Image Denoising using Perona-Malik Diffusion Model

Noaima Bari

1 Problem Statement

Image noise can be seen as defects or unwanted artifacts in an image that are not part of the intended scene. It can result from various sources such as lighting, sensors, compression artifacts, and random fluctuations in the imaging process and can lead to reduced image quality that can affect further downstream tasks. Image denoising refers to the task of eliminating such undesirable noise or artifacts from the image and has been a popular research area in image processing. Several methods have been devised and tested over the years. However, a major challenge in image denoising algorithms has been to remove noise from an image while preserving its original structure and edges. This is the problem we aim to solve in this project.

An interesting way to look at this problem is through the gradient information from the image, which is a measure of how quickly the intensity of an image changes over space. In the context of image processing, the noise can be seen as random variations in pixel values that are not part of the underlying image structure. These random variations can introduce abrupt changes in pixel intensity, which can be detected by the gradient of the image. Smooth regions of an image typically have a low gradient, while areas with rapid changes in intensity have a high gradient. When an image is corrupted by noise, the noise can cause fluctuations in pixel values that result in high gradients even in smooth regions of the image. Therefore, by measuring the gradient of an image, we can identify regions where the gradient is higher than expected, indicating the presence of noise.

Hence, we aim to design a diffusion process that penalizes higher values of gradient more than the lower values, and can smoothen the image progressively by removing noise while at the same time, preserving its edges.

In the next sections, we design such a diffusion process using the Perona-Malik Diffusion Model which is a non linear diffusion model that uses partial differential equations and different diffusion coefficients to model the diffusion process.

2 Mathematical Representation

In the last section, we saw how the gradient information can help us in understanding the presence of noise. The Perona-Malik Diffusion Model models a non linear diffusion process that uses diffusion functions as non decreasing functions of the image gradient. More formally, it defines an energy functional that needs to be minimized in order to achieve the denoised image.

The energy functional is of the form:

$$E(I) = \iint_{\Omega} c(||\nabla I||)\, dx\, dy$$

where $c(||\nabla I||)$ is a non decreasing function of the image gradient ∇I

We aim to minimize this energy functional in order to obtain a smooth image. As discussed in the previous section, upon minimizing a non decreasing function of the gradient of the image, we are essentially penalizing areas having higher gradient values (stronger edges) more than the lower gradient values (smoother regions), hence aiming to preserve the structure of the original image.

In order to denoise the image, several diffusion functions can be designed, as per the different edge preserving requirements.

In this project, 3 different diffusion functions have been utilized and their comparative results are presented towards the end.

$$c(||\nabla I||) = \log(||\nabla I||)$$

$$c(||\nabla I||) = \frac{1}{1 + e^{-||\nabla I||}}$$

$$c(||\nabla I||) = \frac{||\nabla I||^2}{2}$$

These diffusion functions have different relationship with the gradient. To visualize the same, a plot is presented in the figure below:

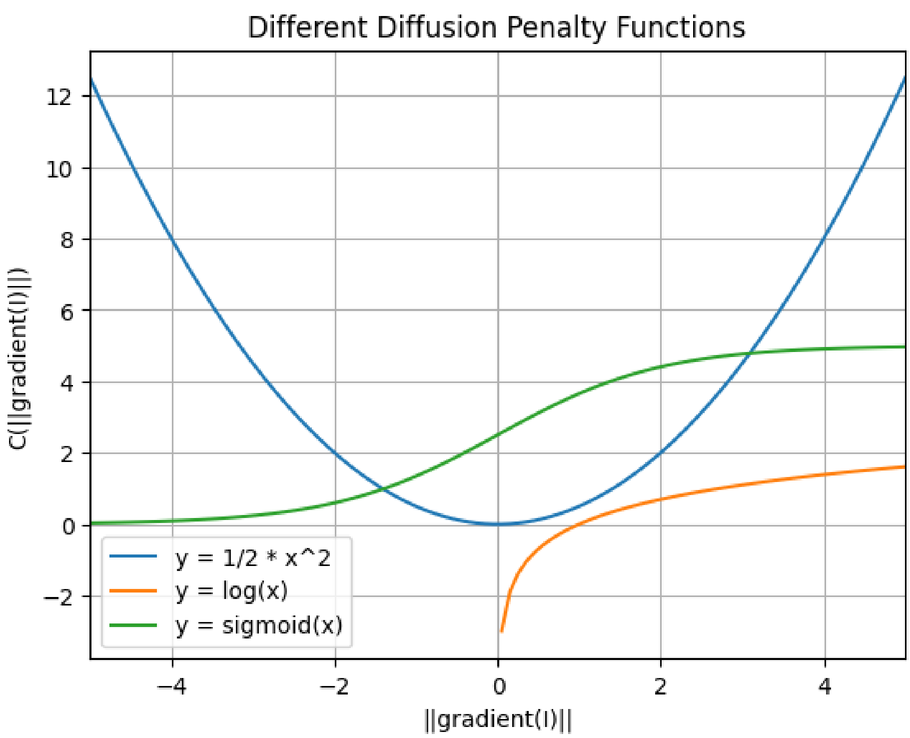

Figure 1: Graphs of diffusion functions

One issue with minimizing the energy functional is to determine when to stop the diffusion process. Running the diffusion process for a long time may lead to the loss of significant edge information from the image. Hence, we introduce a data fidelity term, that penalizes how much the denoised image I can vary from the original image I_o.

In this project, this factor is introduced in the energy functional for the quadratic penalty resulting in the following energy functional:

$$E(I) = \iint_{\Omega} (1-\lambda)\frac{||\nabla I||^2}{2}) + \lambda(I - I_o)^2 \, dx \, dy$$

3 Gradient Descent Partial Differential Equation

The diffusion function inside the energy functional can be seen as a Lagrangian of the form $L(I, I_x, I_y, x, y)$ We then define the gradient as

$$\nabla E = L_I - \frac{\partial L_{I_x}}{\partial x} - \frac{\partial L_{I_y}}{\partial y}$$

We had seen that the minimization of the energy functional was an infinite dimensional optimization problem since it involved finding an entire function as the optimal value. Hence, in order to minimize the same, we saw how we could interpret taking its derivative from a directional derivative standpoint. In finite dimensional case, for the directional derivative to be zero, the gradient must be zero along each dimension. However, in the infinite dimensional case, we had seen that ∇E could in fact be treated as the gradient. When $\nabla E = 0$, all directional derivatives vanish giving us the optimal value of image I. However, solving the same analytically is not always feasible. However, we saw that in order to obtain the optimal value, if we perturbed the image by the gradient, we would be making the most optimal perturbation as it would represent direction of steepest change.

Hence, using gradient descent, we could write the time evolution PDE in the form:

$$I_t = -\nabla E$$

$$I_t = L_I + \frac{\partial L_{I_x}}{\partial x} + \frac{\partial L_{I_y}}{\partial y}$$

We follow this procedure to obtain the PDE in case of each of the diffusion functions:

3.1 Logarithmic Penalty

$$E(I) = \iint_\Omega \log(||\nabla I||)\, dx\, dy$$

For the derivation of the PDE, see Figures 2 and 3.

$$I_t = \frac{1}{(I_x^2 + I_y^2 + \epsilon^2)^2}[(I_x^2 - I_y^2)(I_{yy} - Ixx) - 4I_xI_yI_{xy} + \epsilon^2 \Delta I]$$

3.2 Sigmoid Penalty

$$E(I) = \iint_\Omega \frac{1}{1 + e^{-||\nabla I||}}\, dx\, dy$$

For the derivation of the PDE, see Figure 4.

3.3 Quadratic Penalty

$$E(I) = \iint_\Omega (1 - \lambda)\frac{||\nabla I||^2}{2}) + \lambda(I - I_o)^2\, dx\, dy$$

$$L_I = \lambda(I - I_o)$$
$$L_{I_x} = I_x(1 - \lambda)$$
$$L_{I_y} = I_y(1 - \lambda)$$

$$\frac{\partial L_{I_x}}{\partial x} = I_{xx}(1 - \lambda)$$

$$\frac{\partial L_{I_y}}{\partial y} = I_{yy}(1-\lambda)$$

Gradient:

$$\nabla E = L_I - \frac{\partial L_{I_x}}{\partial x} - \frac{\partial L_{I_y}}{\partial y}$$

Gradient Descent PDE:

$$I_t = -\nabla E$$

$$I_t = -\lambda(I - I_o) + I_{xx}(1-\lambda) + I_{yy}(1-\lambda)$$

4 Discretization and Stability Analysis

4.1 Discretization

For the discretization of the PDEs, the forward time difference and the centered space differnece have been utilized for all the 3 cases. The forward time difference is a suitable choice, since the linear heat equation is ill posed backwards. Since, the above PDEs can be viewed as a modification of the linear heat equation for their stability analysis, a forward time difference is the appropriate choice.

$$I_t = \frac{I(x,y,t+\Delta t) - I(x,y,t)}{\Delta t}$$

We start with using a centered space difference. Upon doing the stability analysis, we will see that this scheme is stable and gives us a CFL condition.

Centered Space Differences:

$$I_x = \frac{I(x+\Delta x,y,t) - I(x-\Delta x,y,t)}{2\Delta x}$$

$$I_y = \frac{I(x,y+\Delta y,t) - I(x,y-\Delta y,t)}{2\Delta y}$$

$$I_{xx} = \frac{I(x+\Delta x,y,t) - 2I(x,y,t) + I(x-\Delta x,y,t)}{\Delta x^2}$$

$$I_{yy} = \frac{I(x,y+\Delta y,t) - 2I(x,y,t) + I(x,y-\Delta y,t)}{\Delta y^2}$$

4.2 Stability Analysis

According to Lax theorem, to prove the convergence of the FDE to the original PDE, ie: to prove that:

$$\lim_{\Delta x \to \infty \Delta y \to \infty} FDE = PDE$$

We need to satisfy consistency and stability. Consistency of the FDE is straighforward since we utilized Taylor series approximations to derive the difference equations.For determining the stability, we use the Von Neumann Analysis The stability analysis of the each of the 3 cases is presented in Figures 5, 6, 7, 8 and 9 respectively. Further, the derivation of the CFL condition for the general case of linear heat is presented in Figure 10.

$$E(I) = \iint_{\Omega} \log \|\nabla I\| \, dx\,dy$$

$$= \iint \log \sqrt{I_x^2 + I_y^2} \, dx\,dy.$$

$$L_I = 0$$

$$L_{I_x} = \frac{1}{\sqrt{I_x^2+I_y^2}} \cdot \frac{1}{2\sqrt{I_x^2+I_y^2}} \cdot 2I_x$$

$$= \frac{I_x}{(I_x^2+I_y^2)}$$

$$L_{I_y} = \frac{I_y}{I_x^2+I_y^2)}$$

$$\frac{\partial}{\partial x} L_{I_x} = \frac{(I_x^2+I_y^2)I_{xx} - I_x(2I_xI_{xx} + 2I_yI_{yx})}{(I_x^2+I_y^2)^2}$$

$$= \frac{I_y^2 I_{xx} - I_x^2 I_{xx} - 2I_xI_yI_{yx}}{(I_x^2+I_y^2)^2}$$

$$\frac{\partial}{\partial x} L_{I_y} = \frac{(I_x^2+I_y^2)I_{yy} - I_y(2I_xI_{xy} + 2I_yI_{yy})}{(I_x^2+I_y^2)^2}$$

$$= \frac{I_x^2 I_{yy} - I_y^2 I_{yy} - 2I_xI_yI_{xy}}{(I_x^2+I_y^2)^2}$$

$$\text{Gradient} = \nabla E = L_I - \frac{\partial}{\partial x} L_{I_x} - \frac{\partial}{\partial y} L_{I_y}$$

Gradient descent PDE:

$$I_t = -\nabla E$$
$$= -L_I + \frac{\partial}{\partial x} L_{I_x} + \frac{\partial}{\partial y} L_{I_y}$$

$$I_t = \frac{1}{(I_x^2+I_y^2)^2}\left[(I_x^2 - I_y^2)(I_{yy} - I_{xx}) - 4I_xI_yI_{xy}\right]$$

From this PDE, we see that there is a third order zero in the numerator and a 4^{th} order zero in the denominator, hence it is numerically unstable when $I_x \to 0$ and $I_y \to 0$. Hence, we introduce a ε term in the energy functional to handle this numerical instability.

Figure 2: Deriving PDE for Log Penalty (without epsilon)

$$E(I) = \iint_{\Omega} \log \|\nabla I\| \, dx\, dy$$

$$= \iint_{\Omega} \log \sqrt{I_x^2 + I_y^2 + \varepsilon^2} \, dx\, dy.$$

$$L_I = 0$$

$$L_{I_x} = \frac{1}{\sqrt{I_x^2 + I_y^2 + \varepsilon^2}} \cdot \frac{1}{2\sqrt{I_x^2+I_y^2+\varepsilon^2}} \cdot 2 I_x$$

$$= \frac{I_x}{(I_x^2 + I_y^2 + \varepsilon^2)}$$

$$L_{I_y} = \frac{I_y}{I_x^2 + I_y^2 + \varepsilon^2}$$

$$\frac{\partial}{\partial x} L_{I_x} = \frac{(I_x^2 + I_y^2 + \varepsilon^2) I_{xx} - I_x (2 I_x I_{xx} + 2 I_y I_{yx})}{(I_x^2 + I_y^2 + \varepsilon^2)^2}$$

$$= \frac{I_y^2 I_{xx} + \varepsilon^2 I_{xx} - I_x^2 I_{xx} - 2 I_x I_y I_{yx}}{(I_x^2 + I_y^2 + \varepsilon^2)^2}$$

$$\frac{\partial}{\partial x} L_{I_y} = \frac{(I_x^2 + I_y^2 + \varepsilon^2) I_{yy} - I_y (2 I_x I_{xy} + 2 I_y I_{yy})}{(I_x^2 + I_y^2 + \varepsilon^2)^2}$$

$$= \frac{I_x^2 I_{yy} + \varepsilon^2 I_{yy} - I_y^2 I_{yy} - 2 I_x I_y I_{xy}}{(I_x^2 + I_y^2 + \varepsilon^2)^2}$$

$$\text{Gradient} = \nabla E = L_I - \frac{\partial}{\partial x} L_{I_x} - \frac{\partial}{\partial y} L_{I_y}$$

Gradient descent PDE:

$$I_t = -\nabla E$$
$$= -L_I + \frac{\partial}{\partial x} L_{I_x} + \frac{\partial}{\partial y} L_{I_y}$$

$$I_t = \frac{1}{(I_x^2 + I_y^2 + \varepsilon^2)^2} \left[(I_x^2 - I_y^2)(I_{yy} - I_{xx}) - 4 I_x I_y I_{xy} + \varepsilon^2 \overbrace{(I_{xx} + I_{yy})}^{\Delta I} \right]$$

$$\boxed{I_t = \frac{1}{(I_x^2 + I_y^2 + \varepsilon^2)^2} \left[(I_x^2 - I_y^2)(I_{yy} - I_{xx}) - 4 I_x I_y I_{xy} + \varepsilon^2 \Delta I \right]}$$

Figure 3: Deriving PDE for Log Penalty (with epsilon)

$$E(I) = \iint_{\Omega} \frac{1}{1+e^{-\sqrt{I_x^2+I_y^2+\varepsilon^2}}}\, dx\, dy.$$

$$L_I = 0$$

$$L_{I_x} = \frac{-1}{\left(1+e^{-\sqrt{I_x^2+I_y^2+\varepsilon^2}}\right)^2} \cdot (-1) \cdot e^{-\sqrt{I_x^2+I_y^2+\varepsilon^2}} \cdot \frac{-1\,(2I_x)}{2(I_x^2+I_y^2+\varepsilon^2)^{3/2}}$$

$$= \frac{-I_x\, e^{-\sqrt{I_x^2+I_y^2+\varepsilon^2}}}{(I_x^2+I_y^2+\varepsilon^2)^{3/2}\left(1+e^{-\sqrt{I_x^2+I_y^2+\varepsilon^2}}\right)^2} = \frac{-I_x\, \sigma\left(\sqrt{I_x^2+I_y^2+\varepsilon^2}\right)\left\{1-\sigma\left(\sqrt{I_x^2+I_y^2+\varepsilon^2}\right)\right\}}{(I_x^2+I_y^2+\varepsilon^2)^{3/2}}$$

$$L_{I_y} = \frac{-I_y\, e^{-\sqrt{I_x^2+I_y^2+\varepsilon^2}}}{(I_x^2+I_y^2+\varepsilon^2)^{3/2}\left(1+e^{-\sqrt{I_x^2+I_y^2+\varepsilon^2}}\right)^2} = \frac{-I_y\, \sigma\left(\sqrt{I_x^2+I_y^2+\varepsilon^2}\right)\left\{1-\sigma\left(\sqrt{I_x^2+I_y^2+\varepsilon^2}\right)\right\}}{(I_x^2+I_y^2+\varepsilon^2)^{3/2}}$$

$$\frac{\partial L_{I_x}}{\partial x} = \Big[-(I_x^2+I_y^2+\varepsilon^2)^{3/2} I_{xx}\,\sigma(\cdot) + 2I_x\,\sigma(\cdot)\sqrt{I_x^2+I_y^2+\varepsilon^2}\,(I_x I_{xx} + I_y I_{yx}) + I_{xx}\,\sigma^2(\cdot)\,(I_x^2+I_y^2+\varepsilon^2)^{3/2} - 2I_x\,\sigma^3(\cdot)\,(I_x I_{xx} + I_y I_{yx})\sqrt{I_x^2+I_y^2+\varepsilon^2}\Big] \Big/ (I_x^2+I_y^2+\varepsilon^2)^3$$

$$\frac{\partial}{\partial y} L_{I_y} = \Big[-(I_x^2+I_y^2+\varepsilon^2)^{3/2} I_{yy}\,\sigma(\cdot) + 2I_y\,\sigma(\cdot)\sqrt{I_x^2+I_y^2+\varepsilon^2}\,(I_x I_{xy} + I_y I_{yy}) + I_{yy}\,\sigma^2(\cdot)\,(I_x^2+I_y^2+\varepsilon^2)^{3/2} - 2I_y\,\sigma^3(\cdot)\,(I_x I_{xy} + I_y I_{yy})\sqrt{I_x^2+I_y^2+\varepsilon^2}\Big] \Big/ (I_x^2+I_y^2+\varepsilon^2)^3$$

Gradient $\nabla E = -L_I + \frac{\partial}{\partial x} L_{I_x} + \frac{\partial}{\partial y} L_{I_y}$; Gradient descent PDE: $I_t = -\nabla E$

$$I_t = \frac{\begin{aligned} &(I_x^2+I_y^2+\varepsilon^2)^{3/2} I_{xx}\,\sigma(\cdot) + 2I_x\,\sigma(\cdot)\sqrt{I_x^2+I_y^2+\varepsilon^2}\,(I_x I_{xx} + I_y I_{yx}) \\ &+ I_{xx}\,\sigma^2(\cdot)\,(I_x^2+I_y^2+\varepsilon^2)^{3/2} - 2I_x\,\sigma^3(\cdot)\,(I_x I_{xx} + I_y I_{yx})\sqrt{I_x^2+I_y^2+\varepsilon^2} \\ &-(I_x^2+I_y^2+\varepsilon^2)^{3/2} I_{yy}\,\sigma(\cdot) + 2I_y\,\sigma(\cdot)\sqrt{I_x^2+I_y^2+\varepsilon^2}\,(I_x I_{xy} + I_y I_{yy}) \\ &+ I_{yy}\,\sigma^2(\cdot)\,(I_x^2+I_y^2+\varepsilon^2)^{3/2} - 2I_y\,\sigma^3(\cdot)\,(I_x I_{xy} + I_y I_{yy})\sqrt{I_x^2+I_y^2+\varepsilon^2} \end{aligned}}{(I_x^2+I_y^2+\varepsilon^2)^3}$$

where $\sigma(\cdot) = \sigma\left(\sqrt{I_x^2+I_y^2+\varepsilon^2}\right)$

Figure 4: Deriving PDE for Sigmoid Penalty

Stability analysis :-

$$c(\|\nabla I\|) = \log \|\nabla I\|$$
$$= \log \sqrt{I_x^2 + I_y^2 + \varepsilon^2}$$

$$c'(\|\nabla I\|) = \frac{1}{\sqrt{I_x^2 + I_y^2 + \varepsilon^2}}$$

$$I_t = \nabla\left(\frac{c'(\|\nabla I\|)}{\|\nabla I\|} \cdot \nabla I\right)$$

$$= \nabla\left(\frac{\nabla I}{(I_x^2 + I_y^2 + \varepsilon^2)}\right)$$

$$= \nabla\left(b \; \nabla I\right) \qquad \text{where } b = \frac{1}{I_x^2 + I_y^2 + \varepsilon^2}$$

$$= b \Delta I$$

Stability criterion : $b \Delta t \leq \frac{1}{4} \Delta x^2$

$$\Delta t \leq \frac{1}{4} \underbrace{(I_x^2 + I_y^2 + \varepsilon^2)}_{} \Delta x^2$$

worst case : $I_x = 0$, $I_y = 0$.

$$\Delta t \leq \frac{1}{4} \varepsilon^2 \Delta x^2$$

$$\boxed{\Delta t \leq \frac{\varepsilon^2 \Delta x^2}{4}}$$ CFL condition.

Figure 5: CFL Condition for Logarithmic Penalty

Stability Analysis of Linear Heat Equation with data fidelity term

Gradient descent PDE: $I_t = -\nabla E$

$$\boxed{I_t = -\lambda(I - I_0) + I_{xx}(1-\lambda) + I_{yy}(1-\lambda)}$$

Discretize the PDE:-

$$I_t = \frac{I(x,y,t+\Delta t) - I(x,y,t)}{\Delta t} \quad \text{: Forward time}$$

$$\left.\begin{aligned} I_{xx} &= \frac{I(x+\Delta x, y, t) - 2I(x,y,t) + I(x-\Delta x, y, t)}{\Delta x^2} \\ I_{yy} &= \frac{I(x, y+\Delta y, t) - 2I(x,y,t) + I(x, y-\Delta y, t)}{\Delta y^2} \end{aligned}\right\} \text{centered time}$$

$$\frac{I(x,y,t+\Delta t) - I(x,y,t)}{\Delta t} = -\lambda(I(x,y,t) - I_0(x,y)) + (1-\lambda)\left[\frac{I(x+\Delta x,y,t) - 2I(x,y,t) + I(x-\Delta x,y,t)}{\Delta x^2}\right] + (1-\lambda)\left[\frac{I(x,y+\Delta y,t) - 2I(x,y,t) + I(x,y-\Delta y,t)}{\Delta y^2}\right]$$

$$I(x,y,t+\Delta t) = I(x,y,t)\left\{1 - \lambda\Delta t - \frac{2(1-\lambda)\Delta t}{\Delta x^2} - \frac{2(1-\lambda)\Delta t}{\Delta y^2}\right\} + (1-\lambda)\Delta t\left\{\frac{I(x+\Delta x,y,t) + I(x-\Delta x,y,t)}{\Delta x^2} + \frac{I(x,y+\Delta y,t) + I(x,y-\Delta y,t)}{\Delta y^2}\right\}$$

Update equation

To do Vonn Neumann Analysis :- Taking DFT

$$I(\omega_x,\omega_y,t+\Delta t) = I(\omega_x,\omega_y,t)\left(1 - \lambda\Delta t - 2\Delta t(1-\lambda)\left[\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2}\right]\right) + \Delta t(1-\lambda)\left\{\frac{I(\omega_x,\omega_y,t)e^{j\Delta x\omega_x} + I(\omega_x,\omega_y,t)e^{-j\omega_x\Delta x}}{\Delta x^2} + \frac{I(\omega_x,\omega_y,t)\cdot e^{j\omega_y\Delta y} + I(\omega_x,\omega_y,t)e^{-j\omega_y\Delta y}}{\Delta y^2}\right\}$$

$$I(\omega_x,\omega_y,t+\Delta t) = I(\omega_x,\omega_y,t)\left(1 - \lambda\Delta t - 2\Delta t(1-\lambda)\left[\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2}\right]\right) + (1-\lambda)\Delta t\, I(\omega_x,\omega_y,t)\left\{\frac{2\cos\omega_x\Delta x}{\Delta x^2} + \frac{2\cos\omega_y\Delta y}{\Delta y^2}\right\}$$

$$= I(\omega_x,\omega_y,t)\left\{1 - \lambda\Delta t - 2(1-\lambda)\Delta t\left[\frac{(1-\cos\omega_x\Delta x)}{\Delta x^2} + \frac{(1-\cos\omega_y\Delta y)}{\Delta y^2}\right]\right\}$$

$$2\sin^2\frac{\omega_x\Delta x}{2} \qquad 2\sin^2\frac{\omega_y\Delta y}{2}$$

$$= I(\omega_x,\omega_y,t)\left\{\underbrace{1 - \lambda\Delta t - 4(1-\lambda)\Delta t\left[\frac{\sin^2\left(\frac{\omega_x\Delta x}{2}\right)}{\Delta x^2} + \frac{\sin^2\left(\frac{\omega_y\Delta y}{2}\right)}{\Delta y^2}\right]}_{\alpha(\omega)}\right\}$$

Figure 6: CFL Condition for Quadratic Penalty (1)

For stability : $|\alpha(\omega)| \leq 1$

$$\cancel{1} - \lambda\Delta t - 4(1-\lambda)\Delta t\left[\frac{\sin^2\left(\frac{\omega_x \Delta x}{2}\right)}{\Delta x^2}\cdot 2 + \frac{2\sin^2\left(\frac{\omega_y \Delta y}{2}\right)}{\Delta y^2}\right]\cdot 2 \leq \cancel{1}$$

$\omega_x \neq 0$

and $\Delta x \neq 0$, $\Delta y \neq 0$

max value $= \frac{1}{\Delta x^2} + \frac{1}{\Delta y^2}$

$$-\lambda\Delta t - 4(1-\lambda)\Delta t\left[\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right]2 \leq 1$$

Assuming $\Delta x = \Delta y$

$$-\Delta t\left[\lambda + \frac{4(1-\lambda)\cdot 4}{\Delta x^2}\right] \leq 1$$

$$\Delta t \geq -\frac{\Delta x^2}{\lambda\Delta x^2 + 8(1-\lambda)}$$

which is true for all Δt

- Now, $\alpha(w) \geq -1$

$$1 - \lambda\Delta t - 4(1-\lambda)\Delta t\left[\frac{\sin^2\left(\frac{\omega_x \Delta x}{2}\right)}{\Delta x^2}\cdot 2 + \frac{2\sin^2\left(\frac{\omega_y \Delta y}{2}\right)}{\Delta y^2}\right]\cdot 2 \geq -1$$

$$\Delta t\left[\lambda + \frac{8(1-\lambda)}{\Delta x^2}\right] \leq 2$$

$$\boxed{\Delta t \leq \frac{2\Delta x^2}{\lambda\Delta x^2 + 8(1-\lambda)}}$$

when $\Delta x = 1$

$$\boxed{\Delta t \leq \frac{2}{8-7\lambda}}$$

Figure 7: CFL Condition for Quadratic Penalty (2)

Stability Analysis for sigmoid Penalty :-

$$c(\|\nabla I\|) = \sigma(\|\nabla I\|) = \frac{1}{1+e^{-\|\nabla I\|}}$$

$$E(I) = \iint_{\Omega} c(\|\nabla I\|)\, dx\,dy$$

$$= \iint_{\Omega} \frac{1}{1+e^{-\|\nabla I\|}}\, dx\,dy$$

$$c'(\|\nabla I\|) = \sigma(\|\nabla I\|)\left\{1-\sigma(\|\nabla I\|)\right\}$$

$$I_t = \nabla\cdot\left(\frac{c'(\|\nabla I\|)}{\|\nabla I\|}\nabla I\right)$$

$$= \nabla\cdot\left(\frac{\sigma(\|\nabla I\|)\left\{1-\sigma(\|\nabla I\|)\right\}}{\|\nabla I\|}\cdot\nabla I\right)$$

$$= \nabla\cdot\left(\frac{e^{-\|\nabla I\|}}{\left(1+e^{-\|\nabla I\|}\right)^2\|\nabla I\|}\nabla I\right)$$

$$\nabla\left(\frac{e^{-\sqrt{I_x^2+I_y^2}}}{\left(1+e^{-\sqrt{I_x^2+I_y^2}}\right)^2\cdot\sqrt{I_x^2+I_y^2}}\nabla I\right)$$

$$b = \frac{e^{-\sqrt{I_x^2+I_y^2}}}{\left(1+e^{-\sqrt{I_x^2+I_y^2}}\right)^2\cdot\sqrt{I_x^2+I_y^2}}$$

$$I_t = \nabla(b\cdot\nabla I)$$

$$= b\,\Delta I \qquad \text{(linear heat eq}^{n}\text{)}$$

For 2D: stability condition was : $b\Delta t \le \frac{1}{4}\Delta x^2$

Figure 8: CFL Condition for Sigmoid Penalty (1)

$$\Delta t \le \frac{\left(1+e^{-\sqrt{I_x^2+I_y^2}}\right)^2 \sqrt{I_x^2+I_y^2}}{e^{-\sqrt{I_x^2+I_y^2}}} \cdot \frac{1}{4}\Delta x^2$$

worst case $I_x=0$, $I_y=0$, $\Delta t \le 0$ Not possible

So, we add a ε^2 in energy functional.

$$E(I) = \iint_{\Omega} c(\|\nabla I\|)\, dx\,dy$$

$$= \iint_{\Omega} \frac{1}{1+e^{-\|\nabla I\|}}\, dx\,dy = \iint_{\Omega} \frac{1}{1+e^{-\sqrt{I_x^2+I_y^2+\varepsilon^2}}}\, dx\,dy$$

$$c'(\|\nabla I\|) = \sigma(\|\nabla I\|)\{1-\sigma(\|\nabla I\|)\}$$

$$I_t = \nabla\cdot\left(\frac{c'(\|\nabla I\|)}{\|\nabla I\|}\nabla I\right)$$

$$= \nabla\cdot\left(\frac{\sigma(\|\nabla I\|)\{1-\sigma(\|\nabla I\|)\}}{\|\nabla I\|}\cdot\nabla I\right)$$

$$= \nabla\cdot\left(\frac{e^{-\|\nabla I\|}}{\left(1+e^{-\|\nabla I\|}\right)^2\|\nabla I\|}\nabla I\right)$$

$$\nabla\left(\frac{e^{-\sqrt{I_x^2+I_y^2+\varepsilon^2}}\,\nabla I}{\left(1+e^{-\sqrt{I_x^2+I_y^2+\varepsilon^2}}\right)^2\sqrt{I_x^2+I_y^2+\varepsilon^2}}\right)$$

$$b = \frac{e^{-\sqrt{I_x^2+I_y^2+\varepsilon^2}}}{\left(1+e^{-\sqrt{I_x^2+I_y^2+\varepsilon^2}}\right)^2\sqrt{I_x^2+I_y^2+\varepsilon^2}}$$

$$I_t = \nabla(b\cdot\nabla I)$$

$$= b\Delta I \quad \text{(linear heat eq}^n\text{)}$$

For 2D: stability condition was: $b\Delta t \le \frac{1}{4}\Delta x^2$

Figure 9: CFL Condition for Sigmoid Penalty (2)

Stability Analysis of 2D linear Heat equation (without data fidelity term)

$I_t = b(I_{xx} + I_{yy})$

using forward time centered space difference :-

$$\frac{I(x,y,t+\Delta t) - I(x,y,t)}{\Delta t} = b\left[\frac{I(x+\Delta x,y,t) - 2I(x,y,t) + I(x-\Delta x,y,t)}{\Delta x^2} + \frac{I(x,y+\Delta y,t) - 2I(x,y,t) + I(x,y-\Delta y,t)}{\Delta y^2}\right]$$

$$I(x,y,t+\Delta t) = I(x,y,t) + \frac{\Delta t b}{\Delta x^2}\left[I(x+\Delta x,y,t) - 2I(x,y,t) + I(x-\Delta x,y,t)\right] + \frac{\Delta t b}{\Delta y^2}\left[I(x,y+\Delta y,t) - 2I(x,y,t) + I(x,y-\Delta y,t)\right]$$

Taking 2D DFT:

$$I(\omega_x,\omega_y,t+\Delta t) = I(\omega_x,\omega_y,t) + b\frac{\Delta t}{\Delta x^2}\left[I(\omega_x,\omega_y,t)e^{j\omega_x\Delta x} - 2I(\omega_x,\omega_y,t) + I(\omega_x,\omega_y,t)e^{-j\omega_x\Delta x}\right] + b\frac{\Delta t}{\Delta y^2}\left[I(\omega_x,\omega_y,t)e^{j\omega_y\Delta y} - 2I(\omega_x,\omega_y,t) + I(\omega_x,\omega_y,t)e^{-j\omega_y\Delta y}\right]$$

$$= I(\omega_x,\omega_y,t)\underbrace{\left[1 - \frac{2\Delta t}{\Delta x^2}b(1-\cos\omega_x\Delta x) - \frac{2\Delta t b}{\Delta y^2}(1-\cos\omega_y\Delta y)\right]}_{\alpha(\omega_x,\omega_y)}$$

Amplification factor: $\alpha(\omega_x,\omega_y)$

for stability $|\alpha(\omega_x,\omega_y)| \le 1$

or $-1 \le \alpha(\omega_x,\omega_y) \le 1$

$$1 - \frac{2\Delta t}{\Delta x^2}b(1-\cos\omega_x\Delta x) - \frac{2\Delta t b}{\Delta y^2}(1-\cos\omega_y\Delta y) \ge -1$$

$$\text{or } \frac{b2\Delta t}{\Delta x^2}\underbrace{(1-\cos\omega_x\Delta x)}_{\text{max value}=2} + \frac{b2\Delta t}{\Delta y^2}\underbrace{(1-\cos\omega_y\Delta y)}_{\text{max value}=2} \le 2$$

Assuming $\Delta x = \Delta y$

$$b\left[\frac{2\Delta t}{\Delta x^2}\cdot 2 + \frac{2\Delta t}{\Delta x^2}\cdot 2\right] \le 2$$

$$8\frac{\Delta t b}{\Delta x^2} \le 2$$

$$\boxed{b\Delta t \le \frac{1}{4}\Delta x^2}$$

23/24

Figure 10: CFL Condition for Linear Heat Equation (Without Data Fidelity Term)

5 Experimental Results

In order to obtain the experimental results, an image of a can was used. The image was introduced with Gaussian noise as well as salt pepper noise. The diffusion process for each of the 3 cases was implemented as results were obtained.

The Peak Signal-to-Noise ratio, Mean Squared Error and the Structural Similarity score for each of the diffusion functions was obtained and is summarized in Table 1.

Table 1: Comparison of diffusion functions

Diffusion Function	**Mean Squared Error (MSE)**	**Peak Signal to Noise Ratio (PSNR)**	**Structural Similarity Score (SSIM)**
$\log \|\|\nabla(I)\|\|$	65.3609	0.9702	29.9776
$\frac{1}{1+e^{-\|\|\nabla(I)\|\|}}$	69.7805	0.9701	29.6935
$\frac{\|\|\nabla(I)\|\|^2}{2}$	78.3914	0.9697	29.1881

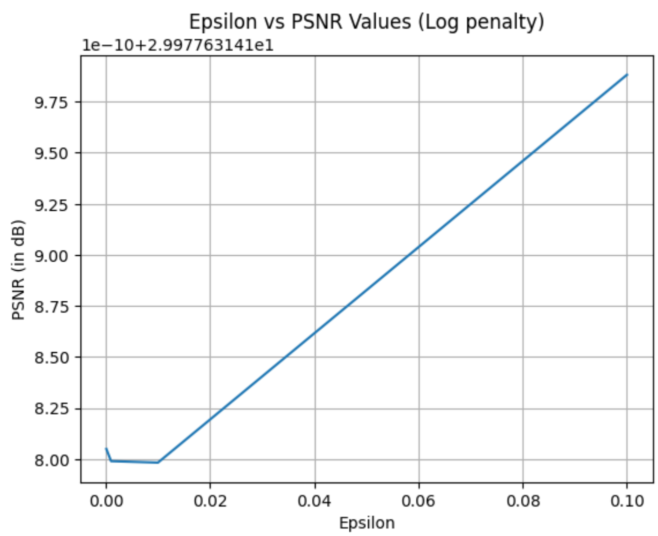

Figure 11: Effect of increasing epsilon on PSNR values (log penalty)

Figure 12: Denoised images obtained upon varying values of epsilon (log penalty)

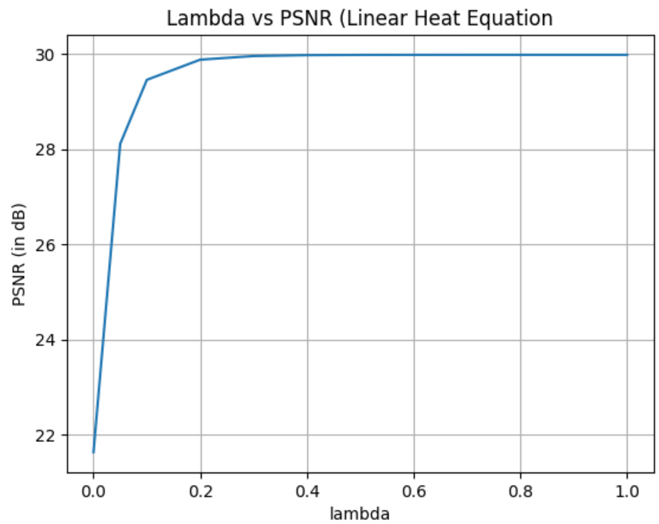

Figure 13: Effect of Lambda on PSNR Values (Linear Heat Equation)

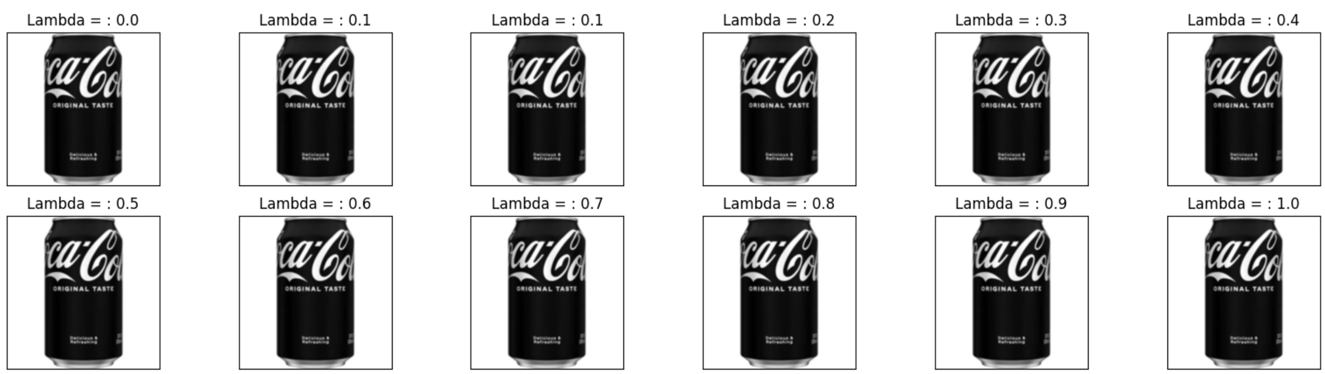

Figure 14: Denoised images obtained by varying lambda in data fidelity term of linear heat equation

6 Conclusions and Discussion

It was observed that the quadratic penalty was the most aggressive for higher values of gradient and hence had more impact on the edges of the image. On the other hand, the logarithmic penalty had a higher edge preserving effect compared to the quadratic penalty. It was interesting to see that the sigmoid function, which becomes almost constant for higher values of gradient, achieved comparable results to the other 2 cases. The sigmoid function has a derivative for values near the gradient and was more effective to remove noise in flat areas with minimal noise effect. However, for more stronger edges, it would not be a suitable choice since it does not have a harsh penalty for higher gradient values. It was also observed that the addition of the data fidelity term leads to the need of higher number of iterations to achieve a smooth image.