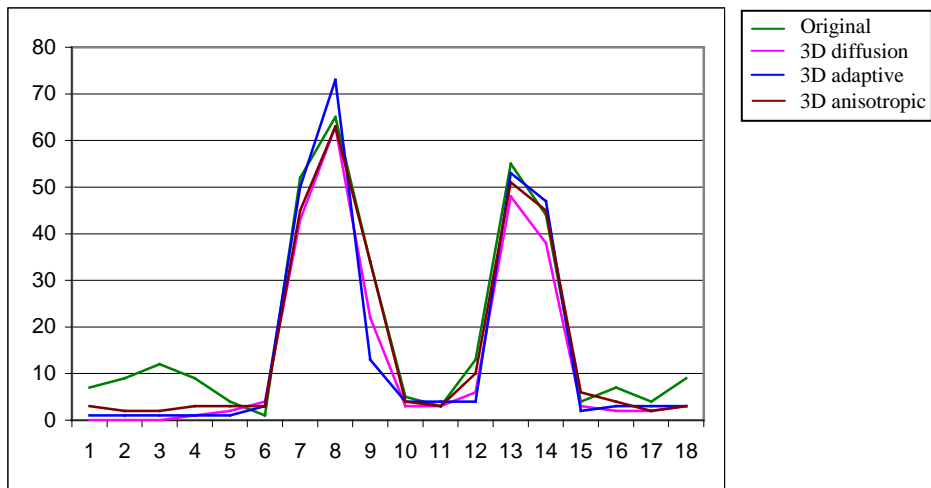
Fig. 1. (a) Slice of the heart MRI dataset. Pixel (b) and (c) gradient intensities are plotted for the highlighted edge illustrated in image (a).
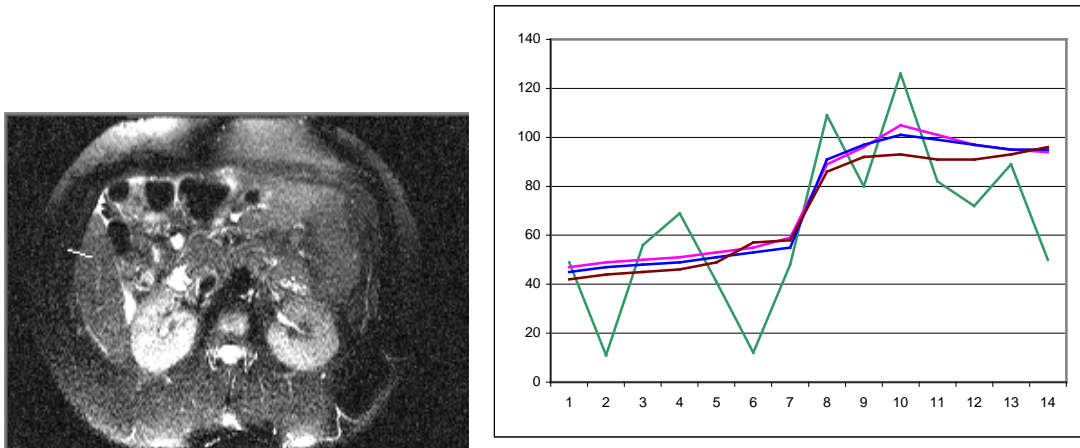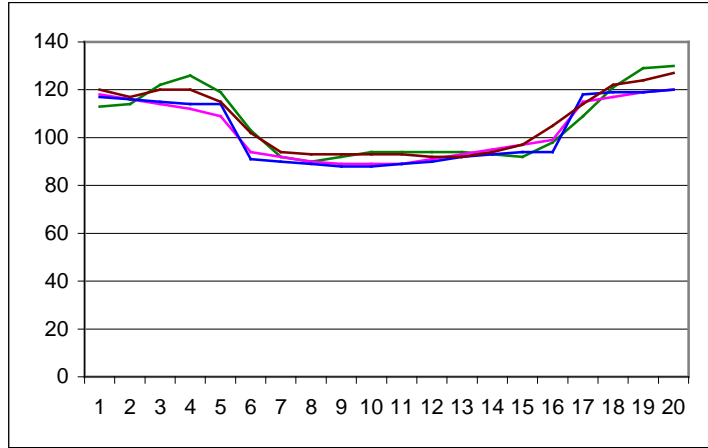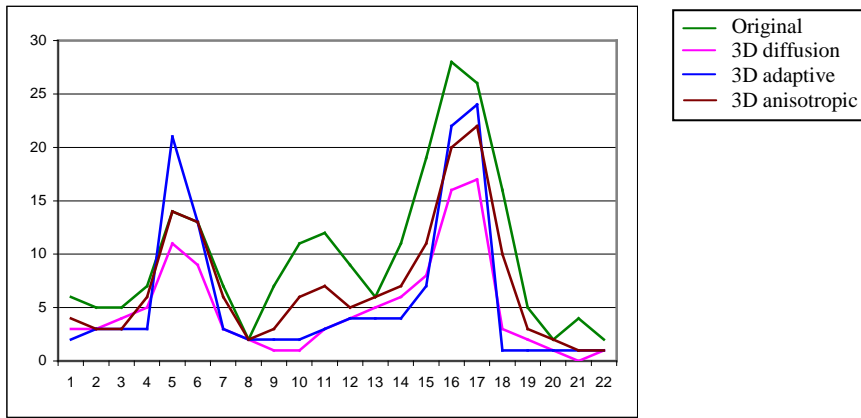
(a)

(b)
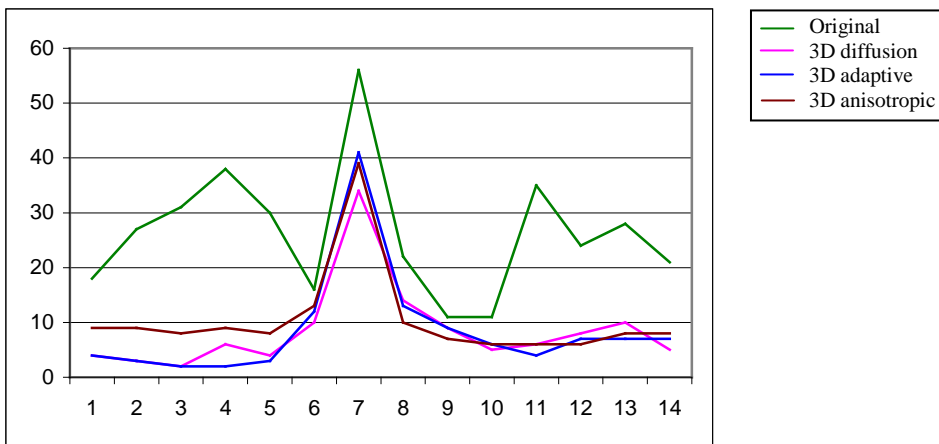
(c)

Fig. 2. (a) Slice of the MRCP dataset. Pixel (b) and gradient intensities (c) are plotted for the highlighted edge illustrated in image (a).
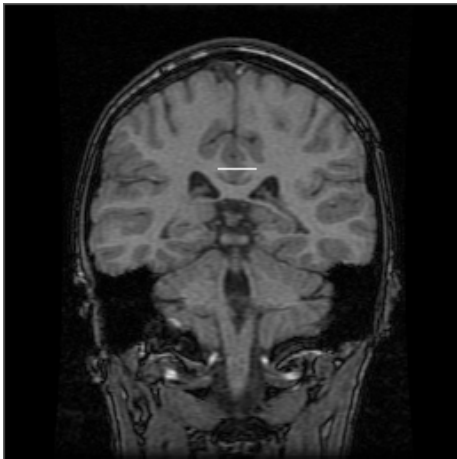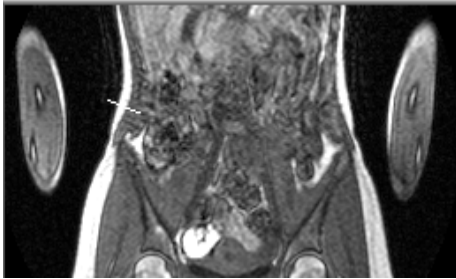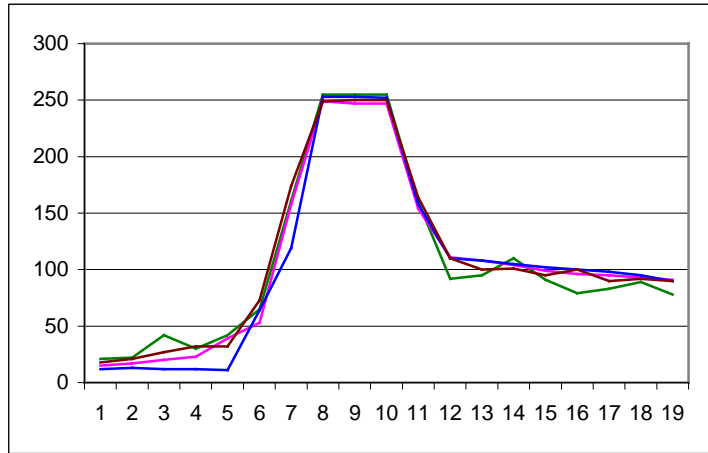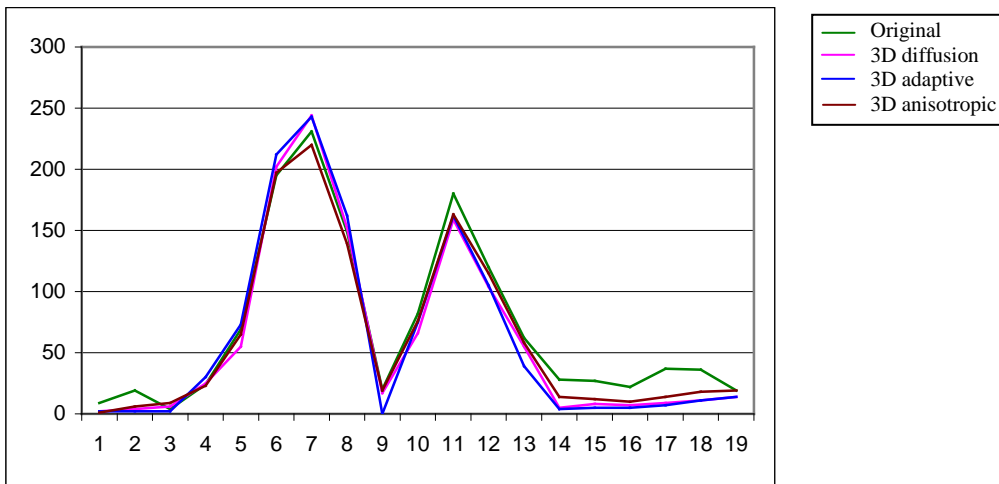
(a)

(b)

(c)

Fig. 3. (a) Slice of the brain MRI dataset. Pixel (b) and gradient intensities (c) are plotted for the highlighted edge in image (a).

(a)

(b)

(c)

Fig. 4. (a) Slice of the whole body MRI dataset. Pixel (b) and gradient intensities (c) are plotted for the highlighted edge illustrated in image (a).
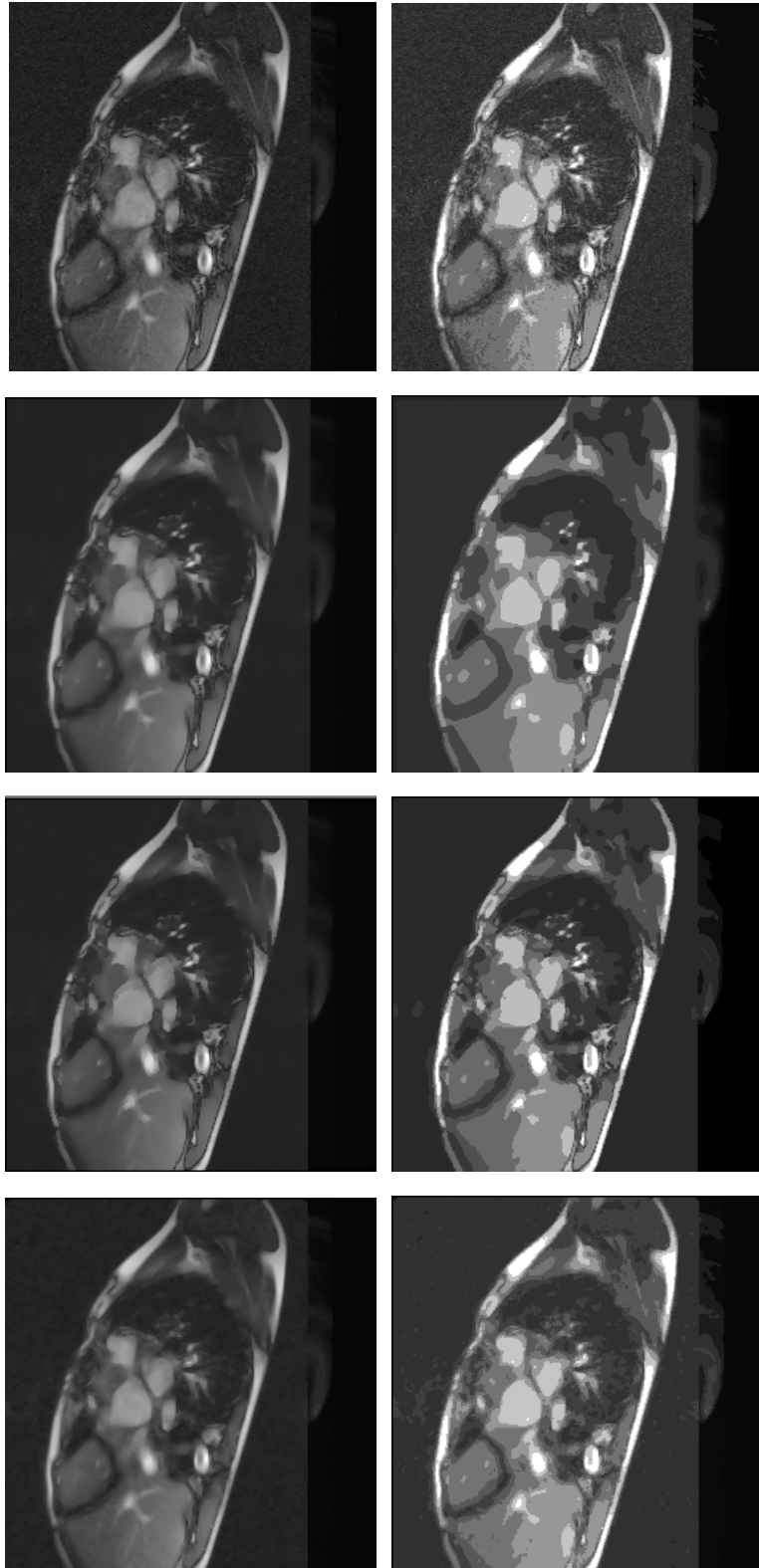
Fig. 5. 3D data clustering results – heart dataset. (First row) Original dataset (slice 9). and corresponding image resulted after clustering. (Second row) 3D diffusion smoothed data (slice 9) and corresponding image resulted after clustering. (Third row) 3D adaptive smoothed data (slice 9) and corresponding image resulted after clustering. (Forth row) 3D anisotropic smoothed data (slice 9) and corresponding image resulting after clustering.
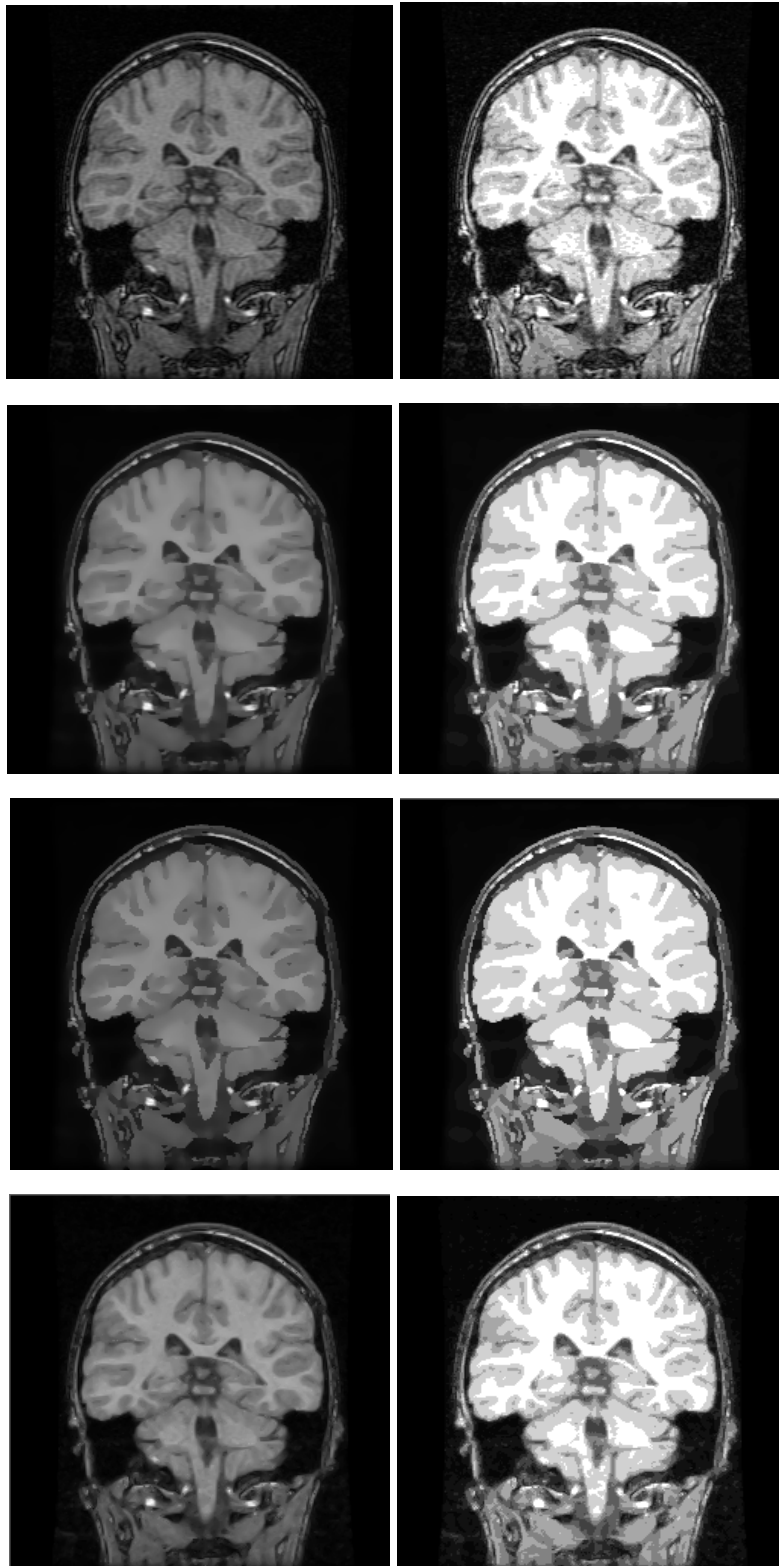
Fig. 6. 3D data clustering results – brain dataset. (First row) Original dataset (slice 4). and corresponding image resulted after clustering. (Second row) 3D diffusion smoothed data (slice 4) and corresponding image resulted after clustering. (Third row) 3D adaptive smoothed data (slice 4) and corresponding image resulted after clustering. (Forth row) 3D anisotropic smoothed data (slice 4) and corresponding image resulting after clustering.
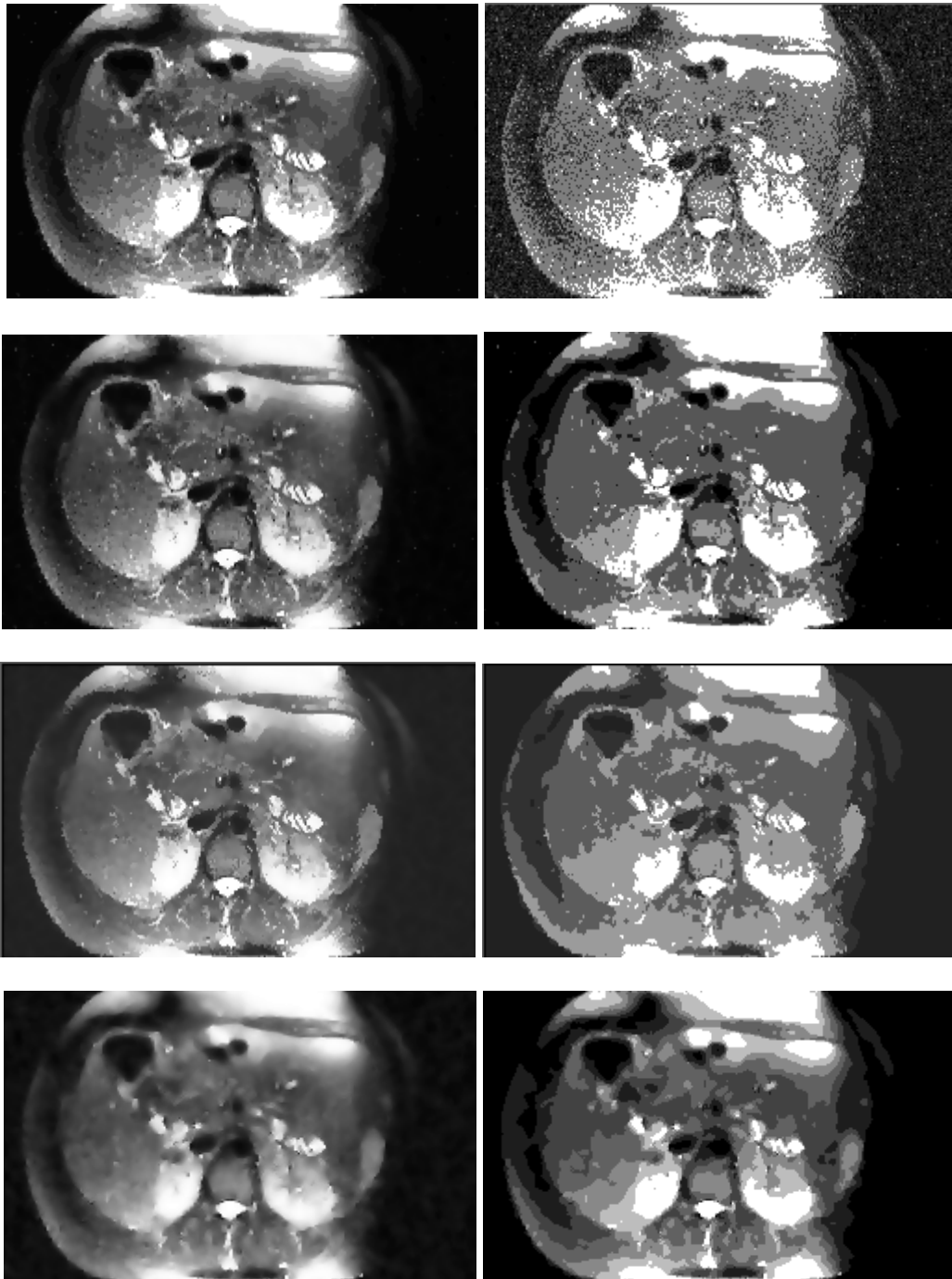
Fig. 7. 3D data clustering results – MRCP dataset. (First row) Original dataset (slice 10). and corresponding image resulted after clustering. (Second row) 3D diffusion smoothed data (slice 10) and corresponding image resulted after clustering. (Third row) 3D adaptive smoothed data (slice 10) and corresponding image resulted after clustering. (Forth row) 3D anisotropic smoothed data (slice 10) and corresponding image resulting after clustering.
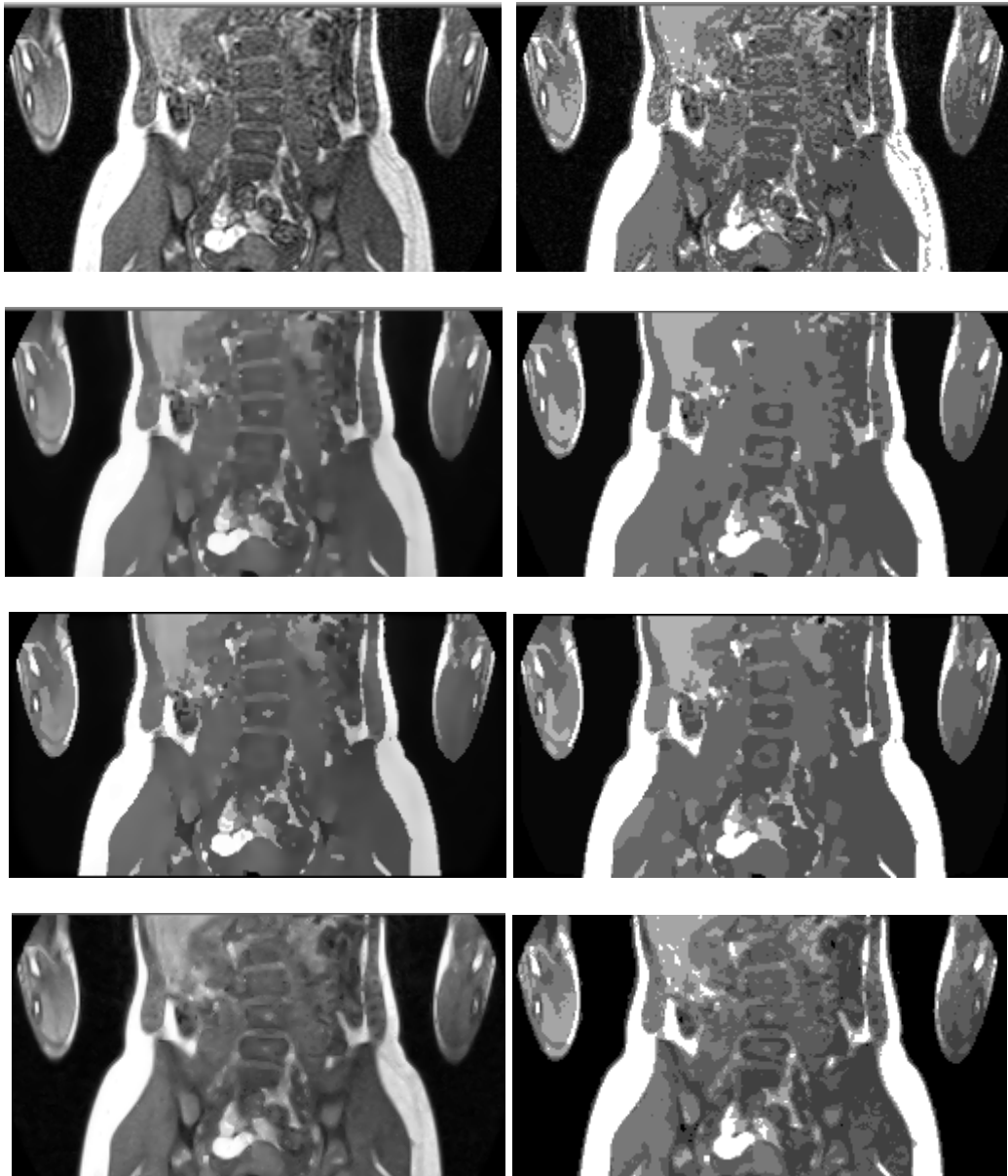
Fig. 8. 3D data clustering results – whole body dataset. (First row) Original dataset (slice 6) and corresponding image resulted after clustering. (Second row) 3D diffusion smoothed data (slice 6) and corresponding image resulted after clustering. (Third row) 3D adaptive smoothed data (slice 6) and corresponding image resulted after clustering. (Forth row) 3D anisotropic smoothed data (slice 6) and corresponding image resulted after clustering.

Noise type	S.Dev. noise	S.Dev. F1	S.Dev. F2	S.Dev. F3.	Greyscale Expected	Greyscale F1	Greyscale F2	Greyscale F3
G-15	13.72	1.91	1.62	2.06	127	127	128	128
G-30	31.93	7.64	3.03	5.57	127	128	129	133
P-15	13.02	1.07	0.76	1.74	127	139	138	138
P-30	26.97	9.60	7.62	3.69	127	141	141	142
W-15	4.63	1.50	0.21	0.69	127	126	127	127
W-30	8.56	1.71	0.60	1.14	127	125	126	127

Table 1. Performance characterisation results when the algorithms have been applied to an artificially created dataset. F1, F2, F3 denote the standard diffusion, adaptive smoothing and anisotropic diffusion respectively.

	Heart	Brain	Whole body	MRCP
Original data	4.95	9.21	20.46	18.80
3D diffusion	1.88	6.28	14.47	10.96
3D adaptive	1.73	6.16	14.05	10.83
3D Anisotropic	2.08	6.48	16.00	9.77

Table 2. The RMS of the standard deviations of the homogenous areas for the original and filtered MRI datasets used in our experiments.