A denoising PDE model based on isotropic diffusion and total variation models

Neda Mohamadi
Department of Mathematics and Statistics, Mashhad Branch, Islamic Azad University, Mashhad, Iran.
E-mail: mohamadi_37@yahoo.com

Ali Reza Soheili*
Department of Applied Mathematics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran.
E-mail: soheili@um.ac.ir

Faezeh Toutounian
Department of Applied Mathematics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran.
E-mail: toutouni@math.um.ac.ir

Abstract
In this paper, a denoising PDE model based on a combination of the isotropic diffusion and total variation models is presented. The new weighted model is able to be adaptive in each region in accordance with the image's information. The model performs more diffusion in the flat regions of the image, and less diffusion in the edges of the image. The new model has more ability to restore the image in terms of peak signal to noise ratio and visual quality, compared with total variation, isotropic diffusion, and some well-known models. Experimental results show that the model is able to suppress the noise effectively while preserving texture features and edge information well.

Keywords. Image denoising, Isotropic diffusion, Total variation, Partial differential equation.

2010 Mathematics Subject Classification. 2010 Mathematics Subject Classifications. 65M06, 68U10.

1. INTRODUCTION

Image denoising is an important subject in computer vision and image processing systems. The aim of all denoising methods is to effectively suppress noise while keeping intact the features of the image. In order to recover a digital image that has been contaminated by noise, a variety of methods have been proposed [34, 37, 38, 39]. Some linear filtering methods [26, 27] have been suggested to remove Gaussian and uniform noise in images. Other commonly used linear filtering methods are Wiener filter [20] and Mean filter [12, 32]. Nonlinear image filters [5, 6] have emerged to improve the effectiveness of linear filters, where the median filter is the most used nonlinear filtering [30]. Various wavelet-based techniques have also been proposed for image denoising [8, 10, 24]. Image denoising techniques based on partial differential
equation and Computational Fluid Dynamics (CFD) have been developed, such as Total Variation (TV) methods [3, 7, 15, 25], level set methods [29], essentially non oscillatory schemes [40], and nonlinear diffusion algorithms [2, 11, 13, 22].

Recently, Partial Differential Equation (PDE) approaches of image denoising have become important. The main idea of PDE-based models is to deform an image with a PDE and achieve the expected image as a solution to this equation. The famous form of diffusion, known as the Gaussian filter, is homogeneous isotropic linear diffusion. The Isotropic Diffusion (ID) performs well in the flat areas of the image. However, it blurs edges in the images and moves their positions. To overcome these shortcomings, nonlinear denoising models have been developed. Perona and Malik (PM) proposed an influential nonlinear anisotropic diffusion (AD) scheme [23]. Numerous denoising approaches derived from their model have been proposed since then [31]. An influential variational denoising technique was developed by Rudin, Osher and Fatemi. Their denoising model, which is known as the TV model, is based on the minimization of the TV norm. TV model is a successful approach to recover images with sharp edges. Nevertheless, TV denoising will produce the block effect when dealing with the flat areas.

Recently, some denoising methods by combining different PDE-based models have been proposed [36]. An image denoising algorithm has been suggested by using stochastic optimization algorithm for combining denoising methods based on partial differential equations [21].

Isotropic diffusion and total variation models are remarkably effective at image smoothing and edge-preserving, respectively. In order to eliminate the noise, and at the same time maintain the edges and the other necessary features in the image, we propose a weighted model based on the ID and TV models. In the proposed model the gradient is used to determine whether the region is the edge or the flat area. The model highlights the role of the ID model in the flat areas of an image, and the role of TV model in the regions which contains more image features (such as edges, etc.).

The remainder of this paper is organized as follows. The isotropic diffusion and total variation models are briefly described in Section 2. The proposed PDE model and the discretization of the model are described in Section 3. The experimental results that confirm the efficiency of the proposed model are presented in section 4. Finally, this paper is concluded in the Section 5.

2. The Isotropic Diffusion and Total Variation Models

2.1. The Isotropic Diffusion Model. The famous and simplest form of diffusion, known as the Gaussian filter, is the linear Isotropic Diffusion (ID) model [31]. In general, the ID model can be defined as finding the minimum of the functional

\[ E(u) = \int_{\Omega} |\nabla u|^2 + \lambda (u - u_0)^2 \, dx \, dy, \]

(2.1)

where \( \Omega \) is an open bounded domain in \( \mathbb{R}^2 \) (the domain of the image), \( \lambda \) is the Lagrange multiplier and \( u_0(x, y) \) is the degraded image.
The Euler-Lagrange equation gives
\[ \nabla \cdot (\nabla u) - \lambda (u - u_0) = 0. \] (2.2)

To prove (2.2), let
\[ E(u) = \int_\Omega F(x, y, u, u_x, u_y) \, dx \, dy, \] (2.3)
the extremizing function \( u(x, y) \) is determined from the solution of the following equation which is known as Euler-Lagrange equation [9]
\[ \frac{\partial}{\partial x} F_{u_x} + \frac{\partial}{\partial y} F_{u_y} - F_u = 0. \] (2.4)

The partial derivatives of the integrand
\[ F(x, y, u, u_x, u_y) = |\nabla u|^2 + \lambda (u - u_0)^2 = u_x^2 + u_y^2 + \lambda (u - u_0)^2, \] (2.5)
are
\[ F_{u_x} = 2u_x, F_{u_y} = 2u_y, F_u = 2\lambda (u - u_0). \] (2.6)

Therefore, the Euler-Lagrange equation of (2.1) gives
\[ 0 = \frac{\partial}{\partial x} (2u_x) + \frac{\partial}{\partial y} (2u_y) - 2\lambda (u - u_0) \]
\[ = 2 \left( \frac{\partial}{\partial x} \frac{\partial}{\partial y} \right) (u_x, u_y) - 2\lambda (u - u_0) \]
\[ 0 = \nabla \cdot (\nabla u) - \lambda (u - u_0). \] (2.7)

The ID model is defined as
\[ \begin{aligned}
\frac{\partial u}{\partial t} &= \nabla \cdot (\nabla u) - \lambda (u - u_0), \\
\frac{\partial u}{\partial n} &= 0 \quad \text{on} \quad \partial \Omega \times (0, T), \\
u(x, y, t) \mid_{t=0} &= u_0(x, y) \quad \text{in} \quad \Omega.
\end{aligned} \] (2.8)

where \( u(x, y, t) \mid_{t=0} = u_0(x, y) \) is the initial condition, \( u(x, y, t) \) is the restored version of the initial degraded image \( u_0(x, y) \), \( \nabla \) is gradient operator with respect to the spatial variables \( x, y \), and \( \Omega \) is an open bounded domain in \( \mathbb{R}^2 \).

The diffusion coefficient of the model is one, thus, the diffusion is the same in all directions. This model usually is used to smooth an image. Nevertheless, the ID model will blur the edge of the image during removing the noise [36].

2.2. The total variation model. The best known variational denoising model is the Total Variation (TV) model proposed by Rudin et al. [25] in 1992. The TV model is a successful approach to recover images with sharp edges. Total variation denoising is a process based on the principle that images with excessive possibly spurious detail have high total variation, that is, the integral of the absolute gradient of the image is high. Therefore, lessening the total variation of the image subject to it being as close as possible to the original image, removes unwanted detail whilst preserving important
features such as edges. The TV model can be defined as finding the minimum of the functional

$$E(u) = \int_{\Omega} \left( |\nabla u| + \frac{\lambda}{2} (u_0 - u)^2 \right) dx dy,$$

(2.9)

and the Euler-Lagrange equation gives

$$\nabla \cdot \left( \frac{\nabla u}{|\nabla u|^2 + \varepsilon} \right) - \lambda (u - u_0) = 0.$$

(2.10)

This equation can be written as

$$\nabla \cdot \left( \frac{\nabla u}{\sqrt{|\nabla u|^2 + \varepsilon}} \right) - \lambda (u - u_0) = 0,$$

(2.11)

where $\varepsilon > 0$ is very small, chosen to avoid division by zero at places where $|\nabla u| = 0$. For numerical implementations, the quantity $|\nabla u|$ is replaced with $\sqrt{|\nabla u|^2 + \varepsilon}$ for some small positive value of $\varepsilon$ such as $10^{-11}$. The value of $\varepsilon$ can be assigned to lowest machine number to avoid divide by zero conditions during implementations.

The TV model can be written as

$$\begin{cases}
\frac{\partial u}{\partial t} = \nabla \cdot \left( \frac{\nabla u}{\sqrt{|\nabla u|^2 + \varepsilon}} \right) - \lambda (u - u_0), \\
\frac{\partial u}{\partial n} = 0 \text{ on } \partial \Omega \times (0, T), \\
u(x, y, t) |_{t=0} = u_0(x, y) \text{ in } \Omega,
\end{cases}$$

(2.12)

where $u(x, y, t) |_{t=0} = u_0(x, y)$ is the initial condition and $\Omega$ is an open bounded domain in $\mathbb{R}^2$.

The TV model has a capacity of handling edges and removing noise in a given image [1]. The model is a successful approach to recover images with sharp edges. Nevertheless the TV model produces a block effect when being applied for the flat areas, thus the local details characteristics of the original image are lost [17, 18].

3. New model

The main problem of image denoising is how to remove noise without blurring edges of the image. Considering the characteristics of the isotropic diffusion and total variation models, we propose a weighted model based on the isotropic diffusion and total variation models. According to what is mentioned above, the isotropic diffusion model has good performance in the flat areas of the image, and the total variation model has good performance in the edges. In the proposed model the gradient is used to determine whether the region is the edge or the flat area. In order to suppress noise while preserving important features of the image, we integrate the isotropic and
total variation models to get a new PDE model as follows

\[
\begin{aligned}
\frac{\partial u}{\partial t} &= w \nabla \cdot (\nabla u) + (1 - w) \nabla \cdot \left( \frac{\nabla u}{\sqrt{|\nabla u|^2 + \varepsilon}} \right) - \lambda (u - u_0), \\
\frac{\partial u}{\partial n} &= 0 \quad \text{on } \partial \Omega \times (0, T), \\
\left. u(x, y, t) \right|_{t=0} &= u_0(x, y) \quad \text{in } \Omega,
\end{aligned}
\]  

(3.1)

where \( u(x, y, t) \big|_{t=0} = u_0(x, y) \) is the initial condition, \( \Omega \) is an open bounded domain in \( \mathbb{R}^2 \), and the weight function \( w \) is defined as follows

\[
w = \frac{1}{\sqrt{1 + |\nabla u|^2}} = \frac{1}{\sqrt{1 + u_x^2 + u_y^2}}, \quad (0 < w < 1).
\]

(3.2)

The new model is able to be adaptive in each region depending on the information of the image.

In the flat areas of image, which contain less image features (such as edges, etc.), \( |\nabla u| \) (the magnitude of the gradient) will be small and based on (3.2), \( w \) will be close to 1, while the value of \( 1 - w \) will be close to 0. Therefore, the new model will highlight the role of isotropic model.

In the regions which contain more image features, \( |\nabla u| \) will be large and based on (3.2), \( w \) will be close to 0, while the value of \( 1 - w \) will be close to 1. Thus, the model will highlight the role of total variation model.

More precisely, by using the weight \( w \) and \( 1 - w \), the model performs more diffusion in the flat areas and less diffusion in the edges of the image.

A PDE problem is said to be well-posed if a solution to the problem exists, the solution is unique, and the solution depends continuously on the problem data. The ID and TV models are well-posed. During the process of the image smoothing, the image gradient \( \nabla u \) keeps changing as the iterative evolution changes. For this reason, the weight function \( w \), which is used to determine whether the region is the edge or the flat area, should not be set fixed, but keeps changing with the number of iterations. The weight function is a continuous function, so at each time step, the value of weight function is a finite number in \([0, 1]\). The finite convex combination of two well-posed models, obviously, satisfies in the above three conditions of well-posed problem and therefore the new model is well-posed.

To solve (3.1) with the finite difference method, we let

\[
I = \nabla \cdot (\nabla u),
\]

(3.3)

\[
T = \nabla \cdot \left( \frac{\nabla u}{\sqrt{|\nabla u|^2 + \varepsilon}} \right).
\]

(3.4)
Then,

\[
I = \nabla \cdot (\nabla u) = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right) \cdot (u_x, u_y) = \frac{\partial}{\partial x} (u_x) + \frac{\partial}{\partial y} (u_y),
\]

\[
T = \nabla \cdot \left( \frac{\nabla u}{\sqrt{|\nabla u|^2 + \varepsilon}} \right) = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right) \cdot \left( \frac{u_x}{\sqrt{u_x^2 + u_y^2 + \varepsilon}}, \frac{u_y}{\sqrt{u_x^2 + u_y^2 + \varepsilon}} \right) = \frac{\partial}{\partial x} \left( \frac{u_x}{\sqrt{u_x^2 + u_y^2 + \varepsilon}} \right) + \frac{\partial}{\partial y} \left( \frac{u_y}{\sqrt{u_x^2 + u_y^2 + \varepsilon}} \right).\]

Assume \(0 \leq x, y \leq L, h = \frac{L}{M}\) and let \(x_i = ih, y_j = jh\), for \(0 \leq i, j \leq M\), be the discrete points (in our numerical calculation, we have \(h = 1\)). By using the notations

\[
w_{i,j}^n \sim u(x_i, y_j, t_n),
\]

\[
\Delta^u_{x} u_{i,j}^n = \pm \left( u_{i+1,j}^n - u_{i,j}^n \right),
\]

\[
\Delta^u_{y} u_{i,j}^n = \pm \left( u_{i,j+1}^n - u_{i,j}^n \right),
\]

\[
\Delta^u_0 u_{i,j}^n = \frac{1}{2} \left( u_{i+1,j}^n - u_{i-1,j}^n \right),
\]

\[
\Delta^u_0 u_{i,j}^n = \frac{1}{2} \left( u_{i,j+1}^n - u_{i,j-1}^n \right),
\]

the discrete forms of \(w, I,\) and \(T\) can be written as follows

\[
w_{i,j}^n = \frac{1}{\sqrt{1 + \left( \frac{\Delta^u_x u_{i,j}^n}{h} \right)^2 + \left( \frac{\Delta^u_y u_{i,j}^n}{h} \right)^2}} = \frac{1}{\sqrt{1 + \left( \frac{u_{i+1,j}^n - u_{i-1,j}^n}{2h} \right)^2 + \left( \frac{u_{i,j+1}^n - u_{i,j-1}^n}{2h} \right)^2}},
\]

(3.5)

\[
I_{i,j}^n = \frac{1}{h} \Delta^u_x \left[ \frac{\Delta^u_x u_{i,j}^n}{h} \right] + \frac{1}{h} \Delta^u_y \left[ \frac{\Delta^u_y u_{i,j}^n}{h} \right] = \frac{u_{i+1,j}^n - 2u_{i,j}^n + u_{i-1,j}^n}{h^2} + \frac{u_{i,j+1}^n - 2u_{i,j}^n + u_{i,j-1}^n}{h^2},
\]

(3.6)
We can approximate the partial differential equation of model \((3.1)\), by

\[
\frac{u_{i,j}^{n+1} - u_{i,j}^{n}}{\Delta t} = u_{i,j}^{n} I_{i,j} + \left(1 - u_{i,j}^{n}\right) T_{i,j}^{n} - \lambda \left(u_{i,j}^{n} - u_{0i,j}\right),
\]

where \(n = 0, 1, 2, \ldots\) is the time level. This equation yields the following explicit scheme for approximating the partial differential equation \((3.1)\) for all interior points \((x_i, y_j)\) such that \(1 \leq i, j \leq M - 1\)

\[
u_{i,j}^{n+1} = u_{i,j}^{n} + \Delta t \left(u_{i,j}^{n} I_{i,j} + \left(1 - u_{i,j}^{n}\right) T_{i,j}^{n} - \lambda \left(u_{i,j}^{n} - u_{0i,j}\right)\right). 
\]

Assuming that the approximate solutions \(u_{i,j}^{n}\), for \(1 \leq i, j \leq M - 1\) have been computed, we can approximate the boundary condition by \(u_{0,j}^{n} = u_{i,j}^{n}, u_{M,j}^{n} = u_{M-1,j}^{n}, u_{i,0}^{n} = u_{i,1}^{n}, u_{i,M}^{n} = u_{i,M-1}^{n}, u_{0,0}^{n} = u_{1,1}^{n}, u_{0,M}^{n} = u_{1,M-1}^{n}, u_{M,0}^{n} = u_{M-1,1}^{n}, u_{M,M}^{n} = u_{M-1,M-1}^{n}.\)

Since the diffusion based denoising models involve a huge amount of data, the explicit schemes are easy to implement and well-suited for implementations.
4. Experimental results and analysis

To show the efficiency of the proposed model, some experiments are carried out to compare the denoising result of the new model with that of some other PDE-based denoising models in terms of the Peak Signal to Noise Ratio (PSNR). The PSNR is defined in decibels for 8-bit gray-scale images, as follows

$$PSNR = 10 \log \frac{255^2 \times M \times N}{\sum_{i=1}^{M} \sum_{j=1}^{N} [I_{or}(i,j) - I_{de}(i,j)]^2},$$

where $I_{or}$ and $I_{de}$ are the original image and the denoised image, respectively. $M$ and $N$ are the numbers of pixels horizontally and vertically, respectively, and 255 is the peak signal with an 8-bit resolution. A higher PSNR usually indicates that the image is of higher quality.

The commonly used 256×256 bit Lena, Rice, and Boat images are taken in figures. The experimental results for Gaussian noise and speckle noise are presented. Figure 1 presents the original Lena image, the corrupted image by the Gaussian noise of standard deviation $\sigma_n = 15$, and the result obtained by the new model. Table 1 contains the PSNRs of the images obtained by different diffusion based schemes, and the new model. It can be seen from Table 1 that the PSNR of the new method is the maximum that means the denoising effect of the method is the best.

**Table 1.** The PSNRs of the images obtained by different algorithms for Lena image

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR</td>
<td>22.0817</td>
<td>29.2988</td>
<td>28.8990</td>
<td>29.5538</td>
<td>29.5632</td>
<td>29.6405</td>
<td><strong>30.5919</strong></td>
</tr>
</tbody>
</table>

For the speckle noise of variance $v = 0.01$, Figure 2 and Figure 3 show the original image, the corrupted image, and the result obtained by the new model for Rice image and Boat image, respectively. Table 2 presents the PSNR for different variances of the
speckle noise for Rice image. The PSNR graph of this models is displayed in Figure 4. From Table 2, we can see that increasing the variance of the noise will decrease the PSNRs of the models. However the PSNR of the new model is the largest among the seven models for the same variance, which means that the denoising effect of the new model is the best with regard to PSNR. The results of Table 2 show that the new model is better in terms of PSNR than the other models especially for high variances of the noise. Finally, from Figure 4, we can see that the proposed model for all noise power has the highest PSNR among the seven models.

**Figure 2.** (A): original Rice image, (B): noisy image, and (C): the result obtained by the new model

![Figure 2](image1)

**Figure 3.** (A): original Boat image, (B): noisy image, and (C): the result obtained by the new model

![Figure 3](image2)

5. Conclusion

In this paper a new PDE-based image denoising model has been presented. After taking into consideration the attributes of the isotropic diffusion and total variation models, we integrated these models to get the new model. In the proposed model the gradient has been used to determine whether the region is the edge or the flat area. The model has been designed to perform more diffusion in the flat areas and less diffusion in the edges of the image, and to effectively suppress noise while keeping intact the features of the image. To illustrate the efficiency of the proposed model, we have used the Peak Signal to Noise Ratio (PSNR) as the subjective criterion.
Table 2. The PSNRs of different algorithms with different variances of speckle noise for Rice image

<table>
<thead>
<tr>
<th>Variance of the noise</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
<th>0.04</th>
<th>0.05</th>
<th>0.06</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID model</td>
<td>22.0765</td>
<td>22.0725</td>
<td>22.0534</td>
<td>22.0188</td>
<td>22.0128</td>
<td>21.9443</td>
</tr>
<tr>
<td>PM model</td>
<td>30.0830</td>
<td>27.5987</td>
<td>25.5351</td>
<td>23.8745</td>
<td>22.5782</td>
<td>21.5451</td>
</tr>
</tbody>
</table>

Figure 4. The PSNR graph of the ID model, PM model, TV model, Ref. [16] model, Ref. [35] model, Ref. [36] model, and new model for different speckle noise level for Rice image

Experimental results confirm the efficiency of the new approach compared with some well-known models.

Acknowledgements

The authors are grateful to anonymous referees for their valuable comments which substantially improved the quality of this paper.
REFERENCES


