

A New Adaptive Diffusive Function for Magnetic Resonance Imaging Denoising Based on Pixel Similarity

Mostafa Heydari, Mohammad Reza Karami

Department of Biomedical Engineering, Faculty of Electrical and Computer Engineering, Babol University of Technology, Mazandaran, Iran

Submission: 02-02-2015

Accepted: 31-08-2015

ABSTRACT

Although there are many methods for image denoising, but partial differential equation (PDE) based denoising attracted much attention in the field of medical image processing such as magnetic resonance imaging (MRI). The main advantage of PDE-based denoising approach is laid in its ability to smooth image in a nonlinear way, which effectively removes the noise, as well as preserving edge through anisotropic diffusion controlled by the diffusive function. This function was first introduced by Perona and Malik (P-M) in their model. They proposed two functions that are most frequently used in PDE-based methods. Since these functions consider only the gradient information of a diffused pixel, they cannot remove noise in noisy images with low signal-to-noise (SNR). In this paper we propose a modified diffusive function with fractional power that is based on pixel similarity to improve P-M model for low SNR. We also will show that our proposed function will stabilize the P-M method. As experimental results show, our proposed function that is modified version of P-M function effectively improves the SNR and preserves edges more than P-M functions in low SNR.

Key words: *Magnetic Resonance Imaging, Noise, Image Processing, Computer Assisted, Diffusion*

INTRODUCTION

Magnetic resonance imaging (MRI) has very important role on current medical and research procedures. However, these images are normally corrupted by random noise from the measurement process. Therefore, removing noise in MRI is one of the major issues in the study of medical imaging. The traditional methods of image denoising use smoothing (such as average filters and Gaussian filtering) for achieving good results, but the edge information and texture may be lost.^[1] However, medical images consist of many details and subtle features such as angular and edge, and the lack of edge information of image brings difficulties for subsequent analysis. From above, in the image restoration field, we extremely need an image restoration process technology which can remove the noise and keep edge information also. So, partial differential equation (PDE)-based methods are one of the best selections to MRI denoising.

In addition to PDE-based methods, there are many methods that are based on nonlocal means (NLM) filter^[2,3] such as multi component NLM and its variants,^[4]

unbiased NLM,^[5] adaptive NLM,^[6,7] nonparametric principal component analysis NLM,^[8] and MRI denoising using nonlocal PCA.^[9]

The main drawback of the NLM-based algorithms is computational burden due to calculate weighted average for all pixels.^[10] Although this complexity value is reduced in some modified version (by blocking image^[6] or using PCA), but still the calculation volume is high. On the other hand, they cannot remove the noise and simultaneously preserve edges, as well as PDE-based methods, also. In these methods, oversmoothing is observed in some regions which results in a loss of edges and fine structures in the image.^[11] In additional, PDE-based methods are one of the best selections to denoising smooth area and preserving edges simultaneously. These filtering are still widely used in the preprocessing of MRIs because of its efficiency and simplicity in implementation.

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

Address for correspondence:

Dr. Mohammad Reza Karami, Department of Biomedical Engineering,
Faculty of Electrical and Computer Engineering,
Babol University of Technology, Mazandaran, Iran.
E-mail: mkarami@nit.ac.ir

How to cite this article: Heydari M, Karami MR, A New Adaptive Diffusive Function for Magnetic Resonance Imaging Denoising Based on Pixel Similarity. J Med Sign Sense 2015;5:201-9.

Several PDE-based denoising methods were proposed such as Perona and Malik (P-M) method.^[12] They proposed a nonlinear equation which replaced isotropic diffusion expressed through a linear heat equation with an anisotropic diffusion (AD). On the other hand, AD is controlled by the diffusion coefficient (also called *conductance* or *diffusive function*). The diffusive function g tends to “0” while the absolute of image gradient ($|\nabla u|$) is large (edges) and “1” while $|\nabla u|$ is small (flat regions). Therefore the P-M model can effectively remove the noise, as well as containing edge information.

Many revised AD models have been proposed done as stated by P-M diffusion model. Alvarez *et al.*^[13] introduced a mean curvature motion approach for AD. This method performs the diffusion process only in the direction of minimum rate of gradient change. Therefore, the image will be smoothed in a direction parallel to edges, not across it. On the other hand, Catte *et al.*^[14] showed that the P-M AD model is ill-posed, that is, very similar images could produce divergent solutions. To solve the problem correctly, they showed that the diffusion coefficient should be calculated from isotropically smoothed versions of the image.

For noisy images with a low signal-to-noise (SNR) ratio, however, noise variations may be comparable to or even greater than edge variations. In this case, the discrimination between edge and noise can be difficult and then the conventional AD usually fails to produce satisfactory results for low SNR images.^[15,16] Therefore, the kernel AD (KAD) is proposed to improve the performance of AD for low SNR images. In this method, the input space is mapped to a higher order feature space, and then discriminates between edges and noise in the feature space.^[17-19]

The classical P-M’s AD model only considers the local gradient information of each pixel in the image. However, a high gradient magnitude is generally a good indication of edges, whereas a low gradient magnitude may not always point to nonedge regions or noise. For example, the important local details along with edges in the image may have low gradient. In order to retain fine details while removing noise, the local gray-level variance along with the gradient is added to a modified AD model.^[20] According to above discussion, it is obvious that the gradient magnitude alone cannot be a good criterion to discriminate edge, smooth area, and noise. Because of this, Wang *et al.*^[21] introduced a pixel similarity measurement into the PDE-based on mean curvature flow of curve. Using this measurement, each pixel can be classified in five categories: Edge, strong noise, median noise, weak noise, and smooth area. In this paper we utilize similarity measurement to determine edge, noise, and smooth area. In this work, we improve P-M diffusive function by pixel similarity. Our novel diffusive function can stable P-M model. Experimental results show that the proposed function can effectively improves the SNR

and preserves both edges and fine details of objects in the restored image even for low SNR.

In Section 2, the AD is introduced. In this section the P-M model and KAD model that is appropriate for low SNR is studied. Section 3 is dedicated to the pixel similarity measurement. In Section 4, we present an improved diffusive function based on pixel similarity with fractional and adaptive power and then the stability of our function is analyzed. Section 5 introduce two stabled AD. The experimental and comparison results are described in Section 6.

ANISOTROPIC DIFFUSION MODEL

Perona and Malik Model

Diffusion algorithms remove noise from an image by modifying the image via aPDE. The diffusion driven by heat equation is isotropic and image features will be blurred by noise removing. One of the advanced properties of AD is to integrate prior knowledge of the image to the diffusion coefficient. Perona and Malik^[12] first introduced the idea of nonlinear diffusion that is preferred within a smooth region to diffusion near an edge. They proposed the enhancement of noisy image $I(x, y)$ by the solution of the following PDE:

$$\frac{\partial}{\partial t} u(x, y; t) = \text{div}(g(|\nabla u(x, y; t)|)\nabla u(x, y; t)) \quad \text{in } \Omega \times (0, T) \quad (1)$$

$$u(x, y; 0) = I(x, y) \quad \text{in } \Omega$$

$$\frac{\partial u}{\partial \vec{\eta}} = 0 \quad \text{on } \partial\Omega$$

where $u(x, y; t)$ is the enhanced version of the image at time t , Δu denotes the local gradient of u , $\Omega \subset R^2$ is typically a square domain, $\partial\Omega$ means the edges of Ω , $\vec{\eta}$ represents outward normal direction to $\partial\Omega$ and $g: (0, +\infty) \rightarrow (0, 1)$ so called conductance or diffusive function is a nonnegative function of the magnitude of local gradient. This function is chosen to be close to zero where gradient is large, so that the diffusion is stopped across edges. g should be large in pixels with low gradient variation, so that the diffusion is maximal within smooth regions. Functions satisfying these assumptions are commonly called edge stopping functions, and examples of such a function were firstly proposed by Perona and Malik,^[12] e.g.,

$$g_1(s) = \exp(-(s/k)^2) \quad (2)$$

$$g_2(s) = \frac{1}{1+(s/k)^2} \quad (3)$$

where k is the gradient magnitude threshold parameter that controls the rate of the diffusion and serves as a soft threshold between the image gradients that are attributed

to noise and those attributed to edges. According to Perona and Malik, Tsotsios and Petrou,^[12,22] the g_1 favors high-contrast edges over low-contrast ones, while the g_2 function favors wide regions over smaller ones.

In fact, above AD equation (Eq. 1) can be obtained by minimizing following functional.

$$E_{p-M}(u) = \iint_{\Omega} w(|\nabla u|^2) dx dy \tag{4}$$

where $\frac{d}{ds} w(s) = g(s)$ ^[23]

Kernel Anisotropic Diffusion

In low SNR image that the noise magnitude is large, it may be that the diffusive function cannot distinguish between noise and edge. Therefore, the conventional AD usually fails to produce satisfactory results for low SNR images.^[15,16] For low SNR images, the discrimination between edges and noise can be viewed as a nonlinearly separable classification problem. According to Cover's theorem on the separability of patterns, nonlinearly separable patterns in an input space are linearly separable with high probability if the input space is transformed nonlinearly to a feature space with high dimensionality.^[19,24,25] While the gradient operation in the original AD is ineffective for low SNR images, the input space can be mapped to a higher order feature space, and then discriminate between edges and noise in the feature space.

If the input data is represented by $U (U \subseteq R)$, and feature space is denoted by $F (F \subseteq R^n)$, the nonlinear mapping function Φ is defined as:

$$\Phi : U \rightarrow F, u \rightarrow \Phi(u) \tag{5}$$

where u is an input in U , which is mapped into a feature space with dimensionality of n . By substituting the gradient magnitude $|\nabla u|$ in the P-M's AD with $\|\nabla\Phi(u)\|$, the KAD can be obtained as following:^[19]

$$\frac{\partial u}{\partial t} = \text{div}(g(\|\nabla\Phi(u)\|)\nabla u) \tag{6}$$

where $\|\nabla\Phi(u)\|$ is the gradient magnitude in feature space that is defined by the mapping function Φ . It is called as a kernelized gradient operator or a kernelized edge detector, and plays a major role in KAD.

Suppose $\|\nabla u\|_p$ denotes the norm of gradient for pixel p , we can write:

$$\|\nabla u\|_p = \sqrt{\frac{1}{|\partial p|} \sum_{p \in \partial p} \|u_p - u_q\|^2} \tag{7}$$

where ∂p represents the spatial neighborhood of pixel p (eight neighboring pixels around p), $|\partial p|$ refers to cardinality of ∂p ($|\partial p| = 8$) and u_p denotes the intensity of p . Similarly, $\|\nabla\Phi(u)\|_p$ can be represented as Yu *et al.*:^[19]

$$\begin{aligned} \|\nabla\Phi(u)\|_p^2 &= \frac{1}{|\partial p|} \sum_{p \in \partial p} \|\Phi(u_p) - \Phi(u_q)\|^2 \\ &= \frac{1}{|\partial p|} \sum_{p \in \partial p} (K(u_p, u_p) + K(u_q, u_q) - 2K(u_p, u_q)) \end{aligned} \tag{8}$$

where K is the kernel function and thus the mapping function Φ is defined by kernel function K . Commonly used kernel functions are polynomial kernels, radial-basis function kernels, and two-layer perceptron kernels:^[25]

$$K(x, y) = (x^T y + 1)^d \tag{9}$$

$$K(x, y) = \exp\left[-\frac{\|x - y\|^2}{2\sigma^2}\right] \tag{10}$$

$$K(x, y) = \tanh(\beta_0 x^T y + \beta_1) \tag{11}$$

where d, σ, β_0 and β_1 are specified a priori by the user.

PIXEL SIMILARITY AND NEIGHBORING INCONSISTENCY

The pixel similarity measures the similarity between two pixels. By calculating the pixel similarity of center pixel with its neighbors, we can obtain the neighboring consistency of the center pixel. Informally, if a pixel is not consistent enough with its surrounding pixels, it can be considered as a noise, but an extremely strong inconsistency suggests an edge. The pixel similarity is a probability measure; its value is between 0 and 1. Accordingly, the pixel similarity can be used to determine *neighboring inconsistency*. It indicates the inconsistency of a pixel with its neighbors. Thus, not only we distinguishing noise and nonnoise pixels, but also we divide pixels into five categories namely smooth area, weak noise, medium noise, strong noise, and edge.^[26]

If $p_i (i = 1, \dots, 8)$ denotes the neighboring pixels of pixel p , the pixel similarity between p and p_i is defined as:

$$f(p, p_i) = \exp(-\lambda |u_p - u_{p_i}|^a) \tag{12}$$

where u_p, u_{p_i} are the intensity of pixels p, p_i and λ, a are constants (typically $\lambda = 1$ and $a = 2$). The pixel similarity measures the similarity between two pixels. Then the degree of consistency of the center pixel with its surrounding pixels can also be calculated. The idea is, if a pixel is not consistent enough with its surrounding pixels; it can be considered as a noise. Therefore, by using a threshold we can obtain the consistency or inconsistency of center pixel

with its neighbors. This threshold that is named confidence level is a predetermined constant ($0 \leq \beta \leq 1$ and typically $\beta = 0.5$). Using this threshold we can define the consistency discrimination function d :

$$d(p, p_i) = \begin{cases} 1 & f(p, p_i) \leq \beta \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

where $d(p, p_i) = 1$ means p and p_i are inconsistent. By defining $M = \sum_{i=1}^8 d(p, p_i)$, the degree of inconsistency of center pixel with its neighbors is determined such that if $M > L$ (L is a predetermined constant and typically $L = 5$), then p is low consistent (high inconsistent) with its neighboring pixels and is considered as a noise. In this case the center pixel p is called a low consistency pixel under the confidence level β .^[26]

Finally, in MRIs, if a pixel is not consistent enough with its surrounding pixels, it can be considered as a noise, but an extremely strong inconsistency suggests an edge. By detecting all kinds of low consistency pixels under different levels of confidence, the following thresholding rule that is shown in Table 1 is found:

DIFFUSIVE FUNCTION BASED ON PIXEL SIMILARITY

In Figure 1 diffusive functions g_1 and g_2 are plotted in terms of $|\nabla u|$. As can be deduced from this figure, the value of g_1 and g_2 , even for not very large values of $|\nabla u|$, sharply (strictly) decrease. Thus, the diffusive process will stop and noise will remain. This result is due to the exponential operator in g_1 and second ordering of $|\nabla u|/K$ in g_2 . According to^[12,22] function g_1 is appropriate for high contrast images, while noisy MRIs especially in high detail regions have not high contrast.

Table 1: Category division for low consistency pixels under different confidence levels

Confidence levels	$0 < \beta \leq 0.3$	$0.3 < \beta \leq 0.4$	$0.4 < \beta \leq 0.5$	$0.5 < \beta \leq 0.6$	$0.6 < \beta \leq 1$
Categories	Edge	Strong noise	Median noise	Weak noise	Smooth area

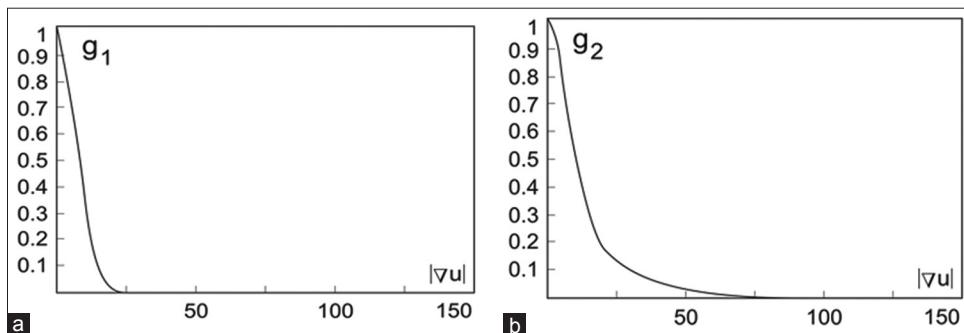


Figure 1: Diffusive functions g_1 (a) and g_2 (b) in terms of $|\nabla u|$

The power of $|\nabla u|/K$ in g_2 is a constant value, whereas this value can be more flexible and can be changed. According to this fact, we propose a new diffusive function with fractional power. In our function the real number ν is used for this power. Thus, function g_2 will change as follows:

$$g(s) = \frac{1}{1+(s/K)^\nu}, \nu \geq 0 \quad (14)$$

If ν is negative, the diffusion process will stop in smooth regions and will fully perform in edges. The value of ν is obtained by stability condition for P-M model that is discussed following.

When we use PDEs to carry through image processing, by supposing that $\nabla u \neq 0$, $\vec{\eta} = \nabla u / |\nabla u|$ and $\vec{\xi} \perp \vec{\eta}$, where $\vec{\eta}$ and $\vec{\xi}$ is unit tangent and normal vector respectively, we can get the following expressions:

$$\vec{\eta} = \frac{\nabla u}{|\nabla u|} = \frac{(u_x, u_y)}{\sqrt{u_x^2 + u_y^2}} \quad (15)$$

$$\vec{\xi} = \frac{\nabla^\perp u}{|\nabla u|} = \frac{(-u_y, u_x)}{\sqrt{u_x^2 + u_y^2}} \quad (16)$$

Then we can get:^[27]

$$\begin{aligned} \frac{\partial u}{\partial t} &= [g'(|\nabla u|)|\nabla u| + g(|\nabla u|)]u_{\eta\eta} + g(|\nabla u|)u_{\xi\xi} \\ &= g_\eta(|\nabla u|)u_{\eta\eta} + g_\xi(|\nabla u|)u_{\xi\xi} \end{aligned} \quad (17)$$

where $u_{\eta\eta}$, $u_{\xi\xi}$ and g_η , g_ξ are the second order derivatives and diffusive functions in $\vec{\eta}$ and $\vec{\xi}$ direction, respectively. The Eq. 17 is called as the global scheme for PDE-based restoration approaches.^[28]

Obviously, $g_\eta(|\nabla u|) = g(|\nabla u|)$ in Eq. 17 is always positive which means diffusion forward and making image smooth. However, $g_\xi(|\nabla u|) = g'(|\nabla u|)|\nabla u| + g(|\nabla u|)$ can be either positive or negative leading to unstable results, this causes instability of the diffusion process.^[27,29,30] Considering P-M functions of Eq. 2 and Eq. 3, it appears that their

corresponding g_η functions, in the global scheme of Eq. 17, can sometimes takes negative values [Figure 2]. This leads to local instabilities of the P-M's PDE which degrades the processed image instead of denoising it. As shown in Figure 2, g_2 is more stable than g_1 , and for this reason; diffusive function g_2 is utilized in this paper.

As mentioned in above, for the stability of P-M model, g_η and g_ξ must be nonnegative. g_ξ is always positive and by doing some mathematical calculations, it can be found that g_η will be nonnegative if:

$$(1 - \nu) (s/K)^\nu \leq 1 \tag{18}$$

This condition is valid for $|\nabla u| < K$ and practical value of ν , but for $|\nabla u| \geq K$ that higher than the previous case occurs, the following relationship must be established:

$$1 \leq \nu \leq 2 \tag{19}$$

As describe in Section 2, the main weakness of P-M model is laid in discriminating between edges and noise such that the noise may be assigned to the edge for low SNR images and then the noise remains. In order to solve this problem, ν can be a function of neighboring inconsistency. So that, if pixel p assigned to edge, ν should be tend to 2. Thus the value of diffusion function more rapidly tends to zero. On the other hand, if p assigned to noise ν should be tending to 1 that prevents from reducing diffusive function value. In this case, since the value of ν decreases, the diffusion process cannot be stopped completely when $|\nabla u|$ is large but the pixel is assigned to noise (in low SNR images).

According to above discussion, we propose the following decreasing function for ν :

$$\nu_p = 1 + \frac{\sum_{i=1}^8 d(p, p_i)}{8} \tag{20}$$

and then our diffusive function is obtained as:

$$g(|\nabla u|_p) = \frac{1}{1 + (|\nabla u|_p / k)^{\frac{\sum_{i=1}^8 d(p, p_i)}{8}}} \tag{21}$$

We name this function as the pixel similarity-based diffusive function. As can be deduced from Eq. 21, the proposed function is not only a function of the gradient magnitude and its value is related to neighboring inconsistency. Therefore, the slope of diffusive function is controlled with neighboring inconsistency. It will prevent from remaining noise in relatively low SNR images.

OTHER STABLED ANISOTROPIC DIFFUSION

As can be deduced from subsection 2.2, the main weakness of P-M model is occurred in high noise level images where the diffusive function cannot distinguish between noise and edge. Therefore, we use the consistency discrimination function to classify pixels into five groups: Smooth area, weak noise, medium noise, strong noise, and edge. This causes noisy pixels and edges to be easily distinguished from each other, even in noisy images with high noise level. On the other hand, we use consistency discrimination function in diffusive function to stable diffusion process. In other words, our proposed method will causes the conventional AD is stable. For this reason, it is essential that the proposed method is compared with the some stabled method, in addition to the classical P-M and KAD models.

Nonlocal Orientation Diffusion Partial Differential Equation

In Qiao and Zou^[28] In order to achieve more stability and yet preserve edge information of the image, the diffusion process is only performed in the direction that is parallel to the image edges and can be limited in gradient direction. In this method, g_η is supposed zero and therefore does not take the negative values. Then the Eq. 1 may be changed to:

$$\frac{\partial u}{\partial t} = \alpha(r) \cdot g(|\nabla u|) \cdot u_{\xi\xi} \tag{22}$$

Where $\alpha(r)$ that is called nonnegative monotony digression function is related with NLM of gradient g_η . This function can control diffusion velocity. When the nonlocal area of image is flat, there may be less nonlocal gradient operator,

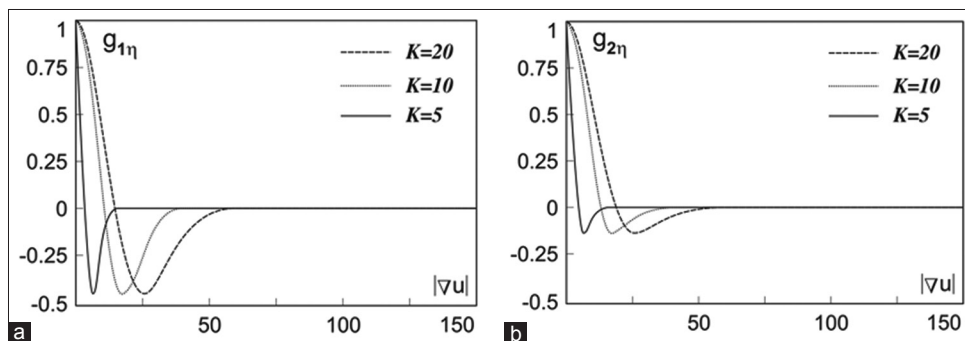


Figure 2: Investigation of the instability of Perona and Malik functions for different value of K: (a) $g_{1\eta}$ for Eq. 2, (b) $g_{2\eta}$ for Eq. 3

and consequently the value of $\alpha(r)$ is increased. This will make to a faster diffusion. When the image contains more edge features in the nonlocal areas, there may be larger nonlocal gradient operator, and it may decrease diffusion velocity to preserve the edge information. By selecting an appropriate value for $\alpha(r)$ to control diffusion velocity in Eq. 22, we can wipe off the image noise better and preserve the edge information effectively.

Diffusive Function Based on Double Well Potential Function

In Histace and Menard,^[31] authors propose a new diffusive function to stable the P-M model:

$$g_{DW}(|\nabla u|) = 1 - \psi(|\nabla u|_N) \quad (23)$$

where $|\nabla u|_N$ denotes the normalized gradient magnitude and $\psi(\cdot)$ is double well potential. In Figure 3, some graphical representations of $\psi(\cdot)$ for different values of α are shown.

Now, if the mathematical expression of $g_{\eta_{DW}}$ is calculated, one can obtain that:

$$g_{\eta_{DW}}(|\nabla u|) = g'_{DW}(|\nabla u|) |\nabla u|_N + g_{DW}(|\nabla u|) \quad (24)$$

Since $|\nabla u|_N \in [0, 1]$ and $g'_{DW}(|\nabla u|)$ is a one-order-less polynomial function than $g_{DW}(|\nabla u|)$, it can be concluded that:^[31]

$$g_{\eta_{DW}}(|\nabla u|) \approx g_{DW}(|\nabla u|) = g_{\zeta_{DW}}(|\nabla u|) \quad (25)$$

By taking Eq. 23, it is clear that corresponding functions in the tangential and orthogonal directions never takes negative values. Thus, diffusive process stays stable for all gradient values of processed image which is of primary importance.

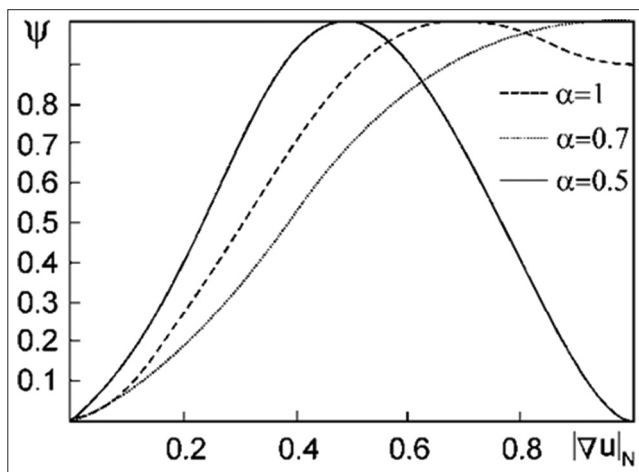


Figure 3: Double well potential function in terms of normalized $|\nabla u|$ for different value of α

According to Eq. 23, $g_{\eta_{DW}}$ and $g_{\zeta_{DW}}$ are almost the same. Therefore, $g_{\eta_{DW}}$ tends to zero when $g_{\zeta_{DW}}$ is close to zero. As a result, by using a right choice for α , it becomes possible to completely stop the diffusion process in the tangential and orthogonal directions of the contours, simultaneously.

RESULTS AND ANALYSIS OF EXPERIMENT

In this section, we compare our proposed diffusive model with P-M, KAD, nonlocal orientation diffusion (NLOD) and double well potential (DWP) models. We perform experiments on T1-weighted MRIs with 8-bit quantization (1000 images) to show the performance of our proposed model. Our samples are real MRIs from McConnell Brain Imaging Centre of the Montreal Neurological Institute. These images are destroyed by additive Gaussian noise with zero mean and various standard deviations.

Noise in MRI can be Gaussian or Rician distributed depending on the image type (real or magnitude data).^[4,32] Since all images in our experiments are real data, they are corrupted with Gaussian-distributed random noise. Although MRIs are usually magnitude data and their noise obeys a Rician distribution^[20,33,34] but at high SNR (let us define SNR as I/σ), the Rician noise will be described by the Gaussian distribution. In other word, If the SNR is sufficiently high ($I/\sigma \geq 2$,^[29] 3^[20,33] or 4^[35] or 5^[30] which are usually sufficient for any practical purposes), the Rician distribution can be approximated by the Gaussian distribution.

To measure the quality of the recovered image, the criterions of SNR and improved SNR (ISNR) are employed here. For any given image u , SNR and ISNR are characterized by:

$$SNR = 10 \log \frac{\sum_{x,y} (A(x,y))^2}{\sum_{x,y} (u(x,y) - A(x,y))^2} \quad (26)$$

$$ISNR = 10 \log \frac{\sum_{x,y} (I(x,y) - A(x,y))^2}{\sum_{x,y} (u(x,y) - A(x,y))^2} \quad (27)$$

where A , I and u are the original, noisy and restored image, respectively. Generally, the larger the SNR value, the better the image quality. Another criterion that we use in this paper is mean square error (MSE) that is defined as:

$$MSE = 10 \log \left[\frac{1}{N} \sum [u(x,y) - A(x,y)]^2 \right] \quad (28)$$

where N is the number of pixels in the image. In order to evaluate the edge-preserving ability, we propose new measurement which calculates the MSE of edge pixels (edge MSE):

$$EMSE = 10 \log \left[\frac{1}{N_B} \sum_{(x,y) \in B} [u(x,y) - A(x,y)]^2 \right] \quad (29)$$

In which, B denotes the edges of original image A and N_B is the number of pixels in B . We also adopt the structure similarity index measure (SSIM) to measure the similarity between two images,^[30] which is defined as:

$$SSIM = \frac{(2\mu_A\mu_u + C_1)(2\sigma_{A,u} + C_2)}{(\mu_A^2 + \mu_u^2 + C_1)(\sigma_A^2 + \sigma_u^2 + C_2)} \quad (30)$$

where μ_A and μ_u are the average of A and u , σ_A and σ_u are the variance of A and u , $\sigma_{A,u}$ is the covariance of A and u , finally c_1, c_2 are two constants to avoid instability.

In Figure 4, we compare the proposed model with P-M, KAD, NLOD, and DWP methods. In this comparison, the gradient threshold and iteration number are selected as $K = 6$ and $T = 10$, respectively. Figure 4a is the original T1-weighted MRIs and Figure 4b presents the noisy Image by random Gaussian white noise with $SNR = 10$ db. Figure 4c shows the restoration result from the P-M model, in which the noise cannot be effectively eliminated. The result of KAD model with kernel function from Eq. 10 is shown in Figure 4d. The result of NLOD model with $\alpha(r) = 1/(1+r^2)$ is shown in Figure 4e. Figure 4f shows the denoising results with DWP function with $\alpha = 0.7$. Figure 4g shows the denoising results with proposed model by setting $a = 2$,

$\lambda = 0.1$ and $\beta = 0.2$. It can be observed that our model performs better than the other methods to eliminate noise and preserve edges. As can be seen in the Figure 4c-g, the ability of the proposed method in edge preserving and noise reduction is much more than the others.

Table 2 summarizes the SNR, ISNR, MSE, EMSE, and SSIM performance of the P-M, KAD, NLOD, DWP, and proposed methods for 1000 MRIs from McConnell Brain Imaging Centre of the Montreal Neurological Institute (MNI, McGill University). These images are composed of 20 different clean (without noise) images with 50 slices that are captured from the same machine. The experimental results show that the proposed model can significantly improve the measurements. From these data, the absolute superiority of the proposed method is obvious. In our method, measurements ISNR, MSE, EMSE, and SSIM are improved more than 1.6 db, 1.6 db, 1 db, and 2%, respectively. However, these improvements are given to about 5.2 db, 5.2 db, 3.8 db and 21% for ISNR, MSE, EMSE, and SSIM, respectively, in comparison with P-M model. In addition, these data indicate that the proposed method can simultaneously reduce noise and preserve edge, much more than the other methods. These advantages are due to the selection of a proper method for stabilizing the diffusion process. In Table 3, we

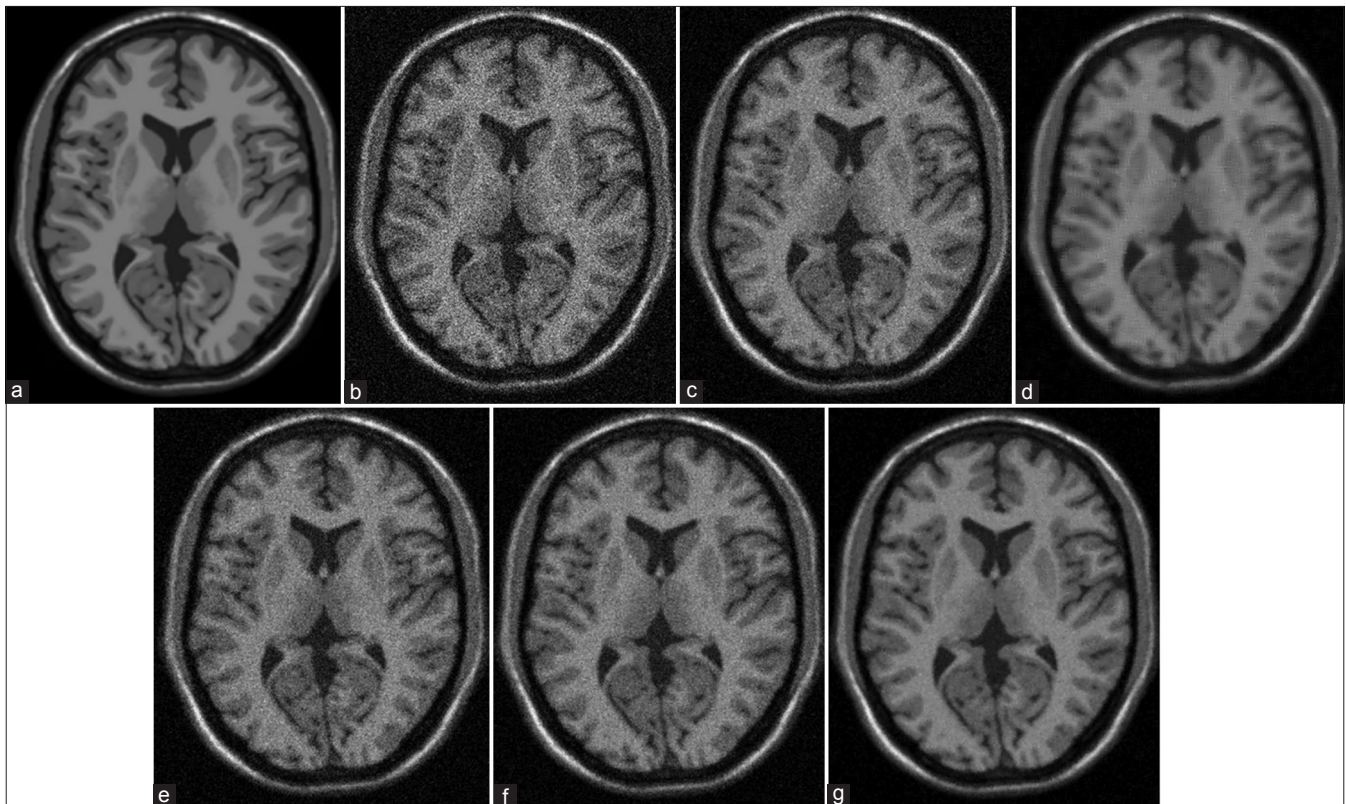


Figure 4: Comparison of proposed diffusive function with Perona and Malik function and kernel anisotropic diffusion, nonlocal orientation diffusion, double well potential methods for $T = 10$ and $K = 6$: Original image (a), noisy image by Gaussian white noise with signal-to-noise = 10 db (b), output of Perona and Malik model (c), restored image by kernel anisotropic diffusion (d), output of nonlocal orientation diffusion method (e), restored image by double well potential function with $\alpha = 0.7$ (f), result of proposed diffusive function (g)

employ P-M, KAD, NLOD, DWP, and proposed models for 1000 MRIs, and then average the values of ISNR, EMSE, and SSIM value for all images. As expected, this table confirms the previous results. We have also performed additional experiments in Table 4 for different σ values and 1000 MRIs to show the efficiency of the proposed models in higher noise levels. In this table, the additive noise is supposed white Gaussian noise. Experimental results show that the

Table 2: Comparison of proposed function with P-M, KAD, NLOD and DWP methods

Figure	4.b	4.c	4.d	4.e	4.f	3.g
SNR (db)	10	12.48	16.09	14.94	16.16	17.73
ISNR (db)	0	2.48	6.09	4.94	6.16	7.73
MSE (db)	30.50	28.01	24.41	25.74	23.83	22.76
EMSE (db)	30.45	29.13	26.32	26.92	26.04	25.38
SSIM (%)	55.67	64.42	83.57	82.19	84.49	85.80

SNR – Signal-to-noise; ISNR – Improved signal-to-noise; MSE – Mean square error; EMSE – Excess mean square error; SSIM – Structure similarity index measure; P-M – Perona and Malik; KAD – Kernel anisotropic diffusion; NLOD - Nonlocal orientation diffusion; DWP – Double well potential

Table 3: Comparison of proposed function with P-M function and KAD, NLOD, DWP methods for 1000 MR images

Figure	Noisy Image	Output of P-M	Output of KAD	Output of NLOD	Output of DWP	Result of proposed function
Averaged ISNR (db)	0	1.72	5.74	4.86	5.80	6.21
Averaged EMSE (db)	31.49	29.70	27.41	27.85	25.38	24.92
Averaged SSIM (%)	52.32	61.97	80.57	80.29	80.83	81.69

EMSE – Excess mean square error; SSIM – Structure similarity index measure; P-M – Perona and Malik; KAD – Kernel anisotropic diffusion; NLOD - Nonlocal orientation diffusion; DWP – Double well potential; ISNR – Improved signal-to-noise; MR – Magnetic resonance

Table 4: Comparison of proposed function with P-M function and KAD, NLOD, DWP methods for high noise levels (these results are the averaged values for 1000 MR images)

Figure	Output of P-M	Output of KAD	Output of NLOD	Output of DWP	Result of proposed function
Noisy image SNR=10					
ISNR (db)	1.72	5.74	4.86	5.80	6.21
EMSE (db)	29.70	27.41	27.85	24.29	24.92
Noisy image SNR=5					
ISNR (db)	0.83	2.43	2.41	2.82	3.58
EMSE (db)	34.76	34.25	34.59	33.76	32.70
Noisy image SNR=0					
ISNR (db)	0.01	0.27	0.20	0.98	1.12
EMSE (db)	40.52	40.18	40.47	40.02	39.41

P-M – Perona and Malik; KAD – Kernel anisotropic diffusion; NLOD - Nonlocal orientation diffusion; DWP – Double well potential; ISNR – Improved signal-to-noise; MR – Magnetic resonance; EMSE – Excess mean square error; SNR – Signal-to-noise

proposed method can effectively remove noise while well preserving the sharp edges and fine details of a noisy image even if the noise level is high.

CONCLUSIONS

In this paper, the instability of classical P-M model is analyzed and an improved diffusive function is proposed. This new function is based on pixel similarity and it is an improved version of P-M function with fractional and adaptive power. In the proposed function, the decreasing rate of diffusive function is controlled with the power of $|\nabla u|/K$. Such that this power (ν) increases in edge and decreases for smooth and noisy areas.

Since the P-M functions are only the function of gradient magnitude and $|\nabla u|$ is large for low SNR images, the value of g tends to zero and then the noise remains. On the one hand, the noisy image is mapped to feature space and therefore, the edges of the image are not available. As can be seen from Figure 4d, the restored image in KAD is blurred.

The proposed method uses two criteria for the classification of pixels: Gradient magnitude and pixel similarity. Unlike the P-M functions that a pixel is classified in two categories, a pixel is classified into five categories in our diffusive function. Therefore, if the gradient magnitude of pixel is large but it is not noise, the diffusive function does not decrease and then the diffusion process is not stopped. Since proposed function is directly performed in input space (nor the feature space), it has much better edge-preserving than KAD. Our function allows the classical P-M model to be stable, also. Experimental results show that the proposed function can effectively improve the ISNR, MSE, and SSIM and preserve both edges and fine details of objects in the restored image. Advantages of the proposed function are apparent even for higher noise levels.

REFERENCES

- Li B, Que D. Medical images denoising based on total variation algorithm. *Procedia Environ Sci* 2011;8:227-34.
- Buades A, Coll B, Morel JM. A review of image denoising algorithms, with a new one. *Multiscale Model Simul* 2005;4:490-530.
- Buades A, Coll B, Morel JM. A non-local algorithm for image denoising. *Proc IEEE Trans IEEE Conf Comput Vis Pattern Recognit* 2005;2:60-65.
- Manjón JV, Thacker NA, Lull JJ, Garcia-Martí G, Martí-Bonmatí L, Robles M. Multicomponent MR image denoising. *Int J Biomed Imaging* 2009;2009:756897.
- Manjón JV, Carbonell-Caballero J, Lull JJ, García-Martí G, Martí-Bonmatí L, Robles M. MRI denoising using non-local means. *Med Image Anal* 2008;12:514-23.
- Manjón JV, Coupé P, Martí-Bonmatí L, Collins DL, Robles M. Adaptive non-local means denoising of MR images with spatially varying noise levels. *J Magn Reson Imaging* 2010;31:192-203.
- Descoteaux M, Wiest-Daesslé N, Prima S, Barillot C, Deriche R. Impact of Rician adapted non-local means filtering on HARDI. *Med Image*

- Comput Comput Assist Interv 2008;11(Pt 2):122-30.
8. Kim DW, Kim C, Kim DH, Lim DH. Rician nonlocal means denoising for MR images using nonparametric principal component analysis. *EURASIP J Image Video Process* 2011;15:1-8.
 9. Manjón JV, Coupé P, Buades A. MRI noise estimation and denoising using non-local PCA. *Med Image Anal* 2015;22:35-47.
 10. Mohan J, Krishnaveni V, Guo Y. A survey on the magnetic resonance image denoising methods. *Biomed Signal Process Control* 2014;9:56-69.
 11. Bhujle HV, Chaudhuri S. Laplacian based non-local means denoising of MR images with Rician noise. *Magn Reson Imaging* 2013;31:1599-610.
 12. Perona P, Malik J. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans Pattern Anal Mach Intell* 1990;12:629-39.
 13. Alvarez L, Lions PL, Morel JM. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM J Numer Anal* 1992;29:845-66.
 14. Catte F, Lions PL, Morel JM, Coll T. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM J Numer Anal* 1992;29:182-93.
 15. Ling J, Bovik AC. Smoothing low-SNR molecular images via anisotropic median-diffusion. *IEEE Trans Med Imaging* 2002;21:377-84.
 16. Torkamani-Azar F, Tait KE. Image recovery using the anisotropic diffusion equation. *IEEE Trans Image Process* 1996;5:1573-8.
 17. Chen S, Zhang D. Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure. *IEEE Trans Syst Man Cybern B Cybern* 2004;34:1907-16.
 18. Takahashi N, Nishi T. Global convergence of decomposition learning methods for support vector machines. *IEEE Trans Neural Netw* 2006;17:1362-9.
 19. Yu J, Wang Y, Shen Y. Noise reduction and edge detection via kernel anisotropic diffusion. *Pattern Recognit Lett* 2008;29:1496-503.
 20. Chao SM, Tsai DM. An improved anisotropic diffusion model for detail- and edge-preserving smoothing. *Pattern Recognit Lett* 2010;31:2012-23.
 21. Wang Y, Shi C, Li C, Wang P, Xia D. Mean curvature flow of curve under fuzzy rule for noise removal. *Comput Eng Appl (Chin)* 2004;33:18-20.
 22. Tsiotsios C, Petrou M. On the choice of the parameters for anisotropic diffusion in image processing. *Pattern Recognit* 2013;46:1369-81.
 23. Zhou B, Mu CL, Feng J, Wei W. Continuous level anisotropic diffusion for noise removal. *Appl Math Model* 2012;36:3779-86.
 24. Cover TM. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Trans Electron Comput (EC)* 1965;14:326-34.
 25. Haykin S. *Neural Networks – A Comprehensive Foundation*. USA: Prentice Hall; 1999.
 26. Jin R, Song E, Zhang L, Min Z, Xu X, Huang CC. Denoising of brain MRI images using modified PDE model based on pixel similarity. *Proc SPIE* 2008;691428.1-8.
 27. Yang M, Liang J, Zhang J, Gao H, Meng F, Xingdong L, *et al*. Non-local means theory based Perona – Malik model for image denoising. *Neurocomputing* 2013;120:262-7.
 28. Qiao N, Zou B. Nonlocal orientation diffusion partial differential equation model for optics image denoising. *Optik* 2013;124:1889-91.
 29. Deriche R, Faugeras O. Les edp en traitements des images et visions par ordinateur. *Traitement Signal* 1996;13:551-78.
 30. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: From error visibility to structural similarity. *IEEE Trans Image Process* 2004;13:600-12.
 31. Histace A, Menard M. MR image enhancement: A PDE-based approach integrating a double-well potential function for thin structure preservation. *INSTICC press* 2010;1:17-21.
 32. Sijbers J, den Dekker AJ. Maximum likelihood estimation of signal amplitude and noise variance from MR data. *Magn Reson Med* 2004;51:586-94.
 33. Shah J. A common framework for curve evolution, segmentation and anisotropic diffusion. *IEEE Conf Comput Vis Pattern Recognit* 1996; 136-42.
 34. Mumford D, Shah J. Optimal approximations by piecewise smooth functions and associated variational problems. *Commun Pure Appl Math* 1998;42:577-685.
 35. Maximov II, Farrher E, Grinberg F, Shah NJ. Spatially variable Rician noise in magnetic resonance imaging. *Med Image Anal* 2012;16:536-48.

BIOGRAPHIES



Mostafa Heydari has received his Bs.c & Ms.c in electrical and electronic engineering 2004 & 2006 respectively. Now he is PhD candidate in Babol Noshirvani Institute of Technology, department of electrical and electronic engineering. He published 9 articles in journals and conferences. His research interests are digital signal processing, biomedical signal processing and image processing.

E-mail: m.heydari@stu.nit.ac.ir



Mohammad Reza Karami has received the Bs.c in electrical and electronic engineering in 1992, Ms.c of signal processing in 1994, and PhD in 1998 in biomedical engineering from I.N.P.L d’Nancy of France. He is now the Associate professor with the Department of Electrical and Computer Engineering, Babol University of Technology. Since 1998 his research is in signal and speech processing. He published more than 100 articles in journals and conferences. He teaches Digital signal processing, biomedical signal processing and speech processing in university. His research interests include Speech, Image and signal processing.

E-mail: mkarami@nit.ac.ir