Visual and Artistic Images Denoising Methods Based on Partial Differential Equation

Zhenhe Ye 1, Xin Li 2, and Ying Li 2
1. College of Mechanical and Electrical Engineering, Agricultural University of HeBei, Baoding, 071000, China
2. College of Arts, Hebei Normal University of Science & Technology, Qinhuangdao, 066000, China
Email: yezhenhe@163.com

Abstract—Partial differential equation has a remarkable effect on image denoising, compression and segmentation. Based on partial differential equations, the denoising experiment is carried out on those artistic images requiring high degree of visual reduction through the application of 3 image-denoising algorithm models including thermal diffusion equation, P-M diffusion equation and the TV diffusion equation. By this experience, the respective characteristics in image-denoising of these 3 methods can be analyzed so that a better way can be chosen in adapting to digitization of artistic images or in dealing with distant signal.

Index Terms—Denoising, thermal diffusion equation, P-M diffusion equation, TV diffusion equation, artistic images

I. INTRODUCTION

The visual art sets a high demand on images’ reduction in brightness, color and texture. However, in digital display and distant transmission, the obtained images are often affected by various noises inevitably, which will certainly affect the accurate artistry in the subsequent images presentation [1, 2, 3, 4]. Therefore, those questions, which have been widely researched, are how to eliminate noise effectively, to improve image quality and to denoise images which are also necessary courses in image processing [5, 6]. Partial differential equation is one of the important parts in mathematical analysis and at present it has been widely used in many fields such as image processing and computer visual digitalization. In its application in image processing, the first step is to build an image partial differential equations model. Assuming that the image is a 2-D function in real number field about x, y, since the image is a continuous piecewise linear (PWL) function, it is feasible that we use function to approach the original image in continuous field except for the edge so as to accurately express intense local content and information of the visual images and represent the subject as well as the visual and mental information which the drawer wants to express [7, 8]. Therefore, this paper will focus on the discussion and improvement of key problems in image denoising process so that this method can be well applied in processing the visual art images [9, 10].

II. THEORETICAL BASIS OF PARTIAL DIFFERENTIAL EQUATION

A. Basic Conception

In a differential equation, if the unknown function to be solved only contains one argument, the equation is called ODE (ordinary differential equation); if the unknown function to be solved contains at least one argument, and there appears partial derivative of each order of multivariate function for different independent variables in the equation, the equation is called partial differential equation [11, 12, 13].

The partial differential equation on unknown function \( u(x_1, x_2, \ldots, x_n) \) is a relation similar to

\[
F(x, u, Du, \ldots, \frac{\partial^n}{\partial x_1^n \partial x_2^n \ldots \partial x_n^n} u) = 0 \quad (1)
\]

where \( x = (x_1, x_2, \ldots, x_n) \), \( Du(u_1, u_2, \ldots, u_n) \), \( F \) is the known function of independent variable \( x \), unknown function \( u \) and several finite partial derivatives on \( u \). \( F \) can be showed that it doesn’t contain the independent variable \( x \) and the unknown function \( u \), but it must contain \( u \)’s partial derivative. The order \( m = m_1 + m_2 + \ldots + m_n \) of the highest derivative is called order of the partial differential equation.

B. Advantages of Partial Differential Equation

The method of partial differential equation has remarking advantages in both theory and calculation, which are revealed as follows: Firstly, it has a stronger local adaptability. Fourier transform method has no localization at all, so it is only applicable to stationary signals processing. However, the images are usually non-stationary. Partial differential equation is based on serial image models, and it makes the changes of the value of a certain pixel only rely on an “infinitesimal” neighborhood of the pixel point at the current time \( t \). In this sense, it can be seen that the partial differential equation method in image processing has “infinite” ability of local adaptability. Secondly, it has high flexibility. If a basic model is set up successfully, and after some modification or improvement for it, a processing method can be obtained with more perfect performance and more wide application. And the modification and improvement are usually direct and easy. For example, two-dimension is upgraded into three-dimension, and a single-valued image processing method into vector diagram. Finally, the partial differential equation provides a analytical mode of
the image in continuous space as well as a better image processing effect[14] [15]. What’s more, the existence, uniqueness and stability of the algorithm solution can be proved available in the unique analysis-theory framework of the partial differential equation. The model is unrelated with the grid size (corresponding with the image pixel) of digital images. When the size of the grid’s meshes tends towards zero, the discrete filter can be understood as the approximation of continuous differential operators in the partial differential equation.

### III. PROPOSED SCHEME

#### A. Thermal Diffusion Equation

Thermal diffusion equation is set up based on the physical model, and the model is the following: the heat of a partial heated iron tube will spread gradually as the heat conduction process until the temperature of the whole iron tube is consistent [16, 17]. When the thermal diffusion equation is applied to image denoising, the noises of the image are equal to the catastrophe point of the temperature. Thermal diffusion will make the gray values of the whole image tend to be consistent, and then the purpose of denoising can be achieved.

Suppose the original image is \( u(x,y,0) \), the smooth image of the time \( t \), and then the partial differential equation of the thermal diffusion is

\[
\frac{\partial u(x,y,t)}{\partial t} = \Delta u(x,y,t) \tag{2}
\]

where \( \Delta u(x,y,t) \) is the Laplace operator of the image, with initial condition being \( u(x,y,0) \), and its solution is

\[
u(x,y,t) = G_t * u(x,y,0)
\]

\[
G_t(x,y) = \frac{1}{4\pi t} \exp\left(-\frac{x^2 + y^2}{4t}\right) \tag{3}
\]

Here \( * \) is the convolution. It can be seen from the above that: The original image and the convolutions of Gaussian filters of different scales are equivalent to the solution of the thermal diffusion equation which is isotropic, and has the same diffusion intensity to all the areas of the image. With the strengthening of smoothing effect, particulars such as edges of the original image, will disappear gradually.

The discrete form of Equation (2) is

\[
u^{t+1} = u^t + \frac{\lambda}{|n_s|} \sum_{p \in n_s} \nabla u_{s,p} \tag{4}
\]

where \( u_s \) is sampling discrete image, \( s \) is pixel coordinate, \( t \) is iteration, \( \lambda \) is the weighting of distribution coefficient reflecting the smoothness, \( n_s \) means all the consecutive points of the pixel \( s \), and \( |n_s| \) represents the number of consecutive points (those four points, directly adjacent to the pixel, are often adopted from the up, down, left and right with edge excepted.)

\[
\nabla u_{s,p} = u_p - u_s, \quad p \in n_s \tag{5}
\]

Namely the gradient value of image along the direction \((s,p)\). When Equation (5) is substituted into Equation (4), the template of Equation (2)’s action on image can be obtained:

\[
\begin{bmatrix}
\lambda \\
\lambda - 4\lambda \\
\lambda - 16\lambda \\
\lambda
\end{bmatrix} / 4 \tag{6}
\]

All the weight of pixel \( s \) acting on each directions are \( \lambda \). Regardless of the location of pixel \( s \), it has the same smoothness intensity to all pixels. When \( \lambda = 1 \), it is the denoising filter of mean value. In this way, the characteristic of image is blurred while being denoised.

#### B. P-M Diffusion Equation

The shortcoming of thermal equation of energy is that the coefficient is a constant and cannot be diffused differently according to different directions [18, 19]. Focusing on the shortcoming of thermal diffusion equation, Perona and Malik improved the model of thermal diffusion. The basic thought is that image is separated into domains of different scales in scale space, and is smoothed in the internal domains of all the scales while weakening the smoothing effect on the edge of domains, namely edge [20, 21, 22].

Perona and Malik put forward nonlinear diffusion in 1990 and introduced function \( c(x,y,t) \) [23] [24], which satisfied the properties put forward above and improved the thermal diffusion equation:

\[
\frac{\partial u(x,y,t)}{\partial t} = d [c(x,y,t) \nabla u(x,y,t)] \tag{7}
\]

The initial condition is \( u(x,y,0) \), \( u \) is the input image, \( \text{div} \) is the divergence operator, \( \nabla \) represents the gradient, and \( c \) is the diffusion coefficient. Under ideal conditions, the corresponding value of \( c(x,y,t) \) is 1 in the internal domain of image, and Equation (7) degenerates into the isotropic thermal diffusion Equation (2) which will do isotropic diffusion in the internal domain of the image; in the edge that is the edge of the image domain, the value of \( c(x,y,t) \) is 0, and does not perform the diffusion. The diffusion coefficient of equation is decided by the spatial location of the image, namely the gray change of it. The coefficient is anisotropic so that the edge feature of the image can be reserved when the diffusion occurs.

However, in the actual application of image processing, the edge of image domains, namely the edge, cannot be known in advance. So function \( E(x,y,t) \), a vector-function, is introduced as evaluation of the edge points. This thesis adopts gradient operator to evaluate the edge points, namely

\[
c(x,y,t) = g(|E(x,y,t)|) = g(|\nabla u(x,y,t)|) \tag{8}
\]

In accordance with the strategy of smoothing mentioned above, \( g(\cdot) \) sets the degree of diffusion, and satisfies the following conditions: If \( g(s) \) is a smooth and non-increasing function. If \( g(0) = 1 \), and \( g(s) \geq 0 \). When \( III \ s \rightarrow \infty \), and \( g(s) \rightarrow 0 \), the gradient of image is large at the edge, \( g(s) \) is small, and the degree of diffusion becomes smaller; on the contrary, the gradient of image is
small, $g(s)$ is large, and the degree of diffusion becomes larger. As a result, the image can be smoothed alternatively—smoothed less at the edge and more in the smooth areas. The functional figure 1 of $g(s)$ is shown as follows:

![Figure 1](image.png)

**Figure 1.** Note how the caption is centered in the column.

Here the function of diffusion coefficient $c(x, y, t)$ becomes a monotonically decreasing function which adopts $\nabla u(x, y, t)$ as the variable. From the above, the following can be obtained:

$$\frac{\partial u(x, y, t)}{\partial t} = d \# [g(\nabla u(x, y, t)] \nabla u(x, y, t)$$  \hspace{1cm} (9)

This is the traditional P-M diffusion model equation. And the two functions of diffusion coefficient are

$$g(\nabla u) = \frac{1}{1 + \left(\frac{\nabla u}{k}\right)^2}$$  \hspace{1cm} (10)

or

$$g(\nabla u) = \exp \left(-\frac{1}{k} \left(\frac{\nabla u}{k}\right)^2\right)$$  \hspace{1cm} (11)

where $k$ is the threshold parameter of gradient controlling the process of diffusion. The larger its value is, the better smoothing effect it will have. And the smaller $k$ value can strengthen the edge. The effects these two functions of diffusion coefficients act on the diffusion of image are different. The previous function will strengthen the high-contrast edges of image and weaken the low-contrast edges of it, and the latter function will reserve the large areas in the image and remove the small ones in it.

Generally, $g$ is the gradient function of image. And with the increasing of gradient, it will decrease monotonically and its value range is restricted to $[0, 1]$, such as the formulas (10), (11). The diffusion coefficient decides the way of diffusion and provides a diffusion-controlling strategy of local adaptability. It makes the diffusion process the location of noise as much as possible, but stops at the edge of image.

The discrete version of P-M equation is

$$u_{i+1}^{t+1} = u_i^t + \frac{\lambda}{|p_i|} \sum_{p_i} c(\nabla u_{i,p}) \nabla u_{i,p}$$  \hspace{1cm} (12)

As a result, the template of P-M equation’s action on the image is

$$c(u_{i+1}^t - u_i^t) + \left(4 - \sum_{p_i} c(\nabla u_{i,p})\right) c(u_{i+1}^t - u_i^t) / 4$$  \hspace{1cm} (13)

Here, the weight value of pixel $s$ is no longer a fixed value, but is related to the spatial location of pixel $s$. If the pixel is at the location with large gradient, the relevant weight value of pixel will be small and the smoothing degree will be weak, because $c(x, y, t)$ is a monotonically decreasing function with $\nabla u(x, y, t)$ as its variable. On the contrary, if the pixel $s$ is at the location with small gradient, the smoothing degree will be intense. Usually, the gradient is larger at the edge of image, and smaller at other areas. So the P-M equation can denoise the image while preserving or even strengthening the edge.

### C. TV Diffusion Equation

A discrete image can be regarded as a real function above $\mathbb{R}^2$ in the bounded closed region, and the essence of denoising those noisy images is to recover the real images from them, so the characteristic information (such as the edge) should be kept as much as possible so it will not be lost during denoising. Let’s suppose $u^0(x, y)$ as the original images and $u(x, y)$ as noisy images.

$$u^0(x, y) = u(x, y) + n(x, y)$$  \hspace{1cm} (14)

$n(x, y)$ is an additive noise with its 0 average and $\sigma^2$ variance, and it is a gaussian noise here. The processing of denoising images is to estimate the original images $u(x, y)$ according to the prior knowledge we have and the statistics of noise, which is to work out $u(x, y)$ by equations. But usually it is infeasible to solve (14) directly in reality, because usually there are errors in actual statistics, and solving directly often result in worse consequence. The way of solving this problem is to adopt regularization, which can solve the equation by restricting its scope. Then the problem of denoising is turned into the problem of optimization.

$$\int_\Omega u(x, y) d\Omega = \int_\Omega u^0(x, y) d\Omega$$

$$\int_\Omega [u(x, y) - u^0(x, y)]^2 d\Omega = \sigma^2$$  \hspace{1cm} (15)

The corresponding form without constrain conditions is:

$$E_s(u) = \int_\Omega [\nabla u(x, y)]^2 d\Omega + \frac{\lambda}{2} \int_\Omega (u(x, y) - u^0(x, y))^2 d\Omega$$  \hspace{1cm} (16)
\( \lambda \) is a Lagrange multiplier, which is a constant given before-hand and decides the intensity of smooth denoising, therefore, the value of it depends on the noise level. Here the function need to be obtained while the value of \( E_J(u) \) being the minimum, which is a functional extremum problem, namely variation.

Through Euler equation, extremum condition of Equation (16) can be obtained as follows:

\[
-\nabla \left( \frac{\nabla u}{|\nabla u|} \right) + \lambda (u - u^0) = 0 \tag{17}
\]

The following is obtained by solving Equation (17) with gradient descent method

\[
\frac{\partial u}{\partial t} = \nabla \left( \frac{\nabla u}{|\nabla u|} \right) + \lambda (u^0 - u) \tag{18}
\]

In actual calculation, such situation may occur that \( \nabla u \) becomes zero in flat areas of image, when the denominator on the right of Equation (18) needs to be adjusted. Generally a small constant \( a^2 \) will be added to the denominator, and

\[
|\nabla u|^2 + a^2 = (19)
\]

The performance of TV model will not be affected as long as the value of \( a^2 \) is small enough.

Here the following can be obtained:

\[
E_{\lambda}(u) = \int_{\Omega} \left[ |\nabla u|^2 + \frac{\lambda}{2} |u - u^0|^2 \right] d\Omega \tag{20}
\]

The discrete version of Equation (20) is:

\[
\frac{u^{n+1}_i - u^n_i}{\tau} = \sum_{(i,j) \in N(i,j)} g_{ij}^u (u^n_{i,j} - u^n_i) + 2\lambda (u^n_i - u^n_{i,j}) \tag{21}
\]

Hereinto, \( \tau \) is time step, and \( h \) is pixel space \( g_{ij}^u = \frac{1}{|\nabla u^n_i|} \). \( \alpha \) or \( \beta \) are used to represent the coordinates of pixel: \((i, j)\) or \((k, l)\), and

\[
w_{ij}^{\alpha\beta} = \frac{g_{\alpha}^u + g_{\beta}^u}{2h^2} \tag{22}
\]

After the arrangement of the above equation, we can obtain

\[
u^{n+1}_\alpha = \tau \sum_{\beta \in N(\alpha)} w_{\alpha\beta}u^n_\beta + \tau\lambda u^n_\alpha + u^n_\alpha \left( 1 - \tau \sum_{\beta \in N(\alpha)} w_{\alpha\beta} - \tau\lambda \right) \tag{23}
\]

There are two parameters in the above equation: Lagrange multiplier \( \lambda \) and time step \( \tau \). Reference [3] provides the optimal estimation of approximation for \( \lambda \)

\[
\lambda = \frac{1}{\sigma^2} \frac{1}{|\Omega|} \sum_{\alpha \in \Omega} \sum_{\beta \in N(\alpha)} w_{\alpha\beta} (u^n_\beta - u^n_\alpha)(u^n_\alpha - u^n_\beta) \tag{24}
\]

\( |\Omega| \) is the area of image zone. From Equation (24), we can see that \( \lambda \) is in inverse proportion to variance of noise.

When smoothing images first, we can choose \( \lambda = \frac{1}{\sigma^2} \), and update the value of \( \lambda \) with Equation (24) after a period of iterations.

Time step can either be constant or calculated according to the data of each step. It affects the convergence of iterative equation directly and Equation (25) provides the calculation method of current iteration step size in accordance with the previous iteration results.

\[
\tau = \frac{1}{\sum_{\beta \in N(\alpha)} w_{\alpha\beta} + \lambda} \tag{25}
\]

Putting Equation (25) into Equation (23), we can obtain:

\[
u^{n+1}_\alpha = \sum_{\beta \in N(\alpha)} h_\alpha\beta u^n_\beta + h_\alpha u^n_\alpha \tag{26}
\]

\[
h_\alpha\beta = \frac{w_{\alpha\beta}}{\lambda + \sum_{\gamma \in \alpha} w_{\alpha\gamma} \lambda} \tag{27}
\]

Thereinto,

\[
h_\alpha + \sum_{\beta \in N(\alpha)} h_\alpha\beta = 1 \tag{28}
\]

The iterative process of denoising image using TV model is equivalent to that of non-linear lowpass filtering. For edge pixels, these pixel values are preserved as the value of \( |\nabla u|^2 \) is bigger, and as a result, the value of \( h_\alpha \) becomes smaller; for pixels in flat areas of image, it is the equal of lowpass filtering for the pixels as the value of \( |\nabla u|^2 \) is smaller while the value of \( h_\alpha \) is bigger. The smooth intensity of TV model for different pixels is different as well as anisotropic. Although the number of iterations increases from 10 to 35, the difference of Lena image after TV diffusion is not as obvious as that after P-M diffusion. This is because TV diffusion equation (16) includes \( \frac{1}{2} \int_{\Omega} |u(x,y) - u^0(x,y)|^2 d\Omega \).

The iteration and denoising for image with TV model is equal to non-linear lowpass filtering, the diffusion degree in different points are different as well as anisotropic and stable. Since TV model is based on the “optimization” of the image’s integral variation model and it has globally optimal solution, it can get stable and correct results even when the image noise is very large.

**IV. EXPERIMENT**

Select a 256*256 grayscale image, and add white Gaussian noise with mean value of 0 and variance of 25 to it. Denoise the noise image with thermal diffusion, P-M diffusion and TV diffusion respectively. And the iterations are 10 and 20, as shown in Figure2, Figure3.
Figure 2. Denoising the image with three methods

Figure 3. Denoising the artistic image with three methods
Observing images denoised through isotropic thermal diffusion equation, we can see that: when the iteration number reaches 10, the edge of image is fairly clear, while there is still lots of noise on the image; when the iteration number reaches 20, the image becomes obscure seriously because the noise is basically eliminated, some important information such as the edge and details of the image gets missing for smooth, though. Isotropic thermal diffusion equation turns image-denoising into a partial differential issue, however, its denoising effect is pretty average for the severe edge information missing, and image-denoising effect of P-M diffusion equation is obviously better than that of thermal diffusion equation. The noise is mostly eliminated with maintenance of the edge to some extent, but obscurity appears on some details of the image. This is because that P-M diffusion equation can seize the changes with large image gradient very well, but it cannot seize the texture and petty edge caused by tiny changes, which is the reason of missing image information. While the noisy points are being reduced, there appear some bigger noisy points, which are caused by the misjudging of some noisy points as edge because of their rather large gradient values. As a result, those noisy points are not removed, on the contrary, are strengthened to some extent. And color block appears, that is, after the P-M diffusion, there appear a lot of areas with the same gray value which produces such visual effect as the image is composed of many areas with different brightness. This is because that P-M diffusion equation is smoothing the image regionally, and in areas with rather small gradient, the value of diffusion coefficient function $g(\nabla u)$ is larger with strong diffusion intensity, thus eliminating the noise effectively, decreasing the gradient values and at the same time inclining gray value into constant in those areas; in areas with rather big gradient, the value of diffusion coefficient function $g(\nabla u)$ is small with weakened diffusion intensity, and along with the diffusion, the gradient values of those areas will evolve towards infinity, finally resulting to a zero value of diffusion coefficient function $g(\nabla u)$ and inactive diffusion on these areas, after which the value in each area tends to be constant and since the diffusion cannot be conducted between areas, there will appear staircase effect for largely different gray values among areas although the edge feature of the image is preserved; while the image denoised through T-V diffusion has no such obvious distinction as that through P-M diffusion even when the number of iterations is increased from 10 to 20. This is due to the constraint $\frac{1}{2} \int_{\Omega} [u(x, y) - u(x, y)]^2 d\Omega$ contained in TV diffusion equation, which controls the degree of deviation between denoised image and original image. Therefore, the distinction led by iterations is not as obvious as that of P-M diffusion, which is hence stable and pathological. But compared with P-M diffusion, the edge preserving effect of TV diffusion is slightly poorer. Although the function of bounded variation (BV) has a favorable ability of expression for image edge information, it cannot represent the information of image texture and details very well. Therefore, it is inevitable to drop out some information like texture and details of the image in the process of denoising.

From the denoised image with three methods mentioned above, it can be seen that the denoising effects after the iteration of 10 times are basically the same. However, after the iteration of 20 times, denoising effects using the method of P-M diffusion is obviously better than by the other two. Besides, when the noise is filtered, the edge can be protected. As for the method of thermal diffusion equation, with the increase of iteration, the noise is filtered, but some important information such as the edge and particulars of the image will be dropped because of the smoothing. Finally, the image has been blurred seriously. Although the isotropic equation of thermal diffusion changes image denoising into partial differentiation, its denoising effect is very common and edgeal information is dropped seriously. The effect of image denoising by P-M diffusion equation is obviously better than that by thermal diffusion equation. Most of the noise is removed and the edge is also protected. The impact of increasing iteration on the image denoised by TV diffusion is not obvious, and the image is stably non-pathological. However, compared with P-M diffusion, the effect of edge-maintaining is slightly inferior.

V. Conclusion

It is unavoidable for works of artistic images to be jammed by noise in collecting and displaying image information through visual hierarchical system and to lose its realistic artistry by shortage of visual effect of artistic images. The denoising of image is crucial and its effect influences directly the degree of accuracy of subsequent analysis and processing. The paper analyzed the application of partial differential equation on denoising algorithmic model of image and conducted contrast experiments for image-denoising through thermal diffusion equation, P-M diffusion equation and TV diffusion equation respectively. Compared with other denoising methods achieved on visual images, although each of these methods has their advantages and disadvantages, they reduce the adverse effects of noise on images and meet the different requirements in real-time processing of artistic images in visual system.

REFERENCES


Zhenhe Ye received his M.E. degree from Hebei University of Technology, China, in 2009. He is a key teacher in mechanical and Electrical Engineering College, Hebei Agricultural University. His research interests include image compression, Computer-graphics design, interaction design and its theories and Ergonomics etc.

Xin Li received her M.A. degree in Art from Hebei University of Technology,China in 2008. She is a key teacher in Art College, Hebei Normal University of Science & Technology and her research interests include environmental art, visual communicati on and CAD, etc.

Ying Li received her M.A. degree in Art from Hebei Normal University, China, in 2005. Her is a key teacher in Art College, Hebei Normal University of Science & Technology. Her research interests include environmental art, visual communication, CAD, etc.